Information Aggregation in Stratified Societies

Marina Agranov,[*] Ran Eilat,[†] and Konstantin Sonin[‡]

February 17, 2025

Abstract

We analyze a model of political competition in which the elite forms endogenously to aggregate information and advise the uninformed median voter which candidate to choose. The median voter knows whether or not the endorsed candidate is biased toward the elites, but might still prefer the biased candidate if the elite's endorsement provides sufficient information about her competence. The elite size and the degree of information aggregation by the elite depend on the extent to which the median voter follows the elite's advice. A higher cost of redistribution minimizes the elite's information advantage, hinders information transmission, and decreases the expected competence of the elected politician.

Keywords: political economy, cheap talk, information club, stratification.
JEL Classification: D72, D83.

[*]Division of Humanities and Social Sciences, Caltech and NBER; 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

Economic progress requires efficient institutions of information aggregation. The idea that the public can benefit from trusting a small group of better informed people – be it politicians, professional public servants, journalists, or academic scholars — in making political decisions is as old as the idea of a representative democracy. Information transmission, however, might be fragile. For example, it breaks down if the informed elites are suspected, rightly or not, that they exploit their power to promote own interests at the expense of the general public. In these cases, the social cohesion, social welfare, and the strength of the democratic system all decline.

The recent wave of populism has been often attributed to the breakdown of trust between elites and voting masses (Algan et al., 2017; Dustmann et al., 2017; Guriev and Papaiannou, 2021). Inglehart and Norris (2016) consider the 2016 Brexit vote as a rejection of the informed elite's advice. In Eichengreen (2018), the breakdown of trust results from a combination of economic insecurity and the inability of the political system to address the demand for change. Guiso et al. (2018) show that populist policies that disregard long-term economic harm emerge when voters 'lose faith' in the institutions and elites.[1] This literature hints that loss of faith or trust in experts impedes the dissemination and transmission of valuable information in society and hinders the ability to make informed economic decisions.

In this paper, we offer a simple political model that explores the above intuition and relates information aggregation by an elite, the inefficiency of redistribution, and the willingness of uninformed voters to follow the elite's advice. The model is composed of two parts. In the first part, the *Elite formation stage*, the heterogeneous population divides into two groups: the Elites minority, which forms endogenously to pool and share information, and the rest of society, referred to as the "Commons". The selection process for Elites membership is stochastic and influenced by the wealth of individuals.

In the second part, called the *political game*, two politicians compete for office, differing in two key aspects: their competence in generating economic resources and their affinity with the Elites. Members of the Elites group observe imperfect signals about the candidates' abilities, share this information among themselves, and endorse a candidate based on the aggregated signals. By pooling individual signals, the Elites gain an informational advantage over the Commons. However, when the uninformed Commons elect a politician, they do not simply accept the Elites' endorsement at face value. Instead, they recognize that the endorsement reflects not only the candidate's competence but also his bias toward the Elites. This bias is crucial because,

[1] In a classic study, Dornbusch and Edwards (1991) emphasized that populist policy "have almost unavoidably resulted in major macroeconomic crises that have ended up hurting the poorer segments of society."

depending on the cost of redistribution, it influences how politicians allocate resources within the economy.[2] Our approach to modeling the information structure in this part is inspired by Argenziano, Severinov and Squintani (2016).

Two key features of our modeling approach are the heterogeneity of agents' wealth and the formation of Elites, which leads to endogenous stratification between Elites and Commons. In the model, the size of the Elites group is determined in equilibrium and is influenced by the expected outcomes of the political game. On the one hand, the Elites benefit from expanding their group size, as it allows for more accurate aggregation of dispersed information, making their advice more valuable to the Commons. On the other hand, a larger Elites group reduces the resource share for each member, limiting the incentive for further expansion. These competing forces balance at the stable Elites size. Once this size is determined, a parameterized process governs how citizens are selected into the Elites group. At one extreme, only the wealthiest citizens join, creating a perpetual wealth-based oligarchy. At the other, all citizens have an equal chance of joining the Elites.

We start by analyzing the political game. We demonstrate that if the cost of redistribution is below a certain threshold, the Commons will follow the Elite's endorsement. However, if the cost of redistribution exceeds the threshold, the Commons will not trust the Elites' advice, resulting in a loss of valuable information. This inverse relationship between the willingness to follow the Elites' endorsement and the cost of redistribution is driven by the information mechanism. Specifically, as the cost of redistribution rises, the benefit the Elites derive from supporting a biased politician grows, making their endorsement less reliable. A similar effect occurs when the bias of the elite-aligned politician increases. In both cases, the overall quality of the elected politician declines.

Next, we analyze the elite-formation stage of the model. We define the notion of a stable Elites club and characterize the stable size of the Elites. This stable size strikes a balance between information aggregation and resource exploitation. We show that this balance ensures that in the ensuing political game, the Commons adhere to the advice of the informed Elites, thereby enhancing the expected competence of the elected politician. The stable size of the Elites' club decreases with the cost of redistribution and the bias of the elite-aligned politician.

As a final step, we analyze the relationship between the inclusiveness of the Elites formation process, the wealth inequality, and the outcomes of the political game.[3] We demonstrate that, in the long run, societies with more inclusive Elites achieve more equal wealth distribution. Fur-

[2]We follow the standard assumption in political economy and interpret the losses of redistribution as dead-weight loss of taxation (Acemoglu and Robinson, 2001).

[3]We use the Lorentz curve as a measure of inequality for wealth distributions (Shorrocks, 1983; Marshall, Olkin and Arnold, 2011).

thermore, if the cost of redistribution increases with wealth inequality, rather than being exogenously fixed, we show that a society with a more inclusive Elites is more likely to end up with a lower cost of redistribution in the future, and, as a result, with more competent politicians.

Relationship to the literature. 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 elite's endorsement is biased, yet tries to take advantage of the information that is contained in it. Myerson (2008) models trust as an equilibrium phenomenon, but the context is very different: trust is what keeps the autocrat's lieutenants abiding his command.

Chakraborty and Ghosh (2016) consider a model of Downsian competition between two office-seeking parties, in which voters that care about both the policy platform and "character" of candidates make a decision based on a media endorsement.[4] The media has its own policy agenda and, though 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. Chakraborty and Yılmaz (2017) analyze a model of two-sided expertise that can be used to evaluate endorsements and elections with multiple informed parties with different interests; Chakraborty, Ghosh and Roy (2020) offer a model of elite endorsement and policy advocacy in a spatial model. In our model, the breakdown of information transmission is akin to the non-existence of influential endorsements when the interests are too divergent.

Chan et al. (2019) study a model in which voters have to choose between two alternatives, where voting for the non-status quo option is costly and the cost varies across agents. A sender who can commit to a signal structure tries to persuade voters to deviate from the status quo. In the part of their analysis closest to ours, the sender communicates publicly with voters and targets the voter with the K^{th} lowest cost, where K is the super-majority required for deviating from the status quo. In contrast, in our model, Elites cannot commit to the signal structure a priori, they are not fully informed, and their preferences are state-dependent.

In Martinelli (2006), voters decide whether to acquire information before making a choice. In Prato and Wolton (2016), successful communication between candidates and voters during the pre-election campaign requires both an effort from the candidates and attention from voters (See also Prato and Wolton, 2018, on populism as political opportunism by incompetent politicians and Pastor and Veronesi, 2020, for an equilibrium model of populism where voters elect

[4] 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.

a populist in response to rising inequality.) In Kartik and van Weelden (2019), uncertainty generates reputationally-motivated policy distortions in office, regardless of 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 elite and votes for the unbiased politician, in which case valuable information is lost.[5]

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 of our club formation process is information aggregation: the benefit of having a larger club is that the aggregated information is based on more independent signals and is, therefore, more precise.

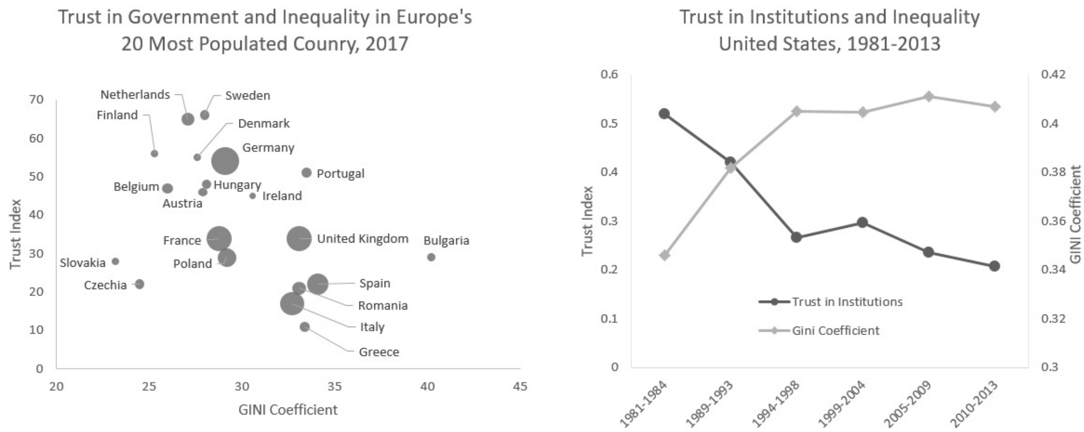
The rest of the paper is organized as follows. Section 2 briefly reviews the key features of anti-elite politics. In Section 3, we introduce our model and outline the political game along with the Elites formation process. Section 4 analyzes the political game under a fixed Elites size. In Section 5, we endogenize the size of the Elites and examine the properties of the stable Elites size. Section 6 investigates how the Elites formation process impacts wealth inequality and societal welfare. Section 7 concludes.

2 Anti-Elite Politics

The notion of the anti-elite politics has perhaps as long pedigree as politics itself. In 1820s, Andrew Jackson rode a horse as the champion of the "common man" against the emerging New England "aristocracy". In 1930s, the populist Louisiana Senator Huey Long threatened the dominance of Franklin Delano Roosevelt within his own Democratic party. Senator McCarthy did not run for president in 1950s, but his anti-elitism was bipartisan — he attacked professionals in both the Democratic and Republican administration – and highly popular at the peak.

In the 21st century, the anti-elite politics is most commonly associated with notion of populism. In fact, the most inclusive definition of populism adopted in the major recent survey by Guriev and Papaiannou (2021) from Mudde (2004) and Mudde and Kaltwasser (2017) defines it as a "thin-centered ideology" that considers society to be ultimately stratified into two homogeneous, antagonistic groups: "the pure people" and "the corrupt elite."

[5] For other models of cheap talk in elections, see Harrington (1992), Panova (2017), Schnakenberg (2016), and Kartik, Squintani and Tinn (2015).



(a) Cross-country data (b) United States data

Figure 1: Relationship between Trust and Inequality

Rodrik (2017) points out that the modern populists often target the new elites, "unelected technocrats running central banks, independent regulatory agencies and international organizations, mainstream media, national and international NGOs, and corporate lobbyists". Rodrik goes on to argue that the solutions that elite offers on immigration, trade, outsourcing, or automation have been often indeed skewed towards the elites' interests. What our theory adds to this picture is that the distrust of the elites and the low quality of these elites are mutually reinforcing. When the people distrust the elites, the elites have low incentives to aggregate information, which leads to even more distrust as the quality of advice worsens.

