When the light shines too much: Rational inattention and pandering

Federico Trombetta*
Department of Economics
The University of Warwick

August 13, 2018

First version: April 13, 2017

Abstract

Should voters always pay attention to politics? I explore the role of endogenous costly attention allocation in politics, combining insights from the growing literature on rational inattention with a standard model of political agency. I show that, for any given level of attention, in equilibrium the voters will focus too much attention on actions and too little on the state of the world, with respect to the optimum, despite the complementarity between the two. This in turn induces too much political pandering. Moreover, a reduction in the total cost of attention will not correct this inefficiency and can even reduce welfare. This model can be a demand-driven explanation of the under-provision of analytical contents by news channels.

Keywords: Pandering, rational inattention, political agency, social media, populism.
JEL Classification: D72, D78

*F.Trombetta@warwick.ac.uk. ESRC support is gratefully acknowledged. I thank my supervisors, Ben Lockwood and Helios Herrera, for their invaluable help and advice. I also thank Dan Bernhardt, Michela Boccola, Guido Merzoni, Pietro Ortoleva, Ayush Pant, Giacomo Ponzetto and Federico Vaccari, and participants at the 2017 Warwick Political Economy & Public Economics Research Day, 2018 Royal Economic Society Conference and 2018 RES Symposium of Junior Researchers. Previous versions of this paper circulated under the title "When the light shines too much. Rational inattention, populism and pandering". All mistakes are mine.

"Contemporary media and the new media in particular are facilitating the growth of populism".

Martinelli (2016)

1 Introduction

1.1 Motivation

Despite concrete doubts over the effectiveness of a wall on the USA-Mexico border, both in terms of reducing illegal immigration and fighting drug cartels[1], President Donald Trump seems to insist on this political choice. One possible explanation is that it strengthens in the voters the idea of a President aligned with America's interests, something that the voters would notice and appreciate. At the same time, maybe, those voters are not paying too much attention to the effectiveness of the wall itself.

When 24 hour news channels and social media provide constant access to political news, what is the role of voters' attention in shaping political decisions? Are voters paying the right amount of attention to politics? And are they paying attention to the right elements? What are the consequences of inefficient allocation of attention? At first, it seems obvious that voters should be motivated to pay as much attention as possible to politics, so that they make better choices and elect better politicians. Moreover, tools that make attention "cheaper" should have an unambiguously positive impact. I show that reality is more complex: the decision about how much attention to pay to politics and how to allocate this attention is non trivial, with equilibrium incentives that motivate voters to pay too much attention to what politicians do (e.g. building a wall on the USA-Mexico border) and too little attention to what politicians *should* do (e.g. given the way illegal immigration and drug dealing work, is the wall the most effective way to achieve the objective?). This is consistent with the media studies literature that looks at the increasingly important news channels and finds that "80-85 per cent of news stories on the news channels contain no context or analytical content at all"[2] Moreover, if news on political actions are easier to put in a headline than an in depth analysis of the cost benefits of a policy, those results are consistent with the fact that many Americans seem to be just headlines readers.[3]

Voters' attention may motivate politicians to do what voters want, irrespective of what is right given the "state of the world", and voters may be paying too much attention to some aspects of the political process and too little to others. Given those possible suboptimal allocations,

[1] See e.g. Driver (2018).
[2] Lewis et al. (2005), p. 470.
[3] Cillizza (2014).

policies intended to improve citizen's attention to politics are actually making things worse.

1.2 Contribution and main results

This is the first paper in political economy to look at the effects of voters rational inattention on pandering in a political agency framework. The model builds on the literature on pandering, under which the politician chooses the action that guarantees his re-election, rather than the optimal one, e.g. Fox 2007, Maskin and Tirole 2004).[4] In that context, this paper endogenizes voters' level of attention to politics, following the recent and growing literature on the political consequences of rational inattention (Matejka and Tabellini 2016, Prato and Wolton 2016 and 2017). On top of that, it takes into account that attention to politics is a manifold concept. Therefore, this paper endogenizes both the overall amount of attention to politics and the way attention is distributed between actions and the state of the world, studying the interaction between the two.

I find novel results based on the fact that the inability to pre-commit to a certain level of attention, together with the direction of the relationship between different types of attention and the politicians' actions, can create important and unexplored trade-offs.

The model shows that voters' equilibrium allocation of attention is generically suboptimal. For any level of attention to politics, we pay too much attention to what politicians are doing and too little to what they should do. There is an equilibrium tendency to under-evaluate the disciplining effect of attention to the state and the potential strategic complementarity between the two. When selection is the only thing that matters, knowing the state on top of the action buys nothing to the voter, in expectation, because she would choose a more nuanced re-election rule, whose politician's best response will make her indifferent more often.[5] On the other hand, the knowledge of state and actions exhibits an important complementarity in disciplining the incumbent. Complementarity that is ignored in the equilibrium allocation of attention. As a consequence of both those inefficiencies, politicians have incentives to exert too much pandering, hurting the voters' welfare. When the total level of attention is also endogenous, an exogenous reduction in the total cost of attention is likely to make things worse, overall, because it induces an even higher level of over-attention to action, and rewards more pandering, without improving the level of attention to the state. I show that, when everything is endogenous (i.e. the total

[4] Pandering can be a good theoretical approximation of the basic essence of populism. This interpretation is not new in economics, e.g. Frisell (2009).

[5] In practice, when the voter observes only the action, she is never indifferent and she confirms the incumbent whenever she sees a certain action. When she observes the state as well, she makes use of this information confirming the incumbent only when she sees the right action and the right state. In all the other cases, she either chooses the challenger or she is indifferent, which is equivalent in terms of period 2 expected utility. When the voter chooses how to allocate attention, however, those two expectations are the same, hence it is worthless to pay any cost to learn the state.

level of attention and the way it is allocated), in equilibrium there is always under-attention to the state of the world and I find sufficient conditions for over-attention to action. Interestingly, when the strategic complementarity of the two types of attention is sufficiently important, the ex ante optimal choice may involve more attention to both the state and the action.

1.3 Suggestive evidence

My results are consistent with the in-depth content analysis of 24 hours news channels in Lewis et al. (2005). They show that "the news channels do not use the enormous time available to them to provide the viewer with a deeper understanding of the news world".[6] One of their case studies is particularly important in terms of attention to the actions vs to the state of the world: "all three news channels covered the governments defence review on 21 July, and yet none did so in the context of a broader analysis of the changing rationale behind military spending, or even by clarifying what that rationale might be".[7] Consistent with the point I make here, all the attention was focused on the government's action (the defence review) and none on the state of the world, i.e. the "rationale behind military spending". Lewis et al. (2005) find an effect on the supply side, but it may well be driven by the demand side allocation of attention by the voters which follows the inefficiencies highlighted in the present paper.

1.4 Application

Those results help to explain the relationship between social media and populism. Interestingly, politicians advocating populist policies are over-represented on social media[8], social media usage is highly correlated with political engagement[9] and newspaper articles are asking whether social media is "empowering populism"[10]. In the framework of this paper, social media reduces the cost of paying attention to politics and politicians, generating more attention[11] but also inducing more pandering (which has a negative effect on voters' welfare), that can be seen as a good proxy for populism. The overall welfare effect depends on how damaging pandering is, but there are parameters where a decrease in the cost of attention has negative welfare consequences overall. Hence, social media can lead to too much attention to what politicians do, too little analysis of the state of the world, and too much populism.

[6] Lewis et al. (2005), p. 473

[7] Lewis et al. (2005), p. 471.

[8] Extreme Tweeting, The Economist, 19th November 2015.

[9] Wihbey (2015) and Barberá (2018) (although evidences are somehow mixed).

[10] Is social media empowering Dutch populism? Financial Times, 14th March 2017.

[11] According to the Pew Research Center, the 2016 US presidential campaign was characterized by very high levels of interest, across parties and age groups (Pew Research Center 2016).

1.5 Related literature

This paper contributes to the growing political economy literature on the effects of voters' cognitive biases on decision making and political outcomes. This is the first contribution that looks at the effect of rational inattention on pandering. As such, it is at the intercept of several different branches of the political economy literature.

First, it is related to the literature on pandering and political agency (Canes-Wrone et al. 2001, Maskin and Tirole 2004, Fox 2007, Besley 2006, Morelli and van Weelden 2013), from which it borrows the basic structure of the model.

Second, it is related to the growing literature on the consequences of behavioural patterns on political choice (Ashworth and Bueno de Mesquita 2014, Levy and Razin 2015, Ortoleva and Snowberg 2015, Alesina and Passitelli 2017, Glaeser and Ponzetto 2017, Lockwood 2016, Lockwood and Rockey 2016).

Third, it relates to the small but growing literature on rational inattention (Sims 2003) and its interaction with political economy (Matejka and Tabellini 2016). In particular, Prato and Wolton (2016) and (2017) look rational inattention in a political agency set up.

Of these, Prato and Wolton (2016) highlight the importance of voters' attention choice in determining political behaviour in a model of competition and reforms. They find a "curse of interested voter", showing that too much interest in the political outcome would harm the voters in equilibrium. This paper differs from them in three important aspects: first of all, their main focus is on the distinction between interest (exogenous) and attention (endogenous), and the voter there does not choose "over-attention" as an equilibrium behaviour.[12] This paper is different as it shows that, in models of pandering, attention to the action is generically too high when compared with the welfare maximizing one. In terms of the populist interpretation, in this model the correlation between attention and populism is positive, while it is negative in theirs. Second, this paper endogenizes attention and its allocation between actions and the state of the world, with important implications in terms of optimality. Finally, theirs is a model of pre-electoral political competition, while this is the first one to look at rational inattention in the framework of standard principal-agent literature on pandering. An interesting corollary is that while the fact that attention can be bad for the voter is true in both frameworks, the implications of the relationship between equilibrium attention and populism are very different.

