A Political Model of Trust

Marina Agranov, Ran Eilat, and Konstantin Sonin

APRIL 2020
A Political Model of Trust

Marina Agranov∗, Ran Eilat,† and Konstantin Sonin‡

April 25, 2020

Abstract

We analyze a simple model of political competition, in which the uninformed median voter chooses whether to follow or ignore the advice of the informed elites. In equilibrium, information transmission is possible only if voters trust the elites’ endorsement of potentially biased candidates. When inequality is high, the elites’ informational advantage is minimized by the voters’ distrust. When inequality reaches a certain threshold, the trust, and thus the information transmission, breaks down completely. Finally, the size of the elite forming in equilibrium depends on the amount of trust they are able to maintain.

Keywords: trust, inequality, political economy, cheap talk, information club.

JEL Classification: D72, D83.

∗Division of Humanities and Social Sciences, Caltech; magranov@hss.caltech.edu.
†Department of Economics, Ben-Gurion University; eilatr@bgu.ac.il.
‡Harris School of Public Policy, University of Chicago; ksonin@uchicago.edu.
Introduction

In the run-up to the US presidential elections in 2016, Hillary Clinton, the Democratic candidate, had an overwhelming advantage in editorial boards endorsements (240 to Trump’s 19), though they would typically tilt slightly Republican in presidential elections. Yet, poll after poll has demonstrated that the majority of voters finds Clinton “untrustworthy”. If we define trust as “the act of inviting someone to be in control of discretionary powers while relying on their goodwill” (Zoega, 2017), the pivotal voter rejected the elites’ advice, voting for and trusting in President Trump.\(^1\) Similarly, voting for Brexit in June 2016, British voters rejected the advice of their elite, which was predominantly pro-Remain (Inglehart and Norris, 2016).

In this paper, we suggest a simple political model that allows us to investigate the relationship between inequality and trust. The population consists of two groups, with the elite minority group, which forms endogenously, having an informational advantage over the majority. There are two politicians running for office: one of them is pro-elite and the other is unbiased; the candidates’ bias is common knowledge. Voters know that elite members have better information about the quality of political candidates, but are as interested as the elite is in electing the candidate of higher quality. Still, when the uninformed majority makes the choice, it takes into account that the elite is interested not only in the high quality of the candidate, but also in a candidate that is tilted towards elite’s interest. Thus, the majority’s trust (i.e., their willingness to follow the advice of the elite) plays a critical role: if there is no trust, elite’s endorsement is ignored, and valuable information is lost.\(^2\)

If cost of redistribution is not too high, in equilibrium the majority follows the elite’s endorsement. As in the classic model by Crawford and Sobel (1982), the equilibrium in which information is (partially) transmitted is welfare-improving. However, when inequality is relatively high, trust breaks down, and valuable information that the elite possesses is not transmitted. The median voter completely ignores the informed elite’s advice. In other words, the negative relationship between trust and inequality is driven by the information mechanism: because of the deadweight losses of taxation, the elite’s relative benefit is increasing in the level of inequality, which makes the elite’s endorsement less informative when inequality is high.

Across countries, there is a negative correlation between levels of political trust and inequality

---

\(^1\) Zoega (2017) further cites the philosopher Annette Baier, who defines trust as “accepted vulnerability to another’s possible but not expected ill (or lack of good will) toward one” (Baier, 1986).

\(^2\) Prato and Wolton (2016) consider a different source of information loss in pre-election communication. In their model, successful communication between candidates and voters during the pre-election campaign requires both an effort from the candidates and attention from voters, which depends on voters’ interest in politics and thus interacts with elections ability to screen politicians.
(a) Cross-country data
(b) United States data

Figure 1: Relationship between Trust and Inequality

in democracies. Figure 1 makes this point in two ways. Panel (1a) shows the negative correlation between inequality levels and trust in governments using data from the 20 most populated countries in Europe in 2017. Panel (1b) presents the evolution of political trust in institutions in the United States from 1981 to 2013, as captured by trust citizens have in parliament, the government, and political parties. The general decrease in trust over the past 40 years is accompanied by a well-documented and steady increase in inequality in the United States (see Piketty and Saez, 2003).

Eichengreen (2017) suggests that failure of trust is activated by a combination of economic insecurity, threats to national identity, and an unresponsive political system. Acemoglu and Robinson (2019) note: “the widespread perception that institutions have failed to address issues such as inequality has been eroding public trust in major institutions since the 1970s.” Guiso et al. (2018) analyse empirically the demand and supply of populist policies roughly defined as those that (a) promote the interests of the masses rather than the elites and (b) disregard the long-term consequences. In our model, (a) translates into electing the unbiased candidate even when informed elites support the other candidate, and (b) results in lower overall societal welfare as a result of (a). Guiso et al. (2018) show that populist policies are likely to emerge

---

3Trust data are taken from the Eurobarometer 88 database. The trust index is the percentage of people who “tend to trust” the national government in each country in 2017. GINI coefficient data and population data are taken from the Eurostat database for 2017. See Dustmann et al. (2017) for more illustrations.

4Trust data are taken from the World Values Survey, which is conducted every five years and asks respondents the following questions: “I am going to name a number of organizations. For each one, could you tell me how much confidence you have in them?” There are four possible answers: (a) A great deal, (b) Quite a lot, (c) Not very much, and (d) None at all. We plot the average fraction of respondents who answered either (a) or (b) when asked about parliament, the government and political parties.
when voters ‘lose faith’ in the institutions. Our model provides a micro-foundation for this logic. (See Prato and Wolton, 2018, on populism as political opportunism by incompetent politicians and Mattozzi, Morelli and Nakaguma, 2020, on effects of populist policies.)

Chakraborty and Ghosh (2016) consider a model of Downsian competition between two parties, in which voters that care about both the policy platform and “character” of candidates make a decision based on a media endorsement. The media has its own policy agenda and, though the voters know that the media’s endorsement is based solely on information about the candidate’s character, candidates in equilibrium pander to the media’s policy preferences. Our result on the breakdown of trust when inequality is high is similar to Chakraborty and Gosh’s flight to populism, when a large policy gap forces voters to ignore an endorsement. (Chakraborty and Yılmaz, 2017, apply this insight to information transmission between a corporate board and management.) In Kartik and van Weelden (2019), uncertainty generates reputationally-motivated policy distortions in office no matter what the policymaker’s true preference, so voters might prefer a “known devil to the unknown angel.” In our setting, a similar outcome occurs via a different mechanism when the pivotal voter ignores the recommendation of the elites and go with the unbiased politician, in which case valuable information is lost.

There is a substantial theoretical literature that focuses on the impact of third-party (e.g., media or special interest group) endorsements following the classic paper by Grossman and Helpman (1999). In our paper, there is no third party: the pivotal voter knows that the informed elite’s endorsement is biased, yet tries to take advantage of the information that is transmitted by the endorsement. In Martinelli (2006), voters decide whether or not to acquire information before making a choice. Myerson (2008) considers trust as an equilibrium phenomenon, but the context is very different: trust is what keeps the autocrat’s lieutenants abiding his command. In Egorov and Sonin (2011), the dictator ignores the advice of a competent subordinate as he expects the latter to be disloyal in equilibrium; this might be interpreted as a breakdown of trust.

