When the light shines too much.
Rational inattention, populism and pandering

Federico Trombetta*
Department of Economics
The University of Warwick

February 20, 2018

First version: April 13, 2017
PRELIMINARY

Abstract

Should voters always pay attention to politicians’ actions? This paper shows that, in equilibrium, too much attention is paid to politics, and this is bad for the voters. Combining insights from the growing literature on rational inattention with a standard model of political agency and pandering, this paper proves that - in general - voters’ attention increases pandering. Hence, if voters cannot commit ex ante to a certain level of attention, their equilibrium choice will is too high with respect to the welfare maximizing one. As a consequence, a reduction in the cost of attention generically increases the equilibrium level of pandering, and it can be overall welfare depressing if pandering is sufficiently bad for the voters. This contributes to the explanation of the relationship between social media and populism: the former seems to help political engagement and to reduce the cost of paying attention to politics, but this is exactly what increases the electoral reward of populist policies.

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 Pietro Ortoleva, Michela Boccola and Guido Merzoni. They all provided very useful comments and suggestions. Of course, all the mistakes are mine.
“Contemporary media and the new media in particular are facilitating the growth of populism”.

Martinelli (2016)

1 Introduction

1.1 Motivation and contribution

What is the right amount of attention that voters should pay to politics? At first, the answer seems obvious: as much as possible, because this will allow them to make better choices and to elect better politicians. Moreover, if this is true, the impact of tools supposed to make attention “cheaper”, like social media, should have an unambiguously positive impact on politics. This paper shows that reality may be more complex than that.

In particular, there are two points that should be taken into account, in order to have a more complete picture. First, voters’ attention is a scarce resource, that has some cost (in terms of opportunity cost, at the very least) and that is (optimally?) allocated by the voters. Secondly, this allocation choice has repercussions not only on the amount of information extracted by each voter, but also on the way politicians react.

In particular, one may suspect that voters’ attention may motivate politicians to do what voters want, irrespective of what is right given the “state of the world”. Moving from the model to reality, this relates with the opening quotation of this paper: is there a connection between social media and populism, defined as the “thin-centred ideology [...] which argues that politics should be an expression of the volonté générale (general will) of the people”?

This paper is the first one, in the political economy literature, that looks at the effects of rational inattention on pandering. It shows that the equilibrium level of attention to politics is suboptimally high, that a reduction in the cost of attention increases pandering and that this, for a certain range of parameters, is overall welfare depressing for the voters. More in detail, the model builds on the literature on pandering (i.e. the politician chooses the action that guarantees his re-election, rather than the optimal one) because it can be seen as a good theoretical approximation of the basic essence of populism. In that set up, this paper endogenizes voters’ level of attention to politics, following the recent and growing literature on the political consequences of rational inattention, finding novel results. In particular, in equilibrium there is a positive relationship between attention and pandering, robust to whether the voters are able to commit ex ante to a certain attention level or they choose attention just before voting. This creates an important trade off, since pandering is welfare depressing but, obviously, attention

---

1 Mudde and Kaltwasser (2017), pag. 6.

2 This interpretation of pandering as populism is not new in economics. E.g. Frisell (2009).
improves political selection.
As a consequence, the welfare optimal level of attention is, under very mild conditions, smaller than the full one, even when attention is costless. Moreover, if \textit{ex ante} commitment is not possible, the equilibrium level of attention chosen by the voter is always inefficiently high. In other words, too much attention is paid to politics, and this generates an excessive level of pandering from the politicians. Hence, when the negative effect of pandering on voter’s welfare is high enough, a decrease in the cost of attention makes the voter worse off, overall.

Those results help in explaining the relationship between social media and populism. Interestingly, politicians advocating populist policies are over-represented on social media\(^3\), social media usage is highly correlated with political engagement\(^4\) and newspaper articles are asking whether social media are “empowering populism”\(^5\). In the framework of this paper, social media reduce the cost of paying attention to politics and politicians, generating more attention\(^6\) but also inducing more pandering (whose overall effect on voters’ welfare is negative), 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, where the spreading of social media leads to “too much” attention and “too much” populism.

Finally, I expand the model allowing for a more complex signal structure (where the voter can pay attention to the action and to the state of the world) showing that the main results are unchanged, that the two types of attention are complements and that there are conditions where a decrease in the cost of attention to the state of the world increases the probability that the good politician chooses to pander.

1.2 Related literature

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

First of all, 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.

---

\(^3\)Extreme Tweeting, The Economist, 19th November 2015.
\(^6\)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).
Secondly, it is related to the growing literature looking at 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, Lockwood 2016, Lockwood and Rockey 2016). Thirdly, it is related with the growing but still small 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. Among those, two papers are the closest to this one. First of all, Prato and Wolton (2016) finds a “curse of interested voter”, showing that too much interest in the political outcome would harm the voters in equilibrium. Note, however, that they distinguish between interest (exogenous) and attention (endogenous) and they do not get “over-attention” in equilibrium, while this model does. Moreover, theirs is a model of political competition and not of pandering. Secondly, the mechanics of the paper is somewhat similar to Lockwood (2016). Both look at “behavioural” consequences in a set up of political agency and pandering. However, Lockwood (2016) studies confirmation bias whereas this one studies 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). Differently from this literature, I focus on the attention choice coming from the voter’s side of the problem, hence highlighting a different and unexplored channel.

The reminder of this paper is as follows. Section 2 describes the set up of the model and discusses the main assumptions. Section 3 solves the model, highlighting the main results. Section 4 presents a version of the model with a more general signal structure. Finally, Section 5 concludes.

2 The model

2.1 Set up

Results are presented using the simplest possible model able to give us the necessary insights. In particular, I build on the models of pandering by Fox (2007) and Besley (2006), but the main results are qualitatively unchanged if the model is in the spirit of Maskin and Tirole (2004), as explained in Appendix B.

The game is a standard two period 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

---

7 As I show in Appendix C, the results are qualitatively unchanged if, instead of a single representative voter, I assume a continuum of voters in a probabilistic voting set up with sincere voting.
$s_t = \{ A, B \}$, known by $P$ but not by $V$, with common prior $Pr(s_t = A) = p = \frac{1}{2}$. 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, that I am going to endogenize using the choice of attention. Finally, 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, instead, 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, his probability density function is $f$ such that $f(0) = 0$. Without loss of generality, it is assumed that $E[u_t] = 1$. Moreover, as in Fox (2007) and Besley (2006), it is assumed that $\bar{u} > \delta(1 + E)$, 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 $\delta$.

So far, the model is unchanged with respect to the standard political agency and pandering literature. This paper builds on this framework endogenizing the choice of attention. As in Prato and Wolton (2016 and 2017), I model rational inattention as the probability $q$ that the voter observes the action of the politician before casting her choice. Formally, the voter does not observe $x_t$ directly, but rather $\tilde{x}_t$, and she chooses $q \in [0, 1]$ at a cost $\alpha C(q)$, where $Pr(\tilde{x}_t = x_t) = q$ and $Pr(\tilde{x}_t = \emptyset) = 1 - q$, $\alpha \geq 0$. In other words, the higher is the (costly) attention $q$, the more likely is the voter to observe the action chosen by the politician, 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 of conditions. Formally, $C'(0) = 0$, $C'(1) = \infty$, $C'(q) > 0 \ \forall q \in (0, 1)$ and $C''(q) \geq 0 \ \forall q \in (0, 1)$.