The mechanics of the paper are somewhat similar to Lockwood (2016). Both papers look at behavioural consequences in a set up of political agency and pandering. However, Lockwood

[12] When the voter would like to put too much attention, the candidates respond by choosing similar platforms, hence reducing the scope for attention.

(2016) studies confirmation bias and I study rational inattention.

Finally, attention is of course related with transparency (e.g. Prat 2005; recent work by Wolton 2017 and Andreottola 2017 highlights the potentially welfare depressing effect of "too informative" media). My focus is on the attention choice from the voter's side of the problem, hence highlighting a different and unexplored channel.

The remainder of this paper is as follows: section 2 introduces the model in its most generic form. Section 3 presents the different results endogenizing attention to the action in 3.1, the allocation of attention between actions and state in 3.2 and both the level and the allocation of attention in 3.3. Section 4 concludes.

2 The Model

2.1 Set up

I use the models of pandering from Fox (2007) and Besley (2006).

The game is a standard two periods political agency model with two players: an incumbent politician P (he) and a representative voter V (she). There is a binary state of the world $s_t = \{A, B\}$, known ex ante by P but not by V (she can decide to pay attention, learning it with some probability), with common prior $Pr(s_t = A) = \frac{1}{2}$.[13] The action space for the politician is binary as well, with $x_t \in \{A, B\}$. The action will be observed by V with some probability, which I am going to endogenize using the choice of attention.

The politician can be of two types, Consonant (C) and Dissonant (D); formally, $\theta \in \{C, D\}$ with $Pr(\theta = C) = \pi$. The type is private information of the politician, while the prior is common knowledge. In terms of payoff, every type of politician derives some utility E from being in office. Moreover, the Consonant incumbent gains u_t when he matches the action with the state of the world. Formally, when in office, a type C incumbent gets $u_t + E$ if $x_t = s_t$, E if $x_t \neq s_t$. A dissonant incumbent is biased toward a particular action (I assume it is A without loss of generality), irrespective of the state of the world. Formally, when in office, D gets $u_t + E$ if $x_t = A$ and E if $x_t \neq A$. It is assumed that both types get 0 when out of office.

The part of P's utility defined by u_t is private information of the politician. Ex ante, it is distributed according to a cumulative density function F with support $[0, \bar{u}]$. F is continuous and strictly increasing in the whole interval, its probability density function is f.

Assumption 1 *F is uniform, hence* $E[u_t] = \frac{\bar{u}}{2}$.

[13] The flat prior guarantees that there is not a popular state of the world, hence the incentives to pander comes only from the equilibrium decision on the level of separation between types of incumbent.

The uniformity assumption is for convenience, as it simplifies the concavity conditions of the ex ante objective function. Moreover, consistently with Fox (2007) and Besley (2006), it is assumed that $\bar{u} > \delta(\frac{\bar{u}}{2} + E)$, so that every type of politician plays both actions with positive probability. Finally, V gets a utility of 1 if $x_t = s_t$ and 0 otherwise, and there is a common discount factor δ.

So far, the model is unchanged with respect to the standard political agency and pandering literature. I build on this framework by endogenizing the choice of attention. As in Prato and Wolton (2016 and 2017),[14] I model rational inattention as the probability q and β that the voter observes the action of the politician and the state of the world respectively before casting her vote. In particular, $q \in [0,1]$ ($\beta \in [0,1]$) is the amount of attention that the voter pays to the incumbent's action (state of the world) and hence the probability of observing the actual action (state).

Formally, the voter does not observe x_t or s_t directly, but rather $\tilde{x}_t$ and $\tilde{s}_t$, and she chooses q and β at a cost $\tau C(q,\beta)$, where $Pr(\tilde{x}_t = x_t) = q$, $Pr(\tilde{x}_t = \emptyset) = 1 - q$, $Pr(\tilde{s}_t = s_t) = \beta$, $Pr(\tilde{s}_t = \emptyset) = 1 - \beta$, $\tau \geq 0$. In other words, the higher is the (costly) attention q (β), the more likely is the voter to observe the action chosen by the politician (state of the world), and to use this information in her belief updating. The cost of attention function is strictly increasing and convex, and satisfies the usual Inada-type conditions. Formally, $C_y^{'}(0,j) = 0$, $C_y^{'}(1,j) = \infty$, $C_y^{'}(y,j) > 0 \ \forall y,j \in [0,1]$ and $C_{yy}^{''}(y,j) \geq 0 \ \forall y,j \in [0,1]$, where of course $C_y^{'}(y,j)$ is the partial derivative with respect of the first argument of the cost function[15].

2.2 Timing

I assume that the voter chooses the attention level after the incumbent's action, but without knowing that.[16] Hence, the timing used to derive the equilibrium attention choice is as follows:

1. Nature chooses θ, u_1, s_1 (private info of P);
2. P chooses x_1;
3. V chooses q and β at a cost $\tau C(q,\beta) \geq 0$;
4. V observes $\tilde{x}_1$, $\tilde{s}_1$ and votes;
5. Period 1 ends and payoffs are paid;

[14]Note that they have attention to the action only.

[15]I will impose additional restrictions whenever they will be necessary for the tractability of the problem

[16]Formally, it is the same as letting them choose at the same time. It basically assumes that the voter cannot commit, ex ante, to a certain level of attention.

6. Degenerate Period 2 (with the exception of the draw of θ if the incumbent has been re-elected) but no elections;[17]

Note that, crucially, payoffs are paid after the elections.

Of course, the ex ante optimal welfare is the one calculated assuming that the voter could commit to a certain level of attention, known to the politician when making his choice. The solution concept I use is the perfect Bayesian Nash equilibrium (PBNE).

3 Results

To better understand the effect of different types of attention on the political equilibrium, their interaction and the role of attention allocation, I present the results step by step. In particular, Section 3.2 considers the benchmark case where V can decide how much attention to pay to the incument's action only, i.e. she cannot learn the state. In section 3.2 I study the optimality of attention allocation by allowing V to choose attention to the state as well, but with the constraint that the sum of q and β has to be equal to a given constant. Finally, in section 3.3 both the total amount of attention and the allocation are endogenous. In all these cases, I compare the equilibrium level of attention with the ex ante optimal one.

3.1 Benchmark: attention to actions only

3.1.1 Features of the equilibrium

In this subsection, I assume $\beta = 0$ and only look at the optimal allocation of q. The game is solved by backward induction. Proposition (1) summarizes some features of the essentially unique equilibrium that arises, noting that this does not depend on whether pre-commitment to a certain level of attention is possible.[18]

Proposition 1 *In every PBNE of this game, irrespective of the possibility of commitment,*

1. $Pr(re-elect|\tilde{x}=A)=0$ *and* $Pr(re-elect|\tilde{x}=B)=1$*;*
2. *C incumbent chooses* $x=B$ *when* $s=B$ *with probability 1 and* $x=B$ *when* $s=A$ *if* $u \leq \delta\left(\frac{\bar{u}}{2}+E\right)q$*;*
3. *D incumbent chooses* $x=B$ *when* $u \leq \delta\left(\frac{\bar{u}}{2}+E\right)q$*;*
4. $Pr(re-elect|\tilde{x}=\emptyset)$ *can be anything;*

[17] It is important to assume that a new state of the world is drawn in period 2. Otherwise, having a bad incumbent when the state is A would be different than having a bad incumbent when the state is B.

[18] Proofs are in Appendix A.

Intuitively, the dissonant incumbent is comparatively more likely to choose action A than the consonant one, hence the voter rewards the choice of action B. This re-election strategy gives incentives to pander, i.e. to choose an action that is different from the one the politician prefers in the short term.[19] There are multiple equilibria but they are payoff-equivalent. This set of equilibria is very similar to the one derived in models without rational inattention, but Proposition 1 is useful because it leads to the following corollary (whose proof is obvious, hence omitted).

Definition 1 *Define "pandering" as the probability* $\gamma = F[\delta\left(\frac{\bar{u}}{2} + E\right)q]$ *that an incumbent chooses the period 1 action that is suboptimal from his point of view.*

Corollary 1 *Irrespective of the availability of commitment, an increase in attention (to the action) increases the probability of pandering.*

This result is intuitive, since the point of pandering is to guarantee re-election, choosing the "popular" action rather than the right one. Therefore, it is crucial that the voter observes this action.

3.1.2 Equilibrium attention choice

Define $V_C = \delta$ and $V_D = \delta\frac{1}{2}$ as the continuation payoff from having a C and a D politician in period 2. Moreover, define $\Gamma = \pi V_C + (1-\pi)V_D$ as the expected payoff in period 2 from electing the challenger, while γ is defined as in Corollary 1.
V uses the re-election strategy outlined in Proposition 1 and solves $max_{q\in[0,1]}W$, where

$$W = q\{[Pr(x = B)](\hat{\pi}V_C + (1-\hat{\pi})V_D) + [Pr(x = A)]\Gamma\} + (1-q)\Gamma - \tau C(q)$$

and $\hat{\pi} = Pr(\theta = C|x = B)$.[20] Intuitively, the voter observes the actual action with probability q. Upon observing action B, she will re-elect the incumbent, and her expected utility is given by $\hat{\pi}V_C + (1-\hat{\pi})V_D$. She elects the challenger when she observes $\tilde{x} = A$. With probability $1-q$ the voter observes $\emptyset$ and hence she is indifferent between re-electing or not.
From the point of view of P, $\gamma = F[\delta(\frac{\bar{u}}{2} + E)q^e]$, where q^e is the conjectured value of q chosen by the voter in the next stage. In equilibrium, of course, $q^e = q^*$, where q^* is the chosen level of attention that emerges as an equilibrium outcome in this version of the game. Proposition (2) defines this formally.