The literature on inequality, including dynamics and implications for politics, is vast. We refer to the excellent synthesis of the early literature in Bénabou (1996). We parameterize inequality by the deadweight loss of redistributive taxation; this is a standard assumption in political economy literature (Acemoglu and Robinson, 2006).

---

5 As defined in Chakraborty and Ghosh (2016), “character” is similar to “valence” (Groseclose, 2001; Aragones and Palfrey, 2002; Banks and Duggan, 2005). Kartik and McAfee (2007) were the first to introduce voters’ uncertainty about valence. Bernhardt, Câmara and Squintani (2011) consider a dynamic citizen-candidates model with candidates that have both ideology and valence characteristics.

6 For other models of cheap talk in elections, see Harrington (1992), Panova (2017), Schnakenberg (2016), and Kartik, Squintani and Tinn (2015).
Finally, our paper is related to the literature on club formation (Tiebout, 1956; Roberts, 2015; Acemoglu, Egorov and Sonin, 2012). As Ray (2011) observes, the literature on endogenous formation of clubs that aggregate information is scarce. In our model, elites form endogenously, with the optimal size satisfying the natural club formation requirements: current members want neither to accept new members nor to expel any of the current ones. The novel feature is that the benefit of having more members is access to more precise information.

The rest of the paper is organized as follows. In Section 2, we introduce our model. Section 3 contains the analysis, while Section 4 concludes.

2 Setup

Consider a democratic society that consists of a large finite number of citizens, $N$. The citizens engage in two sequential interactions: First, they form two social groups, the Elites and the Commons. Second, they participate in a political game in which their interests depend on the group to which they belong. As part of the political game, information about the competence of politicians can be communicated from Elites to Commons. Whether this information affects the voting decisions of Commons defines the level of trust in the society. Our objective is to study the determinants of trust.

**Elite formation.** The formation of social groups is governed by the interests of Elites. Specifically, we assume that the Elites group size, $k$, is determined so as to maximize the utility of its members: (i) it is not worthwhile for Elite members to increase their group size above $k$ by accepting additional members, and (ii) it is not worthwhile for Elite members to decrease their group size by removing a member. All citizens who are not part of the Elites form the group of Commons. We denote the share of Elites in the citizenry by $\lambda = \frac{k}{N}$, and focus on the case that $\lambda < \frac{1}{2}$.

**The political game.** The citizens have to elect a politician to office. Once elected, the politician decides how to divide available resources between the two groups. Being a majority, Commons can unilaterally decide the identity of the elected politician. However, Elites have an advantage over Commons: with some probability, they are better informed regarding the competence of the candidates.

Each citizen acts to maximize his own per-capita share of the distributed resources. Since all citizens within the Elites get the same level of resources, and all citizens within the Commons get the same level of resources, we refer to the two groups as two players called Elites and
Politicians. There are two politicians who run for office and differ across two dimensions: their preferences and their ability to create resources for the economy. The preferences of each politician are observable. These preferences determine how the politician would distribute resources between Elites and Commons were he to be elected to office. We assume that one of the politicians, denoted by $U$, is unbiased and ascribes equal importance to the marginal per-capita consumption he allocates to each of the two groups. The other politician, denoted $B$, is biased towards the Elites. His level of bias is determined by a parameter $\alpha \in \mathbb{R}^+$ that is known to both Elites and Commons. The value of $\alpha$ proxies the strength of ties the biased politician shares with Elites relative to Commons, where larger values capture higher leniency towards Elites.

We denote by $a^j \in \{0, \alpha\}$ the level of bias of politician $j \in \{U, B\}$ and by $x^E \geq 0$ and $x^C \geq 0$ the per-capita consumption of Elites and Commons, respectively. Thus, the objective function of politician $j \in \{B, U\}$ is given by:

$$v(x^C, x^E) = (x^C + a^j)^{1-\lambda} (x^E)^{\lambda}.$$  

(1)

The functional form of (1) reflects a compromise between politician’s egalitarian and utilitarian motives: on the one hand, allocating higher per-capita consumption for each of the two groups is desirable. On the other hand, larger differences in allocation across groups are detrimental. Note that politicians do not gain material benefits from their decisions per se. Rather, the objective function reflects the utility gained from making the “right” decisions (e.g., due to legacy considerations) or from some future post-office perks. The objective function of the unbiased politician is sometimes referred to as the Nash collective utility function (see, e.g., Moulin, 2004, and Kaneko and Nakamura, 1979, for a discussion of some desirable properties of this function). The objective function of the biased politician is similar, with the only difference being that the importance of a marginal unit of Commons’ per-capita consumption is discounted, and this discount is stronger as $\alpha$ increases.

A politician’s competence to create resources depends on a state of the world $\theta$ that is uniformly distributed over $\Theta = [0, 1]$ and is unknown at the outset. We denote the competence of politician $j \in \{B, U\}$ by $\theta^j$ and assume that

$$\theta^B = 1 + \theta,$$
$$\theta^U = 2 - \theta.$$
Thus, the *ex ante* expected qualities of the two politicians are identical; that is, $E[\theta^B] = E[\theta^U] = \frac{3}{2}$. Furthermore, the biased politician is more competent than the unbiased one if and only if $\theta > \frac{1}{2}$, which happens with a probability one-half.

The politician in office distributes the available resources $\theta^j$ among the two groups such that

$$\lambda x^E + (1 - \lambda) x^C \cdot \psi = \theta^j, \quad (2)$$

where the parameter $\psi$ captures the cost of converting a unit of Elites’ consumption $x^E$ into a unit of Commons’ consumption $x^C$. We assume that $\psi = 1$ corresponds to the case of full equality; in this case, the constraint implies that the politician distributes resources among citizens such that the per-capita consumption in each group, weighed by the group’s size, averages to the politician’s competence $\theta^j$. The higher the inequality between the elite and the rest, the higher $\psi$ is. This assumption is equivalent to the standard assumption of efficiency losses in redistributive taxation; see, for example, Acemoglu and Robinson (2001).\(^7\) To simplify our analysis, we assume further that $\alpha \cdot \psi < 1.\(^8\)

Combining the objective function (1) and the constraint (2) yields the maximization problem that the politician in office solves:

$$\max_{(x^C, x^E)} \quad (x^C + \alpha^j)^{1 - \lambda} (x^E)^\lambda \quad \text{(OPT)}$$

$$\text{s.t.} \quad \lambda \cdot x^E + (1 - \lambda) \cdot x^C \cdot \psi = \theta^j$$

**Information structure.** At the outset, the state of the world $\theta$ is unknown (but the fact that it is drawn from a uniform distribution over $\Theta$ is common knowledge). Then, with some probability $p$, Elites can observe the state $\theta$ and learn the competence of both candidates. For simplicity, assume that each of $k$ Elites members learns the competence of both candidates with probability $q$, probabilities being independent across members, and shares it with fellow Elites. Then,

$$p(\lambda) = 1 - (1 - q)^{\lambda N}.$$ 

There is a remaining probability $1 - p$ that Elites remain uninformed. This information structure captures the general intuition that forming political views and assessing the future quality of

\(^7\)Acemoglu and Robinson (2001, 2006) simply assume that redistributive taxation results in welfare losses; that is, they make an assumption equivalent to ours. The microfoundation for this effect is the classic “no distortion at the top” result in contract theory (see, for example, Bolton and Dewatripont, 2005).

