

Digital Populism and Minimalist Democracy*

Elliot Foote[†] Weijia Li[‡]

September 29, 2022

Abstract

We develop a model where a populist produces and distributes anti-“elite” rhetoric, in anticipation of how voters respond to the rhetoric. By doing so, we provide a framework to understand how populism is affected by changes in technology and the institutional environment. The model uncovers a complementarity between digital media and voter suppression in producing populism. The complementarity explains the uneven strength of populism across time and space. The complementarity produces two more implications. First, under strong voter suppression, a more “centrist” society encourages populism. Second, populism in the digital era is intrinsically hostile towards even a “minimalist” democracy, a key insight that explains the recent revival of voter suppression efforts.

*We thank Peter Lorentzen, Gérard Roland, and Yang Xie for especially valuable discussions. We have also benefited from thoughtful comments from Sascha Becker, Arthur Campbell, Sebastiano Della Lena, Hongyi Li, Vai-Lam Mui, Paul Raschky, Chengsi Wang, and Yves Zenou. We thank participants in the seminars at the Department of Economics and Soda Labs of Monash Business School, UC Riverside, the 2021 Australasian Econometric Society Meeting, and the 2022 Annual Meeting of Society for Institutional & Organizational Economics (SIOE) for their valuable comments.

[†]Monash University; elliott.foote@gmail.com.

[‡]Monash University. weijia.li@monash.edu.

1 Introduction

The recent wave of populism raises numerous puzzles. Populists achieve unprecedented electoral victories while the citizenry remain firmly centrist (Barber et al. (2015); Gentzkow (2016)). Average citizens even impose stronger resistance against populist messages (Chafetz and Pozen (2018); Acemoglu and Robinson (2019)). Despite the claim that populism has pervaded “almost all countries ruled by a constitutional democracy” (Urbinati (2019)), populism actually shows uneven strength across time and space, with several democracies notably exempt from its pressure.¹

In this paper, we construct a formal model that offers a unified framework to explain these puzzles. The model also generates novel insights into the relationship between populism and “minimalist democracy” in the digital era.² The model features an upcoming election where an establishment incumbent competes against an opposition. The opposition candidate runs either as another establishment candidate or as a populist candidate. A populist opposition can produce a rhetoric to criticize the incumbent. Rather than modeling generic negative campaigning, we impose assumptions on the rhetoric to capture the essential characteristic of populism: a populist rhetoric contrasts the “pure people” against the “corrupt elite” in a highly *moralized* manner (Müller (2016); Mudde and Kaltwasser (2017); Guriev and Papaioannou (2022)). After producing the rhetoric, the populist distributes his rhetoric through the media. If the populist can only access broad media outlets, he must distribute his rhetoric to the entire national audience. But if the populist can distribute his rhetoric through “targeted media,” he can choose a precise subgroup of voters as the audience for his rhetoric (Peitz and Reisinger (2015); Goldfarb and Tucker (2019); Zuboff (2019)).³ A populist rhetoric can influence a voter only if the voter receives the rhetoric with a sufficient “intensity.” The intensity can be interpreted as the minimal number of advertisements that can persuade a voter.⁴ Lastly, voters decide whether to turn out and if so, which candidate

¹Key examples that saw a recent surge in populism are Western European democracies (Taggart (2017)). Key examples that are exempt from the recent populist waves are East Asian democracies (Hellmann (2017)) as well as Australia (Moffitt (2017); Brett (2019)).

²A polity satisfies the “minimalist” criterion for democracy when the polity holds competitive elections (Schumpeter (2010); Przeworski (1991); Bidner et al. (2014);). But such a polity may not satisfy “liberal” criteria for a full-fledged liberal democracy, criteria that include rights protection and constraints on the executive (Mukand and Rodrik (2020); Li et al. (2022)). As we will discuss in detail later, previous research focuses on how populism erodes liberal institutions, while this paper focuses how populism may threaten electoral institutions.

³Extensively documented by Goldfarb and Tucker (2019), digital media such as Twitter and Facebook allows precise targeting of commercial advertisements, a crucial capacity that has been harnessed by populist politicians (Liberini et al. (2020)).

⁴The “intensity” can also admit various other interpretations, such as political retaliation against populism like protests.

they support. A voter who turns out pays an exogenous cost of voting. A high cost of voting is a sufficient statistic for an ideal-type suppression against all voters in a *uniform* manner.⁵

By producing and distributing his rhetoric, the populist attempts to win the election by swinging enough voters, while minimizing the persuasion cost. The constrained minimization problem yields the optimal populist rhetoric and the optimal audience for the rhetoric. The optimal targeted audience only includes voters who are disillusioned with establishment politics. Specifically, even though the populist is fully capable of persuading voters who are currently satisfied with the incumbent, the populist refuses to do so. This high homogeneity of the populist audience is driven by the essential feature of populism. An individual voter’s attitude is largely fixed towards any rhetoric that is recognized by the voter as populist. Such fixity is driven by the highly moralized nature of populist rhetoric (see Weber (2004) on the “absoluteness” of moral conviction, as well as Haidt (2001), Doris et al. (2020), and Skitka et al. (2021)). At the same time, different voters hold massively different attitudes towards a rhetoric that classifies voters into “morally pure people” and “morally corrupt elite.” Classified as “morally pure,” voters disillusioned with the incumbent is gratified by the populist rhetoric. These disillusioned voters are far more susceptible to the populist rhetoric than voters satisfied with the incumbent, voters who are attacked by the populist as “morally corrupt.” In equilibrium, the populist chooses a highly inflaming rhetoric so that he can win the election by only persuading voters who are disillusioned with the incumbent. The highly inflaming rhetoric only brings a negligible marginal cost to the populist because each individual voter is fixed in his attitude towards any populist rhetoric. Yet the highly inflaming rhetoric allows the populist to win the election without persuading any voters who are satisfied with the incumbent, henceforth avoiding a large jump in persuasion cost. Furthermore, because the optimal audience is highly homogeneous, the optimal audience reinforces the content of the populist rhetoric as a moralized and confrontational demarcation between the “pure people” versus the “corrupt elite.” To summarize, our model indeed captures essential features of populism rather than generic negative campaigning (Lau and Rovner (2009)). Populism and generic negative campaigning both attack the other *candidate*. But for negative campaigning, its far more heterogeneous audience cannot sustain a moralized classification of *voters*.

⁵It is well documented that voter suppression are strongly targeted against voters of color in the United States, especially African Americans (Bentele and O’Brien (2013); Hajnal et al. (2017); Anderson (2018); Fraga (2018)). We focus on uniform suppression as an ideal-type construct because such a theoretical exercise generates insights that could not otherwise be obtained if we only model targeted suppression. Specifically, we will show that even in a hypothetical world where only uniform voter suppression is possible, a digital populist still benefits enormously from repressing all voters. The analysis explains the impulse of a digital populist to repress even voters of his own base, therefore constituting a huge threat to “minimalist” electoral institutions.

The high homogeneity of the optimal populist audience determines the comparative statics in an especially sharp manner. These comparative statics summarize how populism responds to shocks in technology and institutions. First, the opposition is more likely to run as a populist with stronger voter suppression, even for an ideal-type suppression that imposes a uniform cost on all voters. Under stronger uniform suppression, a populist can replace the most moderate voters in his optimal audience with another group of voters who are discontinuously more extreme in their ideology, therefore far more amenable to populist rhetoric. Such a replacement reduces the cost of populist persuasion. In other words, with little voter suppression, the populist must persuade a audience that is still quite moderate, an audience that is skeptical of the populist rhetoric; but with strong voter suppression, the populist only needs to persuade an extreme audience, an audience that is especially susceptible to populist rhetoric. Note that the result is driven by the high homogeneity of the optimal audience for the populist rhetoric. If the optimal audience also included voters satisfied with the incumbent, stronger uniform suppression on voters would induce the opposition candidate to also persuade voters who are even more satisfied with the incumbent. This competing effect under a heterogeneous audience would reduce the desirability to send out the rhetoric.

Second, the opposition is more likely to run as a populist when targeted media is more effective. Effective targeted media allows the populist to send his rhetoric only to the optimal audience, a narrow audience that minimizes the cost of populist persuasion. By contrast, populist persuasion is too expensive on the national media, which does not allow target messaging. Unable to target his message, the populist would rather lose the election than spend a tremendous amount on broad persuasion, most of which would be wasted on irrelevant voters.

Third, targeted media and voter suppression complement each other in producing populism. The populist only benefits from voter suppression when he can target his rhetoric to the ideal audience with its extreme ideology. The populist does not benefit at all from voter suppression when he has to distribute his rhetoric to the national audience. Therefore, under more effective targeted media, stronger voter suppression produces a larger marginal benefit for the populist. On the flip side of the complementarity, targeted media also produces a larger marginal benefit when voter suppression is stronger. The populist especially values the capability to target an audience when the populist can win the election by targeting a more extreme audience, a situation that is secured by stronger voter suppression. Note that the complementarity is also driven by the highly homogeneous audience for the populist rhetoric.

We want to highlight that all these results are derived under a highly “centrist” distri-

bution of voter preferences. Specifically, we focus on unimodal and symmetric distributions, where the “largest” mass of voters cluster around the average voter. We also assume a state of the world where the average voter⁶ is satisfied with the establishment incumbent. Such a moderate environment is traditionally considered to be ideal for a democracy to function well (Lipset (1959); Moore (1993); Persson and Tabellini (2002); Levitsky and Ziblatt (2018); Svobik (2019); Graham and Svobik (2020)). But the traditional wisdom is upended by voter suppression, in which a large cluster of centrist voters becomes irrelevant as they no longer turn out to vote. Our model actually predicts that populism is *boosted* by a more centrist society when voter suppression is strong. These results call attention to an insight that has not received sufficient attention: in the digital age, voter suppression may neutralize or even overturn the capacity of a centrist society to deter populism.

Many other implications are uncovered by the model, especially by the complementarity between targeted media and voter suppression. The complementarity informs one of the most heated debates about populism, its relationship with democracy. Populists themselves declare that they will liberate the “silent majority” from elite domination (Müller (2016)). Leading scholars also describe populism as an “extreme majoritarianism” (Urbinati (2017)). “Populism strongly champions popular sovereignty and majority rule but opposes minority rights and pluralism[...]In short, populism is essentially democratic,” only at odds with “minority rights as well as the ‘institutional guarantees’ that should protect them” (Mudde and Kaltwasser (2017); see also McCormick (2001) and Gerbaudo (2017)). But our model seriously questions that populists can ever speak for the “silent majority” in the digital era. Digital populism actually thrives through strong voter suppression, therefore intrinsically hostile towards even a minimalist democracy (Schumpeter (2010); Przeworski (1999); Fearon (2011); Bidner et al. (2014)). For digital populism, the ideal regime turns out to be a facade democracy where voters are as suppressed as possible, including the populist’s own base. Extreme digital populism in power is indistinguishable from a full dictatorship that disenfranchises almost everyone.

By contributing to the conceptual debate about populism, our model also explains the dramatic revival of voter suppression efforts in the United States. Some recent Republican assaults on voting rights are especially puzzling because they undermine electoral institutions that have strongly improved turnout among conservative voters, such as new restrictions on postal voting (Brett (2019); we discuss in detail how postal voting traditionally improved turnout among Republican voters of Florida in Section 5.2). Our model readily resolves the puzzle. The model shows that after digital populism has captured a party, the party will be

⁶The average voter is also the “median” voter and the “mode” voter under a unimodal and symmetric distribution.

enthusiastic about suppressing even its own base. Such voter suppression would block the more moderate supporters of a populist party, minimizing the cost to produce and distribute inflaming populist rhetoric.

The complementarity also offers a unified explanation for the uneven strength of populism in time and space. Populism was generally weak in the second half of the 20th century, despite draconian voter suppression in some rich democracies (Anderson (2018); Brett (2019)). Our model attributes the weakness of populism in this era to the dominance of national media, a dominance that made populist persuasion prohibitively expensive even when many voters were highly suppressed. But with the recent rise of targeted media, a populist can bypass many citizens who are increasingly resistant to populist rhetoric (Chafetz and Pozen (2018)), instead focusing persuasion efforts on a small and receptive audience. Yet such digital populism only imposes an extraordinary threat to some democracies but not others. Digital populism especially struggles in Australia, a democracy noted for the entire absence of “widespread populist movement, or even a particular party that looks to establish a more permanent populist opposition to the major parties” (Moffitt (2017)). Our model attributes Australia’s resilience to its progressive electoral institutions that reinforce each other and jointly minimize voting cost. The minimal voting cost is reflected by the extremely high turnout rate in Australia, never below 90% for the general election since 1927. The specific mechanism of our model also finds strong support. Brett (2019) summarizes it well in her discussion of the compulsory voting system in Australia. Compulsory voting “tempers the impact of the passionate and committed voters of the base with the votes of the moderate and indifferent. It lowers the emotional temperature of our politics and keeps open the sensible centre. The angry and aggrieved will always be drawn to politics, but compulsory voting ensures that they are not the main occupants of the public square.” The Australian case is of special interest because Australia has long been a pioneer of progressive electoral institutions, including secret ballot, compulsory voter registration, compulsory voting, and non-partisan electoral administration (Acemoglu and Robinson (2012); Acemoglu and Robinson (2013); Brett (2019)). Generalized by our model, the Australian lesson informs current debates about electoral reforms and voting rights in other liberal democracies, highlighting the fundamental role of a low voting cost in resisting digital populism.

More related literature Numerous recent articles explore the many causes of populism. “Demand” side explanations attribute populism to socioeconomic grievances, and a major debate is whether economic grievance or sociocultural grievance plays a more fundamental role (Guiso et al. (2017); Colantone and Stanig (2019); Margalit (2019); Noury and Roland (2020); Dorn et al. (2020); ;Cachanosky et al. (2021); Cachanosky et al. (2022)). On the “supply”

side, the literature explains the rise of populism by institutional causes that reduce the influence of ordinary voters on policy-making, such as campaign finance deregulation and technocratic bureaucracies (Acemoglu et al. (2013); Berman (2021)). There is also substantial attention to the role of media (Guriev and Treisman (2020); Zhuravskaya et al. (2020); Guriev and Papaioannou (2022)). To the best of our knowledge, our paper is the first formal model that focuses on how voter suppression directly affects populism, especially in its interaction with digital media. By identifying voter suppression as a fundamental cause of digital populism, we engage with the central debate on the relationship between populism and (minimalist) democracy (Schumpeter (2010); Przeworski (1999); Bidner et al. (2014)). Because digital populism thrives under strong voter suppression, digital populists in power would threaten the very foundation of minimalist democracy via their intrinsic impulse to exclude as many voters as possible. Therefore, our model highlights that populism is far more dangerous than the evaluation of previous literature, a literature that focuses on how populism erodes *liberal* institutions such as “constraints on the executive, checks and balances, rule of law, and independent bureaucratic agencies” (Guriev and Papaioannou (2022); see also Müller (2016), Urbinati (2019), and Sasso and Morelli (2021); also see the literature on the economic consequences of populism: Grier and Maynard (2016) and Absher et al. (2020)).

As to the specific topic of voter turnout and populism, the most relevant to our paper is Guiso et al. (2017). The pioneering empirical paper shows that economic insecurity causes strong populism by incorporating voter turnout as a key mediator variable. Compared with the existing literature on voter turnout and populism (Guiso et al. (2017); Huber and Ruth (2017); Houle and Kenny (2018); Leininger and Meijers (2021)), our contributions are two-fold. First, we uncover a novel mechanism where voter suppression promotes populism only through targeted media, offering a unified explanation for a few seemingly unrelated puzzles of populism. Second, our paper constructs a formal model of voter suppression and populism, which contributes to a formal literature on populism noted for its paucity (Guriev and Papaioannou (2022)), complementing the more empirical contributions.