[19] In case of a consonant politician, this is always bad for the voters.

[20] W includes only future expected payoffs from P's actions. The full welfare function includes also present expected payoffs, but they do not matter for the determination of the optimal level of attention, given that they are "sunk" when attention plays its role.

Proposition 2 *The equilibrium level of attention q^* is unique, interior and implicitly defined by*

$$\pi(1-\pi)\delta(1-F[\delta(\frac{\bar{u}}{2}+E)q^*])=4\tau C^{'}(q^*) \quad (1)$$

Note that, if attention is costless (i.e. $\tau = 0$), then $q^* = +\infty$. This is because, in this case, the voter does not take into account the fact that q increases the probability of pandering. Hence, she is just considering the informative value of attention, and in an attention-for-free world she would choose the highest possible one.

3.1.3 Equilibrium Comparative Statics

In terms of comparative statics:

Corollary 2 *The equilibrium attention level is increasing in the uncertainty over the type of incumbent ($var(\theta)$) and decreasing in the office rent (E).*

Intuitively, the trade off here is just between the informational value of q and its cost. Uncertainty over θ increases the value of knowing the chosen action, while E, leaving everything else equal, boosts the probability of pandering, making the chosen action a poorer signal of the incumbent's type.

I also consider comparative statics on the equilibrium probability of pandering. Given that in equilibrium $q^e = q^*$, the equilibrium level of pandering is $\gamma^* = F[\delta(\frac{\bar{u}}{2}+E)q^*]$. Corollary 3 follows directly from Proposition 2.

Corollary 3 *The equilibrium level of pandering is increasing in the uncertainty over the incumbent's type and in the office rent*

Intuitively, uncertainty over θ increases the equilibrium level of attention, making pandering more profitable. E has a negative effect on q^*, hence a negative indirect effect on the probability of pandering and a positive direct effect on γ^*, making being in office more attractive, and the second one always dominates.

3.1.4 Ex ante optimal attention

The ex ante expected welfare is

$$
\begin{aligned}
W^{ex} &= \pi\{\frac{1}{2}(1+qV_C)+\frac{1}{2}[(1-\gamma)(1+q\Gamma)+\gamma qV_C]\}+ \\
&+(1-\pi)\{\frac{1}{2}[\gamma(1+qV_D)+(1-\gamma)q\Gamma]+ \\
&+\frac{1}{2}[\gamma qV_D+(1-\gamma)(1+q\Gamma)]\}+(1-q)\Gamma-\tau C(q)
\end{aligned} \tag{2}
$$

Intuitively, the consonant incumbent behaves well in state B, guaranteeing the right action. Moreover, he is re-elected when the action is observed (hence with probability q). In state A, the consonant incumbent will not pander with probability $1-\gamma$. This gives 1 to the voter, but the incumbent is voted out when the action is observed, and the challenger is elected. With probability γ, the consonant incumbent panders in state A, giving 0 to the voter but being re-elected and consequently giving V_C in period 2. The same logic applies to the dissonant incumbent. When the action is not observed, the voter is always indifferent between the challenger and the incumbent, so her expected welfare is just $(1-q)\Gamma$.
To find the optimal level of attention, the voter solves solves $max_{q\in[0,1]}W^{ex}$. Lemma 1 establishes that W^{ex} is strictly concave in q, therefore it is possible to look directly at the first order conditions.

Lemma 1 *When $\beta=0$, the ex ante voter's welfare function is strictly concave in q.*

Solving for the ex ante optimal level of attention and comparing it with the equilibrium one, I can highlight the first important result of the paper.

Proposition 3 *The ex ante optimal choice of attention exists and it is unique. Moreover, it is strictly smaller that the equilibrium choice q^*.*

Intuitively, when the voter chooses ex post, she over-evaluates q, because she does not take into account the effect of q on welfare-depressing pandering.[21] On the other hand, the ex ante welfare calculation takes this into account as well, leading to a lower level of q.
A first consequence is that attention pre-commitment can help the voter to obtain a better political outcome. A second one is a natural question: is it possible that a decrease in the cost of attention (captured by a decrease in τ) is overall welfare depressing, when no pre-commitment is available?

[21] As expected given the existing literature, the effect of q on ex ante welfare can be decomposed in three parts. A positive information availability effect, a negative direct effect of more pandering and a negative effect on selection induced by more pandering. In equilibrium, however, only the first effect is relevant for the choice of the voter, hence the over-attention result.

Proposition 4 *A decrease in τ is welfare depressing if*

$$\frac{\pi((1-\pi)\delta q^* + 2)\delta(\frac{\bar{u}}{2}+E)\frac{1}{\bar{u}}C'(q^*)}{\pi(1-\pi)\delta^2(\frac{\bar{u}}{2}+E)\frac{1}{\bar{u}} + 4\tau C''(q^*)} > C(q^*) \tag{3}$$

Equation (3) is more likely to hold when, at the equilibrium level of attention, the welfare cost imposed by the increased pandering is high. Furthermore, it is more likely to hold as $C'(q^*)$ increases with respect to $C(q^*)$.

Proposition 4 has important policy consequences, since it implies that cheaper access to politicians actions is not necessarily better for the voter. Interestingly, the channel through which this happens is pandering, i.e. a proxy for populism, boosted by the increased level of attention.

3.2 Fixed total level of attention

In this section the voter is allowed to allocate attention between the action and the state. To study inefficiencies in the allocation shutting down effects coming from the total choice of attention, I assume that the total amount of attention is fixed, exogenous and equal to α. Hence, the voter is now allowed to choose q and β with the constraint that $q + \beta = \alpha$.

Before looking at the allocation of attention, I summarize some generic features of the equilibrium. In particular, defining $r_{\tilde{x},\tilde{s}} = Pr(re - elect|\tilde{x}, \tilde{s})$, all the PBNE of the game can be summarized as follows.

Proposition 5 *In every equilibrium with two types of attention*

1. $\gamma_B = 1$, $\lambda_B = F[\delta(\frac{\bar{u}}{2} + E)q]$;

2. $\gamma_A = \lambda_A = F[\delta(\frac{\bar{u}}{2} + E)q(1 - \beta + \beta(r_{B,A} - r_{A,A}))]$.

3. $r_{B,\emptyset} = 1$, $r_{A,\emptyset} = 0$, $r_{B,B} = 1$, $r_{A,B} = 0$. $r_{B,A}$ *and* $r_{A,A}$ *are unconstrained because of indifference.* $r_{\emptyset,\tilde{s}}$ *can be anything.*

Proposition 5 contains several insights. First, there are multiple equilibria, depending on the randomization when the voter observes state A and the action. Second, the effect of q on pandering (i.e. on λ_A, λ_B and γ_A) is the same as before: more attention on the action makes pandering more desirable, for the politician. Finally, β has either no effect or a negative effect on the equilibrium level of pandering. Irrespective of the values taken by $r_{B,A}$ and $r_{A,A}$, $\frac{\partial \gamma_A}{\partial \beta} = \frac{\partial \lambda_A}{\partial \beta} \leq 0$. Again, this makes intuitive sense, since pandering can be punished if the voter learns that the politician is choosing the wrong action. For tractability, I make the following assumption.

Assumption 2 *When indifferent, the voter re-elects the incumbent with the same probability in every information set.*

This implies in particular that $r_{B,A} = r_{A,A}$.
Given the resluts of Proposition 5, it is now possible to focus on the equilibrium attention choice. The voter's problem is now to $max_{q,\beta}\tilde{W}$ subject to $q + \beta = \alpha \leq 1$, where

$$\begin{aligned}\tilde{W} &= q\beta\{Pr(x = B, s = B)(\hat{\pi}_{B,B}V_C + (1 - \hat{\pi}_{B,B})V_D) + Pr(x = A, s = B)\Gamma \\ &+ Pr(x = A, s = A)(r_{A,A}(\hat{\pi}_{A,A}V_C + (1 - \hat{\pi}_{A,A})V_D) + (1 - r_{A,A})\Gamma) \\ &+ Pr(x = B, s = A)(r_{B,A}(\hat{\pi}_{B,A}V_C + (1 - \hat{\pi}_{B,A})V_D) + (1 - r_{B,A})\Gamma) \\ &+ q(1 - \beta)\{Pr(x = B)(\hat{\pi}_{B,\emptyset}V_C + (1 - \hat{\pi}_{B,\emptyset})V_D) + Pr(x = A)\Gamma\} \\ &+ (1 - q)\Gamma\end{aligned}$$

Noticing that in equilibrium $\hat{\pi}_{A,A} = \hat{\pi}_{B,A} = \pi$, and substituting the constraint, this can be rewritten as

$$\begin{aligned}\tilde{W} &= q(\alpha - q)\{Pr(x = B, s = B)(\hat{\pi}_{B,B}V_C + (1 - \hat{\pi}_{B,B})V_D) \\ &+ (1 - Pr(x = B, s = B))\Gamma\} \\ &+ q(1 - (\alpha - q))\{Pr(x = B)(\hat{\pi}_{B,\emptyset}V_C + (1 - \hat{\pi}_{B,\emptyset})V_D) + Pr(x = A)\Gamma\} \\ &+ (1 - q)\Gamma\end{aligned} \tag{4}$$

Proposition 6 clarifies how attention is allocated, in equilibrium.