\(^8\)This assumption guarantees, for example, that the threshold defined in Lemma 1 is interior.
politicians—that is, their effectiveness if they are elected—is a complicated task that requires personal investments, expertise, and interaction with those who possess some private information; these interactions take the form of discussions and arguments among club members. Most of our results extend to the case of any increasing function $p(\lambda)$, where $\lambda$ is share of Elites in the citizenry.

**Endorsements and voting.** While Commons constitute the majority of the population and can effectively decide the identity of the elected politician, Elites are (potentially) better informed. It is therefore in the interest of both groups that information held by Elites is shared with Commons to ensure that the more competent politician is elected.

We assume that Elites cannot credibly share their information with Commons (in other words, they cannot prove that they know the state and what it is) and that they cannot commit *ex ante* to a strategy of information disclosure (for example, they cannot commit to not talking when being uninformed). Thus, information transmission between the two groups takes the form of cheap talk. Having observed the state of the world $\theta$, Elites can send costless and unverifiable messages to Commons, who update their beliefs about $\theta$ and elect their desired politician.

We denote by $M$ the set of possible messages that Elites can send, and assume without loss of generality that $M = \{m_B, m_U\}$. We interpret the message $m_B$ as an *endorsement for the biased politician* and the message $m_U$ as an *endorsement for the unbiased politician*. The strategy of Elites in the endorsement stage is denoted $\sigma_E: \Theta \cup \{\emptyset\} \rightarrow M$, where $\sigma_E(\theta)$ is the endorsement when the state $\theta \in \Theta$ is known to Elites, and $\sigma_E(\emptyset)$ is the endorsement if the state is unknown.

After Commons hear the endorsement, they update their posterior belief regarding the state of the world and the politicians’ competence, and vote for a politician. Thus, a strategy for Commons, denoted $\sigma_C: M \rightarrow \Delta\{B, U\}$, is a mapping between messages and distributions over votes (that is, for each message $m \in M$, the outcome $\sigma_C(m)$ is a lottery over the politician for which Commons vote). Since Commons constitute the majority, the politician for whom they vote is elected into office.

Finally, the elected politician distributes the available resources according to his preferences and competence in a way that maximizes his objective, as described in problem (OPT).

---

9Restricting attention to message sets with no more than two elements is without loss of generality in the following sense: for any equilibrium in the game, there exists another equilibrium in which Elites send at most two messages with positive probabilities such that the distribution over outcomes in both equilibria is the same for almost all states $\theta \in \Theta$ and for the case that Elites are uninformed. The basic idea is that messages that induce the same action for Commons in equilibrium can be “merged” in a way that does not affect the distribution over outcomes.
Our solution concept is the Perfect Bayesian Equilibrium, and in the voting stage, we assume that each citizen votes as if her vote is decisive (which is a weakly undominated strategy).

**Timing.** We conclude this section by recapping the timing of the game. For the analysis that follows, it will be useful to divide the timeline into two stages – the *formation stage* and the *political subgame* – as follows:

**Formation stage:**

1. The Elites’ group of size $k$ (which corresponds to the share $\lambda = k/N$) is formed; the size is optimal for Elite members.

**Political subgame:**

2. Nature determines the state of the world $\theta \in \Theta$. Elites learn the state $\theta$ with probability $p(\lambda) = 1 - (1 - q)^{\lambda N}$, or otherwise stay uninformed.

3. Elites choose which politician to endorse: either $B = (\theta^B, \alpha)$ or $U = (\theta^U, 0)$.

4. Commons observe the Elites’ endorsement and elect one of the two politicians.

5. The elected politician steps into office and distributes resources.

3 Analysis

We divide our analysis into two parts. We start by characterizing the equilibrium in the political subgame played by the Elites and Commons for a pre-determined size of the Elites’ share $\lambda = k/N$ (section 3.1). We then proceed by determining the optimal choice of $\lambda$, taking into account how this choice affects behavior and payoffs in the follow-up political subgame (section 3.2).

3.1 The Political Subgame

We solve the political subgame backwards. First, we derive the actions of the elected politician. Then, we find a pair of endorsement and voting strategies $(\sigma_E, \sigma_C)$ for Elites and Commons, respectively, that constitute an equilibrium in the subgame: each strategy is a best response to the other. We show that when Elites are informed, they use a cutoff strategy for endorsement, and we describe the conditions under which Commons are willing to accept the endorsement. Commons’ beliefs and the level of inequality play a critical role in this analysis.
Politician’s Actions. The actions of the politician elected into office depend on her type $(\theta^j, a^j)$ and are determined by the solution to problem (OPT). Specifically, a politician with type $(\theta^j, a^j)$ chooses $(x^E, x^C)$ such that:

$$
\begin{align*}
    x^E (\theta^j, a^j) &= \theta^j + (1 - \lambda) \cdot a^j \psi, \\
    x^C (\theta^j, a^j) &= \frac{\theta^j}{\psi} - \lambda \cdot a^j.
\end{align*}
$$

Thus, with full equality $(\psi = 1)$, the unbiased politician distributes resources equally among all members of the society, while the biased politician allocates a higher share of the available resources to the members of the Elites. When the level of inequality is positive $(\psi > 1)$, then even the unbiased politician allocates a higher share of resources to the Elites. This is a consequence of the assumption that redistribution is costly.

When the unbiased politician assumes office, the size of the Elites $\lambda$ does not affect allocations. On the other hand, if the biased politician is elected, a larger size of the Elites decreases the per-capita consumption of both Elites and Commons. For Commons, this is because the politician diverts a larger share of the available resources toward the Elites. For Elites, this is because the excess share of resources they get is divided over more group members. Comparative statics with respect to other parameters are as follows: a greater bias of the politician always increases the per-capita consumption of Elites, and decreases that of Commons. Higher inequality $(\psi)$ always decreases the per-capita consumption of Commons, but increases that of Elites only if the biased politician is in office.

Commons Trust and Elites Endorsement. Given a pair of strategies $(\sigma_E, \sigma_C)$, denote by $\sigma_C (m_B) [B]$ the probability that Commons elect the biased politician when Elites send $m_B$.\(^{10}\)

Similarly, denote by $\sigma_C (m_U) [B]$ the probability that Commons elect the biased politician when Elites send $m_U$. Since messages are cheap talk, it is without loss of generality to assume that message $m_B$ leads to a higher probability for electing $B$ than message $m_U$ does; that is, $\sigma_C (m_B) [B] \geq \sigma_C (m_U) [B]$.\(^{11}\)

An equilibrium $(\sigma_C, \sigma_E)$ is said to be responsive if Elites’ endorsements $m_B$ and $m_U$ induce

\(^{10}\)Recall that $\sigma_C (m_B)$ is a probability distribution over $\{U, B\}$. Thus $\sigma_C (m_B) [B]$ is the probability that the particular outcome $B$ is realized.