In 21st-century Europe, populism was fueled primarily by the issues of immigration and increased policy control by technocratic bureaucrats. Nowadays, populist parties represent a significant chunk of voters: the Freedom Party in Austria, the National Rally (formerly the National Front) in France, the League and the Five Star Movement in Italy, the Dutch Party for Freedom and the Forum for Democracy in the Netherlands, the True Finns Party in Finland, the People's Party in Denmark, the UK Independence Party and the Brexit Party in Great Britain. In our theory, there is no political positioning. However, the main force is exactly what drives the anti-elite populism: in an ideal world with full commitment, a competent pro-elite politician would commit to a position that would guarantee information transmission, and, therefore, the election of a more competent candidate. Our model demonstrates how this inability to commit translates into mistrust, which, in turn, leads to a low level of information aggregation.

Keefer, Scartascini and Vlaicu (2019) analyze survey data from 6,000 respondents in seven Latin American countries to demonstrate the critical link between populism, trust, and the quality of government: voters who express low trust are significantly more likely to prefer populist policies, which in turn are determined by the low quality of government.

Another important relationship that arises endogenously in our model is the one between redistribution costs and the willingness of the voters to use the elite's advice, which contains valuable information. Figure 1 illustrates the negative correlation between the level of political trust, a common sociological variable, and wealth inequality, which is positively correlated with redistribution costs, in two ways. Trust, as measured by opinion polls, is an imperfect proxy for the willingness to follow political endorsement; still, this is the best measure available to researchers. Panel (1a) uses data from the 20 most populated countries in Europe in 2017; a similar picture may be obtained if one uses trust in media instead of the trust in governments, both of which are imperfect but reasonable proxies for trust in elites. The simple OLS regression detects negative relation between inequality as measured by the GINI coefficient and any of these two measures of political trust ($p = 0.03$ for trust in media and $p = 0.08$ for trust in governments).[6] Panel (1b) presents the evolution of political trust in institutions in the US from 1981 to 2013.[7] In general, the decrease in trust is accompanied by a steady increase in inequality (Piketty and Saez, 2003). In our model, this correlation arises for pure informational reasons: greater redistribution costs lead to diverging interests among different groups, which impedes the flow of information and decreases trust. Consequently, valuable information is lost and welfare declines. Not surprisingly, the growing inequality contributes to the rise of populism (Pastor and Veronesi, 2020).

3 Setup

Consider a democratic society that consists of a large finite number of citizens, denoted by N. Citizens are heterogeneous with respect to their wealth. The wealth of citizen i is denoted by w_i, and without loss of generality we assume that $w_1 \geq w_2 \geq \ldots \geq w_N$. The citizens engage in two sequential interactions: First, they form two social groups, Elites and Commons. Those citizens who form Elites share their private information among themselves, while those who form Commons do not. Second, they participate in a political game where their interests depend on their group affiliation. In this game, information about the politicians' competence can be communicated from Elites to Commons. Whether this information influences the voting decisions of Commons defines the level of *trust* in the society. In this section, we describe the components

[6] Trust 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.

[7] Trust 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.

of the model and the timing of the game.

Elite formation. In the first stage of the game the group of Elites is formed. The group size, denoted by k^*, is determined *endogenously* to maximize the utility of the group members. In particular, in equilibrium, Elite members do not want to change the group size by accepting or removing members.

We assume that the selection of members into Elites is non-deterministic and depends on the individuals' wealth. Specifically, we capture the stochastic nature of this process with a lottery that randomly awards k^* spots in Elites to citizens with wealth ranking from 1 to M, where $M \in \{k^*, \ldots, N\}$ is a parameter that proxies the degree to which the elite-selection process is skewed towards the richest members of the society. When $M = k^*$, the process forms a persistent wealth-based oligarchy as the k^* richest citizens become the Elites; when $M = N$, all citizens have equal chances to get into Elites.

All citizens who are not part of the Elites form the group of Commons. We denote the share of Elites in the citizenry by $\lambda = k/N$, and focus on the case that Elites is the minority group, $\lambda < \frac{1}{2}$.

The political environment. After the group of Elites is formed, The citizens have to elect one of two politicians into office. Once elected, the politician determines how to allocate the available resources between the two groups. As the majority, the Commons can unilaterally decide the identity of the elected politician. However, Elites have an advantage over the Commons: the information possessed by Elite members aggregates, making them better informed than the Commons about the competence of the candidates. Because all citizens within each group receive the same level of resources, there are no collective action problems within groups.

The two politicians running for office differ along two dimensions: their preferences for resource allocation between the groups, and their ability to create resources for the economy (to which we refer as their *competence*). We assume that one of the politicians, denoted by U, is *unbiased* and assigns equal importance to the marginal per-capita consumption of Elites and Commons. The other politician, denoted by B, is *biased* towards Elites. The level of bias is determined by a parameter $\alpha \in \mathbb{R}^+$, which is common knowledge among Elites and Commons. The value of α represents the strength of ties the biased politician shares with Elites relative to the Commons, where larger values reflect 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. The objective function of politician $j \in \{B, U\}$ is given by:

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

The functional form of Equation (1) reflects a compromise between the politicians' egalitarian and utilitarian motives. 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 different in that the importance of a marginal unit of Commons' per capita consumption is discounted, and this discount is stronger as α increases.

The competence of politicians in creating resources depends on a state of the world, denoted by θ, which is drawn from a uniform distribution over the interval $[0, 1]$. However, the citizens cannot directly observe θ, and instead they observe only noisy signals about it, in a way that is described below. We denote the competence of politician $j \in \{B, U\}$ by θ^j and assume that

$$\theta^B = 1 + \theta, \tag{2}$$

$$\theta^U = 2 - \theta. \tag{3}$$

Consequently, the *ex ante* expected competence of each of the politicians is the same: $\mathbb{E}[\theta^B] = \mathbb{E}[\theta^U] = \frac{3}{2}$. The biased politician is more competent than the unbiased one if and only if $\theta > \frac{1}{2}$, which occurs with a probability of one-half.

The politician in office allocates the available resources θ^j among the two groups such that

$$\lambda x^E + (1-\lambda)\, x^C \cdot \psi = \theta^j. \tag{4}$$

The parameter ψ captures the cost of redistribution in our model, i.e., the cost of converting a unit of Elites' consumption, x^E, into a unit of Commons' consumption, x^C. We assume that redistribution is costly, i.e. $\psi > 1$.

Equation (4) implies that, from the perspective of the elected politician, allocating a unit of resources to Elites is "cheaper" than allocating a unit of resources to Commons. This assumption captures in a stylized way the observation that members of Elites often have better access to resources generated in the economy, and that diverting a unit of resources from Elites to Commons is costly for the politician.[8] In practice, this cost of redistribution can arise from various

[8]The observation that elites have better access to resources is well documented in the literature and can be attributed to various factors. Examples include the increased political influence of the elites, e.g., due to their superior information or ability to form interest groups (see Grossman and Helpman (2001)), their control over economic in-

reasons: the deadweight loss of taxation (see, e.g., Meltzer and Richard (1981); Saez, Slemrod and Giertz (2012)), the political costs of redistributive policies (which might be unpopular even among the poor, see e.g., Benabou and Ok (2001)) or the costs of compliance and enforcement of such policies (e.g., Keen and Slemrod (2017)).

A dual (and formally equivalent) way of interpreting the assumption that $\psi > 1$ is that the effect of allocating a unit of resources to Elites is amplified (compared to a unit allocated to Commons) because of the benefits of being part of the Elites' club. These benefits encompass the social and economic advantages of being in the Elites (the in-group members), such as access to better jobs, new equipment, better assignments, and other tangible and intangible rewards, which are not accessible to the commoners (the out-group members).

To simplify our analysis, we assume that $\alpha \cdot \psi < 1$.

Information structure. To model the information structure in society, we adopt the framework developed in Argenziano, Severinov and Squintani (2016). Specifically, we assume that *after* the group of Elites is formed, and *after* the state θ is realized (but cannot be directly observed by the citizens), each member of Elites conducts a (conditionally) independent experiment that results in either a success or a failure. The probability of success is equal to the true value of θ. Consequently, a successful experiment serves as a signal that θ is high, implying that the biased candidate is the more competent one. Conversely, a failed experiment serves as a signal that θ is low, implying that the unbiased candidate is the more competent one.

We assume that Elites members share the outcomes of their experiments, enabling all members of the club to observe all the outcomes. This assumption captures the general intuition that evaluating politicians' competence is a complex task that requires expertise, time investment, and interaction with others who possess private information. In our model, these interactions are represented through the sharing of information among club members.

Members of the Commons do not conduct experiments. Instead, they receive information about the state of the world through endorsements sent by the Elites as part of the political game. The assumption that Elites have access to superior information compared to the Commons aligns with a large body of literature on the sociology of elites.[9] While the qualitative nature

stitutions (Acemoglu and Robinson (2008)), the network effects of being close to positions of power (Fisman (2001)), or corruption (Reinikka and Svensson (2004)).

[9] For example, Khan (2012) argues that knowledge capital is one of the five significant types of resources typically controlled by the elites (the other four types of resources are political, economic, social, and cultural). To accumulate knowledge capital, which translates into an informational advantage in our model, elites facilitate a network of social connections between group members to transfer information. These connections are created via social institutions such as elite schools and social clubs, which are used both to strengthen the ties between group members and to exclude outsiders. (See also Zimmerman, 2019, and Michelman, Price and Zimmerman, 2021.)

of our results would not change significantly if the commoners were slightly informed (rather than completely uninformed), the analysis would become more complex, and we do not fully explore this version of the model here. Nevertheless, in Proposition 6 below, we take a small step in this direction and show that, under certain mild restrictions on the cost of redistribution and the bias of the pro-elites politician, even if the commoners were capable of conducting their own experiments, but unable to share the outcomes, each commoner would still lack the incentive to conduct an experiment when the size of the Elite group is determined endogenously.[10]

Endorsements and voting. While Commons constitute the majority of the population and can effectively decide who is elected, Elites possess better information. It is thus in the interest of both groups to share the information held by Elites to increase the chance that the more competent politician is elected.

We assume that Elites cannot credibly share their information with the Commons (i.e., they cannot reveal the number of successful experiments) nor can they commit to a strategy of information disclosure in advance. Thus, information transmission between the two groups takes the form of "cheap talk" (Crawford and Sobel, 1982). Specifically, after observing the total number of successful experiments, Elites can send a costless and unverifiable message to the Commons, who update their beliefs about θ and elect their preferred candidate.[11]

We denote by $\mathcal{M}$ the set of possible messages that Elites can send to Commons, and assume without loss of generality that $\mathcal{M} = \{m_B, m_U\}$.[12] 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 by $\sigma_E : L \to \mathcal{M}$, where $\sigma_E(l)$ is interpreted as the endorsement when Elites observe $l \in L \equiv \{0, \ldots, \lambda N\}$ successful experiments.

After Elites endorse one of the candidates, each member of the Commons updates his posterior belief about the state of the world θ (and therefore about the competence of the politicians), and casts his vote. A strategy for a commoner, denoted by $\sigma_C : \mathcal{M} \to \Delta\{B, U\}$, maps each mes-

[10] More precisely, each commoner would be indifferent between conducting an experiment or not, as the information gathered would not influence his actions. For any positive cost of experimentation, Commons would strictly prefer not to acquire information.

[11] An alternative approach to modeling the communication protocol between the two groups could be to allow the Elites to commit to a mapping between their information and the messages they send to the Commons, in line with the Bayesian Persuasion literature (Kamenica and Gentzkow (2011)). Under this protocol, the optimal signal structure would render the "representative" commoner indifferent when voting for the biased candidate, while strictly preferring the unbiased candidate when voting for him. Although this approach would increase the Elites' power and payoff in the political game, their overall payoff may not necessarily rise because the increased communication power might lead to a non-monotonic effect on elite size.