Proposition 6 *In equilibrium, the voter chooses* $q^{**} = \alpha$ *and* $\beta^{**} = 0$.

Intuitively, the point is simple. On the one hand, the observation of the state, together with the action, makes the voter more confident abut the quality of the incumbent she is re-electing. However, because of the way politicians adjust their behaviour, she also gets an expected utility of Γ every time she does not observe $\tilde{x} = B$ and $\tilde{s} = B$. On the other hand, when she observes only the action, she is more likely to keep the incumbent ($\tilde{x} = B$ is enough) but is less sure about his quality. Those two effects cancel out, in expectation. Hence, in equilibrium it is

pointless to allocate any attention to the state, as this buys you nothing in expectation, and all the attention is paid to the action.
In this set up, the ex ante welfare function is given by

$$\tilde{W}^{ex} = \frac{1}{2}[1 - F(\delta(\frac{\bar{u}}{2} + E)q(1 - (\alpha - q))) + \pi + (1 - \pi)F(\delta(\frac{\bar{u}}{2} + E)q)] + \tilde{W} \tag{5}$$

where the first part of the equation is simply the present part of the welfare function, given by the policy choices of the current incumbent. Lemma 2 establishes that the objective function is concave in q.

Lemma 2 *For every* $q \in [0, \alpha]$ $\frac{\partial^2 \tilde{W}^{ex}}{\partial q^2} \leq 0$.

From this I solve the optimization problem, giving the following result:

Proposition 7 *If* $\frac{1}{\bar{u}} > \frac{\pi(1-\pi)(1-F(\delta(\frac{\bar{u}}{2}+E)\alpha))}{\pi(1-\pi)\delta^2(\frac{\bar{u}}{2}+E)\alpha+2\delta(\frac{\bar{u}}{2}+E)(\pi+\alpha)}$ *then the ex ante optimal allocation of attention is such that* $\beta^{ex} > 0$ *and* $q^{ex} < \alpha$.

Intuitively, it may be that - even ex ante - it is optimal to allocate all the attention to the action. However, Proposition 7 establishes that while the extreme allocation always occurs in equilibrium, this may not be the case ex ante, if the positive effect of β in reducing pandering is taken into account and sufficiently big.

3.3 Endogenous level and allocation

In this section everything is endogenous, i.e. q and β are chosen at a cost $\tau C(q, \beta)$. The equilibrium behaviour in the rest of the game is still described by Proposition 5, so I can focus on the voter's attention choice (and allocation).
The maximization problem for the ex ante solution is more complex, because the two variables are unconstrained. It is necessary to impose additional conditions to guarantee the concavity of the ex ante objective function. I make some simplifying assumptions, although none of them are necessary for the existence and uniqueness of the equilibrium allocation of attention. I will discuss ways to relax them in Appendix B. I assume the following:

Assumption 3 $C(q, \beta) = q^2 + \beta^2$.

This assumption makes a closed form solution possible, and highlights the conditions for concavity of the ex ante maximization (it is not an issue in the equilibrium one).

Assumption 4 $4\tau \geq \delta \frac{1}{\bar{u}} \left(\frac{\bar{u}}{2} + E\right)$

Assumption 4 is sufficient to ensure that the Hessian matrix of the ex ante maximization problem is negative semidefinite. However, shown in Appendix B, it is possible to ensure enough concavity even when it is violated.

The solution of the model follows a path very similar to the previous section. The unique equilibrium, in particular, follows exactly the same logic as proposition 6, strengthened by the presence of attention costs. Defining for simplicity $K = \frac{1}{\bar{u}}\left(\frac{\bar{u}}{2} + E\right)$,[22] the following Lemma describes the equilibrium attention choice.

Lemma 3 *The allocation of attention in equilibrium is unique and is given by* $\beta^{***} = 0$ *and*

$$q^{***} = \begin{cases} \frac{\pi(1-\pi)}{8\tau+\pi(1-\pi)\delta K} & \text{if } \tau \geq \bar{\tau} \\ 1 & \text{if } \tau < \bar{\tau} \end{cases}$$

where $\bar{\tau} = \frac{\pi(1-\pi)(1-\delta K)}{8}$

The particular closed form of q^{***} is of course a feature of the assumptions.[23] In general it is the value of q that solves

$$\frac{1}{4}\pi(1-\pi)[1 - F(\delta\left(\frac{\bar{u}}{2} + E\right)q)] = \tau C_q^{'}(q, \beta)$$

In this case the ex ante welfare function is described by

$$\tilde{\tilde{W}}^{ex} = \frac{1}{2}\left[1 - F(\delta\left(\frac{\bar{u}}{2} + E\right)q(1-\beta)) + \pi + (1-\pi)F(\delta\left(\frac{\bar{u}}{2} + E\right)q)\right] +$$

$$+\left[\frac{1}{4}\pi(1-\pi)(1 - F(\delta\left(\frac{\bar{u}}{2} + E\right)q))\right]q + \Gamma + \tau C(q, \beta)$$

to be maximized with respect to q and β. As anticipated,

Lemma 4 *If Assumption 4 is true, then the Hessian matrix of the ex ante maximization is negative semidefinite.*

Thanks to Lemma 4, the ex ante optimal allocation of attention is given by the system of the two partial derivatives, looking at conditions for interior solutions. The main results are:

Proposition 8 *Comparing the ex ante optimal and the equilibrium level of attention:*

[22] By construction $\delta K < 1$.

[23] The quadratic cost violates the assumption that $C_q^{'}(1, \beta) = \infty$, that would guarantee an interior solution. Moreover, if $\delta K > \frac{1}{9}$, then $\bar{\tau} > \frac{1}{4}\delta\frac{1}{\bar{u}}\left(\frac{\bar{u}}{2} + E\right)$, i.e. lower bound of τ assumed in Assumption 4, hence q^{***} is strictly interior as well.

1. *The ex ante optimal level of attention to the state is always weakly higher than the equilibrium;*

2. *If* $1 - \pi - 2\delta K > 0$ *then the ex ante optimal* β *is strictly higher than the equilibrium;*

3. *If* $(1-\pi)\frac{\delta^2 K(1-4\tau\pi)}{2\tau} < \pi(1-\pi)^2 + 16\tau\delta$ *then the ex ante optimal level of attention to the action is strictly smaller than the equilibrium.*

Proposition 8 makes a very straightforward point about the optimal β: as long as parameters do not produce a corner solution where both q and β are zero (note that if the ex ante optimal q is zero then the ex ante optimal β is zero as well), then the equilibrium level of attention to the state is too small. The situation would not improve with a reduction in τ, as the allocation of attention would still be unbalanced toward attention to the action. On the contrary, the ex ante choice allows to take into account not only the beneficial role of attention to the state in reducing pandering without affecting selection (and the detrimental role in attention to the action), but also the reduction in bad pandering that comes from a higher β (i.e. the consonant politician choosing action B in state A: in equilibrium, this happens with probability $F(\delta\left(\frac{\bar{u}}{2}+E\right)q(1-\beta)))$ increasing the likelihood that q has a *positive* effect on present welfare as well, via its increase of the "good pandering" (i.e. the dissonant politician choosing action B in state B).

In other words, the two types of attention are complements, but this complementarity is ignored in the equilibrium choice of the voter, where instead attention to the state is irrelevant, in expectation (because of the different re-election rule and consequent best responses of the politician). The ex ante optimal q may be higher or lower than the equilibrium one, depending on whether the complementarity effect is sufficiently strong to overcome the negative effect of higher bad pandering.

4 Conclusion

Are voters allocating attention to politics optimally? I endogenize the attention choice in an otherwise standard model of pandering, highlighting two levels of inefficiency. The equilibrium level of attention to politicians' actions tends to be too high, making pandering too attractive. Conversely, attention to the state (i.e. to what politicians should do) is too low as its effect on selection is limited, even though it could improve the incumbent's behaviour. Moreover, there is a useful complementarity between attention to the action and to the state, with the latter reducing the consonant politician's incentive to pander.

The results suggest that trying to increase attention to the state by reducing the cost of attention to politics is likely to make things worse, as it will not change the allocation of attention. It would just make attention to the action even more extreme. The results highlight the importance of accounting for allocation of attention and its general equilibrium effect on political behaviour. In terms of applications, if populism is a form of pandering and social media is a technological innovation that reduces the cost of attention, this paper gives a rationale for the observed correlation between populism and social media. Moreover, it highlights a general tendency to care too much about what politicians are doing and too little about what they should do, that would be made worse by a reduction in the cost of attention.