\(^{11}\)The assumption that $\sigma_C (m_B) [B] \geq \sigma_C (m_U) [B]$ is without loss of generality in the following sense: for any equilibrium in which $\sigma_C (m_B) [B] < \sigma_C (m_U) [B]$, there exists a corresponding equilibrium that satisfies $\sigma_C (m_B) [B] \geq \sigma_C (m_U) [B]$ in which, for each state $\theta \in \Theta$, the distribution over outcomes is identical to that of the original equilibrium. Transformation between the two equilibria is achieved by simply “renaming” the messages: whenever Elites sent $m_B$, they now send $m_U$, and vice versa. Also, whatever the Commons’ response to $m_B$ was, it is replaced by their response to $m_U$, and vice versa.
different distributions over Commons’ actions (i.e. if $\sigma_C(m_B) \neq \sigma_C(m_U)$). Otherwise, we call the equilibrium unresponsive. Our next result characterizes the strategy of the Elites in a responsive equilibrium (if such an equilibrium exists):

**Lemma 1** Suppose that $(\sigma_C, \sigma_E)$ is a responsive equilibrium. Then, the Elites’ strategy $\sigma_E$ attains the following threshold structure when Elites know the state $\theta$:

$$
\sigma_E(\theta) = \begin{cases} 
m_B & \text{if } \theta \geq \hat{\theta} \\
m_U & \text{if } \theta < \hat{\theta} 
\end{cases}
$$

where $\hat{\theta} \equiv \frac{1}{2} - \frac{1}{2}(1 - \lambda) \cdot \alpha \psi$. When Elites do not know the state, they always endorse the biased politician; that is, $\sigma_E(\emptyset) = m_B$.

A responsive equilibrium need not necessarily exist. On the other hand, there are always many unresponsive equilibria. In particular, Commons always electing the unbiased politician and Elites randomizing over messages constitutes one such unresponsive equilibrium. In the remainder of this section, we study the properties of a responsive equilibrium and look for necessary and sufficient conditions for its existence.

The condition described in Lemma 1 regarding the strategy of Elites in a responsive equilibrium can be equivalently described as follows: when Elites know the state, they endorse the biased politician (i.e., send $m_B$) if and only if $\theta^B \geq \theta^U - \mu^*$, where

$$
\mu^* = (1 - \lambda) \cdot \alpha \cdot \psi. \tag{5}
$$

Thus, Elites endorse the biased politician even if his competence is smaller than that of the unbiased one, so long as the difference does not exceed a threshold $\mu^*$. Note that a greater inequality level $\psi$, and a larger politician bias $\alpha$, increase the size of $\mu^*$. Intuitively, this is because *ceteris paribus*, the benefit for Elites of electing the biased politician is increasing in these quantities. On the other hand, a greater size of the Elites $\lambda$ decreases $\mu^*$. This is because, as discussed before, a greater size of the Elites decreases the per-capita consumption of each member of the Elites, thus weakening Elites’ incentive to endorse the biased politician.

In a responsive equilibrium, Elites’ endorsements convey information regarding the competence of politicians. The expected competence of each politician, conditional on being endorsed,
is computed as follows:

\[
\mathbb{E}[\theta^B|m_B] = \mathbb{E}[\theta^B|\theta^B > \theta^U - \mu^*] = \frac{7 - \mu^*}{4} \quad (6)
\]

\[
\mathbb{E}[\theta^U|m_U] = \mathbb{E}[\theta^U|\theta^B < \theta^U - \mu^*] = \frac{7 + \mu^*}{4}. \quad (7)
\]

Thus, when \(\mu^*\) increases (for example due to increased inequality), \(m_B\) gives a weaker indication of the competence of \(B\) while \(m_U\) gives a stronger indication of the competence of \(U\). At the same time, \(\mu^*\) also affects the probability of each endorsement in equilibrium:

\[
\Pr(m_B) = \Pr(\theta^B > \theta^U - \mu^*) = \frac{1 + \mu^*}{2}
\]

\[
\Pr(m_U) = \Pr(\theta^B < \theta^U - \mu^*) = \frac{1 - \mu^*}{2}.
\]

Therefore, as \(\mu^*\) increases, the probability of the less informative endorsement \(m_B\) being sent in equilibrium increases and the probability of the more informative endorsement \(m_U\) being sent decreases. The overall effect of \(\mu^*\) (and therefore also of the inequality \(\psi\)) on the expected competence of the endorsed politician is negative:

\[
\mathbb{E}[\theta|j \text{ is endorsed}] = \Pr(m_B) \cdot \mathbb{E}[\theta^B|m_B] + \Pr(m_U) \cdot \mathbb{E}[\theta^U|m_U] = \frac{7 - (\mu^*)^2}{4}. \quad (8)
\]

We now turn to finding the conditions under which Commons accept Elites’ endorsement in a responsive equilibrium \((\sigma_C, \sigma_E)\). Consider first the case that Elites endorse the biased politician, that is \(m = m_B\). Commons prefer to follow the recommendation whenever the expected benefits of electing the biased politician exceed the expected benefits of electing the unbiased one; that is,

\[
(1 - p) \cdot \mathbb{E}[x^C(\theta^B, \alpha)] + p \cdot \mathbb{E}[x^C(\theta^B, \alpha)|\theta^B \geq \theta^U - \mu^*] \geq (1 - p) \cdot \mathbb{E}[x^C(\theta^U, 0)] + p \cdot \mathbb{E}[x^C(\theta^U, 0)|\theta^B \geq \theta^U - \mu^*]. \quad (9)
\]

The first summand in the left-hand side of equation (9) is the expected benefit for Commons of electing the biased politician in the event that Elites were uninformed at the time of endorsement (thus, the computation is conducted using an unconditional mean). The second summand is the expected value of electing the biased politician in the event that Elites were informed at the time of endorsement. Similarly, the first and second summands of the right-hand side of (9) correspond to the payoffs from electing the unbiased politician when Elites were uninformed.
and informed at the time of endorsement, respectively.

Plugging equations (4), (5), and (6) into (9), and using the fact that $E[\theta^U|\theta^B \geq \theta^U - \mu^*] = (5 + \mu^*)/4$, yields the following equivalent condition for Commons accepting an endorsement for a biased politician:

$$\psi \leq \bar{\psi}(\lambda, \alpha) \equiv \frac{1}{\alpha} \cdot \frac{p(\lambda)}{p(\lambda) \cdot (1 - \lambda) + 2\lambda}. \quad (10)$$

Consider now the case that Elites endorse the unbiased politician; that is, $m = m_U$. In a responsive equilibrium, Elites never endorse the unbiased politician when they are uninformed and only do so when they are informed when $\theta^B < \theta^U - \mu^*$ (where $\mu^* \geq 0$). Thus, upon hearing an endorsement for the unbiased politician, Commons deduce that his expected competence is greater than that of the biased politician (that is, $E[\theta^U|\theta^B < \theta^U - \mu^*] \geq E[\theta^B|\theta^B < \theta^U - \mu^*]$). Since, in addition, the unbiased politician distributes resources more equally, Commons always accept an endorsement for the unbiased politician.

Thus, when inequality level $\psi$ exceeds the threshold $\bar{\psi}(\lambda, \alpha)$, a responsive equilibrium does not exist. In other words, in all equilibria of the game, Commons have no political trust in Elites and disregard their advice. This is despite the fact that Elites potentially possess valuable information regarding the competence of the candidates.

In contrast, when the inequality level is lower than $\bar{\psi}(\lambda, \alpha)$, a responsive equilibrium exists. In this equilibrium, Commons follow Elites’ endorsement despite the fact that sometimes Elites recommend a biased politician of lower quality than the unbiased one. Note that since the endorsed politician is also the elected one, then by equation (8) we have that the expected competence of the elected endorsed politician is decreasing in $\mu^*$. Thus, greater inequality erodes the quality of information that is transmitted in equilibrium, leading to the election of politicians who create less resources.