The paper is organized as follows. Section 2 sets up the main model. Section 3 solves the model, characterizing the optimal populist message and the optimal audience for the message. Section 4 implements detailed analysis of comparative statics. Section 5 discusses implications of the model. Section 6 concludes.

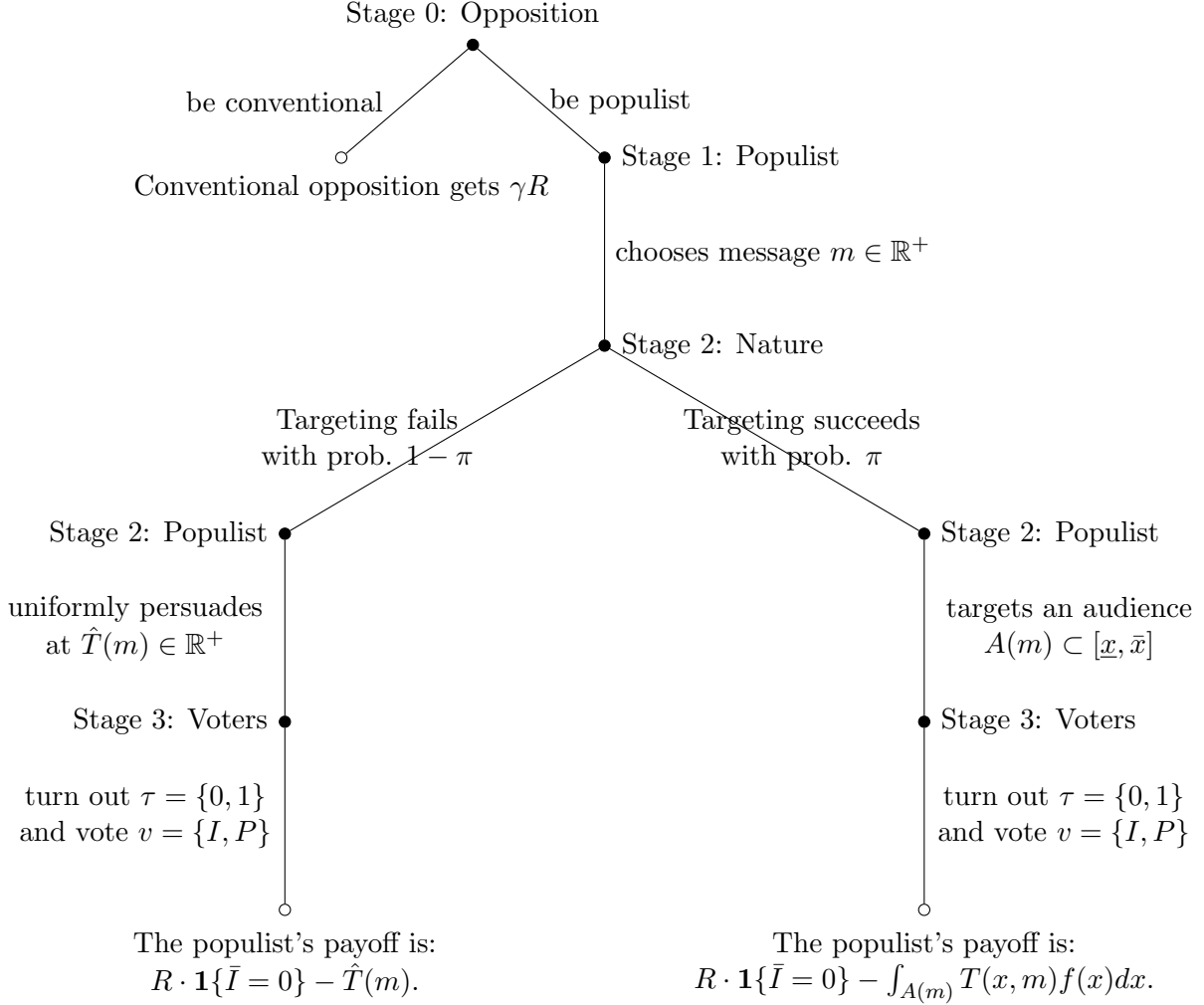


Figure 1: Extensive form of the game

2 Setup of the model

The model is a sequential game. The extensive form is in Figure 1. There are three players, an incumbent candidate, an opposition candidate, and a continuum of voters. A voter's satisfaction with the incumbent is $x \in [\underline{x}, \bar{x}] \subset \mathbb{R}$, $-\infty \leq \underline{x} < 0 < \bar{x} \leq \infty$. Voter satisfaction follows the cumulative distribution function $F(\cdot)$, and the probability density function is $f(\cdot)$. A voter's satisfaction with a populist is normalized to 0.

At Stage 0, the opposition chooses to run as a conventional or populist candidate. If the opposition runs as a conventional candidate, the game ends. The conventional opposition wins the election with an exogenous probability $\gamma \in (0, 1)$.⁷ An electoral victory confers an

⁷By choosing to engage in a conventional electoral competition, the opposition competes with the incumbent on a traditional policy space. Following the logic of Downsian electoral competition, the opposition chooses the same policy platform as the incumbent, and the winner is determined stochastically. We will

exogenous utility of $R \in \mathbb{R}^{++}$ to the opposition. If the opposition runs as a populist, the game continues.

At Stage 1, the populist chooses a message $m \in \mathbb{R}^+$.

At Stage 2, the populist tries to convince voters of the message m . If a voter is convinced of the message m , his satisfaction with the incumbent is reduced to

$$x - m. \tag{1}$$

To convince a voter whose *ex ante* satisfaction with the incumbent is x , the populist must spend a minimum of

$$T(x, m) \in \mathbb{R}^{++} \tag{2}$$

on the voter. The derivative T_x is strictly positive: it is more difficult to persuade a voter who is more satisfied with the incumbent. The other derivative T_m is also strictly positive: it is more difficult to convince any voter of a more inflaming message.

The persuasion protocol depends on $\pi \in (0, 1)$, the effectiveness of a media targeting technology. With probability $1 - \pi$, the targeting technology fails and the populist can only access a broad media outlet (e.g., national media). As such, the populist can only implement uniform persuasion: the populist spends the same amount

$$\hat{T}(m) \in \mathbb{R}^+ \tag{3}$$

on all voters. If $\hat{T} < T(x, m)$, a voter x is unconvinced of the message m and his satisfaction with the incumbent remains at x . If $\hat{T} \geq T(x, m)$, a voter x is convinced of the message m and his satisfaction with the incumbent is reduced to $x - m$.

With probability π , the targeting technology succeeds. The populist can choose a targeted audience for the message m :

$$A(m) \subset [\underline{x}, \bar{x}]. \tag{4}$$

To each voter with $x \in A(m)$, the populist commits to spend $T(x, m)$. The set $A(m)$ can be any Lebesgue measurable subset of $[\underline{x}, \bar{x}]$.

The capacity to choose a precise audience makes the analysis quite challenging because we need to show that the optimal audience yields a higher payoff than any other Lebesgue measurable subsets of “persuadable” voters (see Proposition 2 and its proof). Nevertheless, we believe that it is necessary to endow such a capacity so we can model the full potential of the power of digital targeting technology (Zuboff (2019)). Moreover, we will show that even

assign γ to be a random variable. By changing the support of the random variable, we can further model an incumbent advantage (details in Section 3.5).

with such a flexible capacity to choose any Lebesgue measurable subset of “persuadable” voters, the populist in equilibrium chooses a highly homogeneous audience (Proposition 5). The maximal flexibility in targeting capacity allows us to derive the homogeneity of the populist audience, a defining feature of populism, as an endogenous outcome rather than an assumption. This endogenous homogeneity is the key driver of the comparative statics that uniform voter suppression encourages populism, as well as the complementary between voter suppression and digital media in producing populism.

At Stage 3, a voter decides whether to turn out and if so, which candidate to support. Denote $2c > 0$ as the cost of voting (2 is multiplied to simplify later algebra). The voter follows a simple turnout rule. If a voter is convinced of a message m , the voter turns out if and only if

$$|x - m| \geq 2c. \tag{5}$$

If a voter is unconvinced of any message, the voter turns out if and only if

$$|x| \geq 2c. \tag{6}$$

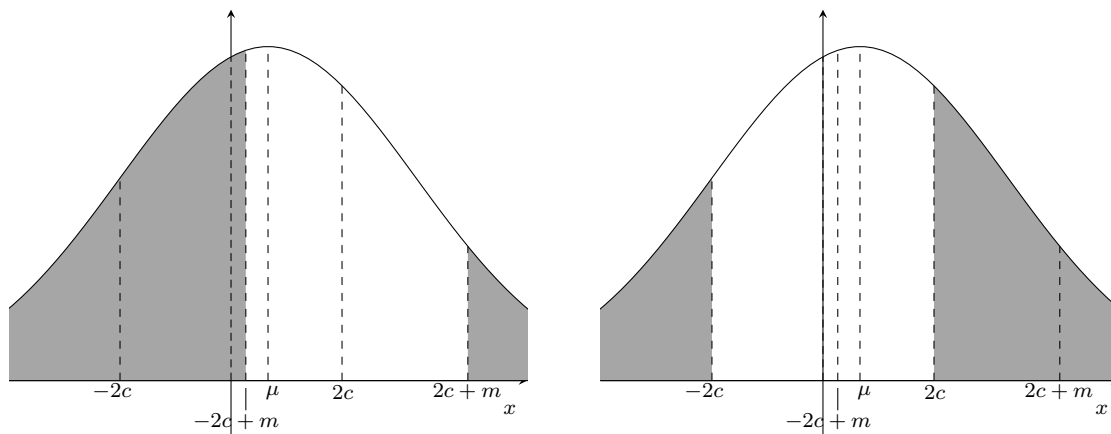
The turnout decision is determined by a standard trade-off between the cost and the benefit of voting. The right-hand side of Equation 5 is the cost of voting. The left-hand side is the difference in utility from electing the incumbent ($x - m$) versus the populist (0), the “benefit” of voting for a voter who is convinced of a message m . For a voter who is unconvinced of any message, her satisfaction with the incumbent remains at x . Her turnout decision is described by Equation 5 by setting $m = 0$, yielding Equation 6.

The decision rule generates a very reasonable turnout behavior. Illustrated in Figure 2, the shaded area are voters who turn out, and the uncolored area are voters who abstain. The left panel is for voters who is convinced of a message m , and the right panel is for a voter who is unconvinced of any message. Voters only turn out if they have relatively “extreme” preferences. A voting citizen either strongly prefers the incumbent over the populist or the other way around. A citizen does not vote if she finds not much of a difference between the incumbent and the populist.

The turnout decision also implies which candidate a voter supports. For the population with $x \geq m + 2c$, they vote for the incumbent because the incumbent in power yields a higher payoff than the populist ($x - m \geq 2c > 0$). Similarly, the population with $x \leq m - 2c < 0$ vote for the populist.

The solution concept is subgame perfect equilibrium. A player maximizes his expected

Figure 2: the turnout decision of voters



payoff. Specifically, if targeting fails, the populist’s payoff is:

$$R \cdot \mathbf{1}\{\bar{I} = 0\} - \hat{T}(m).$$

The function $\mathbf{1}\{\cdot\}$ is the indicator function. The parameter $\bar{I} = 1$ if more than 50% of the voting citizens support the incumbent; otherwise, $\bar{I} = 0$.

Similarly, if targeting succeeds, the populist’s payoff is:

$$R \cdot \mathbf{1}\{\bar{I} = 0\} - \int_{A(m)} T(x, m) f(x) dx.$$

2.1 Assumptions

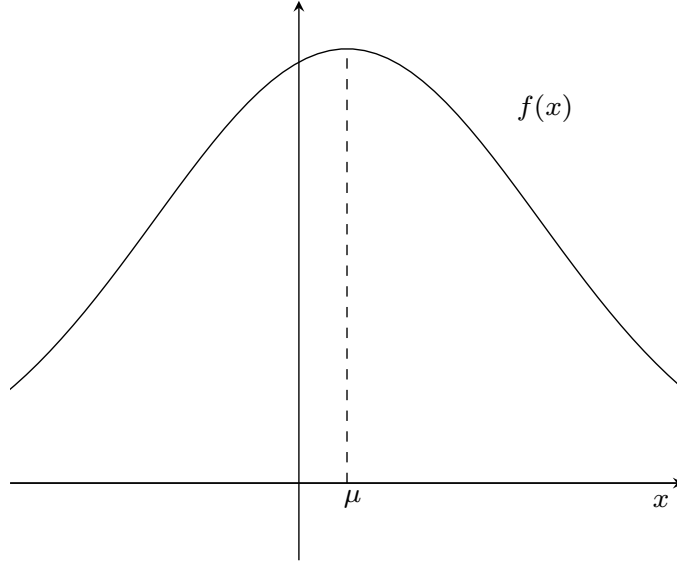
We now state the assumptions that help maintain realistic outcomes. The first assumption looks at the probability density function of voter satisfaction $f(\cdot)$.

Assumption 1. *$f(\cdot)$ has full support on its domain, is uni-modal, and is symmetric around the mean/median/mode $\mu > 0$. In addition, $\sup_x f(x) < \infty$.*

Figure 3 is an example that satisfies Assumption 1. The assumption nests many “nice” distributions, such as uniform distributions and normal distributions. Notice that there are (weakly) more “centrist” voters than extreme voters. The average voter, who is also the defining centrist voter, is more satisfied with the incumbent than a populist ($\mu > 0$).⁸ Such a centrist distribution of voter preferences usually pressures politicians to practice moderate politics, a pressure that stabilizes democracy. But even in such a moderate environment, we will show that digital populism can be menacing.

⁸To repeat, the average voter is also the “median” voter and the “mode” voter under a unimodal and symmetric distribution.

Figure 3: Assumption 1 on a symmetric and uni-modal distribution



For another way to understand Assumption 1, recall that the incumbent actually seeks a re-election in the game. Before the game in Figure 1 begins, the incumbent has served a term as a politician. Under Assumption 1, the incumbent might have spent a huge amount in persuasion so that the public holds a favorable view of the incumbent over the populist before the game in Figure 1 begins. The large persuasion spending by incumbent before the game generates another important consequence. Precisely because the incumbent spent so much on persuasion in the past, voters have been fully “saturated” with the incumbent’s message. So the incumbent cannot further change their view of himself in the current game of Figure 1.

Assumption 2. $2c > \mu$.

This assumption is illustrated in Figure 4. The assumption is necessary to ensure that the turnout rate is less than 100%.

Assumption 3. $\int_{-2c}^{-2c+2\mu} T(x, 2\mu)f(x)dx < R$.