[12] Formally, 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 $\theta \in \Theta$.

sage $m \in \mathcal{M}$ to a probability distribution over the possible voting options (i.e., the biased candidate B, or the unbiased candidate U). Since Commons constitute the majority, the candidate they vote for gets elected into office.

Our solution concept is the Perfect Bayesian Equilibrium, and we assume that each citizen votes as if her vote is decisive, which is a weakly undominated strategy, in the voting stage.

Timing. To facilitate the analysis, we divide the timeline into two stages: the *Elites formation stage* and the *political subgame*, as follows:

ELITES FORMATION STAGE:

1. A group of Elites is formed, with size k^* (corresponding to the share $\lambda^* = k^*/N$) that is optimal for the members of the Elites. Members of Elites are selected randomly, with equal probability, from the M wealthiest citizens.

POLITICAL SUBGAME:

2. Nature determines the state of the world $\theta \in \Theta$.
3. Members of Elites conduct experiments and share their outcomes with each other.
4. Elites endorse one of the politicians, either $B = (\theta^B, \alpha)$ or $U = (\theta^U, 0)$.
5. Commons cast their votes, either accepting or rejecting Elites' endorsement.
6. The elected politician takes office and distributes resources.

4 The Determinants of Trust

Our analysis consists of several parts. This section is the first part of the analysis, in which we characterize the equilibrium in the political subgame for a given exogenous share of Elites' size $\lambda = k/N$. The second part is presented in Section 5. It shows how the optimal size of the Elites λ^* is determined, taking into account how this choice affects behavior and payoffs in the political subgame. After that, in Section 6 we analyze the impact of the elite-selection procedure, parameterized by M, and the political game on the expected wealth distribution.

To solve the political subgame, we work backwards. First, we derive the actions of the elected politician. Then, we find a pair of endorsement and voting strategies (σ_E, σ_C) for Elites and Commons, respectively, that constitute an equilibrium in the subgame. We show that Elites use a cutoff strategy for endorsement, and describe the conditions under which Commons are willing to accept the endorsement.

Actions of the elected politician. The actions of the politician in office depend on her type (θ^j, a^j). Specifically, the politician maximizes the objective given by Equation (1) subject to the constraint given by Equation (4). Solving the maximization problem shows that a politician of type (θ^j, a^j) allocates the per capita consumption of Elites (x^E) and Commons (x^C) as follows:

$$x^E\left(\theta^j, a^j\right) = \theta^j + (1-\lambda)\cdot a^j\psi, \tag{5}$$

$$x^C\left(\theta^j, a^j\right) = \frac{\theta^j}{\psi} - \lambda \cdot a^j. \tag{6}$$

Equations (5) and (6) highlight two useful observations. First, if redistribution were costless (i.e., $\psi = 1$), the unbiased politician ($a^j = a^U = 0$) would distribute resources equally among all citizens, while the biased politician ($a^j = a^B = \alpha$) would allocate a higher per-capita amount of resources to Elites. However, because redistribution is costly (i.e., $\psi > 1$), even the unbiased politician allocates a higher per-capita amount of resources to Elites.[13]

Second, when the unbiased politician assumes office ($a^j = a^U = 0$), the share of Elites in the population (λ) does not affect allocations. In contrast, if the biased politician is elected ($a^j = a^B = \alpha$), a larger share of Elites results in a decrease in the per-capita consumption of *both* Elites and Commons.

Commons' trust and Elites' endorsement. Given a pair of strategies (σ_E, σ_C), denote by $\sigma_C(m_i)[B]$ the probability that a commoner votes for the biased politician when Elites send the message $m_i \in \{m_B, m_U\}$. Since messages are cheap talk, there is no loss of generality in assuming that the message m_B leads to a higher probability of electing B than message m_U, i.e., $\sigma_C(m_B)[B] \geq \sigma_C(m_U)[B]$.[14] We call an equilibrium (σ_C, σ_E) in the political subgame *responsive* if Elites' endorsements m_B and m_U induce different distributions over Commons' actions. Otherwise, we call the equilibrium *unresponsive*.

Recall that l denotes the number of successful experiments that were conducted by the k members of Elites. Thus, given θ and k, the number of successful experiments l is distributed according to the binomial distribution. Specifically, the probability to observe l successes is given by:

$$f(l|k,\theta) = \frac{k!}{l!(k-l)!}\theta^l(1-\theta)^{k-l}, \quad \text{for } 0 \leq l \leq k.$$

[13]These results are consistent with the well-documented fact that policy decisions of elected officials are responsive to the public preference, but in a way that strongly favors the more affluent and well-connected citizens, i.e., the elites. See, e.g., Gilens (2012) and Bartles (2017).

[14]For any equilibrium in which $\sigma_C(m_B)[B] < \sigma_C(m_U)[B]$, one can simply "re-label" the messages to obtain an 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.

The posterior distribution of θ, given l successes in k trials, is a Beta distribution with parameters $l+1$ and $k-l+1$. Its density is given by:

$$\phi(\theta|l,k) = \frac{(k+1)!}{l!(k-l)!}\theta^l(1-\theta)^{k-l}, \quad \text{for } 0 \leq \theta \leq 1. \tag{7}$$

The conditional expectation of θ, after observing l successes in k trials, is therefore given by:

$$\mathbb{E}[\theta|l,k] = \frac{l+1}{k+2}. \tag{8}$$

The conditional expectation given by Equation (8) proves useful for our next result that characterizes the strategy of Elites in a responsive equilibrium (if such an equilibrium exists).

Lemma 1 *Suppose that (σ_C, σ_E) is a responsive equilibrium. Then, Elites' strategy σ_E attains the following threshold structure:*

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

where

$$\hat{l} \equiv \frac{k}{2} - \left(\frac{k}{2}+1\right)\alpha\psi(1-\lambda).$$

In a responsive equilibrium (if one exists), Elites endorse the biased candidate B if and only if they observe at least $\hat{l}$ successful experiments, as defined in Lemma 1. Otherwise, they endorse the unbiased candidate U. It is noteworthy that the threshold value $\hat{l}$ is smaller than $k/2$. That is, Elites endorse the biased candidate even if less than half of the group members observe a successful experiment.

Notice that the threshold $\hat{l}$ decreases with a greater redistribution cost (ψ) or a larger politician bias (α). This means that Elites need fewer successful experiments to endorse the biased politician B when the redistribution cost and/or politician bias are higher. Intuitively, this is because, all else being equal, the benefit to Elites of electing the biased politician increases with these factors.

We assume that if a responsive equilibrium in the political subgame exists, then it is played. However a responsive equilibrium may not necessarily exist.[15] In the remainder of this section we examine the necessary and sufficient conditions for the existence of such an equilibrium, and study its properties.

[15] As is standard in signaling games, an unresponsive equilibrium always exists. For example, Elites always endorsing the biased politician, and Commons always voting for the unbiased one is one such equilibrium.

4.1 Existence of a responsive equilibrium

To study the existence of a responsive equilibrium, we first characterize what Commons learn from endorsements when Elites employ the cutoff strategy defined in Lemma 1. We then examine whether it is in the Commons' interest to follow the endorsement.

In a responsive equilibrium, the Elites' endorsements convey information about the state of the world θ, which determines the competence of the politicians. The expected value of θ, conditional on an endorsement for the biased politician, is:

$$\mathbb{E}(\theta|m_B) = \sum_{\tilde{l}=\hat{l}}^{k} \Pr\left(\tilde{l} \mid \hat{l} \leq l \leq k\right) \cdot \mathbb{E}\left(\theta|\tilde{l}, k\right) = \frac{3-\alpha\psi(1-\lambda)}{4} - \frac{1}{2(k+2)} \tag{9}$$

where $\Pr\left(\tilde{l} \mid \hat{l} \leq l \leq k\right)$ denotes the probability to observe exactly $\tilde{l}$ successes, conditional on the event that the overall number of successes is between $\hat{l}$ and k.[16] Similarly, the expected value of θ, conditional on an endorsement for the unbiased politician, is:

$$\mathbb{E}(\theta|m_U) = \sum_{\tilde{l}=0}^{\hat{l}-1} \Pr\left(\tilde{l} \mid 0 \leq l \leq \hat{l}-1\right) \cdot \mathbb{E}\left(\theta|\tilde{l}, k\right) = \frac{1-\alpha\psi(1-\lambda)}{4}, \tag{10}$$

where $\Pr\left(\tilde{l} \mid 0 \leq l \leq \hat{l}-1\right)$ denotes the probability to observe exactly $\tilde{l}$ successes, conditional on the event that the overall number of successes is between 0 and $\hat{l}-1$.[17]

Recall that, according to Equations (2) and (3), the competence of the biased candidate (θ^B) increases with θ, while the competence of the unbiased candidate (θ^U) decreases with θ. Therefore, as the cost of redistribution (ψ) increases, the endorsement m_B becomes a *weaker* indication of the competence of the biased politician B, whereas the endorsement m_U becomes a *stronger* indication of the competence of the unbiased politician U.

We demonstrate later that as the number of citizens (N) increases, the optimal number of Elite members (k^*) also increases, while their share in the citizenry ($\lambda^* = k^*/N$) converges to zero. Consequently, by Equations (9) and (10), when N is large, the expected value of the state θ conditional on an endorsement for the biased and unbiased politicians, respectively, converges to $(3-\alpha\psi)/4$ and $(1-\alpha\psi)/4$. As a result, the competences of the biased and unbiased politicians, upon being endorsed by Elites, converge to $(7-\alpha\psi)/4$ and $(7+\alpha\psi)/4$, respectively.

We now turn to determine whether and when Commons are inclined to follow the endorsements of Elites, given what they have learned from these endorsements.

[16] The term $\Pr\left(\tilde{l}|\hat{l} \leq l \leq k\right)$ is equal to $1/(k-\hat{l}+1)$ because θ is distributed uniformly.

[17] The term $\Pr\left(\tilde{l}|0 \leq l \leq \hat{l}-1\right)$ is equal to $1/\hat{l}$ because θ is distributed uniformly.

Endorsements for the unbiased politician. Suppose that Elites employ the cutoff strategy defined in Lemma 1 and endorse the unbiased politician (i.e., send the message m_U). It is straightforward to verify that Commons always accept such an endorsement. This is because $\mathbb{E}\left[\theta^U|m^U\right] \geq \mathbb{E}\left[\theta^B|m^U\right]$. Therefore, upon hearing m_U, Commons deduce that the quality of the unbiased politician is higher. Since, in addition, the unbiased politician allocates resources more equally, it is always optimal for Commons to accept an endorsement for the unbiased politician.

Endorsements for the biased politician. Suppose that Elites employ the cutoff strategy defined in Lemma 1 and endorse the biased politician (i.e., send the message m_B). It is optimal for commons to accept this endorsement if, based on the information they learn from the fact that m_b is sent, their expected payoff from electing the biased politician is greater than their expected payoff from electing the unbiased one. Formally, Commons follow an endorsement for the biased politician (m_B) if and only if

$$\mathbb{E}[x^C(\theta^B, \alpha) \mid m_B] \geq \mathbb{E}[x^C(\theta^U, 0) \mid m_B].$$

By Equations (6) and (9), the above condition is satisfied if and only if the cost of redistribution (ψ), does not exceed an upper bound $\overline{\psi}(\lambda, \alpha)$:

$$\psi \leq \overline{\psi}(\lambda, \alpha) = \frac{\lambda N}{\alpha(\lambda + 1)(\lambda N + 2)} \tag{11}$$

Thus, if the redistribution cost (ψ) exceeds the threshold $\bar{\psi}(\lambda, \alpha)$, a responsive equilibrium cannot exist. In this case, Commons do not trust Elites and disregard their advice. By contrast, if the redistribution cost is less 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. The following proposition summarizes the above discussion.