The following Proposition summarizes the above discussion.

**Proposition 1** For any size of the elite $\lambda$ and any bias of the elite candidate $\alpha$, there exists an inequality threshold $\bar{\psi}(\lambda, \alpha)$ such that if $\psi > \bar{\psi}(\lambda, \alpha)$, then Commons disregard Elites’ endorsements and always elect the unbiased politician. If $\psi \leq \bar{\psi}(\lambda, \alpha)$, there exists a responsive equilibrium: Elites recommend the biased politician if and only if they are uninformed or they are informed and $\theta^B > \theta^U - \mu^*$. Commons always accept Elites’ endorsements.

Proposition 1 demonstrates the crucial role that inequality plays in determining the extent of equilibrium information transmission. When inequality is low, Commons tolerate the informational distortions that accompany Elites’ endorsements and accept the recommendations.
When inequality is high, trust breaks and Commons disregard endorsements despite the fact that these recommendations convey informational content. This negative correlation between the equilibrium level of trust that Commons have in the establishment (Elites) and inequality is consistent with the evidence described in the Introduction.

Proposition 1 also allows us to analyze how the politician’s bias \( \alpha \) and the size of the Elites club \( \lambda \) affect the level of trust that transpires in the political game. For the parameter \( \alpha \), the effect is straightforward: when the biased politician is more ‘Elites-oriented’ – that is, when \( \alpha \) is larger – threshold \( \tilde{\psi}(\lambda, \alpha) \) decreases, making Commons less receptive to endorsements. Intuitively, this is because a greater \( \alpha \) decreases Commons’ per-capita consumption when they accept an endorsement to elect the biased politician. Moreover, it also increases the threshold \( \mu^* \), thereby intensifying the distortion of Elites’ endorsements. Thus, greater politician bias is detrimental for trust.

The impact of the Elites share \( \lambda \) is more subtle. First, a greater \( \lambda \) implies lower per-capita consumption for Commons when the biased politician is endorsed and elected. Second, a larger \( \lambda \) increases the probability of Elites being informed, making their endorsements more valuable for Commons. Finally, a larger \( \lambda \) decreases the threshold \( \mu^* \), thereby reducing the distortion of Elites’ endorsements. While the first effect is trust-eroding, the second and third effects are trust-enhancing. The following lemma asserts that the overall effect is negative – a larger size of Elites makes Commons less receptive to endorsements.

**Lemma 2** The inequality threshold \( \tilde{\psi}(\lambda, \alpha) \) defined in (10) is decreasing in \( \alpha \) and in \( \lambda \).

Finally, we can calculate the expected utilities of Elites and Commons in equilibria with and without trust. First, when \( \psi > \tilde{\psi} \) (no trust), Commons always elect the unbiased politician. Thus, the expected utilities of Elites and Commons are given by:

\[
\begin{align*}
\mathbb{E} \left[ x^E \mid \psi > \tilde{\psi} \right] &= x^E \left( \mathbb{E} \left[ \theta^{U} \right], 0 \right) = \frac{3}{2} \\
\mathbb{E} \left[ x^C \mid \psi > \tilde{\psi} \right] &= x^C \left( \mathbb{E} \left[ \theta^{U} \right], 0 \right) = \frac{3}{2\psi}.
\end{align*}
\]

The expected utilities when \( \psi < \tilde{\psi} \) and Commons accept endorsements can be written as
follows:\textsuperscript{12}
\begin{align*}
E \left[ x_E \mid \psi < \bar{\psi} \right] &= x_E \left( E \left[ \theta^B \mid \alpha \right], \alpha \right) + p(\lambda) \cdot \left( \frac{1 - (\mu^*)^2}{4} - \frac{1 - \mu^*}{2} \cdot (1 - \lambda) \alpha \psi \right) \\
E \left[ x_C \mid \psi < \bar{\psi} \right] &= x_C \left( E \left[ \theta^B \mid \alpha \right], \alpha \right) + p(\lambda) \cdot \left( \frac{1 - (\mu^*)^2}{4\psi} + \frac{1 - \mu^*}{2} \cdot \alpha \lambda \right)
\end{align*}
(13) (14)

The expected utility of each group consists of two terms. The first terms on the right-hand side of (13) and (14) are the base per-capita consumption of Elites and Commons, respectively, when Elites are uninformed and endorse the biased politician even though his quality is unknown (this endorsement is accepted by the Commons). The second term corresponds to the extra per-capita consumption that the group gains when Elites are informed and endorse a politician (an event that happens with probability $p(\lambda)$).

We conclude this section by briefly discussing how Elites could affect their payoff in the political subgame if, prior to observing the state, they could choose the bias level of "their" politician, $\alpha$. To see the answer, note that, conditional on equilibrium being responsive, Elites' expected payoff is increasing in $\alpha$.\textsuperscript{13} Therefore, Elites have an incentive to increase the bias level so long as it does not break trust.

Put differently, if Elites have access to a pool of candidates with different levels of $\alpha$, they choose to promote the political career of the candidate with the highest bias among those whose level of bias satisfies

$$\alpha \leq \tilde{\alpha} \equiv \frac{p(\lambda)}{(1 - \lambda)p(\lambda) + 2\lambda} \cdot \frac{1}{\psi},$$

where $\tilde{\alpha}$ is the level of bias which makes equation (10) bind in equality. Thus, when Elites can choose the bias level of their candidate they always preempt the breakdown of trust. Of course, this is only possible when the chosen candidate’s bias is public knowledge. (In Kartik and van Weelden, 2019, politicians strategically use cheap talk to signal their bias; in Acemoglu, Egorov and Sonin, 2013, they have to adopt populist policies to signal their unbiasedness.)

\textsuperscript{12}To see this, write $E \left[ x_E \mid \psi < \bar{\psi} \right] = (1 - p) \cdot \left( E \left[ \theta^E \left( E \left( \theta^B \mid \alpha \right), \alpha \right) \right] + p \cdot Pr \left[ \theta^B > \theta^U - \mu \right] \cdot \left( x_E \left( E \left[ \theta^B \mid \theta^B > \theta^U - \mu \right] \right) \right] + p \cdot Pr \left[ \theta^B < \theta^U - \mu \right] \cdot \left( x_E \left( E \left[ \theta^U \mid \theta^B < \theta^U - \mu \right] \right) \right]$, where the first term in the sum corresponds to the case that uninformed Elites endorse the biased politician, the second term corresponds to the case that informed Elites endorse the biased politician, and the third term corresponds to the case that informed Elites endorse the unbiased politician. Substituting $Pr \left[ \theta^B > \theta^U - \mu \right] = 0.5 (1 + \mu)$ and $E \left[ \theta^B \mid \theta^B > \theta^U - \mu \right] = 0.25 (7 - \mu^*)$ and $E \left[ \theta^U \mid \theta^B < \theta^U - \mu \right] = 0.25 (7 - \mu^*)$ yields (13). The derivation of (14) is similar.

\textsuperscript{13}To see this, plug equations (5) and (3) into equation (13) and simplify to get: $E \left[ x_E \mid \psi < \bar{\psi} \right] = \left( \frac{3}{2} + (1 - \lambda) \alpha \psi + p \cdot \frac{1}{4} (1 - (1 - \lambda) \psi \alpha)^2 \right)$.
3.2 The Optimal Size of the Elites

In this subsection, we analyze the optimal size of the Elites if a responsive equilibrium exists. If information transmission is impossible in equilibrium, then the size of the Elites does not matter: the Commons’ majority will always elect the unbiased candidate.