The current very tractable set up is open to different extensions, suggesting many directions for future research. One would be a model with a multidimensional state of the world. Dimensions that are more salient for voters are likely to invite more pandering from the politician, something that the optimal attention allocation should take into account.
Another option is to microfound voters' attention in a different way. For example with a model of media entry where each outlet can decide whether to provide political news, sports, documentaries etc. If competition increases the provision of entertainment (where it is easier to differentiate the product), this would shift voters' attention away from hard news (or, equivalently, makes attention to politics more costly). The overall effect of this is ambiguous, as the model highlights: it reduces the politician's incentives to pander, but at the same time it worsens political selection.
Finally, it is interesting to study how voters' attention choice affects the way political campaigns are organized. Politicians obviously take into account the (costly) attention allocation of the voters, and may be tempted to focus on simpler messages or policies that can be easily explained. But those are precisely the policies where pandering incentives are stronger.

References

Alesina, A. and Passitelli, F. (2017) Loss Aversion in Politics. *Working paper.*

Andreottola, G. (2017) Flip-flopping and Electoral Concerns. *Working paper.*

Ashworth, S. and Bueno de Mesquita, E. (2014) Is Voter Competence Good for Voters? Information, Rationality, and Democratic Performance. *The American Political Science Review*, 108(3):565-587.

Barberá, P. (2018) The Consequences of Exposure to Disinformation and Propaganda in Online Settings. In Tucker, J. A. et al. (2018) *Social Media, Political Polarization and Political Disinformation*, Hewlett Foundation.

Besley, T. (2006) *Principled Agents? The Political Economy of Good Government.* New York: Oxford University Press.

Canes-Wrone, B., Herron, M. C. and Shotts, K. W. (2001) Leadership and Pandering: A Theory of Executive Policymaking. *American Journal of Political Science*, 45(3): 532-550.

Cillizza, C. (2014) Americans read headlines. And not much else. *washingtonpost.com.* [Accessed 01/08/2018].

Driver, A. (2018) Trump's Mexico wall would be a gift to the drug cartels. *Cnn.com.* [Accessed 01/08/2018].

Fox, J. (2007) Government transparency and policymaking. *Public Choice*, 131(1/2): 23-44.

Frisell, R. (2009) A theory of self-fulfilling political expectations. *Journal of Public Economics*, 93: 715-720.

Glaeser, E. L. and Ponzetto, G. A. M., (2017) Fundamental Errors in the Voting Booth. *working paper.*

Levy, G. and Razin, R. (2015) Correlation Neglect, Voting Behavior, and Information Aggregation. *The American Economic Review*, 105(4): 1634-1645.

Lewis, J., Cushion, S. and Thomas, J. (2005) Immediacy, Convenience or Engagement? An analysis of 24-hour news channels in the UK. *Journalism Studies*, 6(4), 461-477.

Lockwood, B. (2016) Confirmation Bias and Electoral Accountability. *Quarterly Journal of Political Science*, 11(4): 471-501.

Lockwood, B. and Rockey, J. (2016) Negative Voters? Electoral Competition with Loss-Aversion. *Working paper.*

Martinelli, A. (eds) (2016) *Beyond Trump. Populism on the Rise.* Novi Ligure: Edizioni Epoke'.

Maskin, E. and Tirole, J. (2004) The Politician and the Judge: Accountability in Government. *The American Economic Review*, 94(4): 1034-1054.

Matejka, F. and Tabellini, G. (2016) Electoral Competition with Rationally Inattentive Voters. *Working paper.*

Morelli, M. and Van Weelden, R. (2013) Ideology and Information in Policymaking. *Journal of Theoretical Politics*, 25(3): 412-439.

Mudde, C. and Kaltwasser, C. R. (2017) *Populism. A Very Short Introduction.* New York: Oxford University Press.

Ortoleva, P. and Snowberg, E. (2015) Overconfidence in Political Behavior. *The American Economic Review*, 105(2): 504-535.

Pew Research Center (2016) 2016 *Campaign: Strong Interest, Widespread Dissatisfaction.*

Prat, A. (2005) The Wrong Kind of Transparency. *The American Economic Review*, 95(3): 862-877.

Prato, C. and Wolton, S. (2016) The Voters Curses: Why We Need Goldilocks Voters. *American Journal of Political Science*, 60(3): 726-737.

Prato, C. and Wolton, S. (2017) Rational Ignorance, Populism, and Reform. *European Journal of Political Economy*, forthcoming.

Rodrik, D. (2017) Populism and the Economics of Globalization. *working paper.*

Sims, C. A. (2003) Implications of rational inattention. *Journal of Monetary Economics*, 50(3): 665-690.

Wihbey, J. (2015) How does social media use influence political participation and civic engagement? A meta-analysis. *journalistresource.org.* [Accessed 28/08/2017].

Wolton, S. (2017) Are Biased Media Bad for Democracy? *Working paper.*

Appendix A Proofs

Proof of Proposition 1.
In terms of existence, sequential rationality requires that $\hat{\pi}_A \leq \pi$ and $\hat{\pi}_B \geq \pi$, where $\hat{\pi}_{\tilde{x}}$ is the posterior probability of the incumbent being consonant when the voter observes $\tilde{x}$.
Define $r_{\tilde{x}} = Pr(re-elect|\tilde{x})$, $\gamma_s = Pr(C\ plays\ B\ in\ state\ s)$ and $\lambda_s = Pr(D\ plays\ B\ in\ state\ s)$. This equilibrium implies that, from the point of view of V, $\gamma_A = \lambda_A = \lambda_B = F[\delta\left(\frac{\bar{u}}{2}+E\right)q] := \gamma < 1$ and $\gamma_B = 1$.
By Bayes rule,

$$\hat{\pi}_A = \frac{0.5(1-\gamma)\pi}{0.5(1-\gamma)\pi + (1-\gamma)(1-\pi)} < \pi$$

$$\hat{\pi}_B = \frac{(0.5\gamma+0.5)\pi}{(0.5\gamma+0.5)\pi + \gamma(1-\pi)} > \pi$$

hence, from the point of view of V, it is sequentially rational to re-elect after observing B and to vote for the challenger after observing B.
Given this, one needs to check optimality from the point of view of P. Starting from the D type, and given the voter's re-election rules stated above, the expected utility of the dissonant incumbent when he chooses action A, irrespective of the state of the world, is

$$\mathbb{E}U_D(x=A) = u + E + \delta[qr_A\left(\frac{\bar{u}}{2}+E\right) + (1-q)r_\emptyset\left(\frac{\bar{u}}{2}+E\right)]$$

while

$$\mathbb{E}U_D(x=B) = E + \delta[qr_B\left(\frac{\bar{u}}{2}+E\right) + (1-q)r_\emptyset\left(\frac{\bar{u}}{2}+E\right)]$$

Hence, given the fact that in equilibrium $r_A = 0$ and $r_B = 1$ the D incumbent chooses action B, irrespective of the state, when $u \leq \delta\left(\frac{\bar{u}}{2}+E\right)q$.
Moving now to the C incumbent, note first of all that, when $s = B$, all the incentives are aligned and hence he will always choose action B. When $s = A$, instead, the expected utilities are as follows:

$$\mathbb{E}U_C(x=A, s=A) = u + E + \delta[qr_A\left(\frac{\bar{u}}{2}+E\right) + (1-q)r_\emptyset\left(\frac{\bar{u}}{2}+E\right)]$$

and

$$\mathbb{E}U_D(x=B, s=A) = E + \delta[qr_B\left(\frac{\bar{u}}{2}+E\right) + (1-q)r_\emptyset\left(\frac{\bar{u}}{2}+E\right)]$$

and, again, it is optimal for the C incumbent to choose action B in state A when $u \leq \delta\left(\frac{\bar{u}}{2}+E\right)q$. This completes the existence proof. In terms of uniqueness, note that there are actually multiple equilibria, given that $\hat{\pi}_\emptyset = \pi$, hence V is indifferent in that case. However, $r_\emptyset$ does not affect P's equilibrium strategies, hence all the equilibria are identical in terms of outcomes (and strategies, with the obvious exception of $r_\emptyset$).
What needs to be shown, now, is that there are no other PBNE with different re-election probabilities after observing A or B or different strategies for P.
By contradiction, suppose there is a PBNE where $r_A > 0$ and $r_B < 1$.
This is sequentially rational iff $\hat{\pi}(\tilde{x} = A) \geq \pi \Rightarrow \lambda_A + \lambda_B \geq \gamma_A + \gamma_B$ and $\hat{\pi}(\tilde{x} = B) \leq \pi \Rightarrow \lambda_A + \lambda_B \geq \gamma_A + \gamma_B$.
Note that $\gamma_A = F[q\delta\left(\frac{\bar{u}}{2}+E\right)(r_B - r_A)]$, $\gamma_B = 1 - F[-q\delta\left(\frac{\bar{u}}{2}+E\right)(r_B - r_A)]$, $\lambda_A = \lambda_B = F[q\delta\left(\frac{\bar{u}}{2}+E\right)(r_B - r_A)]$.
Hence, $\lambda_A + \lambda_B \geq \gamma_A + \gamma_B$ is equivalent to $F[q\delta\left(\frac{\bar{u}}{2}+E\right)(r_B - r_A)] + 1 - F[-q\delta\left(\frac{\bar{u}}{2}+E\right)(r_B - r_A)] \leq 2F[q\delta\left(\frac{\bar{u}}{2}+E\right)(r_B - r_A)]$.
This simplifies to $F[q\delta\left(\frac{\bar{u}}{2}+E\right)(r_B - r_A)] \geq 1 - F[-q\delta\left(\frac{\bar{u}}{2}+E\right)(r_B - r_A)]$.
Note that, if $r_B \geq r_A$, $F[-q\delta\left(\frac{\bar{u}}{2}+E\right)(r_B - r_A)] = 0$ and hence the equilibrium exists iff $F[q\delta\left(\frac{\bar{u}}{2}+E\right)(r_B - r_A)] \geq 1$, which is impossible because F is strictly increasing, $F[\bar{u}] = 1$ and $\bar{u} > \delta\left(\frac{\bar{u}}{2}+E\right) > \delta\left(\frac{\bar{u}}{2}+E\right)q(r_B - r_A)$.
If instead $r_A > r_B$, $F[q\delta\left(\frac{\bar{u}}{2}+E\right)(r_B - r_A)] = 0$ hence the equilibrium exists iff $F[q\delta\left(\frac{\bar{u}}{2}+E\right)(r_A - r_B)] \geq 1$. Again, this is impossible because $\bar{u} > \delta\left(\frac{\bar{u}}{2}+E\right) > \delta\left(\frac{\bar{u}}{2}+E\right)q(r_A - r_B)$.
Hence, there are no equilibria where $r_A > 0$ and $r_B < 1$. Applying the same logic, it is possible to also rule out equilibria where $r_A = 0$ and $r_B < 1$ and where $r_A > 0$ and $r_B = 1$.