The assumption is illustrated in Figure 5. The assumption says that it is cheap to indoctrinate a populist message when the message can be targeted to a precise audience. The left hand side of Assumption 3,

$$\int_{-2c}^{-2c+2\mu} T(x, 2\mu)f(x)dx, \tag{7}$$

Figure 4: Assumption 2 on a non-trivial voting cost

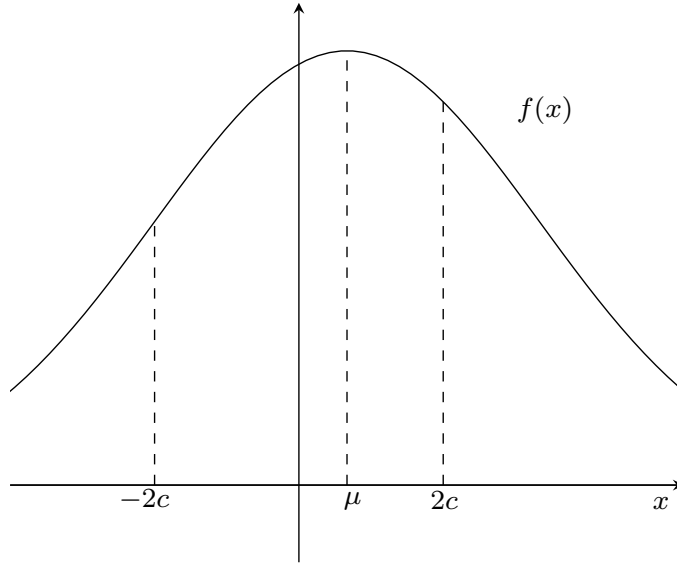
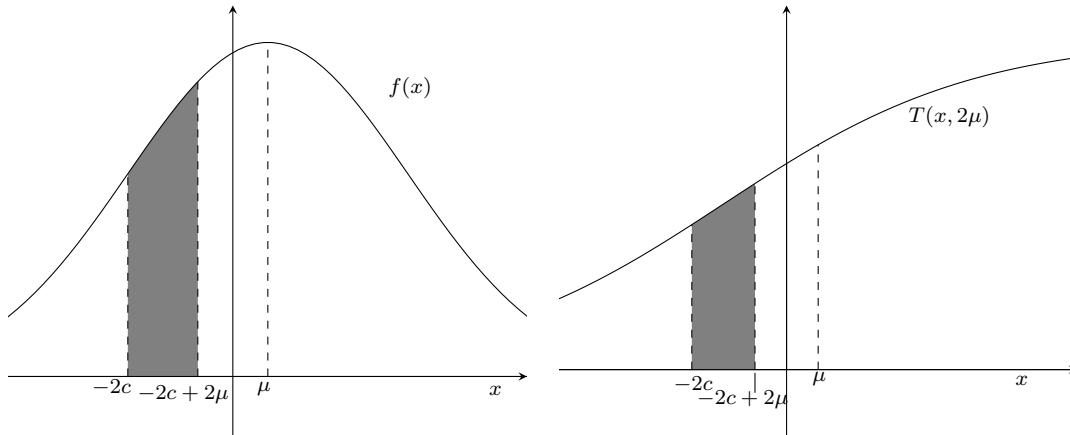


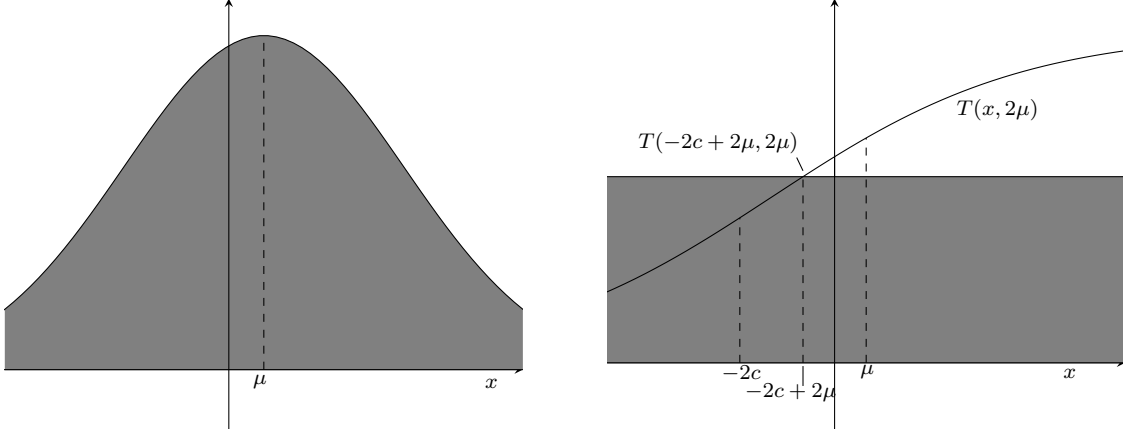
Figure 5: Assumption 3 that targeted persuasion is cheap



is the expenditure to convince voters with $x \in [-2c, -2c + 2\mu]$ of the populist message $m = 2\mu$. Thus, the darkly shaded area in the left panel of Figure 5 includes all convinced voters, and the right panel shows the persuasion spending on each voter. When the populist can target his message, Equation 7 turns out to be the minimal expenditure that secures an electoral victory. Per Assumption 3, the minimal expenditure that secures the election is smaller than R , the exogenous utility of winning the election. Assumption 3 captures a fundamental element of our story: through targeted media, the populist can win the election by only persuading a small and receptive audience, minimizing the persuasion cost.

Assumption 4. $T(-2c + 2\mu, 2\mu) > R$.

Figure 6: Assumption 4 that uniform persuasion is too expensive



The assumption is illustrated in Figure 6. The assumption says that it is expensive to broadcast a populist message uniformly to the national audience. The left hand side of Assumption 4,

$$T(-2c + 2\mu, 2\mu), \tag{8}$$

is the expenditure to broadcast the message $m = 2\mu$ to the entire voter population. The expenditure is just enough to convince a voter with $x = -2c + 2\mu$ of the message $m = 2\mu$, as well as all voters with $x < -2c + 2\mu$ (the right panel of Figure 6). For voters with $x > -2c + 2\mu$, even though they receive advertisements on the national media, they are unconvinced of the message. We later show that Equation 8 is the minimal expenditure that secures an electoral victory when the populist can only broadcast his message on the national media. When the message can only be broadcast on the national media, the expenditure is so high that even an electoral victory cannot justify it. Assumption 4 complements Assumption 3 and captures another key element of our story: when a populist can only access the national media, it is too costly to uniformly broadcast the populist message to such a huge audience. The bulk of the persuasion expenditure is wasted on irrelevant voters because the populist cannot target his audience.

The next assumption warrants more detailed discussion because the assumption captures essential elements of populism.

Assumption 5. 1. *There exists $\underline{T} > 0$ and $m_1 \in (0, \mu)$, for all $x \in [\underline{x}, \bar{x}]$ and for all $m \in [m_1, \infty)$,*

$$T_x(x, m) > \underline{T}. \tag{9}$$

2. *There exists $\tilde{t} > 0$ sufficiently small and $m_2 \in (0, \mu)$, for all $x \in [\underline{x}, \bar{x}]$ and for all*

$m \in [m_2, \infty)$, the following restriction is true.

$$T_m(x, m) < \tilde{t}. \tag{10}$$

Figure 7: Assumption 5 that persuasion cost changes a lot across individuals

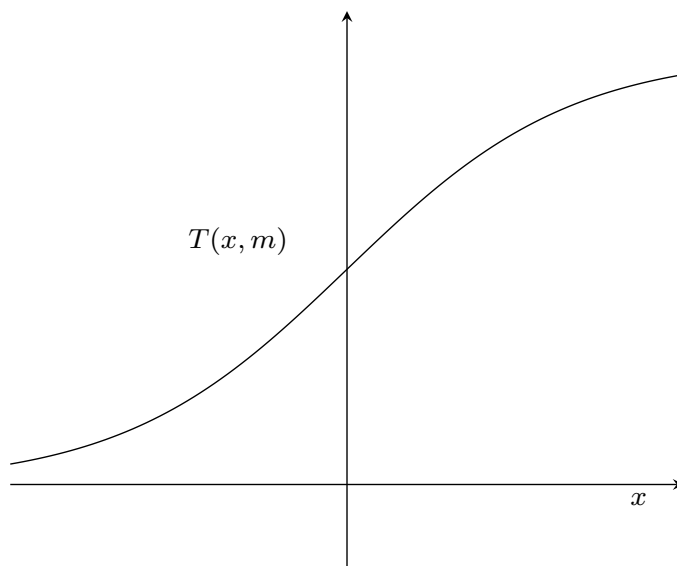
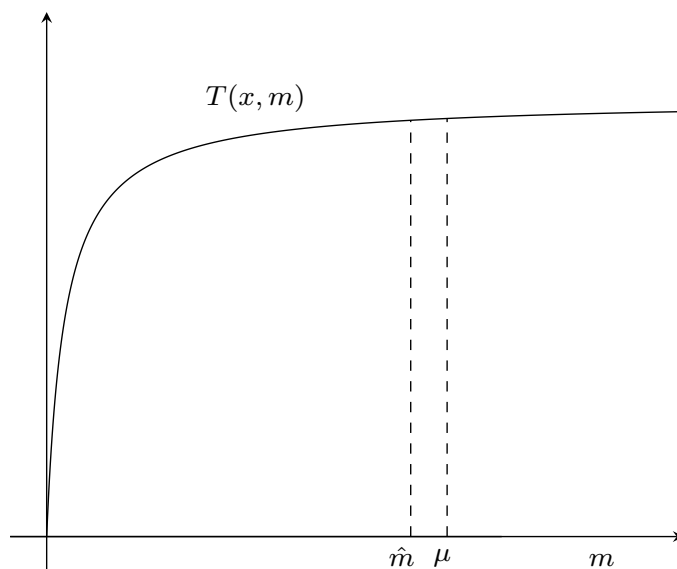


Figure 8: Assumption 5 that persuasion cost is largely “fixed” for inflaming messages



The assumption has two parts. The first part says that for a message sufficiently inflaming, it is strictly more expensive to persuade a voter who is more satisfied with the incumbent by a non-trivial degree (illustrated in Figure 7). The second part says that for a message sufficiently inflaming, any message that is more inflaming only increases the persuasion cost

by a negligible amount (illustrated in Figure 8). Therefore, the persuasion cost changes in qualitatively different manners along the dimension of voters versus the dimension of the message intensity. We now discuss Assumption 5 in the next section.

2.2 Populism as a moralized classification

Assumption 5 captures essential elements of populism. Assumption 5 are imposed on a message that is sufficiently inflaming ($m > m_1$ and $m > m_2$), an attribute that is usually considered necessary for a message to be labelled as “populist” (Guriev and Papaioannou (2022)). Under Assumption 5, each individual is largely fixed in their attitude towards *any* populist messages that are sufficiently inflaming, but different individual holds massively different attitudes towards populism. These two features capture the key elements in the standard definition of populism. Populism is defined as a *moralized* classification of the citizenry into “us” versus “them.” Under a populist rhetoric, its supporters are the *morally* pure “people”, while its opponents are the *morally* corrupt “elite.”⁹ The highly moralized classification should generate strong responses from voters, responses that we attempt to capture by Assumption 5.

Specifically, the populist rhetoric classifies opponents of the incumbent as morally pure. Therefore, if a voter is disillusioned with the incumbent, the voter is morally gratified by the populist rhetoric. Such a voter will probably hold a positive attitude towards populism. More importantly, the attitude is largely fixed because of the highly moralized nature of populist rhetoric. The voter categorizes a highly moralized message itself either as moral or immoral, with little grey space in between. A disillusioned voter continues to identify a populist message as categorically moral, as the message gratifies the disillusioned voter even more flatteringly. The disillusioned voter will fall under the spell of any populist message in an equally swift manner, a key point formalized by Assumption 5.

Similarly, the populist rhetoric classifies supporters of the incumbent as morally corrupt. For a voter who favors the incumbent, he is attacked as immoral by the populist rhetoric. Retaliating the moral attack, such a voter will probably identify a populist message itself as categorically immoral. Suppose that the populist message attacks the “corrupt elite”

⁹Here are two of the most influential definitions of populism.

First, populism, “I suggest, is a particular moralistic imagination of politics, a way of perceiving the political world that sets a morally pure and fully unified—but, I shall argue, ultimately fictional—people against elites who are deemed corrupt or in some other way morally inferior” (Müller (2016)).

Second, “we define populism as a thin-centered ideology that considers society to be ultimately separated into two homogeneous and antagonistic camps, ‘the pure people’ versus ‘the corrupt elite,’ and which argues that politics should be an expression of the *volonté générale* (general will) of the people” (Mudde and Kaltwasser (2017)). This definition is adopted by Guriev and Papaioannou (2022), a comprehensive survey on populism at the *Journal of Economic Literature*.

even more viciously, a group that includes the voter favoring the incumbent. The voter will continue identifying the more vicious message as categorically immoral. The voter will be equally repulsed by any message that falls into the *category* of immoral messages. This key point is again captured by Assumption 5.

To summarize, Assumption 5 is a parsimonious way to capture the two essential, strongly interrelated features of any populist rhetoric: first, the strong moralism; second, the demarcation between “the pure people” and “the corrupt elite.” The second part of Assumption 5 formalizes the strong moralism in populist rhetoric by restricting the voters’ response to populism to be categorical, an essential characteristic of moral reasoning (see Weber (2004) on the “absoluteness” of ethical conviction, along with Weber (1946) on the “rationalization” of value systems; also, see Haidt (2001), Doris et al. (2020), and Skitka et al. (2021)). The first part of Assumption 5 focuses more on the sharp classification of the citizenry, which naturally generates massively different responses from voters who are classified into different groups.

Assumption 5 generates outcomes that reinforce the moralized classification

Here is another way to understand Assumption 5. With Assumption 5, the model generates an equilibrium outcome that is further consistent with the essential feature of populism as a sharply moralized classification of the citizenry. Without Assumption 5, the model generates an equilibrium outcome that can be inconsistent with this essential feature. Therefore, Assumption 5 imposes restrictions on the fundamentals of the model so that the model is indeed about populism.

Specifically, Assumption 5 drives the key result for our paper (Proposition 5). We will give a brief preview of the result, highlighting how the result further reinforces the defining feature of populism as a moralized demarcation between the “good people” and the “evil elite.” Because of Assumption 5, the populist will only persuade disillusioned voters in equilibrium even though the populist is fully capable of persuading voters in favor of the incumbent (Proposition 5). Therefore, the equilibrium audience for the populist rhetoric is a highly homogeneous group of voters. Furthermore, the ideology of the equilibrium audience (measured by x) is very distant from the ideology of the equilibrium voters for the incumbent. Assumption 5 indeed generates two separate groups who are qualitatively different from each other, reinforcing the strong rhetoric that contrasts a group of “good people” with the other group of “evil elite.” Loosely speaking, Assumption 5 generates outcomes that further reinforce the *content* of the populist message.

If Assumption 5 is not satisfied, the opposition may choose to persuade both disillusioned voters and voters in favor of the incumbent. The audience for the message is therefore a

heterogeneous collection of voters. This heterogeneous audience for the message would not be able to sustain a rhetoric of sharp classification of the *voters*. Instead, even though the message criticizes the incumbent, the message will also tend to seek *unity* among the voters in order to communicate with such a heterogeneous audience.

This observation is very important. It implies that Assumption 5 restricts our model to populism rather than a generic form of negative campaigning (surveyed by Lau and Rovner (2009)). With Assumption 5, the opposition sends the message to a highly homogeneous audience, reinforcing the moralized classification in the content of the populist message. Without Assumption 5, the model can be about generic negative campaigning against the incumbent. The opposition will send the message to a much more heterogeneous audience, strengthening the urgency to highlight unity among voters in his message. One can deduce another signature attribute of populism: its strong negativity towards both the incumbent politician and a significant group of *voters* labelled as part of the “morally corrupt elite.” By contrast, negative campaigning only attacks the incumbent politician.

The highly homogeneous audience for the populist message generates the crucial comparative statics that stronger voter suppression boosts populism, indicating that the populist wants to suppress even its own base. This comparative statics further induces the complementarity between voter suppression and digital media in producing populism. Therefore, the peculiar or even pathological features of populism are all traced back to Assumption 5 that formalizes the definition of populism as a moralized classification of the citizenry.