---

8The 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.
2.2 Timing and commitment

For the sake of realism, I assume that the voter chooses the attention level after the incumbent’s action, but without knowing that\(^9\). Hence, the timing used to derive the equilibrium attention choice is as follows:

1. Nature chooses \(\theta, u_1, s_1\) (private info of P);
2. P chooses \(x_1\);
3. V chooses \(q \in [0, 1]\) at a cost \(\alpha C(q) > 0\);
4. V observes \(\tilde{x}_1\) and votes;
5. Period 1 ends and payoffs are paid;
6. Identical period 2 (but no elections);

Note that, crucially, payoffs are paid after the elections. However, in order to better understand the trade off induced by the voter’s attention choice and the role of attention on welfare, it is instructive to look at a benchmark case where the voter is actually able to commit ex ante to a certain level of attention, known to the politician when making his choice. This second time, that I define “ex ante attention choice”, is as follows:

1. Nature chooses \(\theta, u_1, s_1\) (private info of P);
2. V chooses \(q \in [0, 1]\) at a cost \(\alpha C(q) \geq 0\);
3. P observes \(q\) and chooses \(x_1\);
4. V observes \(\tilde{x}_1\) and votes;
5. Period 1 ends and payoffs are paid;
6. Identical period 2 (but no elections);

In both cases, the solution concept I use is the perfect Bayesian Nash equilibrium (PBNE).

3 Results

The first result of this paper are some characteristic of the (essentially unique) equilibrium that apply irrespective of the timing of the game\(^10\).

\(^9\)Formally, it is the same as letting them choose at the same time.
\(^10\)All the proofs are in Appendix A.
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 \( s = A \) when \( u \leq \delta(1 + E)q \);

3. D incumbent chooses \( x = B \) when \( u \leq \delta(1 + E)q \);

4. \( Pr(re - elect|\tilde{x} = \emptyset) \) can be anything;

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\(^{11}\). Note that there is multiplicity of 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 allows to note the following corollary.

Definition 1 Define “pandering” as the probability \( \gamma = F[\delta(1 + E)q] \) that an incumbent chooses the period 1 action that is suboptimal from his point of view.

Corollary 1 Irrespective of when availability of commitment, an increase in attention increases the probability of pandering.

This result makes intuitive sense, since the whole point of pandering is to guarantee re-election, choosing the “popular” action rather than the right one. As a consequence, it is crucial that this action is observed, by the voter.

It is now possible to look at the choice of attention distinguishing between the equilibrium level and the ex ante optimal one.

3.1 Benchmark: ex ante choice of attention

Given the result of Proposition 1, the sole part that remains to be solved is the attention choice. In the second timing, this happens at the beginning of the game, before the politician takes any action. Hence, the voter chooses \( q \) to maximize her expected welfare, taking into account also the effect of \( q \) on pandering.

Define \( V_C = \delta \) and \( V_D = \delta p \) 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

\(^{11}\)Note that, in case of a consonant politician, this is always bad for the voters.
electing the challenger, while \( \gamma \) is defined as in Corollary 1. Finally, assume for the moment that attention is costless, hence \( \alpha = 0 \).

The expected welfare is

\[
W = \pi \{(1 - p)(1 + qV_C) + p[(1 - \gamma)(1 + q\Gamma) + \gamma qV_C]\} + \\
+ (1 - \pi)\{(1 - p)[\gamma(1 + qV_D) + (1 - \gamma)q\Gamma]\} + \\
+ p[\gamma qV_D + (1 - \gamma)(1 + q\Gamma)]\} + (1 - q)\Gamma.
\]

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 then the incumbent is voted out when the action is observed, and the challenger is elected. With probability \( \gamma \), the consonant incumbent panders in state A, giving 0 to the voter but being re-elected and hence giving \( V_C \) in period 2. The same logic applies to the dissonant incumbent. Finally, note that, when the action is not observed, the voter is always indifferent between the challenger and the incumbent, hence her expected welfare in this case is just \( (1 - q)\Gamma \).

In order to find the optimal level of attention, the voter solves

\[
\max_{q \in [0, 1]} W.
\]

Lemma 1 is used to derive sufficient condition for the strict concavity of this maximization:

**Lemma 1** If \( f'(\delta(1 + E)q) \geq 0 \) between 0 and \( \delta(1 + E) \) then \( W \) is strictly concave in \( q \in [0, 1] \).

For the rest of the paper, it is assumed that this condition is verified. Hence, it is possible to solve the maximization just taking the first order conditions:

\[
\frac{\partial W}{\partial q} = \pi(1 - \pi)\frac{1}{2} \left( 1 - \frac{\partial \gamma}{\partial q} q - \gamma \right) (V_C - V_D) - \frac{1}{2} \pi \frac{\partial \gamma}{\partial q}.
\]

(2)

Thanks to equation (2), the effect of attention on ex ante welfare can be decomposed in three different parts:

1. A positive information availability effect, given by \( \pi(1 - \pi)\frac{1}{2}(1 - \gamma)(V_C - V_D) \). The higher is \( q \), the more likely is the voter to observe the action, hence she can cast a more informed vote;

2. A negative direct effect of pandering, given by \( -\frac{1}{2} \pi \frac{\partial \gamma}{\partial q} \), that is higher the higher is \( \pi \) since the pandering of the consonant politician is always bad for the voter;
3. A negative effect on selection, given by $-\pi(1 - \pi)(1 - p)\frac{\partial V}{\partial q}q(V_C - V_D)$, since the higher is $q$ the more likely is the re-election of a D incumbent;

This list clarifies the trade off that attention implies. On the one hand, higher attention guarantees that more information is available (in expectation), allowing for a better choice. On the other hand, higher attention implies a present welfare depressing pandering decision and the increased probability of re-electing a dissonant incumbent.

The consequence of this trade off is that, under fairly mild condition, the optimal level of attention is interior even when it is costless, i.e. when $\alpha = 0$. This is stated in the following proposition.

**Proposition 2** In case of ex ante, costless attention choice, if $f'(\delta(1+E)q) \geq 0$ when $q \in [0, 1]$, $f(0) = 0$ and $E$ large enough, then the optimal level of attention $q^* \in (0, 1)$ is unique and implicitly defined by

$$
(1 - \pi)(1 - F(\delta(1+E)q^*)) = ((1 - \pi)\delta q^* + 2)(1 + E)f(\delta(1+E)q^*)
$$

(3)

From equation (3), it is immediate to note that an increase in the office rent decreases $q^*$ because it shifts down the LHS and up the RHS. Intuitively, high office rents make pandering more desirable, from the point of view of the incumbent. The voter, choosing ex ante, takes this into account and reduces the attention level in order to reduce the negative effects of pandering. The effect of $\pi$, instead, is ambiguous, reflecting the fact that the importance of the selection problem is affected by the variance of $\theta$ (and the effect of $\pi$ on this variance is U-shaped), while a high $\pi$ makes the negative direct effect of pandering bigger.