Equation (5) gives us the value of the Elites endorsement threshold, \( \mu^* \), as a function of Elites’ share \( \lambda \); that is, \( \mu^* (\lambda) = (1 - \lambda) \cdot \alpha \cdot \psi \). In a responsive equilibrium, trust is maintained, and therefore the Elites’ expected utility in the political subgame is given by equation (13), which can be re-written as follows:

\[
\begin{align*}
    u^E (\lambda) &\equiv E \left[ x^E \mid \psi < \bar{\psi} \right] = B (\lambda) + p (\lambda) \cdot V (\lambda),
\end{align*}
\]  

(15)

where \( B (\lambda) = \frac{3}{2} + (1 - \lambda) \alpha \psi \) is the base value of Elites’ payoff and \( V (\lambda) = \frac{1}{4} (1 - (1 - \lambda) \alpha \psi)^2 \) is the Elites’ value of information. Thus, the payoff for an Elites member when Elites are uninformed is given by the base value \( B (\lambda) \). When Elites are informed, the payoff is the base value plus the value of information (i.e. \( B (\lambda) + V (\lambda) = \frac{3}{2} + \frac{1}{4} ((1 - \lambda) \alpha \psi + 1)^2 \)).

While our results are formulated in terms of club size, \( k \), it is easier to discuss the underlying mechanism using the continuous functions of the share \( \lambda = \frac{k}{N} \). The derivative of \( u^E (\lambda) \) with respect to \( \lambda \) is given by:

\[
\begin{align*}
    (u^E)' (\lambda) &= p' (\lambda) \cdot V (\lambda) + p (\lambda) \cdot (B' (\lambda) + V' (\lambda)) + (1 - p (\lambda)) \cdot (B' (\lambda)) .
\end{align*}
\]  

(16)

This derivative consists of three summands. The first summand, which is always positive, captures the change in the expected payoff for Elites associated with the change in the probability of being informed. The second and third summands, which are both negative, capture the changes in Elite’ per-person shares due to the change in \( \lambda \) when Elites are informed and uninformed, respectively. The optimal club size depends on the relative weights of these three effects.

The existence of equilibrium follows from the fact that the Elites’ utility function is single-peaked:

**Lemma 3** For sufficiently large \( N \), the Elites’ utility function \( u^E (\lambda) \) defined in (15) is single-peaked over \([0, \frac{1}{2}]\).

Since all agents are symmetric *ex ante*, Lemma 3 guarantees, generically, the existence of an equilibrium size of the Elites. First, note that if a continuous function \( u^E (\lambda) \) is single-peaked over a domain, then a discrete function \( u^E (k) \) has at most two maxima; in a generic case, it
has a unique maximum. Now, suppose that $k^*$ is this maximum, and the club of $k^*$ members has been formed. Clearly, this club satisfies our equilibrium criteria regardless of the decision-making rule within the club. Every member would prefer neither to accept any more members nor to expel anyone.

Of course, Lemma 3 does not guarantee uniqueness of a stable club. One reason for non-uniqueness is familiar for students of club formation: the instability of a subcoalition makes a large coalition stable (e.g., Acemoglu, Egorov and Sonin, 2012). In our case, suppose that decisions about club membership are accomplished by majority voting, $k^* < \frac{1}{4}$, and suppose that a club of size $2k^*$ is formed. First, observe that this club will not admit any more members as the utility function of each member is single-peaked. Therefore, increasing membership brings down the utility for each member. Second, there will be at least $k^*$ members who would not agree to the removal of a single Elites member. Indeed, if at least one member from the $2k^*$-sized Elites is removed, there is a coalition of $k^*$ members who have the majority to remove the remaining $k^* - 1$ members. Thus, there is a blocking coalition of $k^*$ members that make the $2k^*$-sized Elites stable.\footnote{This argument is admittedly heuristic, as we have not specified any game that leads to Elites formation. Still, given the equilibrium of the continuation game, the payoffs that citizens have \emph{ex ante} satisfy the conditions for a non-cooperative club formation game in Acemoglu, Egorov, and Sonin (2012). Thus, our argument can be made formal at the cost of introducing additional game-theoretic machinery.}

An Elites group that consists of $k^*$ members is a natural outcome of the elite-formation process: this is the club that forms if formation starts, naturally, from the club consisting of one member. The following Proposition 2 states the existence result formally.

**Proposition 2** For sufficiently large $N$, if the continuation payoffs are given by the responsive equilibrium, the Elites form a stable club of size $k^*$ at the elite formation stage.

**The Responsiveness of the Optimal Club Size.** Proposition 2 asserts that there exists a unique (generically) optimal size for the Elites, $k^*$, assuming that the responsive equilibrium is played in the continuation game. It remains to be demonstrated that if the combination of fundamentals is such that a responsive equilibrium is possible, then $k^*$ satisfies the requirements of the responsive equilibrium threshold. (Indeed, as the inequality threshold determined in (10) is a function of the Elite’ size, it is not obviously true that a responsive equilibrium is possible with the Elites of size $k^*$.\footnote{This argument is admittedly heuristic, as we have not specified any game that leads to Elites formation. Still, given the equilibrium of the continuation game, the payoffs that citizens have \emph{ex ante} satisfy the conditions for a non-cooperative club formation game in Acemoglu, Egorov, and Sonin (2012). Thus, our argument can be made formal at the cost of introducing additional game-theoretic machinery.} Through the rest of the argument, we will assume that $N$ is sufficiently large to guarantee the existence of the optimal $k^*$.

We start with a technical observation that by Lemma 2, the threshold $\tilde{\psi}(\lambda, \alpha)$ decreases in $\lambda$, so that a smaller size of the Elites makes Commons more tolerant to inequality. Now, we can
define the possibility of avoiding the breakdown of trust. Specifically, if \( \lim_{\lambda \to 0} \bar{\psi}(\lambda, \alpha) > \psi \), we say that the breakdown of trust is avoidable.

**Proposition 3** If the breakdown of trust is avoidable – that is, \( \lim_{\lambda \to 0} \bar{\psi}(\lambda, \alpha) > \psi \) – then it is avoided in equilibrium: the optimal Elites size \( k^* \) is such that \( \bar{\psi}(k^* / N, \alpha) > \psi \), and the equilibrium in the political game is responsive.

Proposition 3 guarantees that there may be two types of situations. In one, the costs of progressive redistribution, \( \psi \), are so high that there exists no responsive equilibrium. In this case, there is no information transmission between Elites and Commons and, therefore, the Elites can be of any size. The other case is that the breakdown of trust is avoidable and an Elites club forms with the responsive equilibrium in sight. Example 1 illustrates this logic.

**Example 1** When the costs of progressive redistribution are sufficiently high, Elites cannot avoid the breakdown of trust. Let \( q = 0.005, N = 10000, \) and \( \alpha = 1. \) Then, \( \lim_{\lambda \to 0} \bar{\psi}(\lambda, 1) \approx 0.96 \), so if \( \psi = 0.97 \), then the breakdown of trust is unavoidable. At the same time, if \( \psi = 0.75 \), then trust is supported in the responsive equilibrium for any \( \lambda < \frac{1}{3} \). The optimal Elites size is \( \lambda^* \approx 0.001 \), which satisfies \( \bar{\psi}(0.001, 1) > 0.75 \).