Proposition 1 *For any share of Elites λ and any bias of the Elites' candidate α, there exists a redistribution cost threshold $\bar{\psi}(\lambda, \alpha)$, given by Equation (11), such that if $\psi > \bar{\psi}(\lambda, \alpha)$, then Commons disregard Elites' endorsements and always elect the unbiased politician. Conversely, if $\psi \leq \bar{\psi}(\lambda, \alpha)$, then there exists a responsive equilibrium: Elites recommend the biased politician if and only if they observe more than $\hat{\imath}$ successful experiments, and Commons always accept Elites' endorsements.*

Proposition 1 illustrates the critical role played by the cost of redistribution in determining the degree of information transmission in equilibrium. When the redistribution cost is low, the

Commons tolerate the informational distortions that come with Elites' endorsements and follow their recommendations. When the redistribution cost is high, trust breaks down and Commons disregard the endorsements, despite their informative content.

Proposition 1 also allows us to examine how the politician's bias (α) and the Elite's share of the population (λ) affect the level of trust that prevails in the political game. The effect of the parameter α, is clear: when the biased politician is more "Elites-oriented" (i.e., when α is higher) the threshold $\bar{\psi}(\lambda, \alpha)$ decreases, which makes Commons less receptive to endorsements. Intuitively, this is because a higher value of α decreases the per capita consumption of Commons' when they follow an endorsement for the biased politician.

The impact of the Elites share λ is more nuanced. A larger λ implies lower per capita consumption for Commons *and* for Elites when the biased politician is elected. While the former erodes trust, the latter enhances it. Holding the population size N fixed, a larger λ also leads to more experiments conducted by Elites, making their endorsement more informative and increases the willingness of Commons to accept it.

The following proposition summarizes the comparative statics of the redistribution cost threshold $\bar{\psi}(\lambda, \alpha)$:

Proposition 2 *The redistribution cost threshold* $\bar{\psi}(\lambda, \alpha)$*, defined in Equation (11), decreases in the politician's bias* α*. It increases in the elite's share* λ *if* $\lambda < \sqrt{2/N}$*, and decreases in the Elite's share otherwise.*

4.2 Properties of a responsive equilibrium

Suppose that, in the given economy, the cost of redistribution (ψ), the politician's bias (α), and the size of the elite group (k) satisfy Equation (11), and therefore a responsive equilibrium exists. Our next objective is to examine how these parameters affect the societal welfare. Because ψ, α, and k have distributional effects (as reflected in Equations (5) and (6)), welfare improvements cannot be assessed in the Pareto sense. Instead, we measure societal welfare using the expected ex-ante competence of the elected politician, which quantifies the total resources produced in the economy.

To do this, we first compute the probability for each of the Elites' endorsements, m_B and m_U, in a responsive equilibrium. These probabilities are given by:[18]

$$\Pr(m_B) = \sum_{l=\hat{l}}^{k} \Pr(l|k) = \frac{1}{2}\frac{k+2}{k+1}\left(1+\alpha\psi(1-\lambda)\right) \tag{12}$$

$$\Pr(m_U) = \sum_{l=0}^{\hat{l}-1} \Pr(l|k) = 1 - \frac{1}{2}\frac{k+2}{k+1}\left(1+\alpha\psi(1-\lambda)\right). \tag{13}$$

Together with the expected value of the state θ conditional on each endorsement, as given by Equations (9) and (10), these probabilities are key in computing the ex-ante competence of the endorsed politician, who in a responsive equilibrium is also the elected politician. We denote this expected competence by $\mathbb{E}C$. We thus have:

$$\mathbb{E}C(\psi,\alpha,k) = \Pr(m_B)\cdot\mathbb{E}\left[\theta^B|m_B\right] + \Pr(m_U)\cdot\mathbb{E}\left[\theta^U|m_U\right] = \frac{7-\alpha^2\psi^2(1-\lambda)^2}{4} - \frac{(\alpha\psi(1-\lambda)+1)^2}{4(k+1)} \tag{14}$$

Equation (14) shows that the direct effect of a larger redistribution cost ψ on the ex-ante competence of the elected politician is negative. That is, larger redistribution costs lead to less competent politicians in expectation. Similarly, greater leniency of the biased politician towards the Elites, α, also has the same effect. Intuitively, these comparative statics operate through the channel of information transmission: higher ψ and α decrease the overall informativeness of Elites' endorsement and because equilibrium is responsive, they also decrease the competence of the politician who gets elected to office.

Equation (14) also shows that, all other things being equal, a larger Elite size k positively affects the competence of the elected politician. This occurs because a larger Elite group implies that Elites aggregate more information and hence are better informed. Our results in Sections 5 and 6 below, where k is determined endogenously, will highlight an indirect effect of the cost of redistribution on welfare that operates through affecting the size of the Elites.

Our next proposition summarizes the effects of the economy's parameters on welfare in a responsive equilibrium, as measured by the ex-ante competence of the elected politician.

Proposition 3 *Suppose that Equation* (11) *is satisfied, so that a responsive equilibrium exists. Then, lower redistribution costs, a lower politician bias, and a larger Elite all increase the expected competence of the elected politician, provided that Equation* (11) *continues to hold. Formally:*

[18] By the definition of the Beta function, we have $B(l+1,k-l+1) = \int_0^1 \theta^l(1-\theta)^{k-l}\,d\theta$. To derive the above probabilities, we utilize the key property that for any two integers z_1 and z_2, the Beta function is given by $B(z_1,z_2) = \frac{(z_1-1)!\cdot(z_2-1)!}{(z_1+z_2-1)!}$, along with the definition of $\hat{l}$ in Lemma 1.

1. *For any* $\psi' < \psi$*, we have* $\mathbb{E}C(\psi', \alpha, k) > \mathbb{E}C(\psi, \alpha, k)$*.*

2. *For any* $\alpha' < \alpha$*, we have* $\mathbb{E}C(\psi, \alpha', k) > \mathbb{E}C(\psi, \alpha, k)$*.*

3. *For any* $k' < k$ *such that* $\psi < \overline{\psi}(k', \alpha)$*, we have* $\mathbb{E}C(\psi, \alpha, k') < \mathbb{E}C(\psi, \alpha, k)$*.*

The above analysis also allows us to examine the effects of redistribution costs (ψ) and pro-Elite politician bias (α) on the expected per-capita consumption of members in the two groups in a responsive equilibrium.[19] Specifically, computing the expected per-capita consumption in a responsive equilibrium for Elites (Equation 5) and Commons (Equation 6) reveals that the former *increases* with both redistribution costs and pro-Elite politician bias, while the latter *decreases* with these quantities.[20]

We conclude this section by briefly discussing how Elites could affect their payoff in the political subgame *if*, before observing the state, they could choose the bias level of "their" politician, α. On the one hand, a responsive equilibrium is always better for Elites than a non-responsive one. On the other hand, if the equilibrium is responsive, Elites' expected payoff increases in α. Therefore, Elites would prefer to increase the bias level so long as a responsive equilibrium exists.

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

$$\alpha \leq \bar{\alpha} \equiv \frac{\lambda N}{\psi\,(\lambda+1)\,(\lambda N+2)}$$

where $\bar{\alpha}$ is the level of bias which makes Equation (11) bind in equality. Thus, when Elites can choose the bias level of their candidate they always ensure the existence of a responsive equilibrium. Notice that as N grows, $\bar{\alpha}$ converges to $\frac{1}{\psi(1+\lambda)}$. Clearly, this conclusion hinges on the assumption that the chosen candidate's bias is commonly known (By contrast, 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.)

[19] Note, however, that these comparative statics hold as long as the size and composition of the groups remain fixed. In the next sections, where these quantities are determined endogenously, changes in the cost of redistribution and the pro-Elite politician's bias also affect the composition of the groups. We return to this point in the discussion after Proposition 5.

[20] Formally, the expected per-capita consumption for Elites is given by $\mathbb{E}x^E(\psi,\alpha) = \mathbb{E}C(\psi,\alpha) + Pr(m_B)(1-\lambda)\alpha\psi$, with $\partial\mathbb{E}x^E(\psi,\alpha)/\partial\alpha > 0$ and $\partial\mathbb{E}x^E(\psi,\alpha)/\partial\psi > 0$. Similarly, the expected per-capita consumption for Commons is given by $\mathbb{E}x^C(\psi,\alpha) = (\mathbb{E}C(\psi,\alpha)/\psi) - \lambda Pr(m_B)\alpha$, with $\partial\mathbb{E}x^C(\psi,\alpha)/\partial\alpha < 0$ and $\partial\mathbb{E}x^C(\psi,\alpha)/\partial\psi < 0$.

5 The Optimal Size of Elites

In Section 4, we characterized the conditions for the existence of a responsive equilibrium in the political game and studied its properties. In this section, we shift our focus to analyzing the optimal size of the Elites, taking into account how the equilibrium in the ensuing political game influences this optimal size. We start by defining the stable Elites club, and show that such a club size maximizes the ex-ante expected utility of its members. We therefore sometimes refer to this stable club size as optimal. Since the size of the Elites determines the number of (conditionally) independent signals about the state of the world that citizens observe before electing a politician, this is a study of how optimal information aggregation depends on the extent to which Commons follow the Elites' endorsements.

We begin by defining a notion of stability that is appropriate for our environment. Intuitively, we say that Elites group of size k is stable if its members do not wish to add any number of new members to the group or expel any number of existing members. Formally, our definition of stability is reminiscent of the notion of *Closed Clubwise Stability* as analyzed by Fershtman and Persitz (2021).

Definition 1 (Stability) *Elite club size k is stable if:*

1. *Each of the existing members of the club prefers not to add any number of new members to the club, and*
2. *No coalition of $k - j > k/2$ members of the club prefer to exclude the remaining j members.*[21]

To operationalize this definition, we next compute the utility of each Elites member from being part of a group of size k, or equivalently, from being a member of Elites when their share in the citizenry is $\lambda = \frac{k}{N}$. In a responsive equilibrium, this utility is given by:

$$u^E(\lambda) \equiv \mathbb{E}\left[x^E\right] = \Pr(m_B) \cdot x^E\left(\mathbb{E}\left[\theta^B | m_B\right], \alpha\right) + \Pr(m_U) \cdot x^E\left(\mathbb{E}\left[\theta^U | m_U\right], 0\right)$$

where the dependence on ψ and α is omitted on the left-hand side for brevity. The first term on the right-hand side corresponds to the expected per-capita consumption of Elites when a biased politician is endorsed (and elected), and the second term corresponds to their expected per-capita consumption when an unbiased politician is endorsed (and elected). By substituting

[21] An alternative version of this definition, where $k - j > 0$ members of the club can exclude the remaining ones, would not change any of the results below.

the expressions from Equations (5), (9), (10), (12) and (13), we obtain that:

$$u^E(\lambda) = \frac{3}{2} + \frac{\alpha^2\psi^2(1-\lambda)^2 + 2\alpha\psi(1-\lambda)}{4} + \frac{\alpha^2\psi^2(1-\lambda)^2 + \lambda N}{4(\lambda N+1)}. \quad (15)$$

Intuitively, an Elites club with a share λ of the citizenry is stable if it maximizes u^E. To characterize this stable value, it is useful to first treat λ as a continuous variable that takes values in $\left[0, \frac{1}{2}\right]$. Our next lemma characterizes the value of λ that maximizes $u^E(\lambda)$.