3 Solving the Model

3.1 Stage 2: the maximal audience

Prior to this stage, the populist fixed a message m in Stage 1. Before we derive the optimal persuasion strategies of the populist, we characterize the maximal audience that can be swung by the message m . Denote the maximal audience as

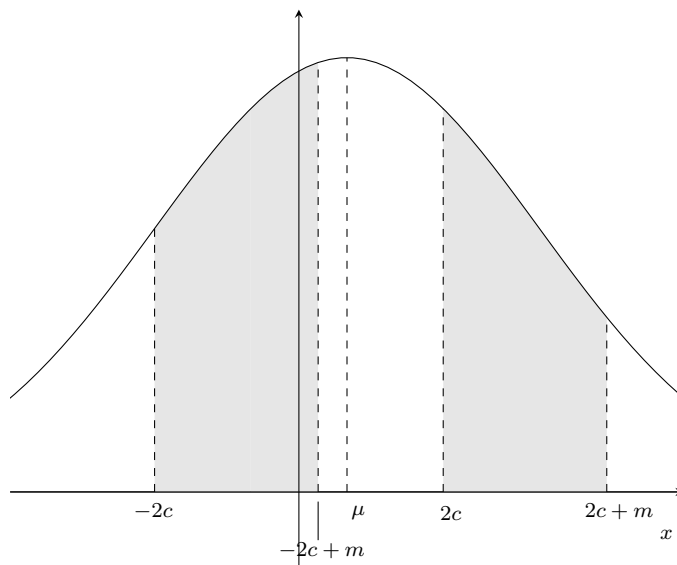
$$\bar{A}(m) \subset [\underline{x}, \bar{x}]. \tag{11}$$

It is straightforward to verify Lemma 1.

Lemma 1. *Suppose that the populist spends enough to convince all voters of a message m . The message m only changes the voting behavior of voters with:*

$$x \in \bar{A}(m) \equiv [-2c, -2c + m] \cup [2c, 2c + m]. \tag{12}$$

Figure 9: the maximum audience that can be swung by a message m



The maximal audience $\bar{A}(m)$ is illustrated in Figure 9, an example when the two intervals do not overlap. Notice that the populist message m can only swing two groups of voters. Voters with $x \in [-2c, -2c + m]$ would not vote for the populist without being convinced of the message m . But if convinced, these voters cast ballots for the populist. We call this group of voters “marginal populist voters.”¹⁰

In the other component of $\bar{A}(m)$, voters with $x \in [2c, 2c + m]$ would vote for the incumbent without being convinced of the message. But if convinced, they would not vote for the incumbent. We call this group of voters “marginal incumbent voters.”¹¹

The populist message cannot influence the other three groups of voters. Voters with $x < -2c$ always vote for the populist and voters with $x > 2c + m$ always vote for the incumbent. Voters with $x \in (-2c + m, 2c)$, if exist, always abstain.

3.2 Stage 2 when targeting fails

With probability $1 - \pi$, the targeted media fails to work. The populist can still access the national media. But on the national media, the populist can only implement uniform persuasion. To convince any voters of the message m , the populist has to spend the *same* amount on *all* voters. The populist’s optimal persuasion spending is characterized by the

¹⁰To further label these voters as disillusioned with establishment politics, one may impose an assumption that $c > \mu$. This extra assumption is entirely for labeling and interpretational convenience. All analysis and results are valid without the assumption.

¹¹If the first group $[-2c, -2c + m]$ intersect with the group $[2c, 2c + m]$, then the marginal incumbent voters are $[2c, 2c + m] \setminus [-2c, -2c + m]$.

following proposition.

Proposition 1. *Suppose that the targeting fails. The populist spends nothing on uniform persuasion:*

$$\text{for any } m \in \mathbb{R}^+, \hat{T}^*(m) = 0. \quad (13)$$

With only access to the national media, the populist spends nothing to persuade voters. So the entire population is unconvinced of the populist message, and the voting population elects the incumbent. To demonstrate Proposition 1, we first characterize the minimum uniform persuasion that secures an election for any message. Then we show that among all possible messages $m \in \mathbb{R}^+$, it is cheapest to broadcast $m = 2\mu$ to win the election. The final step applies Assumption 4, which says that $T(-2c + 2\mu, 2\mu)$, the minimum spending of uniform persuasion that secures an election by broadcasting $m = 2\mu$, is still higher than R , the utility from securing the election. Thus, Proposition 1 formalizes the idea that uniform persuasion of a populist message is too expensive (Figure 6). The populist would rather lose the election than spend a tremendous amount, most of which would be wasted on irrelevant voters.

Proposition 1 explains why populists were rare in democracies before the rise of targeted media. In a world without targeted media, a world where a populist could only win the election by spending a tremendous amount on uniform persuasion through the national media, the opposition always runs as a conventional candidate.

3.3 Stage 2: the optimal audience when targeting succeeds

With probability π , the media targeting technology allows the populist to choose a precise audience for the populist message. We now characterize the optimal audience for any populist message.

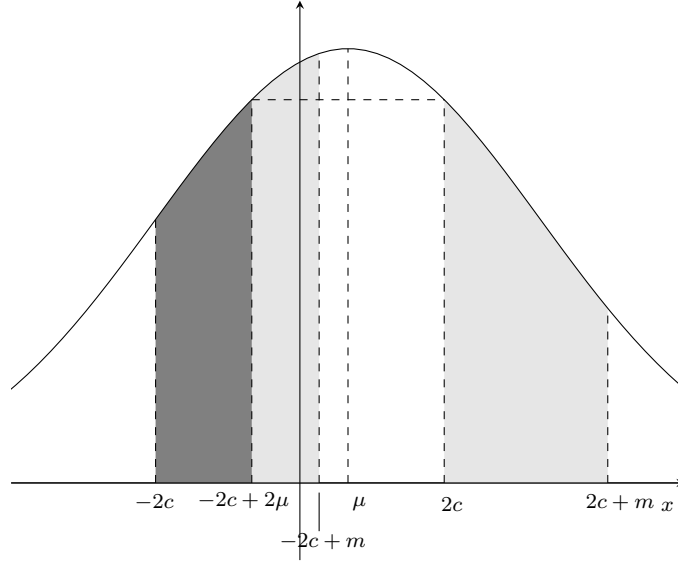
3.3.1 The optimal audience for an inflaming message

Suppose that in Stage 1, the populist has fixed a message that is sufficiently “inflaming:”

$$m \in [2\mu, \infty). \quad (14)$$

The optimal audience for such an inflaming message is characterized in the following proposition.

Figure 10: optimal audience for a message $m \in [2\mu, \bar{m})$



Proposition 2. *Suppose that $m \in [2\mu, \infty)$. There exists a unique $\bar{m} \in (2\mu, \infty]$, such that the optimal audience for the message m is*

$$A^*(m) = \begin{cases} \emptyset & \text{for } m \in [\bar{m}, \infty) \\ [-2c, -2c + 2\mu] & \text{for } m \in [2\mu, \bar{m}) \end{cases}. \quad (15)$$

The populist's payoff is:

$$V(m) = \begin{cases} 0 & \text{for } m \in [\bar{m}, \infty) \\ R - \int_{-2c}^{-2c+2\mu} T(x, m) f(x) dx & \text{for } m \in [2\mu, \bar{m}) \end{cases}. \quad (16)$$

Thus, there is a cutoff level of message intensity \bar{m} . If the message is more inflaming than the cutoff \bar{m} , the populist does not attempt to convince any voters of the message. The unconvinced population elects the incumbent. The losing populist receives a payoff of 0.

If the message is less inflaming than the cutoff \bar{m} , the persuasion strategy is illustrated in Figure 10.¹² The darkly shaded area is the optimal audience $A^*(m)$ and the lightly shaded area includes all other voters in the maximal audience $\bar{A}(m)$. The maximal audience consists of two groups, marginal populist voters $[-2c, -2c + m]$ and marginal incumbent voters $[2c, 2c + m]$. But the optimal audience, $[-2c, -2c + 2\mu]$, only includes marginal populist voters. The persuasion strategy induces all voters with $x \leq -2c + 2\mu$ to support

¹²The figure illustrates the persuasion strategy when the two intervals of the maximal audience do not overlap. The overlapping happens when $m > 4c$, or when the message is extremely inflaming. Proposition 2 is valid with or without overlapping.

the populist. The populist’s vote share is the same as the incumbent’s, securing the election for the populist.¹³ We obtain the maximal payoff for a winning populist:

$$V(m) = R - \int_{-2c}^{-2c+2\mu} T(x, m)f(x)dx. \quad (17)$$

R is the payoff from winning the election. $\int_{-2c}^{-2c+2\mu} T(x, m)f(x)dx$ is the minimal spending to persuade all voters in $[-2c, -2c + 2\mu]$: the populist must spend $T(x, m)$ on each voter, and $f(x)$ is the probability density of voter satisfaction at x .

How did we prove Proposition 2? The basic intuition is also illustrated by Figure 10. The audience $[-2c, -2c + 2\mu]$ secures just enough votes so that the populist wins the election. More importantly, in the maximal audience $\bar{A}(m)$ that can be swung by the message, the group $[-2c, -2c + 2\mu]$ is the least satisfied with the incumbent. Targeting this group minimizes persuasion spending.¹⁴ The minimal spending might still be too large, however, if the populist fixed an extremely inflaming message m in Stage 1. Therefore there is a cutoff \bar{m} . If the message is more inflaming than \bar{m} , it is better to lose the election rather than bear the large persuasion spending that is necessary to win the election. If the message is less inflaming than \bar{m} , it is optimal to spend on persuasion and win the election. Notice that it is possible for the cutoff to be infinite ($\bar{m} = \infty$). In this case, the populist convinces the necessary audience of any message $m \in [2\mu, \infty)$.

To summarize, here is the most important observation of Proposition 2: for any message that is sufficiently inflaming, the non-empty optimal audience only includes marginal populist voters.

3.3.2 The optimal audience for an “intermediate” message

Suppose that in Stage 1, the populist has fixed a message that is “intermediate” in its intensity:

$$m \in [\mu, 2\mu]. \quad (18)$$

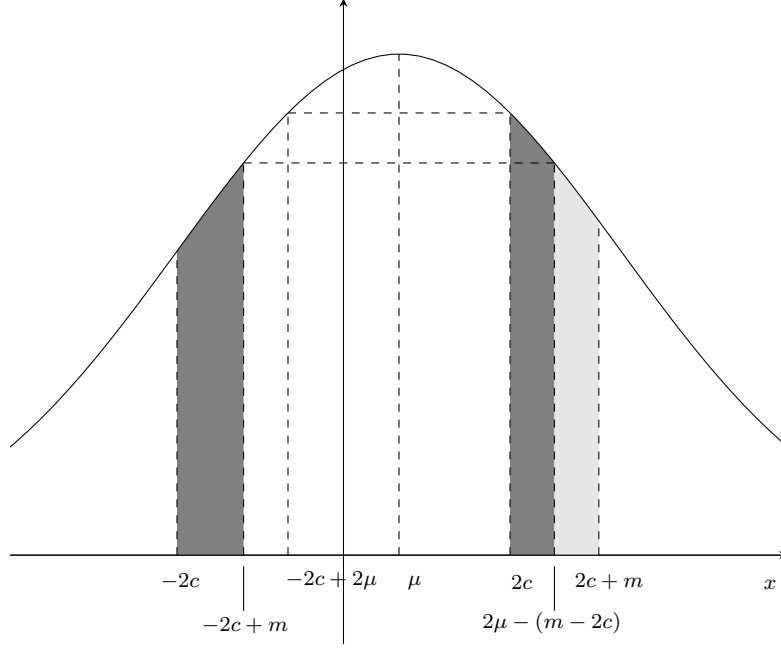
The optimal audience for an intermediate message is characterized as follows.

Proposition 3. *Suppose that $m \in [\mu, 2\mu]$. There exists a unique $\underline{m} \in [\mu, 2\mu)$, such that the*

¹³Vote shares are equal, $F(-2c + 2\mu) = 1 - F(2c)$, because the probability density function f is symmetric around the mean μ .

¹⁴More rigorously, the set $[-2c, -2c + 2\mu]$ belongs to the unique class of optimal audience, where each member of the class can only differ from the set $[-2c, -2c + 2\mu]$ by a measure of zero.

Figure 11: Optimal audience for a message $m \in [\underline{m}, 2\mu]$



optimal audience is

$$A^*(m) = \begin{cases} [-2c, -2c+m] \cup [2c, 2\mu - (m-2c)] & \text{for } m \in [\underline{m}, 2\mu] \\ \emptyset & \text{for } m \in [\mu, \underline{m}] \end{cases}. \quad (19)$$

The populist's payoff is:

$$V(m) = \begin{cases} R - \int_{-2c}^{-2c+m} T(x, m) f(x) dx - \int_{2c}^{2\mu - (m-2c)} T(x, m) f(x) dx & \text{for } m \in [\underline{m}, 2\mu] \\ 0 & \text{for } m \in [\mu, \underline{m}] \end{cases}. \quad (20)$$

Thus, there is another cutoff level of message intensity \underline{m} . If the message is less inflaming than the cutoff, the populist does not attempt to convince any voters of the message. The unconvinced population elects the incumbent. The losing populist receives a payoff of 0.

If the message is more inflaming than the cutoff \underline{m} , the persuasion strategy is illustrated in Figure 11. Again the darkly shaded area is the optimal audience $A^*(m)$ and the lightly shaded area includes all other voters that are in the maximal audience $\bar{A}(m)$. The optimal audience includes all marginal populist voters and some marginal incumbent voters. The persuasion strategy secures the election for the populist, who wins the same vote share as the

incumbent. We obtain the maximal payoff as a winning populist with a message $m \in [\underline{m}, 2\mu]$:

$$V(m) = R - \int_{-2c}^{-2c+m} T(x, m)f(x)dx - \int_{2c}^{2\mu-(m-2c)} T(x, m)f(x)dx. \quad (21)$$

R is the payoff from winning the election. The other two terms are the costs of persuading two different groups of voters.

The main intuition of Proposition 3 is also illustrated in Figure 11. A winning populist persuades voters who are the least satisfied with the incumbent. But if the populist only persuades marginal populist voters, the populist's vote share is still smaller than the incumbent's. To win the election, the populist must also convince some marginal incumbent voters to abstain from the election. The persuasion strategy gives the populist the same vote share as the incumbent (Figure 11), securing the election for the populist.

Again the minimal persuasion spending might still be too large, larger than the utility of winning the election. But different from the above section, it turns out that the persuasion spending is larger when the message is milder. Therefore we obtain a cutoff \underline{m} . Below the cutoff, the populist does not attempt to convince anyone of the message $m < \underline{m}$. Above the cutoff, the populist secures the election by convincing two groups of marginal voters. Notice that it is possible for the cutoff $\underline{m} = \mu$. In this case, the populist convinces a non-empty audience of any message $m \in [\mu, 2\mu]$.

3.3.3 The optimal audience for a weak message

Suppose that in Stage 1, the populist has fixed a weak message:

$$m \in [0, \mu]. \quad (22)$$

The optimal audience is characterized by the following proposition.

Proposition 4. *For any $m \in [0, \mu]$, the optimal audience is*

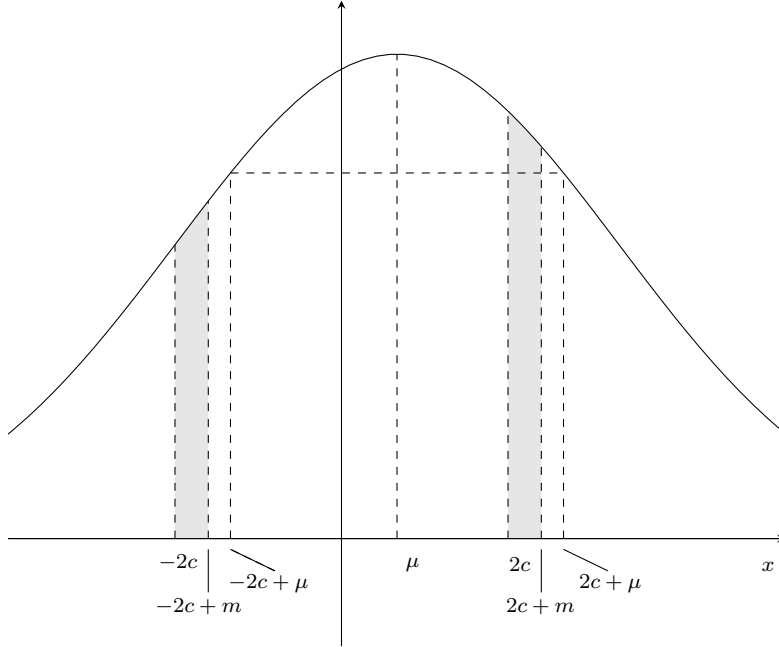
$$A^*(m) = \emptyset. \quad (23)$$

The populist's payoff is:

$$V(m) = 0. \quad (24)$$

The populist does not attempt to convince any voters of a weak message m . The intuition is illustrated in Figure 12, where the shaded area is the maximal audience $\bar{A}(m)$. With such a weak message, the populist cannot win one half of the total votes even if he convinces

Figure 12: Optimal audience for a message $m \in [0, \mu)$



every voter in the maximal audience. Thus, our model encodes the notion that any winning populist must be a sufficiently aggressive one, an important feature of populism (Guriev and Papaioannou (2022)). It is better for the populist to lose the election by spending nothing on persuasion rather than to convince any voters of such a mild message, while still losing the election.