Importantly, under fairly mild conditions it is possible to show that the voter does not want to pay full attention to politics, if she is able to pre-commit to a certain level, even when attention is costless. Finally note that, if $\alpha > 0$, then the level of attention is interior for every value of $E$.

Finally, it is important to point out that the pre-commitment case is not very realistic. Not only because it assumes the existence of some sort of attention-limiting device, but also because it assumes that the voter is able to fully take into account the effect of her level of attention on the decision of the incumbent, and this seems quite unlikely, in reality. However, it is useful as a benchmark for welfare comparison purposes.
3.2 Equilibrium choice of attention

In this section I solve the model in the (more realistic) timing where the voter chooses the level of attention after the politician chooses his action. From the point of view of the incumbent, now \( \gamma = F[\delta(1+E)q^e] \), where \( q^e \) is the expected 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. The voter uses the same re-election strategy, but the welfare function she maximizes is now different. In particular, the voter’s problem is equivalent to 

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

and \( \hat{\pi} = Pr(\theta = C|x = B) \).\(^{12}\) Note that \( \gamma \) is no longer a function of \( q \), because it is determined before the voter chooses her attention. 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.

**Proposition 3** The equilibrium level of attention \( q^{**} \) is unique, interior and implicitly defined by

\[
\pi(1 - \pi)\delta(1 - F[\delta(1+E)q^{**}]) = 4\alpha C'(q^{**})
\]

Note that, if attention is costless (i.e. \( \alpha = 0 \)), then \( q^{**} = 1 \). This is because, in this case, the basic attention trade off highlighted above is no longer true. In other words, the voter now 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 \( q = 1 \).

In terms of comparative statics:

**Proposition 4** The equilibrium attention level is increasing in the uncertainty over the type of incumbent (\( \text{var}(\theta) \)) 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 \( \theta \) increases the value of knowing the chosen action, while \( E \), leaving everything

\(^{12}\)Note that, basically, \( W' \) 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.
else equal, increases the probability of pandering, making the chosen action a poorer signal of
the incumbent’s type.

Finally, it is possible to look at some comparative statics on the equilibrium probability of
pandering as well. In particular, given that, in equilibrium, \( q^e = q^{**} \), the equilibrium level of
pandering is \( \gamma^{**} = F[\delta(1 + E)q^{**}] \). Then, Corollary 2 follows directly from Proposition 3.

**Corollary 2** The equilibrium level of pandering is increasing in the uncertainty over the incum-
ment’s type and in the office rent

Intuitively, uncertainty over \( \theta \) increases the equilibrium level of attention, making pandering
more profitable. \( E \) has a negative effect on \( q^{**} \), hence a negative indirect effect on the pandering
probability, but also a positive direct effect on \( \gamma^{**} \), making being in office more attractive, and
the second one always dominates.

### 3.3 Attention level and voter’s welfare

It is now possible to compare the attention choice that emerges in equilibrium without
commitment with the ex ante optimal attention choice, defined as \( q^{***} \) when there is a positive
attention cost.

**Proposition 5** The equilibrium attention choice \( q^{**} \) is strictly bigger than the ex ante welfare
maximizing choice.

Intuitively, when the voter chooses ex post, she “over-evaluates” \( q \), because she does not
take into account the effect of \( q \) on pandering, that is welfare depressing. On the other hand,
the ex ante welfare calculation takes this into account as well, hence it leads to a lower level of
\( q \).

A first consequence of this is the fact that attention pre-commitment can help the voter in
obtaining 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 \( \alpha \)) is overall welfare depressing,
when no pre-commitment is available?

**Proposition 6** A decrease in \( \alpha \) is welfare depressing if
\[
\frac{\partial W(q^{**})}{\partial q^{**}} - \alpha \frac{\partial C(q^{**})}{\partial q^{**}} \frac{\partial q^{**}}{\partial \alpha} > C(q^{**}).
\]

This requires
\[
\frac{\pi((1 - \pi)\delta q^{**} + 2)\delta(1 + E)f(\delta(1 + E)q^{**})C'(q^{**})}{\pi(1 - \pi)\delta^2(1 + E)f(\delta(1 + E)q^{**}) + 4\alpha C''(q^{**})} > C(q^{**})
\]  
(5)

Equation (5) is more likely to hold when, at the equilibrium level of attention, the welfare
cost imposed by the increased pandering is high (in practice, when not taking into account those
costs is very painful, in utility terms). Moreover, it is more likely to hold the bigger is $C'(q^{**})$ with respect to $C(q^{**})$.

Proposition 6 has important policy consequences, since it implies that a 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.

3.4 Discussion

The assumption on voter’s desire for information needs more discussion. In the model, the unique principal is obviously interested in acquiring information on the action of the unique agent. In reality, however, the number of voters is so high that any costly information acquisition should be ruled out by pivotality considerations. Moreover, in the real world information acquisition on politicians’ actions is usually mediated by mass media, with the well known problems of bias, capture etc. Finally, it is hard to think that individual voters’ attention choices determine politicians’ reactions. Hence: how to justify costly attention in this set up?

A first possibility is to assume that the voter must take an action (before voting) whose return depends on her information about the politician’s action. As a consequence, the voter has an interest in paying costly attention even when she is not pivotal. Secondly, the voter may be motivated by the desire to choose the best possible politician and to vote optimally, even if her individual vote has a very small probability to be the decisive one. In both those cases, the main messages of this paper would not change.

A second concern may be the fact that a big number of voters would not necessarily coordinate to a certain level of attention. In this case, the model would be less tractable but the result would not change much. In fact, under the assumptions of honest voting (weakly dominant with multiple voters) and of the fact that indifferent voters just toss a coin, then the re-election probability of $P$ after choosing action $A$ would be just a decreasing function in the probability that a sufficiently high number of voters pay attention to the action and hence vote against him. The opposite is true when the action is $B$, hence the electoral gains from pandering will still be weighted by a function increasing in the attention of the voters.

4 Extension: two types of attention

A natural extension of the basic model is the one allowing the voter to choose the target of her attention. In reality, in fact, attention may be paid to the state of the world (e.g. to the economic conditions of the country) and to the incumbent’s action (e.g. the type of economic
policy).

4.1 Equilibria of the extended game

Formally, the way to model this extended set up is straightforward. The voter can now choose how much attention to pay to the state of the world \( s \), and to \( x \). Similarly to section 2, define \( \tilde{s} \) as the state of the world known by the voter. Together with \( q \), she chooses \( \beta \), where \( \beta = Pr(\tilde{s}_t = s_t) \), while \( Pr(\tilde{s}_t = \emptyset) = 1 - \beta \). The cost function for \( \beta \) is defined by \( \tau G(\beta) \), with the same characteristics of \( \alpha C(q) \).

It is interesting to study the equilibrium choice of attention when \( \tau > 0 \), \( \alpha > 0 \) and commitment on \( \beta \) and \( q \) is not available. In this way, it is possible to explore the effect of both types of attention on pandering and also the relationship between those.

As before, define \( \gamma_s = Pr(C \text{ plays } B \text{ in state } s) \) and \( \lambda_s = Pr(D \text{ plays } B \text{ in state } s) \). Moreover, \( r_{\tilde{x}, \tilde{s}} = Pr(re - elect|\tilde{x}, \tilde{s}) \). All the PBNE of the game can be summarized as follows.