Hence, the sole re-election strategy in a SPNE of this game is $r_A = 0$ and $r_B = 1$ and as a consequence the sole equilibrium strategy for the incumbent is the one described in Proposition 1. ■

Proof of Proposition 2.
With few algebraic manipulations, $W = \frac{1}{4}\pi(1-\pi)\delta(1-\gamma)]q + \Gamma - \tau C(q)$. Since now γ has already been decided (i.e. it is a function of q^e), the solution is obtained differentiating only with respect to q and setting the first order conditions equal to 0. Note that, given the assumptions about $C(q)$, the problem is concave. Finally, equation (1) is obtained by setting the equilibrium condition $q^e = q^*$.
Uniqueness follows from the fact that the LHS is strictly decreasing in q, the RHS is strictly increasing in q, $LHS(q=0) > RHS(q=0)$ and $LHS(q=1) < RHS(q=1)$. ■

Proof of Corollary 2.
Recalling that the variance of the incumbent's type is $var(\theta) = \pi(1-\pi)$, the implicit function describing q^* is defined as $G(q^*(var(\theta), E), var(\theta), E) = var(\theta)\delta(1 - F[\delta\left(\frac{\bar{u}}{2} + E\right)q^*]) = 4\tau C^{'}(q^*)$.
By implicit function theorem,

$$\frac{dq^*}{dvar(\theta)} = \frac{\delta(1 - F[\delta\left(\frac{\bar{u}}{2}+E\right)q^*])}{var(\theta)\delta^2\left(\frac{\bar{u}}{2}+E\right)f(\delta\left(\frac{\bar{u}}{2}+E\right)q^*) + 4\tau C^{''}(q^*)} > 0$$

$$\frac{dq^*}{dE} = -\frac{var(\theta)\delta^2 q^* f(\delta\left(\frac{\bar{u}}{2}+E\right)q^*)}{var(\theta)\delta^2\left(\frac{\bar{u}}{2}+E\right)f(\delta\left(\frac{\bar{u}}{2}+E\right)q^*) + 4\tau C^{''}(q^{**})} < 0$$

■

Proof of Corollary 3.
Applying the chain rule,

$$\frac{d\gamma^*}{dvar(\theta)} = [\delta\left(\frac{\bar{u}}{2}+E\right)f(\delta\left(\frac{\bar{u}}{2}+E\right)q^*)]\frac{dq^*}{dvar(\theta)} > 0$$

because $\frac{dq^*}{dvar(\theta)} > 0$.

$$\frac{d\gamma^*}{dE} = \delta\left[q^* + \frac{dq^*}{dE}\left(\frac{\bar{u}}{2}+E\right)\right]f(\delta\left(\frac{\bar{u}}{2}+E\right)q^*)$$

Then, $\frac{d\gamma^*}{dE} > 0 \Rightarrow 1 > \frac{\left(\frac{\bar{u}}{2}+E\right)var(\theta)\delta^2 f(\delta\left(\frac{\bar{u}}{2}+E\right)q^*)}{var(\theta)\delta^2\left(\frac{\bar{u}}{2}+E\right)f(\delta\left(\frac{\bar{u}}{2}+E\right)q^* + 4\tau C^{''}(q^*)}$
which is always true. ■

Proof of Lemma 1.
The second derivative of equation (2) with respect to q is

$$\frac{\partial^2 W^{ex}}{\partial q^2} = -\frac{1}{4}\pi(1-\pi)\delta(1+\delta)\left(\frac{\bar{u}}{2}+E\right)\frac{1}{\bar{u}} - \tau C^{''}(q)$$

which is negative in the whole interval. ■

Proof of Proposition 3.
The (ex ante) optimal level of attention is implicitly defined by its first order conditions as

$$\pi(1-\pi)\delta(1 - F(\delta\left(\frac{\bar{u}}{2}+E\right)q)) = 2\delta\left(\frac{\bar{u}}{2}+E\right)\pi\frac{1}{\bar{u}} + (1-\pi)\pi\delta^2\left(\frac{\bar{u}}{2}+E\right)\frac{1}{\bar{u}}q + 4\tau C^{'}(q) \qquad \text{(A.1)}$$

Note that the LHS of (A.1) and (1) is exactly the same. However, the RHS of (A.1) is above the RHS of (1) for every interior q, hence the result stated in the proposition. ■

Proof of Proposition 4.
First q^* is implicitly expressed as a function of τ, i.e. $K(q^*(\tau), \tau) = \frac{1}{4}\pi(1-\pi)\delta(1 - F(\delta\left(\frac{\bar{u}}{2}+E\right)q^*) -$

$\tau C^{'}(q^*)$.
By implicit function theorem,

$$\frac{\partial q^*}{\partial \tau} = -\frac{\frac{\partial K}{\partial \tau}}{\frac{\partial K}{\partial q^*}}$$

$$\frac{\partial q^*}{\partial \tau} = -\frac{4C^{'}(q^*)}{\pi(1-\pi)\delta^2\left(\frac{\bar{u}}{2}+E\right)\frac{1}{\bar{u}} + 4\tau C''(q^*)} < 0$$

Defining the total (ex ante) welfare $W^{ex}(q,\alpha) = G(q) - \tau C(q)$,[24] it is now possible to plug in q^* and to differentiate with respect to τ.
Note that $\frac{dW^{ex}}{d\tau} = [\frac{\partial G(q^*)}{\partial q^*} - \tau\frac{\partial C(q^*)}{\partial q^*}]\frac{\partial q^*}{\partial \tau} - C(q^*)$.
The content in parenthesis is equal to zero at $q = q^* < q^{ex}$ and the function is concave, so it is now negative. Moreover, $\frac{\partial q^*}{\partial \tau} < 0$.
Substituting the parameters, the condition can be expressed as:

$$\frac{[\pi((1-\pi)\delta q^* + 2)\delta\left(\frac{\bar{u}}{2}+E\right)\frac{1}{\bar{u}} + 4\tau C^{'}(q^*) - \pi(1-\pi)\delta(1-F(\delta\left(\frac{\bar{u}}{2}+E\right)q^*)]C^{'}(q^*)}{\pi(1-\pi)\delta^2\left(\frac{\bar{u}}{2}+E\right)\frac{1}{\bar{u}} + 4\tau C''(q^*)} > C(q^*)$$

However, note that, when $q = q^*$, $\pi(1-\pi)\delta(1-F(\delta\left(\frac{\bar{u}}{2}+E\right)q^*) = 4\alpha C^{'}(q^*)$. Hence, the inequality becomes

$$\frac{\pi((1-\pi)\delta q^* + 2)\delta\left(\frac{\bar{u}}{2}+E\right)\frac{1}{\bar{u}}C^{'}(q^*)}{\pi(1-\pi)\delta^2\left(\frac{\bar{u}}{2}+E\right)\frac{1}{\bar{u}} + 4\tau C''(q^*)} > C(q^*)$$

■

Proof of Proposition 5.

As a reminder, $\gamma_s = Pr(C \text{ plays } B \text{ in state } s)$ and $\lambda_s = Pr(D \text{ plays } B \text{ in state } s)$. Moreover, $r_{\tilde{x},\tilde{s}} = Pr(re-elect|\tilde{x},\tilde{s})$.