3.4 Stage 1: the optimal message and its optimal audience

Proposition 1 to Proposition 4 fully characterize the expected payoff for the populist as a function of the message m . With probability $1 - \pi$, targeting fails and the populist receives a payoff of 0 (Proposition 1). With probability π , targeting succeeds and the populist's payoff $V(m)$ are characterized by Equations 16, 20, and 24. Therefore, we obtain the populist's expected payoff $\bar{V}(m) = \pi V(m)$ as a function of the message m :

$$\bar{V}(m) = \begin{cases} \pi[R - \int_{-2c}^{-2c+2\mu} T(x, m)f(x)dx] & \text{if } m \in [2\mu, \bar{m}) \\ \pi[R - \int_{-2c}^{-2c+m} T(x, m)f(x)dx - \int_{2c}^{2\mu-(m-2c)} T(x, m)f(x)dx]. & \text{if } m \in [\underline{m}, 2\mu] \\ 0 & \text{if } m \in [0, \underline{m}) \cup [\bar{m}, \infty] \end{cases} \quad (25)$$

Figure 13: The optimal message and its optimal audience

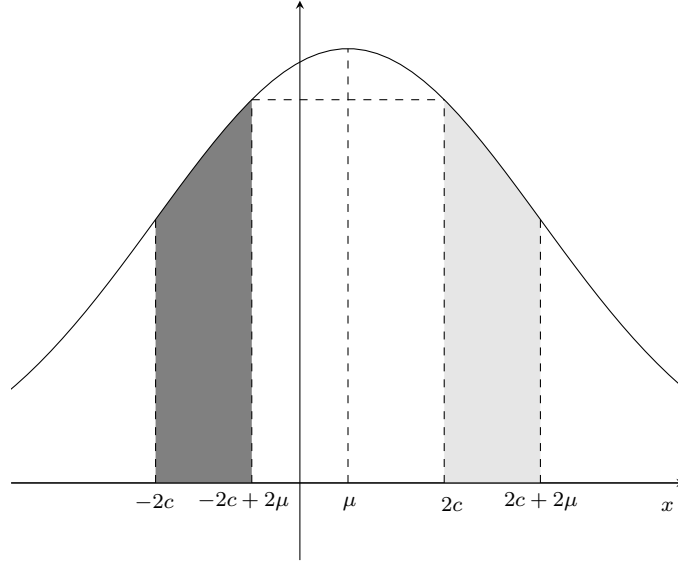
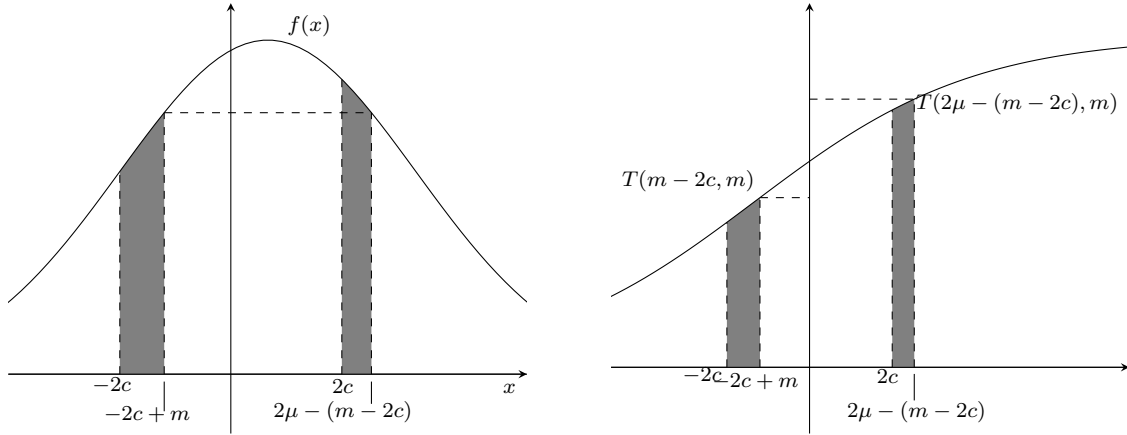


Figure 14: The replacement effect of a stronger message



With the expected payoff function $\bar{V}(\cdot)$, we are ready to derive the optimal message m^* . In Stage 1, the populist chooses the message m^* to maximize $\bar{V}(\cdot)$:

$$m^* = \arg \max_m \bar{V}(m). \quad (26)$$

The optimal message and the equilibrium path are as follows.

Proposition 5. *The optimal message is:*

$$m^* = 2\mu. \quad (27)$$

With probability π , the message $m^* = 2\mu$ convinces the optimal audience

$$A^*(m^*) = [-2c, -2c + 2\mu]. \quad (28)$$

With probability $1 - \pi$, the message $m^* = 2\mu$ convinces no voters.

The Proposition is illustrated in Figure 13. The darkly shaded area is the optimal audience for the optimal message $m^* = 2\mu$. The lightly shaded area includes all other voters in the maximal audience. Notice that the optimal message $m^* = 2\mu$ is the least inflaming message that allows a winning populist to only persuade marginal populist voters, ignoring all marginal incumbent voters.

Proposition 5 is the central proposition that prepares the subsequent comparative statics. Thus, we offer a detailed sketch of the proof in text. First, we show that $m^* = 2\mu$ yields a higher payoff than any message that is more inflaming. For $m \in [2\mu, \bar{m})$, the payoff function is (Equation 25):

$$\bar{V}(m) = \pi[R - \int_{-2c}^{-2c+2\mu} T(x, m)f(x)dx]. \quad (29)$$

The payoff function decreases in $m \in [2\mu, \bar{m})$:

$$\bar{V}'(m) = -\pi[\int_{-2c}^{-2c+2\mu} \overbrace{T_m(x, m)}^{>0} f(x)dx] < 0. \quad (30)$$

Any sufficiently inflaming message $m \in [2\mu, \bar{m})$ targets the same audience $[-2c, -2c + 2\mu]$, marginal populist voters who are the most dissatisfied with the incumbent. Therefore, the message 2μ saves on the persuasion spending than any message that is more inflaming.

Second, we show that $m^* = 2\mu$ yields a higher payoff than any message that is less inflaming. For $m \in [\underline{m}, 2\mu]$, The derivative $\bar{V}'(m)$ is more complicated. But there are essentially only two effects. The same as before, a more inflaming message is more costly to indoctrinate on any voter, reducing the populist's payoff $\bar{V}(m)$. This is a standard effect of a more inflaming message. But a more inflaming message generates another effect. A more inflaming message expands the group of marginal populist voters, shrinking the group of marginal incumbent voters who must be persuaded. In other words, a more inflaming message allows the populist to replace a group of marginal incumbent voters with a group of marginal populist voters, the latter being much more amenable to the populist message. Thus, the replacement effect reduces the persuasion spending and increases the payoff $\bar{V}(m)$. The key observation is that the replacement effect must dominate the persuasion cost effect. The persuasion cost effect is qualitatively "small" because of the second part of Assumption 5, which restricts the marginal persuasion spending for an already inflaming message.

Meanwhile, the replacement effect must be qualitatively large because the marginal populist voters are *discontinuously* easier to persuade than marginal incumbent voters (reinforced by the first part of Assumption 5). Thus, $m^* = 2\mu$ yields a higher payoff than any message that is less inflaming. In equilibrium, the populist chooses $m^* = 2\mu$, a message inflaming enough so that the populist does not need to convince any marginal incumbent voters.

To see the two effects more clearly, we need to explicitly solve the derivative of the payoff function $\bar{V}(m)$ for $m \in [\underline{m}, 2\mu]$. Recall that the payoff function $\bar{V}(m)$ for $m \in [\underline{m}, 2\mu]$ is (Equation 25)

$$\bar{V}(m) = \pi[R - \int_{-2c}^{-2c+m} T(x, m)f(x)dx - \int_{2c}^{2\mu-(m-2c)} T(x, m)f(x)dx]$$

Apply Leibniz formula, the derivative is:

$$\begin{aligned} \bar{V}'(m) = & -\pi \left\{ T(-2c + m, m)f(-2c + m) + \int_{-2c}^{-2c+m} T_m(x, m)f(x)dx \right. \\ & \left. - T(2\mu - (m - 2c), m)f(2\mu - (m - 2c)) + \int_{2c}^{2\mu-(m-2c)} T_m(x, m)f(x)dx \right\}. \end{aligned} \quad (31)$$

Because the probability density function $f(\cdot)$ is symmetric around μ , $f(m - 2c) = f(2\mu - (m - 2c))$ (the left panel of Figure 14). Also, Assumption 5 states that marginal persuasion spending is small for an already inflaming message $m \geq \mu$. Thus, the effect of persuasion cost is small:

$$\int_{-2c}^{-2c+m} T_m(x, m)f(x)dx + \int_{2c}^{2\mu-(m-2c)} T(x, m)f(x)dx < \tilde{t}.$$

Equation 31 implies:

$$\bar{V}'(m) \geq \pi \{ f(2\mu - (m - 2c)) [T(2\mu - (m - 2c), m) - T(-2c + m, m)] - \tilde{t} \}. \quad (32)$$

Notice that

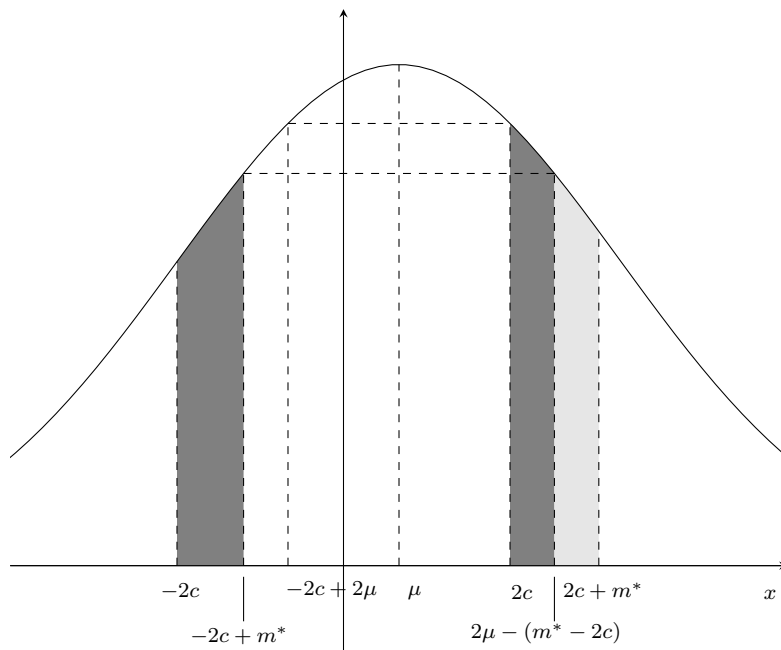
$$f(2\mu - (m - 2c)) [T(2\mu - (m - 2c), m) - T(-2c + m, m)]$$

is the replacement effect. When the message m becomes slightly more inflaming, the populist replaces a $f(2\mu - (m - 2c))$ mass of marginal incumbent voters with the same mass of marginal populist voters (the left panel of Figure 14). These replaced incumbent voters are discontinuously more difficult to persuade than the marginal populist voters: $T(2\mu - (m - 2c), m) - T(-2c + m, m) \gg 0$ (the right panel of 14). \tilde{t} is an upper bound on the persuasion cost effect, a second-order effect because the message is already difficult enough to persuade

(Assumption 5).

In equilibrium, the populist chooses a message inflaming enough so that he only needs to advertise the message to populist voters; the populist does not advertise the message to any incumbent voters. The optimal audience is a connected set, henceforth highly homogeneous. The homogeneous audience reinforces the *content* of the populist message as a sharp demarcation between two groups, the “morally superior people” and the “morally corrupt elite.”

Figure 15: For a generic negative message, the optimal audience is heterogeneous.



Populism and generic negative campaigning Before we investigate the comparative statics in detail, we briefly look at a situation where the optimal audience for the message is qualitatively more heterogeneous. This will help us to further appreciate the significance of the homogeneous audience.

Suppose that the second part of Assumption 5 does not hold. For instance, the partial derivative $T_m(x, m)$ is also bounded away from zero. In this case, the message is presumably not highly moralized because voters respond to the message in a manner that is qualitatively more flexible. Then the optimal message m^* might be an interior solution:

$$m^* \in (\mu, 2\mu).$$

In this case, the optimal audience includes two sets of voters (Figure 15):

$$[-2c, -2c + m^*] \cup [2c, 2\mu - (m^* - 2c)].$$

Notice that the two sets of voters are *disconnected*, so the optimal audience is quite heterogeneous. Presumably such a heterogeneous audience would not be receptive to a message that sharply demarcates between the morally good people and the morally corrupt elite. The opposition can only implement a generic negative campaigning against the incumbent politician, but not against voters who are satisfied with the incumbent.

To summarize, the audience in Proposition 5 is highly homogeneous for two reasons. First, each individual is essentially fixed in his attitude towards a message that is highly moralized. Second, attitudes towards a populist message are discontinuously different between the two pivotal groups of marginal voters, one group presumably classified as “morally pure people” and the other group as “morally corrupt elite.” The two groups are so different also because they are separated by a non-voting population, its existence due to the sufficiently high cost of voting. The highly homogeneous audience for the populist message leads to the key result of the model, the complementarity between targeted media and voter suppression in producing populism.¹⁵

3.5 Stage 0: populism or conventional politics

Proposition 5 immediately yields the maximal payoff for the opposition to be a populist. He will choose the optimal message $m^* = 2\mu$ and obtains:

$$\bar{V}(m^* = 2\mu) = \pi[R - \int_{-2c}^{-2c+2\mu} T(x, 2\mu)f(x)dx]. \quad (33)$$

To understand the payoff 33, first note that the media targeting technology succeeds with probability π . The populist obtains $[R - \int_{-2c}^{-2c+2\mu} T(x, 2\mu)f(x)dx]$, the payoff as a populist who wins the election by persuading marginal populist voters. With probability $1 - \pi$, the media targeting technology fails. The populist does not attempt to convince anyone of the message $m^* = 2\mu$ and loses the election, receiving 0.

If the opposition runs as a conventional candidate, his payoff is γR , $\gamma \sim U[0, 1]$.¹⁶ Thus,

¹⁵In Appendix B, we show that with a heterogeneous audience, we cannot unambiguously derive the complementarity.

¹⁶One can easily incorporate more incumbent advantage by assigning $\gamma \sim U[0, \bar{\gamma}]$, with small $\bar{\gamma}$ indicating a large incumbent advantage when the opposition runs as a conventional candidate. The parameter $\bar{\gamma}$ can be a function of μ , the incumbent advantage if the opposition runs as a populist. All main results stay the same when $\bar{\gamma}$ is a function of μ .

the opposition runs as a populist if and only if

$$\pi[R - \int_{-2c}^{-2c+2\mu} T(x, 2\mu)f(x)dx] \geq \gamma R. \quad (34)$$

We obtain the key object of the model, the probability that the opposition runs as a populist:

$$\Pi \equiv \frac{\pi[R - \int_{-2c}^{-2c+2\mu} T(x, 2\mu)f(x)dx]}{R}. \quad (35)$$

4 Comparative Statics

4.1 Targeted media and voter suppression

The next proposition is the key result of the paper. It states comparative statics on Π , the probability of populism, with respect to targeted media and voter suppression.