**Proposition 7** In every equilibrium with two types of attention

1. \( \gamma_B = 1, \lambda_B = F[\delta(1 + E)q] \);
2. \( \gamma_A = \lambda_A = F[\delta(1 + E)q(1 - \beta + \beta(r_{B,A} - r_{A,A}))] \).
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 7 contains several insights. First of all, there are multiple equilibria, depending on the randomization when the voter observes state \( A \) and the action. Secondly, the effect of \( q \) on pandering (i.e. on \( \lambda_A, \lambda_B \) and \( \gamma_A \)) is the same as before: more attention on the action makes pandering more desirable, for the politician. Finally, note that \( \beta \) 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.

4.2 Voter’s attention choice

It is now possible to solve for the attention choice of the voter, keeping in mind that it happens without knowing the politician’s action.

The voter’s problem is now to \( \max_{q \in [0,1], \beta \in [0,1]} W'' \) where
\[ W'' = q\beta \{ Pr(x = B, s = B)(\hat{\pi}_{B,B}V_C + (1 - \hat{\pi}_{B,B})V_D) + Pr(x = A, s = B)\Gamma \\
+ Pr(x = A, s = A)(r_{A,A}(\hat{\pi}_{A,A}V_C + (1 - \hat{\pi}_{A,A})V_D) + (1 - r_{A,A})\Gamma) \\
+ Pr(x = B, s = A)(r_{B,A}(\hat{\pi}_{B,A}V_C + (1 - \hat{\pi}_{B,A})V_D) + (1 - r_{B,A})\Gamma) \\
+ q(1 - \beta)\{ Pr(x = B)(\hat{\pi}_{B,\emptyset}V_C + (1 - \hat{\pi}_{B,\emptyset})V_D) + Pr(x = A)\Gamma \} \\
+ (1 - q)\Gamma - \alpha C(q) - \tau G(\beta) \]

Noticing that in equilibrium \( \hat{\pi}_{A,A} = \hat{\pi}_{B,A} = \pi \), this can be rewritten as

\[ W'' = q\beta \{ Pr(x = B, s = B)(\hat{\pi}_{B,B}V_C + (1 - \hat{\pi}_{B,B})V_D) \\
+ (1 - Pr(x = B, s = B))\Gamma \} \\
+ q(1 - \beta)\{ Pr(x = B)(\hat{\pi}_{B,\emptyset}V_C + (1 - \hat{\pi}_{B,\emptyset})V_D) + Pr(x = A)\Gamma \} \\
+ (1 - q)\Gamma - \alpha C(q) - \tau G(\beta) \]

Given the assumptions on the cost functions, the objective function is strictly concave and the solution is interior. Hence, it is enough to look directly at the the FOCs. Differentiating with respect to \( \beta \) and \( q \) while keeping in mind that \( \gamma_s \) and \( \lambda_s \) are already determined, the following conditions can be derived.

**Corollary 3** The equilibrium levels of attention \( q^* \) and \( \beta^* \) are implicitly defined by the following two equations:

\[ \left\{ \frac{1}{2}\Gamma + \frac{1}{2}\pi V_C + \frac{1}{2}\lambda_B^* (1 - \pi)V_D + \frac{1}{2}(1 - \pi)(1 - \lambda_B^*)\Gamma \right\} + \beta^* \frac{1}{2}\pi \delta[\pi(1 - \lambda_B^*) + \lambda_B^*] - \alpha C'(q^*) = 0 \]

and

\[ q^* \frac{1}{2}\pi \delta[\pi(1 - \lambda_B^*) + \lambda_B^*] - \tau G'(\beta^*) = 0 \]

where \( \lambda_B^* = F[\delta(1 + E)q^*] \).

Corollary 3, together with the complementarity between \( q \) and \( \beta \), leads to an additional interesting result. Under some conditions, a decrease in the cost of attention to the state of the world makes the equilibrium level of pandering of the consonant politician higher. For tractability, it is assumed that \( r_{B,A} = r_{A,A} \).
Proposition 8 There are parameters where a reduction in $\tau$ increases $\gamma^*_A$. In particular, the necessary conditions are

$$
\frac{0.5\pi\delta[\pi(1 - \lambda^*_B) + \lambda^*_B] + 0.5q^*\pi(1 - \pi)\delta^2(1 + E)f(\delta(1 + E)q^*)}{\tau G''(\beta^*)} < 
$$

and

$$
\frac{\alpha C''(q^*) - \left(\frac{\beta^*}{2} - \frac{1}{4}\right)\pi(1 - \pi)\delta^2(1 + E)f(\delta(1 + E)q^*)}{0.5\pi\delta[\pi(1 - \lambda^*_B) + \lambda^*_B]} < \frac{(1 - \beta^*)}{q^*}
$$

Despite the complicated conditions, the intuition behind Proposition 8 is quite straightforward. Basically, the point is that $q$ and $\beta$ are complements. Hence, decreasing $\tau$ increases $\beta^*$. But this also increases the desirability of $q$, increasing $q^*$ and, when this effect is strong enough, making pandering more likely. Moreover, note that I am looking at the type of pandering that is always welfare depressing, for the voter, because it is the pandering of the consonant incumbent.

5 Conclusion

Are social media and populist policies related? If populism can be seen as a form of pandering and social media are a technological innovation able to reduce the cost of attention, then this paper gives a positive answer. Moreover, from a policy perspective, it shows that a reduction in the cost of attention can be welfare depressing, precisely through its effect on the probability of pandering. Hence, at least on this respect, it is not obvious that easier/greater attention to politics is making the voters better off.

More generally, this paper is the first one looking at the effect of rational inattention on pandering, highlighting a novel trade off: higher attention improves political selection, but it also increases pandering. Hence, full attention is likely not to be welfare optimal even when it is costless, and the equilibrium level of attention is too high if compared with the $ex$ $ante$ optimal one. Interestingly, those results are robust to different specifications and different models of pandering.

Finally, in a more general model, this paper shows that an exogenous shock that makes attention to the state of the world cheaper can boost pandering as well, thanks to the complementarity of the two forces.

The current, very tractable set up is open to different possible extensions, suggesting many directions for future research. First of all, it is interesting to look at a model where the state of
the world is multidimensional. Those dimensions that are more salient for the voters are likely to imply more pandering from the politician, something that the optimal attention allocation should be taken into account.

Secondly, it is possible to microfound voters’ attention in a different way. One possibility is to use 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 contents (where it is easier to differentiate the product), this would shift voters’ attention away from political “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 it worsens political selection as well.

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 as a consequence they may be tempted to focus on simpler messages or on policies that do not present technical complications and can be easily explained. But those are precisely the policies where pandering incentives are stronger.