**The Comparative Statics.** Once we have established that an optimal equilibrium size of the Elites exists, a natural question is: what is the effect of inequality on the optimal size? Proposition 4 provides comparative statics results.

**Proposition 4** The optimal size of the Elites club \( k^* \) is decreasing in the bias \( \alpha \) of the pro-elite candidate and in the cost of progressive redistribution \( \psi \).

The comparative statics results of Proposition 4 are intuitive: when the politician’s bias, \( \alpha \), is high, the marginal contribution of each Elite member to the club is low. Indeed, let us examine how \( \alpha \) affects the derivative \( (u^E)'(\lambda) \) in equation (16). As above, although our results are formulated in terms of club size, \( k \), the underlying intuition is easier to discuss using the continuous functions of the share \( \lambda = \frac{k}{N} \). Notice first that increasing \( \alpha \) intensifies the negative effect of the elite share, \( \lambda \), on Elites’ per-person consumption both when they are informed and when they are uninformed. Indeed, increasing the bias \( \alpha \) decreases \( B'(\lambda) = -\alpha \psi \) and decreases \( B'(\lambda) + V'(\lambda) = -\frac{1}{2} \alpha \psi ((1 - \lambda) \alpha \psi + 1) \). The critical element is the breakdown of trust: with higher bias, the range of parameters for which Commons follow the Elites’ endorsement narrows. Furthermore, increasing \( \alpha \) also decreases the value of information, \( V(\lambda) \). Thus, increasing the
politician’s bias decreases the value of \((u^E)'(\lambda)\) for every \(\lambda\). Since the function \(u^E\) is single-peaked, and consequently the derivative \((u^E)'(\lambda)\) crosses the \(x\)-axis from above, the value \(\lambda^*\) for which \((u^E)'(\lambda^*) = 0\) necessarily decreases.

The effect of an increase in the cost \(\psi\) of progressive redistribution is similar. As with \(\alpha\), increasing \(\psi\) makes the negative effect of the Elites share, \(\lambda\), on Elites’ per-person share stronger in both cases: when Elites are informed and when they are uninformed. Indeed, for the uninformed Elites, we have \(\frac{\partial^2}{\partial \lambda \partial \psi} B(\lambda) = -\alpha < 0\); for the informed ones, we have \(\frac{\partial^2}{\partial \lambda \partial \psi} (B(\lambda) + V(\lambda)) = -\frac{1}{2} \alpha ((1 - \lambda) \alpha \psi + 1) - \frac{1}{2} \alpha (1 - \lambda) \alpha \psi < 0\). Finally, \(\frac{\partial}{\partial \psi} V(\lambda) = -\frac{1}{2} (1 - \lambda) \alpha (1 - (1 - \lambda) \alpha \psi) < 0\).

Thus, when the measure of inequality increases, the value of \((u^E)'(\lambda)\) is lower for every \(\lambda\). For the single-peaked function \(u^E\), the derivative \((u^E)'(\lambda)\) crosses the \(x\)-axis from above, and the optimal size of the club \(\lambda^*\) decreases. Put simply, higher inequality results in a lower trust, which decreases the value of information that a potential member of Elites contributes, and thus the optimal Elites size.

4 Conclusion

Recently, there has been a noticeable decline in trust, both as measured by opinion polls and by surges of support for anti-elite, populist politicians and parties. To analyze the link between rising inequality and this decline in trust, we construct a political model in which the elite has an information advantage over the rest of society, and the median voter chooses the president using the elite’s endorsement. When inequality is relatively low, the interests of the elite and median voter in electing a competent leader are aligned, and valuable information is successfully transmitted in equilibrium. In contrast, when inequality is relatively high, there is a complete breakdown of trust, which results in no information transmitted from the elites to the median voter.
References


Appendix

Proof of Lemma 1

Suppose that \((\sigma_C, \sigma_E)\) is a responsive equilibrium, and suppose that Elites are informed. Then, endorsing the biased politician is a best response for Elites in state \(\theta \in \Theta\) whenever:

\[
\sigma_c(m_B)[B] \cdot x_E(\theta^B, \alpha|\theta) + (1 - \sigma_c(m_B)[B]) \cdot x_E(\theta^U, 0|\theta) 
\geq \sigma_c(m_U)[B] \cdot x_E(\theta^B, \alpha|\theta) + (1 - \sigma_c(m_U)[B]) \cdot x_E(\theta^U, 0|\theta).
\]

Plugging in the expressions for \(x_E(\theta^B, \alpha|\theta)\) and \(x_E(\theta^U, 0|\theta)\) from equation (3) and substituting \(\theta^B = 1 + \theta\) and \(\theta^U = 2 - \theta\) yields:

\[
(\sigma_c(m_B)[B] - \sigma_c(m_U)[B]) \cdot (2\theta - 1) 
\geq - (\sigma_c(m_B)[B] - \sigma_c(m_U)[B]) \cdot (1 - \lambda) \cdot \alpha \psi.
\]

In a responsive equilibrium, \(\sigma_c(m_B)[B] > \sigma_c(m_U)[B]\), and the above inequality is satisfied for all \(\theta > \hat{\theta} \equiv \frac{1}{2} - \frac{(1-\lambda) \cdot \alpha \psi}{2}\). Thus, \(\hat{\theta}\) is the state above which endorsing \(B\) is optimal for Elites and below which endorsing \(U\) is optimal.

When Elites are uninformed, they seek to maximize the probability that the biased politician is elected. Since in a responsive equilibrium, \(\sigma_c(m_B)[B] > \sigma_c(m_U)[B]\), then endorsing \(B\) is optimal. ■

Proof of Lemma 2

Inspection of equation (10) immediately reveals that \(\bar{\psi}\) is decreasing in \(\alpha\). Therefore, to prove the lemma, it only remains to inspect how \(\lambda\) affects \(\bar{\psi}\). To do this, let us treat \(\lambda\) as a continuous variable and show that \(\frac{\partial \bar{\psi}}{\partial \lambda} > 0\). This implies that \(\bar{\psi}\) is decreasing in \(\lambda\) also when \(\lambda\) is discrete.

We compute the partial derivative \(\frac{\partial \bar{\psi}}{\partial \lambda}\) and write it as follows:

\[
\frac{\partial \bar{\psi}}{\partial \lambda} = \frac{1}{\alpha} \cdot \frac{(1-q)^{N\lambda} \left( (1-q)^{N\lambda} - 2 \ln (1-q)^{N\lambda} \right) - 1}{\left( \lambda + \lambda (1-q)^{N\lambda} - (1-q)^{N\lambda} + 1 \right)^2}.
\]

Denote \(s \equiv (1-q)^{N\lambda}\), so that the numerator can be written as \(g(s) \equiv s(s - 2 \ln s) - 1\). Note that \(s\) is a positive number that is smaller than 1 (it goes to 1 as \(\lambda \to 0\)). We now show that \(g(s) < 0\) for all values of \(s \in (0,1)\).
Begin by taking the first, second, and third derivatives of $g(s)$:

\[
\begin{align*}
g'(s) &= 2s - 2\ln s - 2, \\
g''(s) &= 2 - \frac{2}{s}, \\
g'''(s) &= \frac{2}{s^2}.
\end{align*}
\]

The function $g'(s)$ attains a minimum at $s = 1$, where its value is zero; i.e., $g'(1) = 0$. To see why, note that $g''(1) = 0$ and $g'''(s)$ is positive (hence, $g'(s)$ is a convex function). Thus, $g'(s) > 0$ for all values of $s \in (0, 1)$, and the function $g(s)$ is increasing on $s \in (0, 1)$. Since $g(1) = 0$, it immediately follows that $g(s) < 0$ for all values of $s \in (0, 1)$.