Proposition 6. *1. $\Pi_\pi > 0$: the opposition is more likely to run as a populist if the targeting technology (π) improves.*

2. $\Pi_c > 0$: the opposition is more likely to run as a populist if voter suppression is stronger (c increases).

3. The cross derivative $\Pi_{\pi c} > 0$: in a democracy with stronger voter suppression, the marginal effect of targeted media in producing populism is larger.

Targeted media promotes populism The first result formalizes the notion that targeted media generates populism. When a populist can only access the national media, he can never target his message. The populist would have wasted a tremendous amount of persuasion spending on irrelevant voters. The populist would not attempt to persuade any voters, losing the election (Proposition 1). Thus, the opposition runs as a conventional candidate rather than a populist. When the targeting technology is effective, the populist targets a narrow and receptive audience, minimizing the persuasion spending (Proposition 5). Through targeted media, populist persuasion can secure the election in a cost-efficient manner, inducing the opposition to run as a populist.

Figure 16: Voter suppression narrows the optimal audience

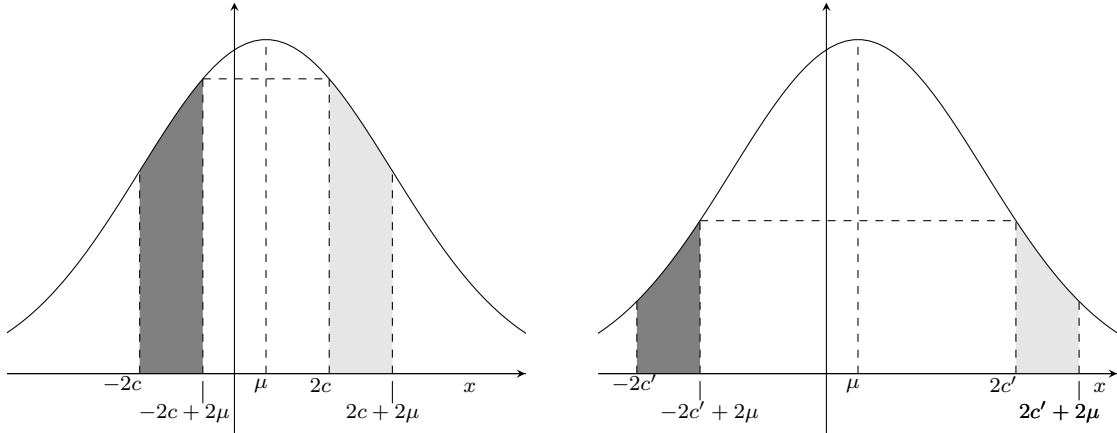
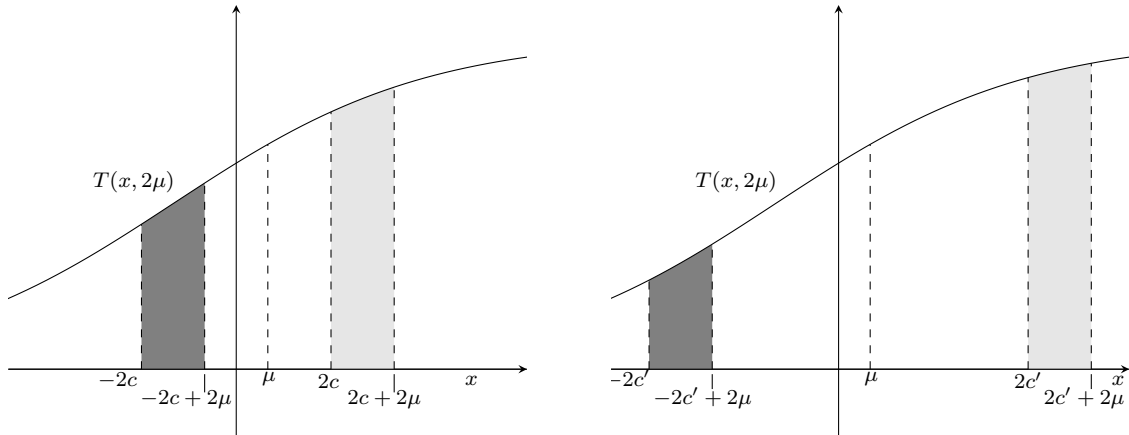


Figure 17: Voter suppression replaces an “angry” audience with an even “angrier” one



Voter suppression promotes populism The second result is more novel: voter suppression also produces strong populism. There are two effects that work in the same direction. First, as in Figure 16, a higher voting cost ($c' > c$) narrows the optimal audience for the populist message. As before, the darkly shaded area is the optimal audience and the lightly shaded area includes other voters in the maximal audience. When it is more costly to vote, the populist only needs to target an even smaller optimal audience to secure the election, further reducing the persuasion spending. This effect is driven by the assumption that the distribution of voter preference is “centrist,” with most voters centering around the average voter (Assumption 1). In a crucial twist, with stronger voter suppression, the society’s intrinsic centrism contributes to populism. We discuss more on the role of centrism in the next section (Section 4.2).

The second effect is illustrated in Figure 17. A higher voting cost replaces an already angry audience with another audience that is even angrier with the incumbent. It becomes

cheaper for the populist to convince the angrier audience, further reducing the persuasion spending. Notice that the second effect is driven by the key result of Proposition 5, the result that a populist only targets his message to marginal populist voters, never attempting to persuade marginal incumbent voters. Indeed, if the optimal audience includes marginal incumbent voters as in Figure 15, a higher voting cost dictates the populist to also persuade a group of voters who are even more enthusiastic supporters of the incumbent, a competing effect that increases the persuasion cost (see detailed analysis in Appendix B). The second effect only unambiguously benefits a populist when the optimal audience is highly homogeneous.

With a homogeneous audience, these two effects work together that exacerbate each other. Under stronger voter suppression, the optimal audience becomes smaller and the smaller audience is more disillusioned with the incumbent. Therefore, we have uncovered the mechanism behind our key claim, the claim that digital populism is intrinsically hostile towards even a minimalist conception of democracy. This mechanism would also drive a digital populist in power to invest in endogenous voter suppression, as demonstrated by Appendix C.

The complementarity between targeted media and voter suppression The third result codifies another set of key predictions. The inequality $\Pi_{\pi c} > 0$ shows that more effective targeting technology increases the marginal effect of voter suppression in producing populism. Voter suppression reduces the persuasion cost only when the populist can precisely target his message. Thus, when targeted media is highly effective, populists are especially enthusiastic about voter suppression.

For the other side of the inequality $\Pi_{\pi c} > 0$, stronger voter suppression increases the marginal effect of the targeted media in producing populism. With stronger voter suppression, more effective targeting helps the populist save more on persuasion spending than the case with weaker voter suppression. Thus, in democracies with strong voter suppression, populists are also especially enthusiastic about targeted media. The complementarity explains the notable resistance of a few democracies against digital populism (Kaltwasser et al. (2017); Moffitt (2017); Brett (2019)). In these democracies, a low cost of voting induces a large and moderate audience that a populist must persuade. Even armed with targeted media, it is costly to persuade such a large and moderate audience. Note that because the complementarity relies on $\Pi_c > 0$ (voter suppression promotes populism), the complementarity is also driven by the high homogeneity of the optimal audience.

4.2 Voter polarization and populism

All our results are obtained under a symmetric, unimodal distribution of voter preference. Such a distribution is highly centrist, with “most” voters clustering around the mean (also the mode) and few voters with extreme preference. Thus, the model has already predicted that even a centrist society is not immune to populist forces. The prediction engages with the large literature on voter polarization (Fiorina and Abrams (2008); Barber et al. (2015); McCarty (2019)), which finds that even though American political elites are becoming highly polarized, American citizens have long remained centrist and moderate. Therefore, our model offers an explanation for the meteoric rise of digital populism in a largely centrist society.

In this section, we further analyze how a *change* in voter polarization influences populism. We obtain a result that is even more counter-intuitive: a more centrist distribution of voters may actually encourage populism. We characterize the condition for this paradoxical result, showing that the key driver is the voting cost.

To describe how centrist a society is in a general way, we look at a space of symmetric, unimodal probability density functions with full support on the real line ($x \in R$). The space is indexed by α :

$$\{f(x; \alpha)\}_{\alpha \in \mathbb{R}^+}.$$

The density functions are ordered: for any $\alpha_1 > \alpha_2 > 0$,

$$\text{for all } x \in \mathbb{R} \setminus \{\mu\}, |f'(x; \alpha_1)| > |f'(x; \alpha_2)|. \quad (36)$$

Such an ordering implies that $\{f(x; \alpha)\}$ is ranked by second-order stochastic dominance. A larger α means that voters are more clustered around the mean μ . If we restrict the space to normal distributions $\{f(x; \mu, \sigma^2)\}$ with mean μ and variance σ^2 , we can assign α to be $1/\sigma^2$, the inverse of the variance. But for more general distributions, the order α may capture extra information than variance or second-order stochastic dominance.

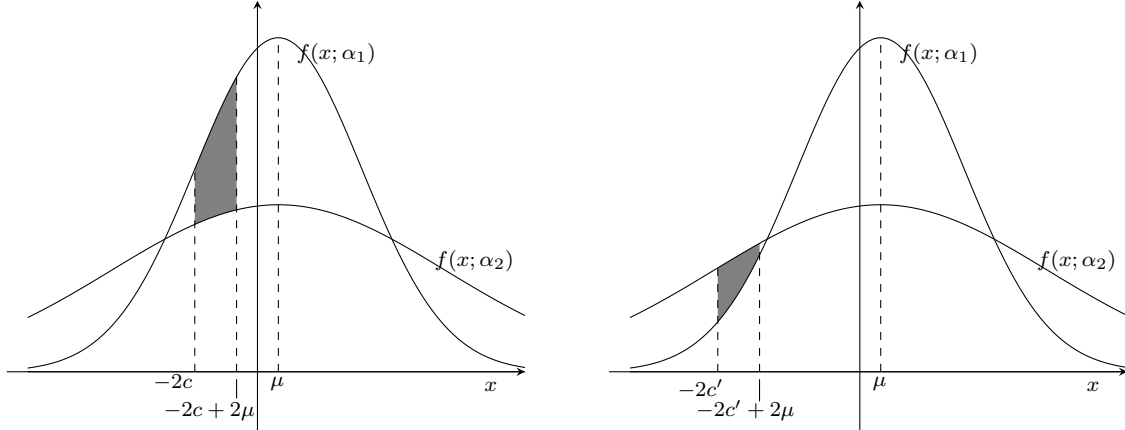
For the domain $x \in (-\infty, \mu]$, the probability density functions $\{f(x; \alpha)\}$ satisfies a version of the single crossing property. And because these functions are symmetric probability density functions, any two of them must cross exactly once on $x \in (-\infty, \mu]$. This intuitive result is proved in the following proposition, allowing us to sharply characterize the diametrically different effects of a centrist society on populism under different voting costs.

Proposition 7. *For any $\alpha_1 > \alpha_2$, there exists a “cutoff” $\mu^* \in (-\infty, \frac{\mu}{2}]$ so that the followings are true.*

1. *If $c < -\mu^*$, $\Pi(\alpha_1) < \Pi(\alpha_2)$. When the voting cost is sufficiently low, a more centrist society further deters digital populism.*

2. If $c > \mu - \mu^*$, $\Pi(\alpha_1) > \Pi(\alpha_2)$. When the voting cost is sufficiently high, a more centrist society further encourages digital populism.

Figure 18: Under low/high voting cost, a more centrist distribution of voters deters/encourages populist forces.



The results are illustrated in Figure 18. The shaded area is the *change* of optimal targeted audience when the probability density function varies from $f(x; \alpha_2)$ to $f(x; \alpha_1)$. We focus on the domain $x \in (-\infty, \mu]$ where all interesting actions happen. As voters become more clustered around the “average voter,” there must be a lower probability density of voters somewhere else. Under the ordering 7, a more centrist distribution must increase the probability density above a threshold, while simultaneously shrinking the probability density below a threshold. In other words, as the society becomes more centrist, the local effect on probability density is exactly the opposite for preferences below and above the threshold. When the populist can target his message, what matters is the local probability density of his ideal audience. The populist pays no attention to probability density elsewhere. Persuasion is more expensive only when the local probability density of his ideal audience increase.

How local probability density behaves depends on the voting cost. When the voting cost is sufficiently low, all voters in the targeted audience are more moderate than the threshold voter. A more centrist distribution raises the size of the targeted audience, making populist persuasion more expensive. But when the voting cost is sufficiently high, all voters in the targeted audience are more extreme than the threshold voter. A more centrist distribution shrinks the size of the targeted audience, making populist persuasion even cheaper. Therefore, with a high voting cost, a more centrist society produces populism, and this complementarity only exists when the populist can target his message.

Proposition 7 questions traditional wisdom on the social foundations of democracy. Leading scholars have long intuited that a strong consensus among voters induces a majoritarian democracy to produce desirable outcomes (Tocqueville (2000); Lipset (1959); Moore (1993); Svobik (2019)). The simple intuition is appealing, but our model shows that the intuition may not work at all in the digital age. Rather than neutralizing many threats to democracy, a reduction in voter polarization may even produce stronger populism, especially when reforms are stalled for electoral institutions and the media market.

For completeness, we also analyze how the average voter satisfaction (μ) affects populism. The results are in Appendix D. With some caveats, the model confirms that a less satisfied voter population produces strong populism. This insight is important to explain why a handful of developing countries pioneered populism even though they impose compulsory voting. Because the insight is more conventional, we leave it to the appendix.

5 Implications and Relevance of the Theory

5.1 Digital populism is hostile to even a minimalist democracy

There is a heated debate on the relationship between populism and democracy. Scholars have conjectured that even though populism introduces chaos into a liberal democracy, it may produce some important benefits. “Populism is seen as a threat but also as a potential corrective for a politics that has somehow become too distant from ‘the people’ ” (Müller (2016)). Populists might lay bare “the truth about a liberal democracy that has become forgetful about its founding ideal of popular sovereignty” (Müller (2016)). As summarized by Urbinati (2019), this perspective treats populism as “democracy at its best, because the will of the people is constructed through the people’s direct mobilization and consent. It is also politics at its best, because it employs only discursive devices and the art of persuasion” (see also Canovan (1999) and Laclau (2005)).

Our model shows that this intuitive reasoning can be problematic. Our model demonstrates that the very nature of digital populism dictates that it is fundamentally against enfranchisement. Digital populists in power would attempt to exclude as many voters as possible from casting their ballots, even including their own base. Such an exclusion impulse echoes the diagnosis of leading political theorists. For Müller (2016), “populists are always anti-pluralist.” They “treat their political opponents as ‘enemies of the people’ and seek to exclude them altogether.” Indeed, “*the* problem with populists is that they exclude” (italic added by us). Müller’s highlight on populist exclusion is echoed by Urbinati (2019). Citing Canovan (1981), Urbinati argues that for populism, the “claim of sovereignty is made by

only a part of the people, excluding another part that, *ex ante*, is defined as violation of the people.” To the best of our knowledge, this paper is the first to formalize and micro-found this exclusion impulse of populism from its intrinsic anti-establishment nature, especially so under effective targeted media.