Lemma 2 *For sufficiently large values of N, the expected payoff of Elites $u^E(\lambda)$ given by Equation (15) is single-peaked in λ and has a unique maximum $\lambda^* = \lambda^*(N) \in \left(0, \frac{1}{2}\right)$. Furthermore, $\lambda^*(N)$ is asymptotically bounded below by $\underline{\gamma}N^{-0.5}$ and above by $\overline{\gamma}N^{-0.5}$ for some positive constants $\underline{\gamma} < \overline{\gamma}$.*

Lemma 2 guarantees, generically, the existence of a stable share $\lambda^* \in \left\{0, \frac{1}{N}, \frac{2}{N}, \ldots, \frac{1}{2}\right\}$ of Elites. Since $u^E(\lambda)$ is single-peaked over a domain when λ is continuous, it has at most two maxima when λ is discrete; in a generic case, it has a unique maximum. Now, suppose that λ^* is this maximum, and the club of $k^* = N\lambda^*$ members is formed. This club is stable according to Definition 1 as every member would prefer neither to accept any more members nor to expel anyone. The next proposition formally states the existence result.

Proposition 4 *For sufficiently large values of N, Elites forms a stable club of size k^* in the elite formation stage. Moreover, for this club size, the condition for the existence of a responsive equilibrium given by Equation (11) is satisfied.*

Proposition 4 and Lemma 2 imply that, according to our definition of stability, the share λ^* (or the size k^*) is uniquely stable. Note, however, that this reasoning implicitly relies on the (standard) assumption that players do not consider sequential deviations. If, instead, players were far-sighted, Lemma 2 would not guarantee the uniqueness of a stable club. A well-known explanation for this is that when players are far-sighted, the instability of a sub-coalition can render a large coalition stable (e.g., Acemoglu, Egorov and Sonin, 2012).[22] Nevertheless, even without uniqueness, an elite group consisting of k^* members remains a natural outcome of the elite-

[22]In our case, suppose that decisions regarding club membership are determined by majority voting, and $k^* < \frac{N}{4}$. Suppose a club of size $2k^*$ is formed. First, observe that this club will not admit additional members because the utility function of each member is single-peaked, implying that increasing membership would reduce each member's utility. Second, there will be at least k^* members who would oppose the removal of a single elite member. Indeed, if at least one memberfrom the $2k^*$-sized elite group is removed, a coalition of k^* members could use their majority to remove the remaining $k^* - 1$ members. Thus, there exists a blocking coalition of k^* members that stabilizes the $2k^*$-sized elite group. This argument is admittedly heuristic, as the precise game that governs elite formation is not fully specified here. Nonetheless, given the equilibrium of the continuation game, the citizens' *ex ante* payoffs satisfy the conditions for a non-cooperative club formation game, as in Acemoglu, Egorov and Sonin (2012). Therefore, our argument could be formalized, albeit at the cost of introducing additional game-theoretic machinery.

formation process, where this club gradually emerges from an initial small group that steadily grows.

Proposition 4 and Lemma 2 also imply that when N is sufficiently large, the number of members in Elites grows asymptotically as $\sqrt{N}$. Thus, as the size of the population grows, the optimal number of members in the Elites club grows without bound (k increases), but their proportion in the population goes to zero (i.e. $\lambda \to 0$).

After establishing the existence of an optimal equilibrium size for Elites, a natural question arises: what is the effect of the redistribution cost and the bias of the pro-elites candidate on the optimal size of Elites? Proposition 5 provides comparative statics results. These results follow from the analysis of the derivative of $u^E(\lambda)$, which is cubic in λ and has a single-peak on the interval $\left[0, \frac{1}{2}\right]$.

Proposition 5 *The optimal size of the Elites club k^* decreases with both the bias of the pro-elite candidate (α) and the cost of redistribution (ψ).*

Proposition 5 presents intuitive comparative statics results. One critical element is the breakdown of trust: with higher politician bias (α) and redistribution cost (ψ), the range of parameters for which Commons follow Elites' endorsement narrows. Additionally, increasing α and ψ decreases the value of information that a potential member of Elites contributes, reducing the benefit of having a large club of Elites.

Proposition 5 also highlights an *indirect* negative effect of redistribution cost (ψ) and politician bias (α) on societal welfare, as measured by the expected competence of the elected politician. This effect is in addition to the direct negative effects demonstrated in parts (1) and (2) of Proposition 3. Specifically, when the size of the elite group is determined endogenously, Proposition 5 implies that higher redistribution costs or greater politician bias reduce the size of this group. Consequently, by part (3) of Proposition 3, this reduction further decreases societal welfare.

Optimal Elites' size and Commons' experimentation. Our analysis so far has assumed that commoners cannot conduct experiments. Our next result demonstrates that a small deviation from this assumption – allowing each commoner to be individually informed – does not alter the model's outcomes. Specifically, we show that if the size of Elites is determined optimally, then under a mild assumption regarding the value of $\alpha\psi$ (which captures the magnitude of the divergence in interests between Elites and Commons), commoners will not have an incentive to conduct experiments, provided they cannot share the results:

Proposition 6 *Suppose that $\alpha\psi < 0.5$. When Elite's share is optimally determined to be λ^*, even if each commoner could conduct an experiment, but not share the results with others, he would have no incentive to do so.*

We view this result as a modest "robustness check" for our model. Intuitively, the result holds because whenever the outcome of a commoner's experiment contradicts the Elites' endorsement, it is actually in the commoner's best interest to disregard his own signal. Naturally, if commoners could share the outcomes of their experiments efficiently, they might have an incentive to acquire information. However, as long as Commons remain collectively "informationally disadvantaged" relative to Elites, our model's qualitative key conclusions would not change significantly.

6 Inclusiveness, Inequality, and Welfare

Suppose that the size of Elites is determined optimally, as described in Section 5, and recall that the actual members of this group are selected according to the non-deterministic elite-selection process described in Section 3. We now turn to examine the relationship between the inclusiveness of the elite-selection process, wealth inequality, and the outcomes of the political game.

To compare different wealth distributions we use Lorenz curves, a standard measure of inequality of wealth distributions (Shorrocks, 1983). Specifically, we employ the concept of *majorization* (Marshall, Olkin and Arnold, 2011) to partially order wealth distributions according to the inequality they exhibit, as described below. In the following discussion, we say that a vector of positive real numbers $(z_1, \ldots, z_N)$ is *descending* if $z_1 \geq z_2 \geq \cdots \geq z_N$.

Definition 2 (Majorization, (Marshall, Olkin and Arnold, 2011)) *A descending vector $w = (w_1, \ldots, w_N)$ is* majorized *by a descending vector $w' = (w'_1, \ldots, w'_N)$ if for any $j = 1, \ldots, N$,*

$$\sum_{i=1}^{j} w_i \leq \sum_{i=1}^{j} w'_i,$$

and

$$\sum_{i=1}^{N} w_i = \sum_{i=1}^{N} w'_i.$$

If w' majorizes w, we denote this relationship by $w' \succ w$.

Definition 3 (More equal than) *Given two descending vectors w and w' that represent wealth distributions, we say that the wealth distribution w is* more equal than *the wealth distribution w' if $w \prec w'$.*

Recall that the elite-selection process, as described in Section 3, is governed by the parameter $M \in \{k^*, k^*+1, \ldots, N\}$. Specifically, for each citizen $i \in \{1, \ldots, M\}$, there is an equal probability (k^*/M) of being selected into the Elites. The case of $M = k^*$ corresponds to a self-perpetuating wealth oligarchy: in each period, the k^* richest citizens form the Elites. Conversely, $M = N$ describes an egalitarian environment, where every citizen has an equal chance of joining the Elites.[23]

Consider two societies with identical parameters: the same redistribution cost parameter ψ, the same pro-elites politician bias α, the same size of Elites k^*, and the same initial wealth of citizens that is given by a descending wealth vector $w = (w_1, \ldots, w_N)$. The only difference between these societies is their elite-selection process. In the first society, citizens are selected into the Elites using a process with $M = M_1$, while in the second society, citizens are selected using a process with $M = M_2$. We assume $M_1 < M_2$, implying that the elite-selection process in the first society is less inclusive than in the second.

We are interested in comparing the wealth distributions of these two societies after the political game is played. Due to the stochastic nature of the elite-selection process, in the *short run* (i.e., after the game is played once or a small number of times), it is not possible to definitively determine a priori which society will have a more equal wealth distribution. However, as our next result demonstrates, in the *long run*—that is, if the game is played repeatedly over a sufficiently large number of times—the more inclusive society will end up having a more equal wealth distribution with a probability arbitrarily close to one. Formally:

Proposition 7 *Consider two societies that differ solely in the inclusiveness of their elite-selection process. Let $w^{(1)}(T)$ and $w^{(2)}(T)$ denote the random variables that correspond to the wealth distribution in the first (less inclusive) and second (more inclusive) societies, respectively, after the political game is played for T times. Then, for all $\varepsilon > 0$, there exists $\overline{T}$ such that*

$$\Pr\left(w^{(1)}(T) \succ w^{(2)}(T)\right) > 1 - \varepsilon \qquad \forall T > \overline{T}.$$

The core idea behind the proof is that as the game is repeated sufficiently many times, the average amount of resources each citizen receives converges to the *expected* amount of resources

[23] Our elite-selection process inherently favors the wealthy citizens in society, with the value of M controlling the intensity of this favoritism. One could also envision an elite-selection process that favors the poor. For instance, suppose the process of selection into the Elites was governed by a parameter $m \in 0, 1, \ldots, N - k^*$, such that the first m richest members of society (in descending order) are excluded from the elite-selection process. E.g., if $m = 1$, then everyone except the richest citizen has an equal chance of being selected into the Elites.

Note, however, that the comparative statics with respect to m is largely a theoretical exercise. Historically, very few societies have disenfranchised the rich, with the exception of socialist countries in their early days, such as the Soviet Union in the 1920s. Naturally, a higher m would lead to a more equal expected wealth distribution. In a society that disenfranchises the rich, the elite-selection process would result in lower inequality.

he receives when the game is played once. Specifically, for each of the M_i-richest citizens in society $i \in \{1,2\}$, this mean value is given by

$$\frac{k^*}{M_i}\mathbb{E}[x^E] + \left(1 - \frac{k^*}{M_i}\right)\mathbb{E}[x^C],$$

and for each of the other citizens, this mean value is $\mathbb{E}[x^C]$, where $\mathbb{E}[x^E]$ and $\mathbb{E}[x^C]$ represent the expected per-capita resources that each member of the Elites and Commons receives in a responsive equilibrium. We then demonstrate that after the game is played sufficiently many times, the wealth distribution in the first (less inclusive) society will majorize that of the second (more inclusive) society with very high probability. In other words, the more inclusive society will end up having a more equal wealth distribution.

The reasoning behind the result stated in Proposition 7 implicitly assumes that as the game is repeated many times, the parameters of the society remain unchanged. But what happens if the parameters of the society respond to the changing wealth distributions between consecutive instances of the game? Our next results address this question and establish a connection between the inclusiveness of the society in the initial period and welfare in a future period. Specifically, we focus on the channel influenced by the cost of redistribution.