References


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 $\hat{\pi}_x$ is the posterior probability on the incumbent being consonant when the voter observes $x$.
Define $r_2 = \Pr(relect|\bar{x})$, $\gamma_A = \Pr(C\ plays\ B\ in\ state\ s)$ and $\gamma_B = \Pr(D\ plays\ B\ in\ state\ s)$. This equilibrium implies that, from the point of view of $V$, $\gamma_A = \lambda_A = \lambda_B = F[\delta(1+E)q] := \gamma < 1$ 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 the 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
$$EU_D(x = A) = u + E + \delta[q\lambda_A(1+E) + (1-q)r_\emptyset(1+E)]$$
while
$$EU_D(x = B) = E + \delta[q\lambda_B(1+E) + (1-q)r_\emptyset(1+E)]$$
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 $u \leq \delta(1+E)q$.
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:
$$EU_C(x = A, s = A) = u + E + \delta[q\lambda_A(1+E) + (1-q)r_\emptyset(1+E)]$$
and
$$EU_D(x = B, s = A) = E + \delta[q\lambda_B(1+E) + (1-q)r_\emptyset(1+E)]$$
and, again, it is optimal for the C incumbent to choose action $B$ in state $A$ when $u \leq \delta(1+E)q$. 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 $\hat{\pi}(\bar{x} = A) \geq \pi \Rightarrow \lambda_A + \lambda_B \geq \gamma_A + \gamma_B$ and $\hat{\pi}(\bar{x} = B) \leq \pi \Rightarrow \lambda_A + \lambda_B \geq \gamma_A + \gamma_B$.
Note that $\gamma_A = F[q\delta(1+E)(r_B-r_A)], \gamma_B = 1 - F[-q\delta(1+E)(r_B-r_A)], \lambda_A = \lambda_B = F[q\delta(1+E)(r_B-r_A)]$. Hence, $\lambda_A + \lambda_B \geq \gamma_A + \gamma_B$ is equivalent to $F[q\delta(1+E)(r_B-r_A)] + 1 - F[-q\delta(1+E)(r_B-r_A)] \leq 2F[q\delta(1+E)(r_B-r_A)]$.
This simplifies to $F[q\delta(1+E)(r_B-r_A)] \geq 1 - F[-q\delta(1+E)(r_B-r_A)]$. Note that, if $r_B \geq r_A$, $F[-q\delta(1+E)(r_B-r_A)] = 0$ and hence the equilibrium exists iff $F[q\delta(1+E)(r_B-r_A)] \geq 1$, which is impossible because $F$ is strictly increasing, $F[\bar{u}] = 1$ and $\bar{u} > \delta(1+E) > \delta(1+E)\bar{q}(r_B-r_A)$.
If instead $r_A > r_B$, $F[q\delta(1+E)(r_B-r_A)] = 0$ hence the equilibrium exists iff $F[q\delta(1+E)(r_A-r_B)] \geq 1$. Again, this is impossible because $\bar{u} > \delta(1+E) > \delta(1+E)\bar{q}(r_B-r_A)$.
Hence, there are no equilibria where $r_A > 0$ and $r_B < 1$. Applying the same logic, it is possible to rule out also 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 Lemma 1.
The second derivative of equation (1) with respect to \( q \) is

\[
\frac{\partial^2 W}{\partial q^2} = \pi(1 - \pi) \frac{1}{4} \left( -2 \frac{\partial \gamma}{\partial q} - \frac{\partial^2 \gamma}{\partial q^2} q \right) - \frac{1}{2} \pi \frac{\partial^2 \gamma}{\partial q^2}
\]

Noticing that \( \frac{\partial q}{\partial \gamma} = \delta(1 + E)f(\delta(1 + E)q) > 0 \), a sufficient condition for \( \frac{\partial^2 W}{\partial q^2} < 0 \) is \( \frac{\partial^2 q}{\partial \gamma} = (\delta(1 + E))^2 f'(\delta(1 + E)q) \geq 0 \), that requires \( f'(\delta(1 + E)q) \geq 0 \), as stated in the lemma. ■

Proof of Proposition 2.
First of all, note that, following Lemma 1, if \( f'(\delta(1 + E)q) \geq 0 \) when \( q \in [0, 1] \) then \( W \) is concave, hence it is possible to use the FOC. Hence, setting \( \frac{\partial W}{\partial q} = 0 \) and rearranging, the following equation is derived:

\[
(1 - \pi)(1 - F(\delta(1 + E)q^*)) = (1 - \pi)\delta q^* + 2(1 + E)f(\delta(1 + E)q^*)
\]

To solve for the fixed point of this problem, note that \( LHS(q = 0) = (1 - \pi) > LHS(q = 1) > 0 \) and that \( \frac{\partial LHS}{\partial q} < 0 \). Moreover, if \( f(0) = 0 \) then \( RHS(q = 0) = 0 < RHS(q = 1) \) and \( \frac{\partial RHS}{\partial q} > 0 \).

For an interior \( q^* \) one needs \( RHS(q = 1) > LHS(q = 1) \). Note that \( E \) increases \( RHS(q = 1) \) and decreases \( LHS(q = 1) \), hence it is always possible to find an \( E \) large enough that satisfies this condition. For example, a sufficient condition is \( RHS(q = 1) > LHS(q = 0) \), that requires \( f'(\delta(1 + E)) \geq \frac{1 - \pi}{(1 - \pi)\delta(1 + E)} \) ■

Proof of Proposition 3.
With few algebraic manipulation, \( W' = \frac{1}{2} \pi(1 - \pi)\delta(1 - \gamma)q + \Gamma - \alpha C(q) \). Since now \( \gamma \) has already been decided (i.e. it is a function of \( q^* \)), the solution is obtained differentiating only with respect to \( q \) and setting the first order conditions equal to 0. Note that, given the assumptions on \( C(q) \), the problem is concave. Finally, equation (4) is obtained by setting the equilibrium condition \( q^* = q^** \).

Uniqueness follows from the fact that the LHS is strictly decreasing in \( q \), the RHS is strictly increasing in \( q \), \( LHS(q = 0) > RHS(q = 1) \) and \( LHS(q = 1) < RHS(q = 1) \). ■

Proof of Proposition 4.
Recalling that the variance of the incumbent’s type is \( var(\theta) = \pi(1 - \pi) \), the implicit function describing \( q^** \) is defined as \( G(q^*, var(\theta), E, var(\theta), E) = var(\theta)\delta(1 - F[\delta(1 + E)q^*]) = 4\alpha C'(q^**) \).