First of all, the voter re-elects the incumbent if $\hat{\pi}_{\tilde{x},\tilde{s}} > \pi$, chooses the challenger if $\hat{\pi}_{\tilde{x},\tilde{s}} = \pi$ and is indifferent otherwise.
For given λ_s and γ_s, it is easy to see that

$$\hat{\pi}_{\emptyset,\tilde{s}} = \pi$$

hence $r_{\emptyset,\tilde{s}} \in [0,1]$.

$$\hat{\pi}_{A,\emptyset} > \pi \Leftrightarrow \lambda_A + \lambda_B > \gamma_A + \gamma_B$$

hence

$$r_{A,\emptyset} = \begin{cases} 0 & \text{if } \lambda_A + \lambda_B < \gamma_A + \gamma_B \\ [0,1] & \text{if } \lambda_A + \lambda_B = \gamma_A + \gamma_B \\ 1 & \text{if } \lambda_A + \lambda_B > \gamma_A + \gamma_B \end{cases}$$

$$\hat{\pi}_{B,\emptyset} > \pi \Leftrightarrow \lambda_A + \lambda_B < \gamma_A + \gamma_B$$

hence

$$r_{B,\emptyset} = \begin{cases} 1 & \text{if } \lambda_A + \lambda_B < \gamma_A + \gamma_B \\ [0,1] & \text{if } \lambda_A + \lambda_B = \gamma_A + \gamma_B \\ 0 & \text{if } \lambda_A + \lambda_B > \gamma_A + \gamma_B \end{cases}$$

$$\hat{\pi}_{A,A} > \pi \Leftrightarrow \lambda_A > \gamma_A$$

[24] Where of course $G(q) = \pi\{\frac{1}{2}(1+qV_C) + \frac{1}{2}[(1-\gamma)(1+q\Gamma) + \gamma q V_C]\} + (1-\pi)\{\frac{1}{2}[\gamma(1+qV_D) + (1-\gamma)q\Gamma] + \frac{1}{2}[\gamma q V_D + (1-\gamma)(1+q\Gamma)]\} + (1-q)\Gamma$.

hence

$$r_{A,A} = \begin{cases} 1 & \text{if } \lambda_A > \gamma_A \\ [0,1] & \text{if } \lambda_A = \gamma_A \\ 0 & \text{if } \lambda_A < \gamma_A \end{cases}$$

$$\hat{\pi}_{B,B} > \pi \Leftrightarrow \gamma_B > \lambda_B$$

hence

$$r_{B,B} = \begin{cases} 0 & \text{if } \lambda_B > \gamma_B \\ [0,1] & \text{if } \lambda_B = \gamma_B \\ 1 & \text{if } \lambda_B < \gamma_B \end{cases}$$

$$\hat{\pi}_{A,B} > \pi \Leftrightarrow \lambda_B > \gamma_B$$

hence

$$r_{A,B} = \begin{cases} 1 & \text{if } \lambda_B > \gamma_B \\ [0,1] & \text{if } \lambda_B = \gamma_B \\ 0 & \text{if } \lambda_B < \gamma_B \end{cases}$$

$$\hat{\pi}_{B,A} > \pi \Leftrightarrow \gamma_A > \lambda_A$$

hence

$$r_{B,A} = \begin{cases} 0 & \text{if } \lambda_A > \gamma_A \\ [0,1] & \text{if } \lambda_A = \gamma_A \\ 1 & \text{if } \lambda_A < \gamma_A \end{cases}$$

Moving to the best response correspondences of P, for conjectured levels of attention β^e and q^e, he will compare $EU_\theta(x = B, s)$ with $EU_\theta(x = A, s)$. Therefore,

$$\lambda_A = Pr(u \leq \delta\left(\frac{\bar{u}}{2} + E\right) q^e \left[(1-\beta^e)(r_{B,\emptyset} - r_{A,\emptyset}) + \beta^e(r_{B,A} - r_{A,A})\right]) \tag{A.2}$$

$$\lambda_B = Pr(u \leq \delta\left(\frac{\bar{u}}{2} + E\right) q^e \left[(1-\beta^e)(r_{B,\emptyset} - r_{A,\emptyset}) + \beta^e(r_{B,B} - r_{A,B})\right]) \tag{A.3}$$

$$\gamma_A = Pr(u \leq \delta q^e \left(\frac{\bar{u}}{2} + E\right) \left[(1-\beta^e)(r_{B,\emptyset} - r_{A,\emptyset}) + \beta^e(r_{B,A} - r_{A,A})\right]) \tag{A.4}$$

$$\gamma_B = Pr(u + \delta q^e \left(\frac{\bar{u}}{2} + E\right) \left[(1-\beta^e)(r_{B,\emptyset} - r_{A,\emptyset}) + \beta^e(r_{B,B} - r_{A,B})\right]) \geq 0 \tag{A.5}$$

Clearly, in equilibrium $\lambda_A = \gamma_A$, hence $r_{A,A}$ and $r_{B,A}$ are unconstrained because of indifference. The same is true for $r_{\emptyset,\tilde{s}}$ since no relevant information is transmitted.
I claim that, in every equilibrium, it must be that $r_{B,B} = 1$, $r_{A,B} = 0$, $r_{B,\emptyset} = 1$ and $r_{A,\emptyset} = 0$. Given that $\lambda_A = \gamma_A$, those are sequentially rational voting strategies if $\gamma_B > \lambda_B$. Replacing them in equations (A.5) and (A.3) the result, as required, is

$$\gamma_B = 1 > \lambda_B = F\left[\delta\left(\frac{\bar{u}}{2} + E\right) q^e\right]$$

Replacing in (A.4) and (A.2) the result is that

$$\gamma_A = \lambda_A = F[\delta\left(\frac{\bar{u}}{2} + E\right) q(1 - \beta + \beta(r_{B,A} - r_{A,A}))]$$

To show that there are no other equilibria I use contradiction. First, suppose that there exists an equilibrium where $\gamma_B < \lambda_B$, and as a consequence $r_{B,B} = 0$, $r_{A,B} = 1$, $r_{B,\emptyset} = 0$ and $r_{A,\emptyset} = 1$. Replacing those values in equations (A.5) and (A.3) the result is

$$\gamma_B = Pr(u \geq \delta q^e \left(\frac{\bar{u}}{2} + E\right)) > 0 = \lambda_B$$

hence a contradiction.
Finally, suppose that $\gamma_B = \lambda_B$. Defining $\delta\left(\frac{\bar{u}}{2} + E\right) q^e \left[(1-\beta^e)(r_{B,\emptyset} - r_{A,\emptyset}) + \beta^e(r_{B,B} - r_{A,B})\right] := D$, this requires $Pr(u \leq D) = Pr(u \geq -D)$, which never holds.

■

Proof of Proposition 6.

While choosing the equilibirum level of attention, the voter differentiates (4) keeping as a constant the pandering probabilities of the incumbent. Noticing that, once replacing the actual values and taking into account that $\gamma_B = 1$ and $\gamma_A = \lambda_A$, $Pr(x = B, s = B)(\hat{\pi}_{B,B}V_C + (1-\hat{\pi}_{B,B})V_D) + (1 - Pr(x = B, s = B))\Gamma = Pr(x = B)(\hat{\pi}_{B,\emptyset}V_C + (1 - \hat{\pi}_{B,\emptyset})V_D) + Pr(x = A)\Gamma$, the first order condition simplifies to

$$\frac{1}{4}\pi(1-\pi)(1-\lambda_B) = 0$$

which leads to a corner solution where q^{**} is as big as possible, i.e. equal to α. As a consequence, $\beta^{**} = 0$.

■

Proof of Lemma 2.

The second derivative of equation (5) with respect to q is given by

$$-\frac{1}{\bar{u}}\delta\left(\frac{\bar{u}}{2}+E\right) - \frac{1}{4}\pi(1-\pi)\delta\left(\frac{\bar{u}}{2}+E\right)\frac{1}{\bar{u}} - \frac{1}{4}\pi(1-\pi)\delta^2\left(\frac{\bar{u}}{2}+E\right)\frac{1}{\bar{u}} < 0$$

as required.

■

Proof of Proposition 7.

Using lemma 2, it is sufficient to look at the first order conditions of (5) for maxima. Hence, the ex ante optimal q solves

$$\frac{1}{4}\pi(1-\pi)(1-F(\delta\left(\frac{\bar{u}}{2}+E\right)q)) = \frac{1}{4}\pi(1-\pi)\delta^2\left(\frac{\bar{u}}{2}+E\right)\frac{1}{\bar{u}}q + \frac{1}{2}\frac{1}{\bar{u}}\delta\left(\frac{\bar{u}}{2}+E\right)(\pi-\alpha+2q)$$

As the LHS is strictly decreasing in q and the RHS is strictly increasing in q, the maximum is unique.

To have equilibrium misallocation of attention, it is enough that the maximum is below α, and this happens whenever $LHS(q = \alpha) < RHS(q = \alpha)$. Replacing values from above and writing everything as a function of $\frac{1}{\bar{u}}$, the condition simplifies to

$$\frac{1}{\bar{u}} > \frac{\pi(1-\pi)(1-F(\delta\left(\frac{\bar{u}}{2}+E\right)\alpha))}{\pi(1-\pi)\delta^2\left(\frac{\bar{u}}{2}+E\right)\alpha + 2\delta\left(\frac{\bar{u}}{2}+E\right)(\pi+\alpha)}$$

■

Proof of Lemma 3.

The voter chooses to maximize $\tilde{\tilde{W}}$ with respect to q and β, taking as given the pandering decision of the politician. In particular, after some simplifications,

$$\tilde{\tilde{W}} = q\left(\frac{1}{4}\pi(1-\pi)(1-\lambda_B)\right) + \left(\pi\delta + (1-\pi)\delta\frac{1}{2}\right) - \tau(q^2+\beta^2) \tag{A.6}$$

From equation (A.6), β enters only negatively in the objective function, and hence its solution is $\beta^{***} = 0$. Replacing above, noticing the concavity of the objective function and differentiating with respect to q, the solution is given by

$$\left(\frac{1}{4}\pi(1-\pi)(1-\lambda_B)\right) = 2\tau q$$

Combing this with $\lambda_B = \delta K q^e$ and the equilibrium condition $q^e = q^{***}$, it turns out that q^{***} is the fixed point of

$$\left(\frac{1}{4}\pi(1-\pi)(1-\delta K q^{***})\right) = 2\tau q^{***}$$

As the LHS is decreasing in q and the RHS is increasing in q, the solution is unique. It will be interior as long as $LHS(q=1) \leq RHS(q=1)$, noticing that of course $LHS(q=0) > RHS(q=0)$. This happens when $\tau \geq \frac{\pi(1-\pi)(1-\delta K)}{8}$, and this concludes the proof. ■

Proof of Lemma 4.