To do this, we enrich the model and assume that ψ is determined by the citizens' wealth distribution. Specifically, we assume that $\psi : \mathbb{R}^N \to \mathbb{R}$ is a function that maps each wealth distribution in the society to a level of redistribution cost. We further assume that $\psi(\cdot)$ is increasing in the inequality (partial) order defined above. That is, given two wealth distributions w and w', if $w' \succ w$ then $\psi(w') \geq \psi(w)$.[24] This way, because the level of inclusiveness of the elite-selection process affects the resulting wealth distribution, it will also affect the resulting cost of redistribution and hence the outcomes of the game if it is played again.

The function $\psi(\cdot)$ summarizes a wealth distribution in one real number. As before, because the elite selection process is stochastic, we cannot know for sure which of the two societies will have a higher or lower value of $\psi(\cdot)$ after the game is played once. However, because $\psi(\cdot)$ is a real function, we can describe the relationship between the level of inclusiveness of the elite-selection process and the cost of redistribution using the notion of first-order stochastic dominance. Formally, a real random variable A *first-order stochastically dominates* a real random variable B if $F_A(x) \leq F_B(x)$ for all x, with strict inequality at some x, where F_A and F_B are the cumulative probability distributions of A and B, respectively.

Our next result shows that after the game is played, the distribution over the resulting cost-of-redistribution parameter in the first (less inclusive) society first-order stochastically dominates

[24] Formally, the function $\psi(\cdot)$ is said to be Schur-convex.

the distribution over the cost-of-redistribution parameter in the second (more inclusive) society. This means that for each level of redistribution cost $\hat{\psi} > 0$, the probability of getting a cost-of-redistribution less than $\hat{\psi}$ is higher in the more inclusive society compared to the less inclusive one.

To formally state this result, we define $F_{\psi}^{(1)}$ and $F_{\psi}^{(2)}$ to be the distributions of the cost-of-redistribution parameter after the game is played in the first (less inclusive) and second (more inclusive) societies, respectively. Formally, given the random variable $w^{(i)}$ that represents the resulting wealth distribution in society $i \in \{1,2\}$ after the game is played once, define $\psi^{(i)}$ to be the pushforward distribution obtained by computing $\psi(w^{(i)})$, and let $F_{\psi}^{(i)}$ be the cumulative distribution function that corresponds to $\psi^{(i)}$. We then have:

Proposition 8 *Consider two societies that differ solely in the inclusiveness of their elite-selection process. After the political game is played, the distribution of the cost-of-redistribution parameter in the first (less inclusive) society first-order stochastically dominates that in the second (more inclusive) society:*

$$F_{\psi}^{(1)} \geq_{FSD} F_{\psi}^{(2)}.$$

The proof of this result proceeds as follows. First note that in both societies, at the end of the political game, all citizens receive $\mathbb{E}[x^C]$, and the elite selection process determines which k individuals among the M_1-richest citizens (in the less inclusive society) or the M_2-richest citizens (in the more inclusive society) receive an additional amount of $\mathbb{E}[x^E] - \mathbb{E}[x^C]$. For simplicity, we identify a citizen by his initial wealth.

Now, consider the following two procedures for selecting Elites in the two societies:

1. In the first (less inclusive) society, choose k citizens from $(w_1, \ldots, w_{M_1})$ to be Elites.

2. In the second (more inclusive) society, start by choosing k citizens from $(w_1, \ldots, w_{M_1})$ to form a *temporary* group of Elites. Then, choose a number $j \in \{0, \ldots, \min(k, M_2 - M_1)\}$ at random, with probability $\Pr(j) = \binom{M_1}{k-j}\binom{M_2-M_1}{j}/\binom{M_2}{k}$. If $j = 0$, the process ends, and the temporary Elites become permanent. Otherwise, remove j members from the temporary Elites with an equal probability of $1/\binom{k}{j}$ for each possible set of j members. Finally, add j new members from $(w_{M_1+1}, \ldots, w_{M_2})$ to the elites, with an equal probability of $1/\binom{M_2-M_1}{j}$ for each possible set of j members.

By definition, Procedure 1 generates a selection process that induces a uniform distribution over sets of size k citizens from $(w_1, \ldots, w_{M_1})$. Similarly, a computation shows that Procedure 2 generates a selection process that induces a uniform distribution over sets of size k citizens from $(w_1, \dots, w_{M_2})$.[25] Note, however, that in Procedure 2, whenever $j > 0$, the resulting wealth distribution among the (final) Elite members is more equal according to Definition 3, compared to the temporary Elites.

This argument shows that for any wealth distribution w that is an outcome of the M_1 selection process, the probability of obtaining the same wealth outcome w in the M_2 process is smaller. Specifically, it requires drawing the same Elite members in the temporary Elites stage and then having $j = 0$. Moreover, in the remaining cases (i.e., when the same Elite members are drawn in the temporary stage and $j > 0$), the resulting wealth distributions are more equal than w. Hence, each of these realizations leads to a (weakly) lower value of the redistribution cost ψ.

Consequently, the distribution of the resulting cost-of-redistribution under the (less inclusive) M_1 process first-order stochastically dominates the distribution of the resulting cost-of-redistribution under the (more inclusive) M_2 process, which is what we set out to demonstrate.

Taken together with our previous results, Proposition 8 can shed light on how the inclusiveness of the society's elite-selection process may affect the trajectory of the society's parameters and economic outcomes if the game is played more than once. Specifically, by Proposition 8, a less inclusive society is more likely to end up with a less equal wealth distribution after the game is played, and hence a higher cost of redistribution in the next period. This implies that in the next period the range of parameters for which a responsive equilibrium at the political stage is possible shrinks. Moreover, by Proposition 5, this also means that the size of elites in the next period is more likely to be smaller in the less inclusive society. Consequently, by Proposition 3, the less inclusive society is going to elect a politician with a lower expected quality, and, with that, a lower expected total wealth. The next proposition summarizes these observations:

Proposition 9 *Suppose that after the game is played once, the cost of redistribution is updated according to the resulting wealth distribution. Then, in the next period, the less inclusive society is more likely to have:*

[25]To verify this, consider a vector of Elite members consisting of $k - j$ members from among the citizens in $(w_1, \dots, w_{M_1})$ and j members from among the citizens in $(w_{M_1+1}, \dots, w_{M_2})$. Without loss of generality, suppose this vector is $(w_1, \dots, w_{k-j}, w_{M_1+1}, \dots, w_{M_1+j})$. The probability that this vector is selected according to Procedure 2 can be computed as follows:

$$\frac{\binom{M_1-k+j}{j}}{\binom{M_1}{k}} \cdot \frac{\binom{M_1}{k-j}\binom{M_2-M_1}{j}}{\binom{M_2}{k}} \cdot \frac{1}{\binom{k}{j}} \cdot \frac{1}{\binom{M_2-M_1}{j}} = \frac{1}{\binom{M_2}{k}}.$$

The first term on the left-hand side is the probability that the citizens $(w_1, \dots, w_{k-j})$ are selected into the k temporary Elites, along with j additional members. The second term is the probability of choosing that j members are to be removed from the temporary Elites. The third term is the probability of choosing exactly the j members other than $(w_1, \dots, w_{k-j})$ to remove from the temporary Elites. The fourth term is the probability of choosing the set $(w_{M_1+1}, \dots, w_{M_1+1+j})$ to join the elites. A computation shows that the product of these terms equals the right-hand side of the equation, which is the probability of selecting any set of k Elite members at random from $(w_1, \dots, w_{M_2})$.

1. *a higher cost of redistribution,*
2. *a smaller range of parameters for which a responsive equilibrium exists,*
3. *a smaller size of elites,*
4. *a lower expected quality of elected politicians and a lower expected total wealth.*

While our analysis throughout the paper focuses mainly on a static version of the model, Proposition 9 provides a possible informational micro-foundation for the celebrated result that inequality is harmful for economic growth (Alesina and Rodrik, 1994; Persson and Tabellini, 1994): High inequality, combined with a self-perpetuating elite, results in inefficient policy and growing inequality. While developing a full-fledged dynamic model is beyond the scope of this paper, we hope that our results will open the door for future research in this direction.

7 Conclusion

Recently, there has been a noticeable decline in voters' willingness to follow the elites' advice, as measured by opinion polls and the surge in support for anti-elite, populist politicians and parties. We propose a political model in which the endogenously formed elite has an information advantage over the rest of society, and the median voter elects a politician after considering the elite's endorsement. Our model demonstrates that the ability of elites to effectively transmit valuable information to the broader public is critically influenced by the cost of redistribution and the inherent bias of political candidates. When redistribution costs are high, or the leniency of the biased politician toward the elites is stronger, the trust between elites and commoners deteriorates, leading to the potential election of less competent leaders.

Our model also highlights the delicate balance that guides the formation of the elite group. While a larger elite group can aggregate more precise information, it also faces diminishing returns due to the reduced share of resources per member. The stable elite size thus emerges as a trade-off between these competing forces. Furthermore, the inclusiveness of the elite formation process plays a crucial role in determining the resulting wealth distribution and societal welfare after the political game is played. More inclusive processes lead to higher chances of a more equitable final distribution of wealth, which in turn improves the quality of political outcomes. We hope that our findings will encourage future research that can explore the dynamics of elite formation and trust between them and the rest of society, providing a deeper understanding of the mechanisms that sustain democratic systems.

References

Acemoglu, Daron, and James A Robinson. 2001. "A theory of political transitions." *American Economic Review*, 91(4): 938–963.

Acemoglu, Daron, and James A. Robinson. 2006. *Economic Origins of Dictatorship and Democracy.* Cambridge University Press.

Acemoglu, Daron, and James A Robinson. 2008. "Persistence of Power, Elites, and Institutions." *American Economic Review*, 98(1): 267–293.

Acemoglu, Daron, Georgy Egorov, and Konstantin Sonin. 2012. "Dynamics and stability of constitutions, coalitions, and clubs." *American Economic Review*, 102(4): 1446–76.

Acemoglu, Daron, Georgy Egorov, and Konstantin Sonin. 2013. "A Political Theory of Populism." *The Quarterly Journal of Economics*, 128(2): 771–805.

Alesina, Alberto, and Dani Rodrik. 1994. "Distributive politics and economic growth." *The quarterly journal of economics*, 109(2): 465–490.

Algan, Yann, Sergei Guriev, Elias Papaioannou, and Evgenia Passari. 2017. "The European Trust Crisis and the Rise of Populism." *Brookings Papers on Economic Activity*, 2: 309–400.

Aragones, Enriqueta, and Thomas R Palfrey. 2002. "Mixed equilibrium in a Downsian model with a favored candidate." *Journal of Economic Theory*, 103(1): 131–161.

Argenziano, Rossella, Sergei Severinov, and Francesco Squintani. 2016. "Strategic Information Acquisition and Transmission." *American Economic Journal: Microeconomics*, 8(3): 119–55.

Baier, Annette. 1986. "Trust and antitrust." *Ethics*, 96(2): 231–260.

Banks, Jeffrey S, and John Duggan. 2005. "Probabilistic voting in the spatial model of elections: The theory of office-motivated candidates." In *Social choice and strategic decisions.* 15–56. Springer.

Bartles, Larry, M. 2017. *Unequal Democracy: The Political Economy of the New Gilded Age.* Princeton University Press.