Our model, however, goes far beyond the standard diagnosis for populism by political philosophers, uncovering an exclusion impulse that is qualitatively more extreme. Leading works in political philosophy argue that even though populism attempts to exclude establishment and ethnic minorities, populism does promote the maximal engagement of its core base, “a moral and homogeneous people.” Urbinati (2019) notes that “the populist interpretation of the people stresses the inclusion of the ‘ordinary’ many.” Even though “this inclusion occurs through a parallel process of exclusion,” it is nevertheless an inclusion. “Populism expresses at the same time the denunciation of exclusion and the construction of a strategy of inclusion by means of exclusion (of the establishment).” Indeed, most analysis still focuses on how populism erodes *liberal* institutions and practice (Rummens (2017)) rather than minimalist electoral institutions. But our model shows how populism can seriously undermine the minimalist majoritarian promise of democracy. Indeed, a populist will also attempt to exclude as many voters as possible from his own base, voters who are labelled as the “morally pure people.” Therefore populism not only threatens a liberal democracy (Müller (2016)); it threatens the essential electoral institutions, thus destructive to even a minimalist democracy (Przeworski (1991); Fearon (2011)). Our model provides a precise mechanism for Urbinati (2019)’s conjecture that populism “can stretch constitutional democracy toward its extreme borders and open the door to authoritarian solutions and even dictatorship.” Indeed, Urbinati (2019) believes that despite its authoritarian tendency, populism still must operate in a minimalist democratic framework. Our model shows that even such a cautious optimism about populism is unfortunately misplaced. Even the base of a populist politician would be suppressed so the populist only needs to communicate with a group of extreme supporters. Such suppression against his own base is ironically driven by the populist rhetoric that gratifies the base as “morally pure.”

To summarize, we propose a parsimonious way to capture the defining characteristic of populism as a moralized classification of the citizenry into the “morally pure people” and the “morally corrupt elite.” The existing qualitative literature analyzes this definition, uncovering the fundamental intuition that populism attempts to exclude a substantial group of voters who are labelled as the “morally corrupt elite.” But by formalizing the definition, our model goes far beyond the qualitative literature. The model shows that the populist exclusion of voters is way more extreme than previously thought, especially under highly effective targeted media. Our model construct a novel mechanism through which the essential

characteristic of populism directly contributes to the decay of a (minimalist) democracy. Does our main insight have a bite to explain empirical puzzles? In Section 5.2 and Section 5.3, we show how this main insight resolves puzzles about populism.

5.2 Renewal of voter suppression in the United States

Since the 2020 American Presidential Election, one of the most worrisome development in American politics is the renewed efforts in voter suppression. Voter ID laws have been dramatically tightened, polling stations have been strategically distributed, and convenience voting has been restricted (Epperly et al. (2020)). Against this development, voting rights are returning to the central stage of political debates, as manifested by high-profile efforts to introduce new bills on voting rights in Congress.

Numerous researches have proven that many suppression efforts are targeted towards voters of color, especially African Americans (see summaries in Anderson (2018); also see (Bentele and O'brien (2013), Hajnal et al. (2017) and Fraga (2018))). Still, there remains an important puzzle. At least some of the voter suppression efforts would suppress voters in a uniform manner (Barber et al. (2015); Riccardi and Calvan (2021)), sometimes even putting conservative voters at a disadvantage (Allen (2021); Gardner (2021); Riccardi and Calvan (2021)). Uniform voter suppression is not well studied in the academic literature, which focuses on targeted suppression.

The large literature on targeted suppression is complemented by our model, a model that investigates an ideal-type suppression applying to each voter with exactly the same force. Such a uniform suppression can still produce strong populist forces through the exclusion of centrist voters. This detrimental effect of uniform voter suppression only emerges after the dominance of targeted media. The model offers an explanation for the renewed efforts in voter suppression by the Republican Party since 2020 American Presidential Election. If the Republican Party can target suppression against Democratic voters, the party would surely prefer targeted suppression over uniform suppression. But our model predicts that a populist party would also be enthusiastic in promoting voter suppression even if such suppression has to apply uniformly to all voters. Such uniform suppression would still massively benefit the populist party by blocking the more moderate voters among the conservative population. Consequently, only the most extreme voters would turn out, meaning that the populist party only needs to target its rhetoric to a small audience that is highly extreme in its ideological preference, henceforth especially amenable to populist rhetoric. In any case, studying uniform suppression is conceptually important because a precise targeting of voter suppression is technologically more challenging than a precise targeting of a message on

social media.

An especially illuminating case is the new restrictions on convenience voting by Republicans in Florida. Convenience voting in Florida has consistently contributed to electoral successes of Republicans. “Virtually every narrow Republican victor of the past generation — and there have been many, including two of [Florida’s] current top officeholders, Gov. Ron DeSantis and Sen. Rick Scott — owes their victory, at least in part, to mail voting” (Gardner (2021)). “[The] proposal [to restrict mail-in voting] is notable because for years Republicans dominated vote by mail in Florida, and it was embraced by the state’s many older voters” (Riccardi and Calvan (2021)). Therefore it is especially puzzling that the Republican Party has introduced voter suppression that depresses the turnout of its own base. Our model resolves this puzzle. These suppression efforts are introduced by a party that has been captured by its populist leader, who prefers the turnout of the most extreme voters among the Republican base.¹⁷ Our analysis uncovers a mechanism behind the widely shared belief that voter suppression efforts by Republicans are caused by the dominance of their populist leader: “Florida Republicans’ push to change voting laws shows how deep false claims about mass voter fraud spread by Trump have burrowed into Republicans’ belief system” (Riccardi and Calvan (2021)).

To summarize, we take a different approach from existing researches on voter suppression, researches that attempt to empirically identify whether voter suppression has been targeted to a demographic group. We show that even in a hypothetical world where a perfect enforcement excludes the possibility of targeted suppression, a purely uniform suppression would still be a powerful weapon for a digital populist. By excluding other forces from our model, we uncover one underlying mechanism behind the flooding of voter suppression efforts across a large number of diverse constituencies in United States, the mechanism that a digital populist prefers only the most extreme supporters.

5.3 Voting cost and populism in Australia

Our model predicts that the rise of targeted media is far less likely to induce populist movements in a democracy with low voting cost. Our model also highlights a specific mechanism. To win the election, a populist must persuade a group of marginal voters. But with a low voting cost, this group would be large and politically moderate, which would make populist persuasion a difficult task. Both the prediction and its specific mechanism are well illustrated by the experience of Australia. The quotes in this section are from Brett (2019),

¹⁷Scholars have long noted that intense party competition pressures the incumbent politician to suppress supporters of the opposition (Keyssar (2009); Epperly et al. (2020)). But the literature pays much less attention to potential temptation to implement uniform suppression.

unless otherwise noted.

An array of electoral institutions in Australia keeps her voting cost one of lowest among advanced democracies. Indeed, the defining Australian electoral institution is compulsory voting, which imposes a negative voting cost.¹⁸ The introduction of compulsory voting in 1924 produces an impact that was sizable and persistent: federal turnout rate rose from 60% in 1922 to 91% in 1925; since then, the federal turnout rate has always been above 90%.¹⁹ Compulsory voting is complemented by other institutions to keep voting cost at a minimal level, such as compulsory registration and Saturday voting.²⁰ These electoral institutions work together to produce one of the highest turnout rates among advanced democracies.

While targeted media produces menacing populist movements in other advanced democracies with lower turnouts, targeted media fail to facilitate successful populism in Australia. “When it comes to concrete political outcomes, what is striking about Australian populists like Hanson, Katter, Palmer, and Lambie is that although they tend to be quite prominent political actors – both because of their outsized media profiles and their tendency to hold crucial votes in a tightly contested parliamentary system – the fact remains that they have had little electoral success overall. Indeed, apart from the example of Bjelke-Peterson, populist actors and parties have not enjoyed anything approaching the sustained electoral success of their European counterparts” (Moffitt (2017)). “Although there are a number of populist parties on the electoral market – as well as populist actors within the major parties – there is no widespread populist movement, or even a particular party that looks to establish a more permanent populist opposition to the major parties in Australia” (Moffitt (2017)). The failure of Australian populism is especially notable since 21st century populism is manifested by its “intensity and pervasiveness[...]Populist movements have appeared in almost every democracy” (Urbinati (2019)). The failure of Australian populism is also reflected by

¹⁸Even with compulsory voting as a negative voting cost, there are other costs that could not be realistically eliminated. Indeed, there have never been a 100% turnout rate in Australia.

¹⁹Using a difference-in-difference approach, Fowler (2013) provides causal evidence on the effect of compulsory voting on turnout in Australia. Fowler (2013) exploits a natural experiment where different Australian states introduced compulsory voting at different years for state elections. Fowler (2013) estimates that compulsory voting increased voter turnout by 24%, a figure close to simple descriptive statics from compulsory voting in Australian federal elections.

²⁰To see the importance of Saturday voting, an “American study of reasons for not voting found that the most common was the inability to get away from work or study commitments.” Indeed, “New Zealand shows the benefits of voting on Saturday: even without compulsory voting, its turnouts are generally well above 75 per cent, though it does have the booster of compulsory registration.”

Here is a summary of many other institutions that reduces voting cost in Australia. “Australia was the first to establish a national non-partisan electoral machinery. It paid close attention to potential barriers to voting of distance, literacy and mobility. It made it compulsory to be on the electoral roll, legislated for Saturday polling days and introduced preferential voting. In his 1921 comparative study of modern democracies the British liberal political historian James Bryce wrote that this newest of all the democracies has travelled farthest and fastest along the road which leads to the unlimited rule of the multitude.”

the public’s persistently high trust with the bureaucracy, the defining establishment institution and the arch-enemy of any populist movement (pp. 44-46, Müller (2016); Brett (2019); Sasso and Morelli (2021)). The struggle of digital populism in Australia is consistent with the main prediction of our model.

The specific mechanism of our model is also well supported. Australian media and politicians have long recognized the danger of a polarized electorate under a low turnout regime. As early as 1876, *the Melbourne Herald* published the following editorial. “The present state of affairs, as illustrated in any election contest, shows the necessity for compulsion. In few instances do more than one-half of the electors go to the poll; and in some cases only a third. And these, as a rule, are the violent partizans of one side or the other. Those who might be expected to give an impartial vote for the best man simply remain away.” Meanwhile, John Mackey, the Victorian chief secretary from 1908 to 1909, “believed that it was voters chiefly concerned with the general public interest who stayed at home on voting day, leaving the polls to the zealots.” Indeed, it is widely argued that compulsory voting ensures “the votes of moderate and respectable citizens.” Here is a more recent quote from the Labor MP Daryl Melham: “[A] voluntary ballot is flawed. It encourages political opponents to spend time drumming up support amongst their own supporters rather than amongst the general community – particularly those who are yet to make up their minds. If the Prime Minister is serious about building a consensus across the Australian people either for or against the republic, then it is these middle of the road swinging voters who must be convinced to vote one way or the other. It is this very dynamic that is at work in our compulsory parliamentary elections. It is what determines who wins and loses government.”

Thus, the intuition that compulsory voting moderates politics is shared among generations of political practitioners. Judith Brett also argues that compulsory voting facilitates the resistance against populist forces, which is also consistent with our model:

“Australian commentators frequently suggest that we are heading down the same path, that Australia too is subject to polarising populist forces from both left and right, and is becoming impossible to govern. Most weekends you can find such sentiments in the opinion pages of News Corp papers. I don’t agree. Not only does religious belief have less political influence here than in the United States, but compulsory voting tempers the impact of the passionate and committed voters of the base with the votes of the moderate and indifferent. It lowers the emotional temperature of our politics and keeps open the sensible centre. The angry and aggrieved will always be drawn to politics, but compulsory voting ensures that they are not the main occupants of the public square. Australian elections are won and lost in the centre, and compulsory voting is the reason.”

Brett also compares the resistance against populism in Australia with the explosion of

populist forces in America, highlighting the danger that low turnout regime could produce strong polarization. “We have seen in recent elections in the United States that the need to motivate voter turnout encourages highly targeted and divisive campaigns appealing to fear or moral outrage on issues about which a minority feel strongly, such as abortion, same-sex marriage or asylum seekers.”

Brett’s analysis, however, begs a question. The turnout rate for United States presidential election has always been very low, mostly below 65% for the last one hundred years. Why did American populism only become such a menacing threat in the last decade? Our model shows that the above informal reasoning is incomplete. Populists could only exploit high voting cost after the rise of targeted media. With only access to national media, a high voting cost could not facilitate populist persuasion. Our model relies on the fundamental intuition that high voting cost could polarize the electorate, an intuition that has been iterated by many. But our model also goes beyond the intuition by precisely identifying the digital media environment under which high voting costs could produce populism.

Our model agrees with Brett (2019) that progressive electoral institutions has so far constrained populist forces by minimizing the cost of voting in Australia. But our model differs from the prediction of Brett (2019) that Australian democracy can rely on her electoral institutions to perpetually constrain populism. The protection of progressive electoral institutions cannot last indefinitely, especially as targeted media become more and more effective (Zuboff (2019)). Even with a minimal voting cost, more effective targeted media still promotes populism (Proposition 6). The optimal targeted audience under a minimal voting cost is still much smaller than the audience for national media. In contrast to Brett (2019), progressive electoral institutions cannot perpetually shield a democracy from populism. Fundamental reforms on the media market are necessary, especially for a democracy whose media market has long been dominated by right-wing media (Goldsmith (2013)).

6 Conclusion

Populism appeals to voters by its promise to renew popular sovereignty. Populist leaders claim that they will help ordinary voters to regain the control of their democracy, a democracy that presumably has been lost to corrupt elite. Populism’s claim to empower ordinary voters is questioned by our paper. Our model formulates a novel mechanism that turns populist claim on its head. Far from being the resurrector of popular sovereignty, populism in the digital era thrives through voter suppression. Concurrent with its exclusion of establishment elite, a populist regime would further deprive ordinary citizens their minimalist voting rights.

Through a model on the production and distribution of populist rhetoric, we also make a methodological contribution to the literature on populism, a literature that is noted for its paucity of formal models. We believe that the model is general and flexible enough to study other questions about populism. The adaptability and flexibility of the model is partially demonstrated through a few extensions in our paper. Among others, we extend the model by quantifying changes in voter polarization (Section 4.2). The simple extension predicts that voter suppression would dramatically change how populism responds to a rise in voter polarization, a sharp and novel insight that is difficult to articulate and clarify without our model.

The flexible model might be able to address many other questions on populism, questions that we leave for future research. We highlight two directions. Scholars have speculated that populism can be deterred by robust civic culture, but this speculation is under-theorized by formal models.²¹ Specific to the framework of our model, it might be fruitful to formalize how the turnout decision is affected by the culture and value of “civic duty” in voting. The “civic duty” is recognized as first order in the turnout decision (Feddersen (2004); Herrera et al. (2016)), but is not adequately modelled in the small formal literature on populism. By integrating “civic duty” into our framework through either the cultural transmission model (Bisin and Verdier (2001); Bisin and Verdier (2011); Della Lena and Panebianco (2021)) or the replicator dynamics (Besley and Persson (2019)), such a model would be able to analyze how civic culture interact with electoral institutions and media environment in affecting populism.

For another potential direction, one can adapt our model to investigate whether populism can be deterred by rational, non-partisan electoral administration, an institution that is hailed as a pillar of advanced democracy (Brett (2019)). One can integrate our model with Sasso and Morelli (2021), a groundbreaking paper on bureaucracy and populism. But instead of looking at a general bureaucracy, one should focus on the administration of elections, with its many unique features in contrast to a typical government bureau. For one thing, most government bureaus are delegated by the legislature to executive policy decisions (Gailmard and Patty (2012)). But a government bureau that manages elections cannot avoid its influence over election process. Therefore, it might be especially interesting to formalize and investigate how politicians delegate electoral administration to civil servants.

²¹Major exceptions include Besley and Persson (2019), Besley (2020), and Gorodnichenko and Roland (2021). But none of them focus on populism.