**Proof of Lemma 3**

The function $u^E(\lambda)$ is continuously differentiable. Its first, second, and third derivatives are given by:

\[
\begin{align*}
\frac{d}{d\lambda} u^E(\lambda) &= -\frac{1}{4} \cdot (1 - q)^\lambda \cdot \left( \left( \ln (1 - q)^\lambda \right) \cdot (1 - \alpha \psi (1 - \lambda)) + 2\alpha \psi \right) \cdot (1 - \alpha \psi (1 - \lambda)) \\
&\quad - \frac{1}{2} \alpha \psi (\alpha \psi - \alpha \lambda \psi + 1) \\
\frac{d^2}{(d\lambda)^2} u^E(\lambda) &= -\frac{1}{4} \cdot (1 - q)^\lambda \cdot \left( \left( \ln (1 - q)^\lambda \right) \cdot (1 - \alpha \psi (1 - \lambda)) + 2\alpha \psi \right)^2 - 2 (\alpha \psi)^2 + 0.5 (\alpha \psi)^2 \\
\frac{d^3}{(d\lambda)^3} u^E(\lambda) &= -\frac{1}{4} \cdot (1 - q)^\lambda \cdot \left( \left( \ln (1 - q)^\lambda \right) \cdot (1 - \alpha \psi (1 - \lambda)) + 3\alpha \psi \right)^2 - 3 (\alpha \psi)^2 \cdot \left( \ln (1 - q)^\lambda \right).
\end{align*}
\]

Note that $\lim_{N \to \infty} \left( \ln (1 - q)^N \right) = -\infty$ and $\lim_{N \to \infty} \left( (1 - q)^{0.5N} \ln (1 - q)^N \right) = 0$. These two facts, and our assumption that $1 - \psi \alpha > 0$, have three important implications. First, for large enough $N$, the function $u^E$ is increasing at 0 and decreasing at 0.5; i.e., $\frac{d}{d\lambda} u^E(0) > 0$ and $\frac{d}{d\lambda} u^E(0.5) < 0$. Next, for large enough $N$, the function $u^E$ is concave in the neighbourhood of zero; that is, $\frac{d^2}{(d\lambda)^2} u^E(0) < 0$. Finally, for large enough $N$, the third derivative is always positive in the interval $\lambda \in [0, 0.5]$. This last fact implies that the second derivative can be zero at most one time, which means that the function $u^E$ can switch from concavity to convexity once but cannot switch back to concavity.

Suppose that $N$ is large enough that the properties we mention above hold. Since the function $u^E$ is continuous, increasing at 0 and decreasing at 0.5, then it must have at least one (local) maximum at some value $\lambda' \in (0, 0.5)$. We claim that the function cannot attain a minimum in $(0, 0.5)$. To see why, suppose that it attains a minimum at some value $\lambda''$. Obviously, because
the function is continuous and increasing at 0, it must be the case that $\lambda' < \lambda''$ (if there are many minima and maxima, pick two such points that satisfy this condition). It must be the case that somewhere between $\lambda'$ and $\lambda''$, the function $u^E$ switches from concavity to convexity. Since the function cannot switch back to concavity, it must be increasing over $[\lambda'', 0.5]$. However, this is a contradiction of the fact that the function is decreasing at $\lambda = 0.5$.

Thus, since $u^E$ is continuous, attains a maximum at $\lambda'$, and attains no minima over the range $(0, 0.5)$, we conclude that $\lambda'$ is single-peaked over the interval. ■

**Proof of Proposition 3**

Fix $\alpha$ and $q$ and $\psi$, and assume that the break of trust is avoidable. Let $\bar{\lambda}$ be the size of the Elites that satisfies equation (10) with equality. Because the breakdown of trust is avoidable, $\bar{\lambda} > 0$. If $\bar{\lambda} \geq 0.5$, we are done because the optimal size of the Elites is interior — that is, $\lambda^* \in (0, 0.5)$ — and therefore, $\bar{\psi}(\alpha, \lambda^*) > \psi$. Suppose, then, that $\bar{\lambda} \in (0, 0.5)$. Denote $\bar{s} \equiv (1 - q)^{N\bar{\lambda}}$ and rewrite equation (10) as follows:

$$\psi \alpha = \frac{(1 - \bar{s})}{(1 - \bar{s})(1 - \bar{\lambda}) + 2\bar{\lambda}}. \quad (17)$$

We now show that the first derivative of $u^E$, evaluated at $\bar{\lambda}$, is negative; i.e., $\frac{du^E}{d\lambda}(\bar{\lambda}) < 0$. Plug equation (17) into the expression of $\frac{du^E}{d\lambda}$ that we computed in the proof of Lemma (3) and simplify to get:

$$\frac{du^E}{d\lambda}(\bar{\lambda}) = -\frac{(2\bar{s}\bar{\lambda} - 2\bar{s} - 2\bar{s}^2\bar{\lambda} + \bar{s}^2 + \bar{s}\bar{\lambda}\ln \bar{s} + 1)}{\left(\bar{\lambda} - \bar{s} + \bar{s}\bar{\lambda} + 1\right)^2}. \quad (18)$$

Denote $g(s) \equiv (2s\bar{\lambda} - 2s - 2s^2\bar{\lambda} + s^2 + s\bar{\lambda}\ln s + 1)$ so that the numerator in the right-hand side of equation (18) is $g(\bar{s})$. To show that $\frac{du^E}{d\lambda}(\bar{\lambda}) < 0$, it suffices to show that $g(s) > 0$ for all values of $s \in (0, 1)$. To see this, write the first and second derivatives of $g(s)$:

$$g'(s) = 2s + 3\bar{\lambda} - 4s\bar{\lambda} + \bar{\lambda}\ln s - 2$$
$$g''(s) = \frac{1}{s} \left(2s^2 - \bar{\lambda} - 4s\bar{\lambda}\right).$$

Note that $g'(1) < 0$ and $g''(s) > 0$ for all values of $s \in (0, 1)$ and $\bar{\lambda} \in (0, 0.5)$. Hence, $g'(s) < 0$ for all $s \in (0, 1)$, so $g(s)$ is a decreasing function on $(0, 1)$. Since $g(1) = 0$, $g(s) > 0$ for all values of $s \in (0, 1)$.

We have therefore established that $\frac{du^E}{d\lambda}(\bar{\lambda}) < 0$. Since $u^E$ is single-peaked over $(0, 0.5)$, and since it attains a maximum at $\lambda^* \in (0, 0.5)$, it must be the case that $\lambda^* < \bar{\lambda}$. Therefore, by
Lemma 2, we have that $\bar{\psi}(\alpha, \lambda^*) > \bar{\psi}(\alpha, \bar{\lambda}) = \psi$. ■