Benabou, Roland, and Efe A Ok. 2001. "Social Mobility and the Demand for Redistribution: The POUM Hypothesis." *The Quarterly Journal of Economics*, 116(2): 447–487.

Bernhardt, Dan, Odilon Câmara, and Francesco Squintani. 2011. "Competence and Ideology." *Review of Economic Studies*, 78(2): 487–522.

Bolton, Patrick, and Mathias Dewatripont. 2005. *Contract Theory.* The MIT Press.

Bénabou, Roland. 1996. "Inequality and Growth." In *NBER Macroeconomics Annual 1996, Volume 11. NBER Chapters*, 11–92. National Bureau of Economic Research, Inc.

Chakraborty, Archishman, and Bilge Yılmaz. 2017. "Authority, Consensus, and Governance."

The Review of Financial Studies, 30(12): 4267–4316.

Chakraborty, Archishman, and Parikshit Ghosh. 2016. "Character endorsements and electoral competition." *American Economic Journal: Microeconomics,* 8(2): 277–310.

Chakraborty, Archishman, Parikshit Ghosh, and Jaideep Roy. 2020. "Expert-Captured Democracies." *American Economic Review,* 110: 1713–51.

Chan, Jimmy, Seher Gupta, Fei Li, and Yun Wang. 2019. "Pivotal persuasion." *Journal of Economic Theory,* 180: 178–202.

Crawford, Vincent P, and Joel Sobel. 1982. "Strategic information transmission." *Econometrica,* 1431–1451.

Dornbusch, Rudiger, and Sebastian Edwards. 1991. *The Macroeconomics of Populism in Latin America.* University of Chicago Press.

Dustmann, Christian, Barry Eichengreen, Sebastian Otten, André Sapir, Guido Tabellini, and Gylfi Zoega. 2017. *Europe's Trust Deficit: Causes and Remedies. Monitoring International Integration 1,* CEPR Press.

Egorov, Georgy, and Konstantin Sonin. 2011. "Dictators and their viziers: Endogenizing the loyalty–competence trade-off." *Journal of the European Economic Association,* 9(5): 903–930.

Eichengreen, Barry. 2018. *The Populist Temptation: Economic Grievance and Political Reaction in the Modern Era.* Oxford University Press.

Fershtman, Chaim, and Dotan Persitz. 2021. "Social Clubs and Social Networks." *American Economic Journal: Microeconomics,* 13(1): 224–251.

Fisman, Raymond. 2001. "Estimating the Value of Political Connections." *American Economic Review,* 91(4): 1095–1102.

Gehlbach, Scott, Konstantin Sonin, and Milan W Svolik. 2016. "Formal models of nondemocratic politics." *Annual Review of Political Science,* 19: 565–584.

Gilens, Martin. 2012. *Affluence and Influence: Economic Inequality and Political Power in America.* Princeton University Press.

Groseclose, Tim. 2001. "A model of candidate location when one candidate has a valence advantage." *American Journal of Political Science,* 862–886.

Grossman, Gene M., and Elhanan Helpman. 1999. "Competing for Endorsements." *American Economic Review,* 89(3): 501–524.

Grossman, Gene M, and Elhanan Helpman. 2001. *Special Interest Politics.* MIT Press.

Guiso, Luigi, Helios Herrera, Massimo Morelli, and Tommaso Sonno. 2018. "Populism: Demand and Supply." *CEPR Discussion Paper No. DP11871.*

Guriev, Sergei, and Ellias Papaiannou. 2021. "The Political Economy of Populism." *Journal of Economic Literature*, forthcoming.

Harrington, Joseph E. 1992. "The Revelation Of Information Through The Electoral Process: An Exploratory Analysis." *Economics & Politics*, 4(3): 255–276.

Herrera, Helios, and Cesar Martinelli. 2006. "Group formation and voter participation." *Theoretical Economics*, 1(4): 461–487.

Inglehart, Ronald, and Pippa Norris. 2016. "Trump, Brexit, and the Rise of Populism: Economic Have-Nots and Cultural Backlash." *HKS Faculty Research Working Paper Series*, RWP16-026.

Kamenica, Emir, and Mathew Gentzkow. 2011. "Bayesian Persuasion." *American Economic Review*, 101(6): 2590–2615.

Kaneko, Mamoru, and Kenjiro Nakamura. 1979. "The Nash social welfare function." *Econometrica*, 423–435.

Kartik, Navin, and Richard van Weelden. 2019. "Informative Cheap Talk in Elections." *Review of Economic Studies*, 86(2): 755–784.

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

Kartik, Navin, Francesco Squintani, and Katrin Tinn. 2015. "Information Revelation and Pandering in Elections." *Working paper.*

Keefer, Philip, Carlos Scartascini, and Razvan Vlaicu. 2019. "Social Trust and Electoral Populism: Explaining the Quality of Government." Inter-American Development Bank IDB Publications (Working Papers) 10031.

Keefer, Philip, Sergio Perilla, and Razvan Vlaicu. 2021. "Trust, Collaboration, and Policy Attitudes in the Public Sector." Inter-American Development Bank IDB Publications (Working Papers) 11252.

Keen, Michael, and Joel Slemrod. 2017. "Optimal Tax Administration." *Journal of Public Economics*, 152: 133–142.

Khan, Rahman Shamus. 2012. "The Sociology of Elites." *Annual Review of Sociology*, 38(1): 361–377.

Knack, Stephen, and Philip Keefer. 1997. "Does social capital have an economic payoff? A cross-country investigation." *The Quarterly Journal of Economics*, 112(4): 1251–1288.

Marshall, Albert W, Ingram Olkin, and Barry C Arnold. 2011. "Inequalities: theory of majorization and its applications."

Martinelli, Cesar. 2006. "Would rational voters acquire costly information?" *Journal of Economic Theory*, 129(1): 225–251.

Mattozzi, Andrea, Massimo Morelli, and Marcos Nakaguma. 2020. "Populism and Wars." *CEPR Discussion Paper No. DP14501.*

Meltzer, Allan H, and Scott F Richard. 1981. "A Rational Theory of the Size of Government." *Journal of Political Economy*, 89(5): 914–927.

Michelman, Valerie, Joseph Price, and Seth D Zimmerman. 2021. "Old Boys' Clubs and Upward Mobility Among the Educational Elite*." *The Quarterly Journal of Economics*, 137(2): 845–909.

Moulin, Hervé. 2004. *Fair division and collective welfare.* MIT press.

Mudde, Cas. 2004. "The Populist Zeitgeist." *Government and Opposition*, 39(4): 541–563.

Mudde, Cas, and Cristóbal Rovira Kaltwasser. 2017. *Populism: A Very Short Introduction. Very short introductions*, Oxford University Press.

Myerson, Roger B. 2008. "The autocrat's credibility problem and foundations of the constitutional state." *American Political Science Review*, 102(1): 125–139.

Panova, Elena. 2017. "Partially Revealing Campaign Promises." *Journal of Public Economic Theory*, 19(2): 312–330.

Pastor, Lubos, and Pietro Veronesi. 2020. "Inequality Aversion, Populism, and the Backlash Against Globalization." *Chicago Booth Research Paper No. 20-11, Fama-Miller Working Paper.*

Persson, Torsten, and Guido Tabellini. 1994. "Is Inequality Harmful for Growth?" *The American Economic Review*, 600–621.

Piketty, Thomas, and Emmanuel Saez. 2003. "Income inequality in the United States, 1913–1998." *The Quarterly Journal of Economics*, 118(1): 1–41.

Prato, Carlo, and Stephane Wolton. 2016. "The Voters' Curses: Why We Need Goldilocks Voters." *American Journal of Political Science*, 60(3): 726–737.

Prato, Carlo, and Stephane Wolton. 2018. "Rational ignorance, populism, and reform." *European Journal of Political Economy*, 55(C): 119–135.

Ray, Debraj. 2011. *A game-theoretic perspective on coalition formation.* Oxford University Press.

Reinikka, Ritva, and Jakob Svensson. 2004. "Local Capture: Evidence from a Central Government Transfer Program in Uganda." *The Quarterly Journal of Economics*, 119(2): 679–705.

Roberts, Kevin. 2015. "Dynamic voting in clubs." *Research in Economics*, 69(3): 320–335.

Rodrik, Dani. 2017. "Populism and the Economics of Globalization." National Bureau of Economic Research, Inc NBER Working Papers.

Saez, Emmanuel, Joel Slemrod, and Seth H Giertz. 2012. "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review." *Journal of Economic Literature*, 50(1): 3–50.

Schnakenberg, Keith E. 2016. "Directional Cheap Talk in Electoral Campaigns." *The Journal of*

Politics, 78(2): 527–541.

Shorrocks, Anthony F. 1983. “Ranking income distributions.” *Economica*, 50(197): 3–17.

Tiebout, Charles M. 1956. “A pure theory of local expenditures.” *Journal of political economy*, 64(5): 416–424.

Zimmerman, Seth D. 2019. “Elite Colleges and Upward Mobility to Top Jobs and Top Incomes.” *American Economic Review*, 109(1): 1–47.

Zoega, Gylfi. 2017. “Increase trust to stem the rise of populist movements.” *Vox EU.org*.

A Proofs

Proof of Lemma 1

Suppose that (σ_C, σ_E) is a responsive equilibrium. Elites endorse the biased politician if

$$\begin{aligned}&\sigma_c(m_B)[B]\cdot x_E\left(\mathbb{E}\left[\theta^B|l,k\right],\alpha\right)+(1-\sigma_c(m_B)[B])\cdot x_E\left(\mathbb{E}\left[\theta^U|l,k\right],0\right)\\ \geq&\sigma_c(m_U)[B]\cdot x_E\left(\mathbb{E}\left[\theta^B|l,k\right],\alpha\right)+(1-\sigma_c(m_U)[B])\cdot x_E\left(\mathbb{E}\left[\theta^U|l,k\right],0\right).\end{aligned}$$

Plugging in the expressions for $x_E\left(\theta^B,\alpha\right)$ and $x_E\left(\theta^U,0\right)$ from Equation (5), and the expressions for θ^B and θ^U from Equations (2) and (3), we can rewrite the above condition as follows:

$$(\sigma_c(m_B)[B]-\sigma_c(m_U)[B])\cdot(2\mathbb{E}(\theta|l,k)-1+(1-\lambda)\cdot\alpha\psi)\geq 0$$

In a responsive equilibrium, $\sigma_c(m_B)[B]>\sigma_c(m_U)[B]$. Therefore, the above inequality condition is satisfied for all $\mathbb{E}(\theta|l,k)\geq\frac{1}{2}-\frac{(1-\lambda)\cdot\alpha\psi}{2}$, or equivalently when $l\geq\hat{l}=\frac{k}{2}-\left(\frac{k}{2}+1\right)\alpha\psi(1-\lambda)$.

Proof of Proposition 3

The negative effect of ψ and α on $\mathbb{E}C$ follows directly from Equation (14). To see the positive effect of k on $\mathbb{E}C$, compute the partial derivative:

$$\frac{\partial\mathbb{E}C}{\partial k}=\frac{1}{4N^2(k+1)^2}\left((2\beta+1)N^2+2\beta N+\beta^2(N-k)\left(N+5k+2k^2+4\right)\right).$$

This partial derivative is positive because $k<N$.