Ex ante, the function to be maximized is

$$\tilde{\tilde{W}}^{ex} = \frac{1}{2}(1-\delta K q(1-\beta)+\pi+(1-\pi)\delta K q)+q\left(\frac{1}{4}\pi(1-\pi)(1-\delta K q)\right)+\left(\pi\delta+(1-\pi)\delta\frac{1}{2}\right)-\tau(q^2+\beta^2)$$

To determine the Hessian matrix, note that

$$\frac{\partial^2 \tilde{\tilde{W}}^{ex}}{\partial q^2} = -\delta K - \frac{1}{4}\pi(1-\pi)\delta K - \frac{1}{4}\pi(1-\pi)\delta^2 K - 2\tau < 0$$

$$\frac{\partial^2 \tilde{\tilde{W}}^{ex}}{\partial \beta^2} = -2\tau < 0$$

$$\frac{\partial^2 \tilde{\tilde{W}}^{ex}}{\partial \beta \partial q} = \frac{\partial^2 \tilde{\tilde{W}}^{ex}}{\partial q \partial \beta} = \frac{1}{2}\delta K$$

As $\frac{\partial^2 \tilde{\tilde{W}}^{ex}}{\partial q^2}$ and $\frac{\partial^2 \tilde{\tilde{W}}^{ex}}{\partial \beta^2}$ are both negative, the remaining condition for the negative semidefinitiveness of the Hessian is

$$\frac{\partial^2 \tilde{\tilde{W}}^{ex}}{\partial q^2} * \frac{\partial^2 \tilde{\tilde{W}}^{ex}}{\partial \beta^2} - \left(\frac{\partial^2 \tilde{\tilde{W}}^{ex}}{\partial \beta \partial q}\right)^2 \geq 0$$

that simplifies to

$$\pi(1-\pi)\delta K(1+\delta)2\tau + (4\tau)^2 \geq (\delta K)^2$$

As a consequence, $4\tau \geq \delta K$ is a sufficient condition for this to be true.
■

Proof of Proposition 8.

Thanks to Lemma 4 and the fact that the constraints are all linear, Kuhn-Tucker conditions are necessary and sufficient for a maximum.

Hence, the Lagrangian function is

$$L = \tilde{\tilde{W}}^{ex} + \lambda_1(1-q) + \lambda_2(1-\beta)$$

The conditions are
$q \geq 0$ $q\frac{\partial L}{\partial q} = 0$ $q \leq 1$ $\frac{\partial L}{\partial q} \leq 0$
$\beta \geq 0$ $\beta\frac{\partial L}{\partial \beta} = 0$ $\beta \leq 1$ $\frac{\partial L}{\partial \beta} \leq 0$
$\lambda_1 \geq 0$ $\lambda_2 \geq 0$
$\lambda_1(1-q) = 0$ $\lambda_2(1-\beta) = 0$
and there are multiple cases to be considered.

Case 1: $\beta = 0$ and $q = 0$.
This requires $\lambda_1 = 0$ and $\lambda_2 = 0$. To be a maximum, it must be that $\frac{\partial L}{\partial q} \leq 0$ that requires, after substitution, $1 - \pi - 2\delta K \leq 0$.

Case 2: β and q both interior, hence $\lambda_1 = 0$ and $\lambda_2 = 0$. The solution is given by the system of $\frac{\partial L}{\partial q} = 0$ and $\frac{\partial L}{\partial \beta} = 0$. The solution is given by

$$q = \frac{2\tau\pi(1-\pi-2\delta K)}{2\pi(1-\pi)K\tau(\delta+1)+16\tau^2-(\delta K)^2}$$

$$\beta = \frac{\delta K\pi(1-\pi-2\delta K)}{4\pi(1-\pi)K\tau(\delta+1)+32\tau^2-2(\delta K)^2}$$

which requires

$$1-\pi-2\delta K \geq 0$$

Note that a sufficient condition for $\frac{2\tau\pi(1-\pi-2\delta K)}{2\pi(1-\pi)K\tau(\delta+1)+16\tau^2-(\delta K)^2} < 1$ is $8\tau > \frac{\tau+2(\delta K)^2}{4\tau}$, which is always true because $4\tau \geq \delta K$ and $\delta K < 1$. $\frac{\delta K\pi(1-\pi-2\delta K)}{4\pi(1-\pi)K\tau(\delta+1)+32\tau^2-2(\delta K)^2} < 1$ is also realized, as it is a weaker condition than the previous one.[25]

Case 3: q interior and $\beta = 1$.
This requires $\lambda_1 = 0$ and $\lambda_2 > 0$. Note that, whatever $q < 1$ can be found, when substituting into $\frac{\partial L}{\partial \beta} = 0$ the latter requires $2\lambda_2 = \delta K q - 4\tau < 0$, hence a contradiction.

Case 4: $q = 1$ and β interior.
This requires $\lambda_1 > 0$ and $\lambda_2 = 0$. The solution for β is $\frac{\delta K}{4\tau}$ however, when replacing into $\frac{\partial L}{\partial q} = 0$ it simplifies into $\lambda_1 = \frac{1}{4}\pi(1-\pi)(1-\delta K) - \frac{1}{2}\left(\pi - \frac{\delta K}{4\tau}\right) - \frac{1}{4}\pi(1-\pi)\delta^2 K - 2\tau$, which is negative because $8\tau > \frac{\tau+2(\delta K)^2}{4\tau}$, hence a contradiction.

Case 5: $q = \beta = 1$.
This requires $\lambda_1 > 0$ and $\lambda_2 > 0$. When substituting into $\frac{\partial L}{\partial \beta} = 0$ the latter requires $2\lambda_2 = \delta K - 4\tau < 0$, hence a contradiction.

Case 6: $q = 0$ and β interior.
This requires $\lambda_1 = 0$ and $\lambda_2 = 0$. When replacing, however, the solution for β is only 0, hence a contradiction.

Case 7: q interior and $\beta = 0$.
This requires $\lambda_2 = 0$. However, $\frac{\partial L}{\partial \beta} \leq 0$ only if $q = 0$, hence a contradiction.

Case 8: $q = 0$ and $\beta = 1$.
Then, $\frac{\partial L}{\partial \beta} = 0$ implies a negative λ_2, hence a contradiction.

Case 9: $q = 1$ and $\beta = 0$.
This requires $\lambda_2 = 0$ and $\frac{\partial L}{\partial q} = 0$. However, when substituting, the latter is possible only if $\lambda_1 < 0$, hence a contradiction.

To sum up, the ex ante maximization problem has just two solutions. $q = \beta = 0$ if $1-\pi-2\delta K \leq 0$ and

$$q = \frac{2\tau\pi(1-\pi-2\delta K)}{2\pi(1-\pi)K\tau(\delta+1)+16\tau^2-(\delta K)^2}$$

$$\beta = \frac{\delta K\pi(1-\pi-2\delta K)}{4\pi(1-\pi)K\tau(\delta+1)+32\tau^2-2(\delta K)^2}$$

[25] $\frac{2\tau\pi(1-\pi-2\delta K)}{2\pi(1-\pi)K\tau(\delta+1)+16\tau^2-(\delta K)^2} < 1$ requires $2\tau\pi(1-\pi-2\delta K) < 2\pi(1-\pi)K\tau(\delta+1)+16\tau^2-(\delta K)^2$, while $\frac{\delta K\pi(1-\pi-2\delta K)}{4\pi(1-\pi)K\tau(\delta+1)+32\tau^2-2(\delta K)^2} < 1$ requires $\frac{\delta K}{2}\pi(1-\pi-2\delta K) < 2\pi(1-\pi)K\tau(\delta+1)+16\tau^2-(\delta K)^2$. Obviously, $\frac{\delta K}{2} \leq 2\tau$.

if $1-\pi-2\delta K>0$. As the equilibrium level of β is always 0, $1-\pi-2\delta K>0$ guarantees equilibrium (strict) under-attention to the state of the world.

The last part of Proposition 8 is a simple comparison between $q^{***}=\frac{\pi(1-\pi)}{8\tau+\pi(1-\pi)\delta K}$ and the interior ex ante optimal q, looking for parameters where the latter is smaller. Note that it is possible to find a sufficient condition for equilibrium over-attention to the action, i.e. $4\tau\pi>1$.

■

Appendix B Relaxing the assumptions of Section 3.3

If the explicit functional form assumed in 3 is replaced with a generic $\tau C(q,\beta)$, the condition for the negative semidefiniteness of the Hessian becomes

$$\frac{1}{4}\pi(1-\pi)\delta(1+\delta)K\tau C^{''}_{\beta\beta}+\tau^2 C^{''}_{\beta\beta}C^{''}_{qq}-\left(\frac{1}{2}\delta K-\tau C^{''}_{\beta q}\right)^2\geq 0$$

which holds as long as the second derivative of $C(q,\beta)$ is sufficiently big everywhere. Obviously, it is not possible to characterize closed form solutions for a generic cost function.

Assumption 4 is a sufficiency one: it can be relaxed without losing the concavity. Note that concavity is violated for sufficiently small values of τ. In this case, even though the full analysis is complicated, it seems likely that the solution would move toward $\beta=q=1$.