By implicit function theorem,

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

\[
\frac{dq^{**}}{dE} = -\frac{var(\theta)\delta^2 q^{**}f(\delta(1 + E)q^{**})}{var(\theta)\delta^2(1 + E)f(\delta(1 + E)q^{**}) + 4\alpha C'(q^{**})} < 0
\]

■

Proof of Corollary 2.
Applying the chain rule,

\[
\frac{d\gamma^{**}}{dvar(\theta)} = [\delta(1 + E)f(\delta(1 + E)q^{**})] \frac{dq^{**}}{dvar(\theta)} > 0
\]

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

\[
\frac{d\gamma^{**}}{dE} = \delta \left[ q^{**} + \frac{dq^{**}}{dE}(1 + E) \right] f(\delta(1 + E)q^{**})
\]

Then, \( \frac{d\gamma^{**}}{dE} > 0 \Rightarrow 1 > \frac{(1 + E)\var(\theta)\delta^2 f(\delta(1 + E)q^{**})}{\var(\theta)\delta^2(1 + E)f(\delta(1 + E)q^{**}) + 4\alpha C'(q^{**})} \)

which is always true. ■
Proof of Proposition 5.
Note that, from an ex ante perspective, V’s welfare is expressed by \( W - \alpha C(q) \). Hence, the (ex ante) optimal level of attention (assuming the same conditions as in timing 1) is now implicitly defined by

\[
\pi(1 - \pi)\delta(1 - F(\delta(1 + E)q^{**})) = \pi((1 - \pi)\delta q^{**} + 2)\delta(1 + E)f(\delta(1 + E)q^{**}) + 4\alpha C'(q^{**})
\] (A.1)

where obviously \( q^{**} < q^* \).
Note that the \( LHS \) of (A.1) and (4) is exactly the same. However, the \( RHS \) of (A.1) is above the \( RHS \) of (4) for every interior \( q \), hence \( q^{**} > q^{***} \). ■

Proof of Proposition 6.
First \( q^{**} \) is implicitly expressed as a function of \( \alpha \), i.e. \( K(q^{**}(\alpha), \alpha) = \frac{1}{4}\pi(1 - \pi)\delta(1 - F(\delta(1 + E)q^{**}) - \alpha C'(q^{**}) \).

By implicit function theorem,

\[
\frac{\partial q^{**}}{\partial \alpha} = -\frac{\frac{\partial K}{\partial q} + \frac{\partial C}{\partial q} \delta q^{**}}{\pi(1 - \pi)\delta^2(1 + E)f(\delta(1 + E)q^{**}) + 4\alpha C''(q^{**})} < 0
\]

Defining the total (ex ante) welfare \( TW(q, \alpha) = W(q) - \alpha C(q) \), it is now possible to plug in \( q^{**} \) and to differentiate with respect to \( \alpha \).

Note that \( \frac{dTW}{d\alpha} = [\frac{\partial W}{\partial q^{**}} - \alpha \frac{\partial C}{\partial q^{**}}] \frac{\partial q^{**}}{\partial \alpha} - C(q^{**}) \).
The content in parenthesis is equal to zero at \( q = q^{***} < q^{*} \) and the function is concave, hence it is now negative. Moreover, \( \frac{\partial q^{**}}{\partial \alpha} < 0 \).

Substituting the parameters, the condition can be expressed as:

\[
\frac{\pi((1 - \pi)\delta q^{**} + 2)\delta(1 + E)f(\delta(1 + E)q^{**}) + 4\alpha C'(q^{**}) - \pi(1 - \pi)\delta(1 - F(\delta(1 + E)q^{**}))C'(q^{**})}{\pi(1 - \pi)\delta^2(1 + E)f(\delta(1 + E)q^{**}) + 4\alpha C''(q^{**})} > C(q^{**})
\]

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

\[
\frac{\pi((1 - \pi)\delta q^{**} + 2)\delta(1 + E)f(\delta(1 + E)q^{**})C'(q^{**})}{\pi(1 - \pi)\delta^2(1 + E)f(\delta(1 + E)q^{**}) + 4\alpha C''(q^{**})} > C(q^{**})
\]

■

Proof of Proposition 7.
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, \( r_{\tilde{z}, \tilde{s}} = Pr(relect|\tilde{z}, \tilde{s}) \).

First of all, the voter re-elects the incumbent if \( \hat{\pi}_{\tilde{z}, \tilde{s}} > \pi \), chooses the challenger if \( \hat{\pi}_{\tilde{z}, \tilde{s}} = \pi \) and is indifferent otherwise.

For given \( \lambda_s \) and \( \gamma_s \), it is easy to see that

\[
\hat{\pi}_{\phi, \tilde{s}} = \pi
\]

hence \( r_{\phi, \tilde{s}} \in [0, 1] \).

\[
\hat{\pi}_{A, \phi} > \pi \iff \lambda_A + \lambda_B > \gamma_A + \gamma_B
\]

hence

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

\[
\hat{\pi}_{B, \phi} > \pi \iff \lambda_A + \lambda_B < \gamma_A + \gamma_B
\]
Moving to the best response correspondences of \( P \), it is easy to see that, for conjectured levels of attention \( \beta^c \) and \( q^c \), he will compare \( EU_B(x = B, s) \) with \( EU_A(x = A, s) \). Hence,

\[
\lambda_B = Pr(u \leq \delta(1 + E)q^c [(1 - \beta^c)(r_{B,\theta} - r_{A,\theta}) + \beta^c(r_{B,A} - r_{A,A})]) \\
\lambda_A = Pr(u \leq \delta(1 + E)q^c [(1 - \beta^c)(r_{B,\theta} - r_{A,\theta}) + \beta^c(r_{B,B} - r_{A,B})]) \\
\gamma_B = Pr(u \leq \delta q^c(1 + E) [(1 - \beta^c)(r_{B,\theta} - r_{A,\theta}) + \beta^c(r_{B,B} - r_{A,B})]) \\
\gamma_A = Pr(u \leq \delta q^c(1 + E) [(1 - \beta^c)(r_{B,\theta} - r_{A,\theta}) + \beta^c(r_{B,A} - r_{A,A})]) \\
\]

(A.2) (A.3) (A.4) (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 obviously true for \( r_{B,3} \) since no relevant information is transmitted.

I claim that, in every equilibrium, it must be that \( r_{B,B} = 1, r_{A,B} = 0, r_{B,\theta} = 1 \) and \( r_{A,\theta} = 0 \). Given that \( \lambda_A = \gamma_A \), those are sequentially rational voting strategy if \( \gamma_B > \lambda_B \). Replacing in equations (A.5) and (A.3) the result, as required, is

\[
\gamma_B = 1 > \lambda_B = F[\delta(1 + E)q^c] \\
\]

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

\[
\gamma_A = \lambda_A = F[\delta(1 + E)q(1 - \beta + \beta(r_{B,A} - r_{A,A}))] \\
\]

In order to show that there are no other equilibria, I proceed by contradiction. First, suppose that there exists an equilibrium where \( r_{B,B} = 0, r_{A,B} = 1, r_{B,\theta} = 0 \) and \( r_{A,\theta} = 1 \). This requires \( \gamma_B < \lambda_B \) but, replacing in equations (A.5) and (A.3) the result is

\[
\gamma_B = Pr(u \geq \delta q^c(1 + E)) > 0 = \lambda_B \\
\]

21
hence a contradiction. Finally, suppose that $\gamma_B = \lambda_B$. Defining $\delta(1+E)q^* \left[ (1-\beta^*)(r_{B,B} - r_{A,B}) + \beta^*(r_{B,B} - r_{A,B}) \right] := D$, this requires $Pr(u \leq D) = Pr(u \geq -D)$, that never holds. 

**Proof of Corollary 3.**
The corollary follows directly from taking the FOCs on (6) and from Proposition 7. ■

**Proof of Proposition 8.**
First of all, note that in equilibrium $\gamma^*_A = F[\delta(1+E)q^*(1-\beta^*)]$, and that $q^*$ and $\beta^*$ are implicitly defined by (7) and (8).