Proof of Lemma 2

We calculate and examine the first, second, and third derivatives of u^E, and draw the following implications. First, for large enough N, the function u^E is increasing at 0 and decreasing at $\frac{1}{2}$, i.e. $\frac{d}{d\lambda}u^E(0) > 0$ and $\frac{d}{d\lambda}u^E\left(\frac{1}{2}\right) < 0$. Next, for a sufficiently large N, the function u^E is concave in the neighbourhood of zero, $\frac{d^2}{(d\lambda)^2}u^E(0) < 0$. Finally, for a sufficiently large N, the third derivative is always positive in the interval $\lambda \in \left[0, \frac{1}{2}\right]$. This last observation implies that the second derivative can be zero at most once, which means that the function u^E can switch from concavity to convexity once, but cannot switch back to concavity.

Suppose that N is sufficiently large so the above three properties hold. Since the function u^E is continuous, increasing at 0 and decreasing at $\frac{1}{2}$, then it must have at least one (local) maximum at some value $\lambda' \in \left[0, \frac{1}{2}\right]$. To show that this local maximum is unique, it suffices to show that the function cannot have a local minimum. If it did, then there should be a point, at which the continuous function u^E switches from concavity to convexity, which is impossible as argued above.

Denote the unique maximum $\lambda^* = \lambda^*(N)$. Evaluating $u^{E\prime}(\cdot)$ at $\lambda^* N^{-\frac{1}{2}}$, we get an expression whose sign is determined by the term $1 - \alpha\psi\left(1 + 2(\lambda^*)^2\right)$. Thus, for a small $\varepsilon > 0$ and a sufficiently large N, we have that $\left(\sqrt{\frac{1}{2\alpha\psi} - \frac{1}{2}} - \varepsilon\right) N^{-\frac{1}{2}} < \lambda^*(N) < \left(\sqrt{\frac{1}{2\alpha\psi} - \frac{1}{2}} + \varepsilon\right) N^{-\frac{1}{2}}$. ■

Proof of Proposition 4

The first part of the proposition follows from Lemma 2. To prove that a responsive equilibrium exists when the Elites' share is $\lambda^* = k^*/N$, we rewrite the condition in Equation (11) as follows:

$$\frac{-N\alpha\lambda^2\psi + N\lambda - 2\alpha\lambda\psi - N\alpha\lambda\psi - 2\alpha\psi}{\alpha(\lambda + 1)(N\lambda + 2)} \geq 0.$$

The numerator is a quadratic function with two real roots, $\underline{\lambda}$ and $\overline{\lambda}$. A responsive equilibrium exists whenever $\lambda^* \in \left[\underline{\lambda}, \overline{\lambda}\right]$. This follows from the asymptotic boundedness of λ^* established in Lemma 2. ■

Proof of Proposition 5

Fix $\beta \equiv \alpha\psi$. Inspection of the derivative $\partial u^E(\lambda)/\partial\lambda$, where $u^E(\lambda)$ is given by Equation (15), reveals that for sufficiently large N this derivative has a single root in the interval $[0, 0.5]$, which we denote by λ^*. Moreover, the derivative $\partial u^E(\lambda)/\partial\lambda$ is positive for all $\lambda \in [0, \lambda^*)$ and negative for all $\lambda \in (\lambda^*, 0.5]$. Thus, $u^E(\lambda)$ is single-peaked over the interval $[0, 0.5]$ and attains a maximum at

λ^*.

Consider now the cross derivative $\partial^2 u^E / \partial\lambda\partial\beta$. Computation shows that the sign of this cross derivative is determined by (and is the same as) the sign of the following polynomial:

$$h(\lambda) \equiv 4N^2\lambda^3\beta - \left(2N\beta(2N-5) + 2N^2\right)\lambda^2 - (8\beta(N-1) + 4N)\lambda - 8\beta - 2 - 2N\beta.$$

Suppose that $N \geq 3$. Then, by Descartes' Rule of Signs, the polynomial $h(\lambda)$ has a single positive root. Since $h(0.5) < 0$, and in addition $h(\lambda) > 0$ as λ tends to ∞, we deduce that the single positive root of $h(\lambda)$ is greater than 0.5. Consequently, $h(\lambda)$ is negative for all values of $\lambda \in [0, 0.5]$. Therefore, as $\beta = \alpha\psi$ increases, the derivative $\partial u^E(\lambda)/\partial\lambda$ decreases *pointwise* in the interval $[0, 0.5]$. This implies that as β increases, the maximizer of $u^E(\lambda)$, i.e., λ^*, decreases.

Proof of Proposition 6

Suppose first that a commoner conducts one experiment that fails. By Equation (7), the density function of his posterior belief about θ is given by $\hat{f}(\theta|\text{one failure observed}) = 2(1-\theta)$. From the commoner's perspective, the probability of observing l successes when k more experiments are conducted is given by:

$$\Pr(l \mid k,\text{one failure observed}) = \int_0^1 2(1-\theta)\frac{k!}{l!(k-l)!}\theta^l(1-\theta)^{k-l}\,d\theta = \frac{2(k+1-l)}{(k+1)(k+2)}.$$

By Lemma 1, Elites endorse the biased politician if they observe at least $\hat{l}$ successes. From the commoner's perspective, the probability that exactly l successes are observed by Elites, conditional on the event that Elites observe at least $\hat{l}$ successful experiments, and that he observed one failed experiment, is then given by

$$\frac{\Pr(l \mid k,\text{one failure observed})}{\sum_{j=\hat{l}}^{k}\Pr(j \mid k,\text{one failure observed})} = \frac{\frac{2(k+1-l)}{(k+1)(k+2)}}{\sum_{j=\hat{l}}^{k}\frac{2(k+1-j)}{(k+1)(k+2)}} = \frac{2k+2-2l}{\left(k-\hat{l}+1\right)\left(k-\hat{l}+2\right)}.$$

Denote the conditional expectation of θ as a function of k by $H_F(k)$. We then have that:

$$H_F(k) = \sum_{l=\hat{l}}^{k}\frac{2k+2-2l}{\left(k-\hat{l}+1\right)\left(k-\hat{l}+2\right)}\cdot\mathbb{E}[\theta|l,k+1] = \frac{k+2\hat{l}+3}{3(k+3)}. \qquad (16)$$

The commoner votes for the biased politician whenever $2H_F(k) - 1 - \alpha\lambda\psi \geq 0$. Using Equation 16, the expression for $\hat{l}$ (as defined in Lemma 1), and the fact that $k = N\lambda$ we rewrite the above inequality as follows:

$$\frac{1}{3(N\lambda+3)}\left(-N\alpha\lambda^2\psi+\left(-5\alpha\psi+\left(\frac{1}{2}-\alpha\psi\right)2N\right)\lambda-(4\alpha\psi+3)\right)\geq 0.$$

Lemma 2 implies that for N sufficiently large, the sign of the left-hand side of the above inequality is determined by the sign of $\left(\frac{1}{2}-\alpha\psi\right)$. Since $\frac{1}{2}>\alpha\psi$, the commoner votes for the biased politician even though his experiment failed. A similar argument shows that if a commoner conducts a successful experiment, but Elites endorse the unbiased politician, the commoner finds it optimal to follow the advice of Elites.

Finally, the club size k^* is optimal for Elites even if Commons can conduct experiments. This is because, for sufficiently large N, Elites are always worse off when Commons acquire information and decide the outcome of the elections rather than follow Elites' recommendation. To see this, notice that by Equation (15), when N is sufficiently large and the club size is λ^*, the expected utility of an Elite member converges to $7/4+(\alpha\psi)/2+(\alpha^2\psi^2)/4$. When Commons vote based on their own signal, the quality of the elected politician is bounded above by $7/4$, the probability of electing the biased politician is bounded above by $1/2$, and the expected utility of an Elite member is therefore bounded above by $7/4+(\alpha\psi)/2$, according to Equation (5).

Consequently, when N is sufficiently large, a club of size λ^*, which is optimal when Commons cannot, or do not have an incentive to, acquire information, is better for Elites compared to any smaller club size that potentially induces Commons to conduct experiments. ■

Proof of Proposition 7

Suppose the game is played T times. In each round, the per-capita resources that each member in the Elites and Commons get are given by x^E and x^C, respectively, as defined in Equations (5) and (6). Also, in each round, the probability of each of the M_i-richest individuals in society $i\in\{1,2\}$ being selected into the Elites is given by $\frac{k^*}{M_i}$. With the remaining probability, the individual remains with the Commons. Thus, in expectation, each of the M_i-richest individuals in society i receives $R^{(i)}=\frac{k^*}{M_i}\cdot\mathbb{E}[x^E]+\left(1-\frac{k^*}{M_i}\right)\cdot\mathbb{E}[x^C]$ in every round. Any individual who is not among the M_i-richest individuals in society i receives $\mathbb{E}[x^C]$ in each round.

For each society $i\in\{1,2\}$ and $T>0$, define the vector:

$$w^{(i)}(T)=\left(w_1^{(i)}(T),...,w_N^{(i)}(T)\right)=w+T\Big(\overbrace{R^{(i)},...,R^{(i)}}^{M_i\text{ first indices}},\overbrace{\mathbb{E}[x^C],\ldots,\mathbb{E}[x^C]}^{n-M_i\text{ last indices}}\Big).$$

where $w=(w_1,\ldots,w_N)$ is the original wealth distribution, which is identical in both societies.

Because the selection into the Elites in each society is independent across rounds, by the weak law of large numbers, the average amount of resources that an individual who started with wealth w_k in society i receives after T rounds converges in probability to $w_k^{(i)}(T)$. This implies that for any $\varepsilon > 0$, there exists a number of rounds $\overline{T}$ such that if the game is played for more than $\overline{T}$ rounds, the probability that the wealth of each individual k in society i is within ε of $w_k^{(i)}(T)$, is greater than $1 - \varepsilon$.

Therefore, it remains to show that the wealth distribution $w^{(1)}(T)$ majorizes the wealth distribution $w^{(2)}(T)$, i.e. $w^{(2)}(T) \prec w^{(1)}(T)$. First, for any $j \in \{1, ..., M_1\}$, because $R^{(1)} > R^{(2)}$ and $M_1 < M_2$ we have that

$$\sum_{i=1}^{j} w_i^{(2)}(T) = \sum_{i=1}^{j} w_i + jTR^{(2)} < \sum_{i=1}^{j} w_i + jTR^{(1)} = \sum_{i=1}^{j} w_i^{(1)}(T).$$

Next, for any $j \in \{M_1 + 1, ..., M_2\}$,

$$\sum_{i=1}^{j} w_i^{(2)}(T) = \sum_{i=1}^{j} w_i + jTR^{(2)} \leq \sum_{i=1}^{j} w_i + T\left(k^*\mathbb{E}[x^E] + (M_1 - k^*)\mathbb{E}[x^C] + (j - M_1)\,\mathbb{E}[x^C]\right) = \sum_{i=1}^{j} w_i^{(1)}(T).$$

Finally, for $j \in \{M_2 + 1, ..., N\}$, it is

$$\begin{aligned}\sum_{i=1}^{j} w_i^{(2)}(T) &= \sum_{i=1}^{j} w_i + T\left(k^*\mathbb{E}[x^E] + (M_2 - k^*)\mathbb{E}[x^C] + (j - M_2)\,\mathbb{E}[x^C]\right) \\ &= \sum_{i=1}^{j} w_i + \left(k^*\mathbb{E}[x^E] + (M_1 - k^*)\mathbb{E}[x^C] + (j - M_1)\,\mathbb{E}[x^C]\right) = \sum_{i=1}^{j} w_i^{(1)}(T),\end{aligned}$$

which completes the proof. ∎