By chain rule, $\frac{\partial \gamma^*_A}{\partial \tau} < 0$ implies

$$\frac{\partial \beta^*}{\partial \tau} q^* > \frac{\partial q^*}{\partial \tau} (1-\beta^*)$$

(A.6)

Then, the following system is obtained by total differentiation of (7) and (8) with respect to $\tau$, after some re-arrangement.

$$Q \frac{\partial q^*}{\partial \tau} + \frac{1}{2} \pi \delta [\pi + (1-\pi)\lambda^*_B] \frac{\partial \beta^*}{\partial \tau} = 0$$

$$Z \frac{\partial q^*}{\partial \tau} - \tau G''(\beta^*) \frac{\partial \beta^*}{\partial \tau} = G'(\beta^*)$$

where

$$Q = \pi(1-\pi)\delta^2(1+E)f(\delta(1+E)q^*) \left( \frac{1}{2} \beta^* - \frac{1}{4} \right) - \alpha C''(q^*)$$

and

$$Z = \frac{1}{2} \pi \delta [\pi + (1-\pi)\lambda^*_B] + \frac{1}{2} \pi(1-\pi)\delta^2(1+E)f(\delta(1+E)q^*)q^*$$

Solving the system, by implicit function theorem,

$$\frac{\partial q^*}{\partial \tau} = -\frac{\frac{1}{2} \pi \delta [\pi + (1-\pi)\lambda^*_B] G'(\beta^*)}{-Q\tau G''(\beta^*) - Z \frac{1}{2} \pi \delta [\pi + (1-\pi)\lambda^*_B]}$$

$$\frac{\partial \beta^*}{\partial \tau} = \frac{Q G'(\beta^*)}{-Q\tau G''(\beta^*) - Z \frac{1}{2} \pi \delta [\pi + (1-\pi)\lambda^*_B]}$$

Clearly, $Q < 0$ is necessary for (A.6) to be true. Moreover, the denominator has to be positive, hence it is needed that

$$-Q\tau G''(\beta^*) > Z \frac{1}{2} \pi \delta [\pi + (1-\pi)\lambda^*_B]$$

(A.7)

Finally, (A.6) has to be true once the values are plugged in.

After some manipulations, $Q < 0$ implies $\alpha C''(q^*) > \left( \frac{\beta^*}{\pi} - \frac{1}{4} \right) \pi(1-\pi)\delta^2(1+E)f(\delta(1+E)q^*)$, as in the proposition.

Moreover, $\frac{\partial \beta^*}{\partial \tau} q^* > \frac{\partial q^*}{\partial \tau} (1-\beta^*)$ can be simplified as

$$\frac{1}{2} \pi \delta [\pi + (1-\pi)\lambda^*_B] (1-\beta^*) > -Qq^*$$

(A.8)

The first condition of Proposition 8 is obtained considering together (A.7) and (A.8), and plugging in the values. ■
Appendix B  Maskin-Tirole (2004) set up

In this Appendix I present an alternative set up of the model, closely related with Maskin and Tirole (2004). It is easy to see that the main results of the paper are unchanged: higher attention implies a higher level of pandering and, if the effect of pandering on welfare is negative, the equilibrium level of attention is higher than the ex ante optimal one.

In this set up, the state of the world is \(s_t \in \{A, B\}\) with \(Pr(s = A) = p > \frac{1}{2}\). Hence, state A is the most popular one, according to the public opinion.

The action space for the politician is binary as well, with \(x_t \in \{A, B\}\). Finally, the politician can be of two types, Consonant (C) and Dissonant (D); formally, \(\theta = \{C, D\}\) 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 \(u_B + E\) if \(x_t = s_t\), \(E\) if \(x_t \neq s_t\). A dissonant incumbent, instead, derives utility from not matching the state. Formally, a type \(D\) incumbent gets \(u_B + E\) if \(x_t \neq s_t\), \(E\) if \(x_t = s_t\). 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 cdf \(F\) with support \([0, u]\). \(F\) is strictly increasing in the whole interval. Without loss of generality, it is assumed that \(E[u_t] = 1\). Moreover, it is assumed that \(u > \delta(1 + E)\), 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 \(\delta\).

I model rational inattention as the probability \(q\) that the voter observes the action of the politician before casting her choice. Formally, the voter does not observe \(x_t\) directly, but rather \(\hat{x}_t\), and she chooses \(q \in [0, 1]\) at a cost \(\alpha C(q)\), where \(Pr(\hat{x}_t = x_t) = q\) and \(Pr(\hat{x}_t = \emptyset) = 1 - q\), \(\alpha \geq 0\). The cost of attention function is strictly increasing and convex, and satisfies the usual Inada-type of conditions. Formally, \(C'(0) = 0, C'(1) = \infty, C'(q) > 0\ \forall q \in (0, 1)\) and \(C''(q) > 0\ \forall q \in (0, 1)\).

As above, \(r_{\hat{x}}\) defines the re-election probability, chosen by V, upon observing action \(\hat{x}\). Moreover, \(\gamma_s = Pr(C\text{ chooses }A\text{ in state }s)\) and \(\lambda_s = Pr(D\text{ chooses }A\text{ in state }s)\). In terms of timing, in this appendix I just compare the equilibrium choice of attention with the ex ante optimal one.

I claim that there exists a set of equilibria where, for every \(q, r_A = 1, r_B = 0, r_{\emptyset} \in [0, 1]\), \(\gamma_A = \lambda_B = 1\) and \(\gamma_B = \lambda_A = F(\delta q(1 + E))\).

To see this, note that \(\gamma_A = \lambda_B = 1\) are obvious, given that P faces no trade off between re-election and his favourite policy in period 1. Moreover, note that

\[
\hat{\pi}_A > \frac{(p + \gamma_B(1-p))\pi}{(p + \gamma_B(1-p))\pi + (\lambda_A p + 1-p)(1-\pi)}
\]

implies \((1 - \lambda_A)p > (1 - \gamma_B)(1-p)\), which is always true since \(\lambda_A = \gamma_B\) and \(p > \frac{1}{2}\). Similarly, \(\hat{\pi}_B < \pi\), hence the re-election strategies are optimal. Finally, note that

\[
\gamma_B = Pr(E + \delta(q(1+E) + (1-q)r_{\emptyset}(1+E)) \geq E + u + \delta(1-q)r_{\emptyset}(1+E)]
\]

that simplifies to \(\gamma_B = F(\delta q(1 + E))\), and the same reasoning applies to \(\lambda_A\). For simplicity, from now on I define pandering as \(\gamma = \gamma_B = \lambda_A = F(\delta q(1 + E))\). Note that the higher is \(q\), the higher is \(\gamma\).

Moving now to V’s equilibrium attention choice, noticing that it cannot affect \(\lambda_s\) and \(\gamma_s\) because they are function of the expected level of \(q\), it is easy to see that the maximization problem is equivalent to \(\max_{q \in [0,1]} W_{MT}^\gamma\), where

\[
W_{MT}^\gamma = q[Pr(x = B)]V_C' + Pr(x = A)(\hat{\pi}_AV_C' + (1 - \hat{\pi}_A)V_D'\]) + (1 - q)[1' - \alpha C(q)]
\]

where \(Pr(x = B) = (\pi(1-p)(1-\gamma) + (1-\pi)p(1-\gamma))\), \(Pr(x = A) = (\pi(p+(1-p)\gamma) + (1-\pi)((1-p)+p\gamma))\), \(V_C' = \delta, V_D' = 0\) and \(1' = \pi \delta\).

It is easy to show via FOC that the equilibrium attention choice of V is implicitly defined by

\[
\pi \delta (1 - \gamma)(1 - \pi)(2p - 1) = \alpha C'(q)
\]

Finally, since in equilibrium the expected level of attention must match the optimal one, \(q_{MT}^\gamma\) is unique, interior and implicitly defined by

23
\[
\pi \delta (1 - F(\delta q_{MT}^O(1 + E)))(1 - \pi)(2p - 1) = \alpha C'(q_{MT}^O) \tag{B.1}
\]

The ex ante optimal level of attention \(q_{MT}^O\) is the one that solves \(\max_{q \in [0, 1]} W_{MT}\), where

\[
W_{MT}' = \pi \{p(1 + qV_C') + (1 - p)[\gamma(0 + qV_C') + (1 - \gamma)(1 + q\Gamma')] + (1 - \pi)[p[\gamma(1 + qV_D') + (1 - \gamma)(0 + q\Gamma')] + (1 - p)(0 + qV_D')\} + (1 - q)\Gamma' - \alpha C(q)
\]

noticing that in this case \(\frac{\partial q}{\partial q}\) must be taken into account. Solving, it is easy to see that \(q_{MT}^O\) is unique, interior and implicitly defined by

\[
\pi \delta (1 - F(\delta q_{MT}^O(1 + E)))(1 - \pi)(2p - 1) = \alpha C'(q_{MT}^O) + \frac{\partial q}{\partial q}\bigg|_{q=q_{MT}^O} ((2p - 1)\delta q_{MT}^O (1 - \pi)\pi + \pi - p) \tag{B.2}
\]

It is interesting to compare (B.1) and (B.2). First of all, note that the LHS of both, expressed as a function of \(q\), is the same. Secondly, note that the RHS of (B.2) is the same as the one of (B.1) increased by \(\frac{\partial q}{\partial q}\bigg|_{q=q_{MT}^O} ((2p - 1)\delta q_{MT}^O (1 - \pi)\pi + \pi - p)\).

Noticing that \(\text{sign}(\frac{\partial q}{\partial q}) > 0\), it is clear that \(\pi > p\) is a sufficient condition for \(q_{MT}^O < q_{MT}^E\), i.e. to show that the equilibrium level of attention is inefficiently high. A weaker condition, noticing that \((2p - 1)\delta q(1 - \pi)\pi + \pi - p = 0\) is strictly increasing in \(q\), requires that \((2p - 1)\delta q(1 - \pi)\pi + \pi - p = 0\) for some \(q < q_{MT}^E\) or, equivalently, that \(q_{MT}^O > \frac{(2p - 1)\delta q(1 - \pi)\pi}{(2p - 1)\delta q(1 - \pi)\pi - p}\).

**Appendix C Continuum of voters with probabilistic voting**

In this appendix I show that the result is qualitatively robust to a model with a continuum of voters, as long as it is assumed that they vote sincerely. In the framework of the equilibrium choice of attention, assume that the representative voter is replaced by a continuum of voters uniformly distributed between 0 and 1. Each of them is like the representative voter, hence the optimal choice of attention is efficiently high. A weaker condition, noticing that (2p - 1)\(\delta q(1 - \pi)\pi + \pi - p = 0\) for some \(q < q_{MT}^E\) or, equivalently, that \(q_{MT}^O > \frac{(2p - 1)\delta q(1 - \pi)\pi - p}{(2p - 1)\delta q(1 - \pi)\pi - p}\).

Looking now at the incumbent decision, note that in this case the number of informed voters is a random variable, defined as \(\rho\). Without loss of generality, I focus on the dissonant incumbent. The expected utility from choosing action A is now given by

\[
\mathbb{E}U_D(x = A) = u + E + \delta \mathbb{E} \left[ \frac{1}{2} Pr\left(1 - \rho \left(\frac{1}{2} + \frac{1}{2}\tau\right) \geq \frac{1}{2}\right) + \frac{1}{2} Pr\left(1 - \rho \left(\frac{1}{2} + \frac{1}{2}\tau\right) \geq \frac{1}{2}\right) \right] (1 + E)
\]

The expectation is taken over \(\rho\). \(Pr((1 - \rho)(\frac{1}{2} + \frac{1}{2}\tau) \geq \frac{1}{2})\) is the probability of victory when the popularity shock is pro-challenger, noticing that informed people (i.e. a fraction \(\rho\) of the voters) are already voting for the incumbent with probability 0. Hence, the vote share among uninformed people is \((\frac{1}{2} + \frac{1}{2}\tau)\). However, note that \(Pr((1 - \rho)(\frac{1}{2} + \frac{1}{2}\tau) \geq \frac{1}{2}) = 0\). Similarly, \(Pr((1 - \rho)(\frac{1}{2} + \frac{1}{2}\tau) \geq \frac{1}{2})\) is the probability of victory in case of a pro-incumbent popularity shock, where the vote share from both set of voters is increased by \(\frac{1}{2}\tau\). By the distributional assumption on \(\tau\), \(Pr((1 - \rho)(\frac{1}{2} + \frac{1}{2}\tau) \geq \frac{1}{2}) = 1 - \rho\). As a consequence,

\[
\mathbb{E}U_D(x = A) = u + E + \delta \mathbb{E} \left[ \frac{1}{2} (1 - \rho) \right] (1 + E)
\]

Moving now to the expected utility from choosing action B, it is expressed as follows:
EU_D(x = B) = E + \delta E \left[ \frac{1}{2} \Pr \left( \frac{1}{2} \rho + \frac{1}{2} \tau \geq \frac{1}{2} \right) + \frac{1}{2} \Pr \left( \frac{1}{2} \rho + \frac{1}{2} \tau \geq \frac{1}{2} \right) \right] (1 + E)

Similarly to the previous case, \Pr(\rho + (1 - \rho) \frac{1}{2} - \frac{1}{2} \tau \geq \frac{1}{2}) is the probability of victory in case of a pro-challenger popularity shock (that negatively affects the incumbent vote share among both groups), and given the distributional assumption on \tau it is equal to \rho. Moreover, \Pr(\rho + (1 - \rho)(\frac{1}{2} + \frac{1}{2} \tau) \geq \frac{1}{2}) is the victory probability in case of a pro-incumbent popularity shock, noticing that the vote share among informed voters, since \(x = B\), is already 1 and cannot increase any further. Given the distributional assumption on \tau, this probability is 1. As a consequence,

EU_D(x = B) = E + \delta E \left[ \frac{1}{2} \rho + \frac{1}{2} \right] (1 + E)

Hence, action B is for \(u\) such that EU_D(x = B) \geq EU_D(x = A), i.e. whenever \(u \leq \delta (1 + E)\). Since \(\mathbb{E}(\rho)\) is just equal to the conjectured level of attention chosen by the voters, the problem is identical to the case of a single representative voter. Note that the same logic applies to the case of the consonant politician.