References

- Absher, Samuel, Kevin Grier, and Robin Grier. 2020. The economic consequences of durable left-populist regimes in Latin America. *Journal of Economic Behavior & Organization* **177**: 787–817.
- Acemoglu, Daron, Georgy Egorov, and Konstantin Sonin. 2013. A political theory of populism. *The Quarterly Journal of Economics* **128**: 771–805.
- Acemoglu, Daron, and James A Robinson. 2012. *Why Nations Fail: The Origins of Power, Prosperity and Poverty*. Crown Publishers New York.
- Acemoglu, Daron, and James A Robinson. 2013. Economics versus politics: Pitfalls of policy advice. *Journal of Economic Perspectives* **27**: 173–92.
- Acemoglu, Daron, and James A Robinson. 2019. *The Narrow Corridor: How Nations Struggle for Liberty*. Penguin UK.
- Allen, Greg. 2021. Even as Trump denounces vote by mail, GOP in Florida and elsewhere relies on it. *National Public Radio (NPR)* .
URL <https://www.npr.org/2020/04/11/831978099/even-as-trump-denounces-vote-by-mail-gop-in-florida-and-elsewhere-relies-on-it>
- Anderson, Carol. 2018. *One Person, No Vote: How Voter Suppression Is Destroying Our Democracy*. Bloomsbury publishing USA.
- Barber, Michael, Nolan McCarty, Jane Mansbridge, and Cathie Jo Martin. 2015. Causes and consequences of polarization. *Political Negotiation: A Handbook* **37**: 39–43.
- Bentele, Keith G, and Erin E O’Brien. 2013. Jim crow 2.0? why states consider and adopt restrictive voter access policies. *Perspectives on Politics* **11**: 1088–1116.
- Berman, Sheri. 2021. The causes of populism in the West. *Annual Review of Political Science* **24**: 71–88.
- Besley, Timothy. 2020. State capacity, reciprocity, and the social contract. *Econometrica* **88**: 1307–1335.
- Besley, Timothy, and Torsten Persson. 2019. Democratic values and institutions. *American Economic Review: Insights* **1**: 59–76.

- Bidner, Chris, Patrick Francois, and Francesco Trebbi. 2014. A theory of minimalist democracy .
URL <https://www.nber.org/papers/w20552>
- Bisin, Alberto, and Thierry Verdier. 2001. The economics of cultural transmission and the dynamics of preferences. *Journal of Economic theory* **97**: 298–319.
- Bisin, Alberto, and Thierry Verdier. 2011. The economics of cultural transmission and socialization. In *Handbook of social economics*, volume 1. Elsevier, 339–416.
- Brett, Judith. 2019. *From Secret Ballot to Democracy Sausage: How Australia Got Compulsory Voting*. Text Publishing.
- Cachanosky, Nicolas, Alexandre Padilla, and Alejandro Gómez. 2021. Immigration and institutional change: Did mass immigration cause Peronism in Argentina? *Journal of Economic Behavior & Organization* **184**: 1–15.
- Cachanosky, Nicolás, Alexander W Salter, and Ignacio Savanti. 2022. Can dollarization constrain a populist leader? The case of Rafael Correa in Ecuador. *Journal of Economic Behavior & Organization* **200**: 430–442.
- Canovan, Margaret. 1981. *Populism*. Houghton Mifflin Harcourt P.
- Canovan, Margaret. 1999. Trust the people! populism and the two faces of democracy. *Political studies* **47**: 2–16.
- Chafetz, Josh, and David E Pozen. 2018. How constitutional norms break down. *UCLA Law Review* **65**: 1430.
- Colantone, Italo, and Piero Stanig. 2019. The surge of economic nationalism in Western Europe. *Journal of Economic Perspectives* **33**: 128–51.
- Della Lena, Sebastiano, and Fabrizio Panebianco. 2021. Cultural transmission with incomplete information. *Journal of Economic Theory* **198**: 105373.
- Doris, John, Stephen Stich, Jonathan Phillips, and Lachlan Walmsley. 2020. Moral Psychology: Empirical Approaches. In Zalta, Edward N. (ed.) *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2020 edition.
- Dorn, David, Gordon Hanson, and Kaveh Majlesi. 2020. Importing political polarization? the electoral consequences of rising trade exposure. *American Economic Review* **110**: 3139–83.

- Epperly, Brad, Christopher Witko, Ryan Strickler, and Paul White. 2020. Rule by violence, rule by law: Lynching, Jim Crow, and the continuing evolution of voter suppression in the US. *Perspectives on Politics* **18**: 756–769.
- Fearon, James D. 2011. Self-enforcing democracy. *The Quarterly Journal of Economics* **126**: 1661–1708.
- Feddersen, Timothy J. 2004. Rational choice theory and the paradox of not voting. *Journal of Economic perspectives* **18**: 99–112.
- Fiorina, Morris P, and Samuel J Abrams. 2008. Political polarization in the American public. *Annual Review of Political Science* **11**: 563–588.
- Fowler, Anthony. 2013. Electoral and policy consequences of voter turnout: Evidence from compulsory voting in Australia. *Quarterly Journal of Political Science* **8**: 159–182.
- Fraga, Bernard L. 2018. *The turnout gap: Race, ethnicity, and political inequality in a diversifying America*. Cambridge University Press.
- Gailmard, Sean, and John W Patty. 2012. Formal models of bureaucracy. *Annual Review of Political Science* **15**: 353–377.
- Gardner, Amy. 2021. Florida Republicans rushed to curb mail voting after Trump’s attacks on the practice. now some fear it could lower GOP turnout. *The Washington Post* .
URL https://www.washingtonpost.com/politics/florida-republicans-mail-voting/2021/05/02/4c133920-a9bf-11eb-8c1a-56f0cb4ff3b5_story.html
- Gentzkow, Matthew. 2016. Polarization in 2016. *Toulouse Network for Information Technology Whitepaper* : 1–23.
- Gerbaudo, Paolo. 2017. *The Mask and the Flag: Populism, Citizenism, and Global Protest*. Oxford University Press.
- Goldfarb, Avi, and Catherine Tucker. 2019. Digital economics. *Journal of Economic Literature* **57**: 3–43.
- Goldsmith, Ben. 2013. Factcheck: does Murdoch own 70% of newspapers in Australia? *The Conversation* **8**.
- Gorodnichenko, Yuriy, and Gerard Roland. 2021. Culture, institutions and democratization. *Public choice* **187**: 165–195.

- Graham, Matthew H, and Milan W Svobik. 2020. Democracy in America? Partisanship, polarization, and the robustness of support for democracy in the United States. *American Political Science Review* **114**: 392–409.
- Grier, Kevin, and Norman Maynard. 2016. The economic consequences of Hugo Chavez: A synthetic control analysis. *Journal of Economic Behavior & Organization* **125**: 1–21.
- Guiso, Luigi, Helios Herrera, Massimo Morelli, Tommaso Sonno, et al. 2017. *Demand and supply of populism*. IGIER, Bocconi University.
URL <https://ftp.igier.unibocconi.it/wp/2017/610.pdf>
- Guriev, Sergei, and Elias Papaioannou. 2022. The political economy of populism. *Journal of Economic Literature* **60**: 753–832.
- Guriev, Sergei, and Daniel Treisman. 2020. A theory of informational autocracy. *Journal of Public Economics* **186**: 104158.
- Haidt, Jonathan. 2001. The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological review* **108**: 814.
- Hajnal, Zoltan, Nazita Lajevardi, and Lindsay Nielson. 2017. Voter identification laws and the suppression of minority votes. *The Journal of Politics* **79**: 363–379.
- Hellmann, Olli. 2017. Populism in East Asia. In Kaltwasser, Cristóbal Rovira, Paul A Taggart, Paulina Ochoa Espejo, and Pierre Ostiguy (eds.) *The Oxford Handbook of Populism*, chapter 8. Oxford: Oxford University Press.
- Herrera, Helios, Massimo Morelli, and Salvatore Nunnari. 2016. Turnout across democracies. *American Journal of Political Science* **60**: 607–624.
- Houle, Christian, and Paul D Kenny. 2018. The political and economic consequences of populist rule in Latin America. *Government and Opposition* **53**: 256–287.
- Huber, Robert A, and Saskia P Ruth. 2017. Mind the gap! Populism, participation and representation in Europe. *Swiss Political Science Review* **23**: 462–484.
- Kaltwasser, Cristóbal Rovira, Paul A Taggart, Paulina Ochoa Espejo, and Pierre Ostiguy (eds.) . 2017. *The Oxford Handbook of Populism*. Oxford University Press.
- Keyssar, Alexander. 2009. *The right to vote: The contested history of democracy in the United States*. Basic Books.

- Laclau, Ernesto. 2005. *On populist reason*. Verso.
- Lau, Richard R, and Ivy Brown Rovner. 2009. Negative campaigning. *Annual review of political science* **12**: 285–306.
- Leininger, Arndt, and Maurits J Meijers. 2021. Do populist parties increase voter turnout? evidence from over 40 years of electoral history in 31 European democracies. *Political Studies* **69**: 665–685.
- Levitsky, Steven, and Daniel Ziblatt. 2018. *How Democracies Die*. Broadway Books.
- Li, Weijia, Gérard Roland, and Yang Xie. 2022. Hobbesian wars and separation of powers. *working paper* .
- Liberini, Federica, Michela Redoano, Antonio Russo, Angel Cuevas, and Ruben Cuevas. 2020. Politics in the Facebook Era - evidence from the 2016 US Presidential Elections . URL <https://dx.doi.org/10.2139/ssrn.3584086>
- Lipset, Seymour Martin. 1959. Some social requisites of democracy: Economic development and political legitimacy. *American Political Science Review* **53**: 69–105.
- Margalit, Yotam. 2019. Economic insecurity and the causes of populism, reconsidered. *Journal of Economic Perspectives* **33**: 152–70.
- McCarty, Nolan. 2019. *Polarization: What Everyone Needs to Know*®. Oxford University Press.
- McCormick, John P. 2001. Machiavellian democracy: controlling elites with ferocious populism. *American Political Science Review* **95**: 297–313.
- Moffitt, Benjamin. 2017. Populism in Australia and New Zealand. In Kaltwasser, Cristóbal Rovira, Paul A Taggart, Paulina Ochoa Espejo, and Pierre Ostiguy (eds.) *The Oxford Handbook of Populism*, chapter 6. Oxford: Oxford University Press.
- Moore, Barrington. 1993. *Social Origins of Dictatorship and Democracy: Lord and Peasant in the Making of the Modern World*. Beacon Press.
- Mudde, Cas, and Cristóbal Rovira Kaltwasser. 2017. *Populism: A Very Short Introduction*. Oxford University Press.
- Mukand, Sharun W, and Dani Rodrik. 2020. The political economy of liberal democracy. *The Economic Journal* **130**: 765–792.

- Müller, Jan-Werner. 2016. *What Is Populism?* University of Pennsylvania Press.
- Noury, Abdul, and Gérard Roland. 2020. Identity politics and populism in Europe. *Annual Review of Political Science* **23**: 421–439.
- Peitz, Martin, and Markus Reisinger. 2015. The economics of Internet media. In *Handbook of Media Economics*, volume 1. Elsevier, 445–530.
- Persson, Torsten, and Guido Tabellini. 2002. *Political economics: explaining economic policy*. MIT press.
- Przeworski, Adam. 1991. *Democracy and the market: Political and economic reforms in Eastern Europe and Latin America*. Cambridge university press.
- Przeworski, Adam. 1999. A minimalist conception of democracy. *Democracy's value* : 23–55.
- Riccardi, Nicholas, and Bobby Calvan. 2021. Florida is a model for voting. The GOP wants change anyway. *Associated Press* .
 URL <https://apnews.com/article/donald-trump-ron-desantis-legislation-florida-elections-a9737fcc3eab5aedc572fbcc7c19b984>
- Rummens, Stefan. 2017. Populism as a threat to liberal democracy. In Kaltwasser, Cristóbal Rovira, Paul A Taggart, Paulina Ochoa Espejo, and Pierre Ostiguy (eds.) *The Oxford Handbook of Populism*, chapter 28. Oxford: Oxford University Press.
- Sasso, Greg, and Massimo Morelli. 2021. Bureaucrats under populism. *Journal of Public Economics* **202**: 104497.
- Schumpeter, Joseph A. 2010. *Capitalism, Socialism and Democracy*. Routledge.
- Skitka, Linda J, Brittany E Hanson, G Scott Morgan, and Daniel C Wisneski. 2021. The psychology of moral conviction. *Annual Review of Psychology* **72**: 347–366.
- Svolik, Milan W. 2019. Polarization versus democracy. *Journal of Democracy* **30**: 20–32.
- Taggart, Paul. 2017. Populism in Western Europe. In Kaltwasser, Cristóbal Rovira, Paul A Taggart, Paulina Ochoa Espejo, and Pierre Ostiguy (eds.) *The Oxford Handbook of Populism*, chapter 13. Oxford: Oxford University Press.
- Tocqueville, Alexis de. 2000. *Democracy in America*. The University of Chicago Press. Harvey C. Mansfield and Delba Winthrop (trans.).

- Urbinati, Nadia. 2017. Populism and the Principle of Majority. In Kaltwasser, Cristóbal Rovira, Paul A Taggart, Paulina Ochoa Espejo, and Pierre Ostiguy (eds.) *The Oxford Handbook of Populism*, chapter 29. Oxford: Oxford University Press.
- Urbinati, Nadia. 2019. Political theory of populism. *Annual Review of Political Science* **22**: 111–127.
- Weber, Max. 1946. The Social Psychology of World Religions. In Gerth, Hans Heinrich, and C Wright Mills (eds.) *From Max Weber: essays in sociology*, chapter 11. Oxford University Press.
- Weber, Max. 2004. *The vocation lectures*. Hackett Publishing.
- Zhuravskaya, Ekaterina, Maria Petrova, and Ruben Enikolopov. 2020. Political effects of the internet and social media. *Annual Review of Economics* **12**: 415–438.
- Zuboff, Shoshana. 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. Profile books.

A Proofs

Lemma 1. *Suppose that the populist spends enough to convince all voters of a message m . The message m only changes the voting behavior of voters with:*

$$x \in \bar{A}(m) \equiv [-2c, -2c + m] \cup [2c, 2c + m]. \quad (12)$$

Proof. Before the population is convinced of the message m , the voting behavior is:

$$\begin{cases} \text{vote for the populist} & \text{if } x \in [\underline{x}, -2c) \\ \text{abstain} & \text{if } x \in [-2c, 2c) \\ \text{vote for the incumbent} & \text{if } x \in [2c, \bar{x}] \end{cases} \quad (37)$$

After the population is convinced of the message m , the voting behavior is:

$$\begin{cases} \text{vote for the populist} & \text{if } x \in [\underline{x}, -2c + m] \\ \text{abstain} & \text{if } x \in (-2c + m, 2c + m] \\ \text{vote for the incumbent} & \text{if } x \in (2c + m, \bar{x}] \end{cases} \quad (38)$$

Suppose that $-2c + m < 2c$, or $m < 4c$. We have five groups of voters:

$$\left\{ \begin{array}{ll} \text{always vote for the populist} & \text{if } x \in [\underline{x}, -2c) \\ \text{abstain} \rightarrow \text{populist vote} & \text{if } x \in [-2c, -2c + m] \\ \text{always abstain} & \text{if } x \in (-2c + m, 2c) \\ \text{incumbent vote} \rightarrow \text{abstain} & \text{if } x \in [2c, 2c + m] \\ \text{always vote for the incumbent} & \text{if } x \in (2c + m, \bar{x}] \end{array} \right. \quad (39)$$

The group that changes its voting behavior is:

$$\bar{A}(m) = [-2c, -2c + m] \cup [2c, 2c + m]. \quad (40)$$

Suppose that $-2c + m \leq 2c$, or $m > 4c$. We also have five groups of voters:

$$\left\{ \begin{array}{ll} \text{always vote for the populist} & \text{if } x \in [\underline{x}, -2c) \\ \text{abstain} \rightarrow \text{populist vote} & \text{if } x \in [-2c, 2c] \\ \text{incumbent vote} \rightarrow \text{populist vote} & \text{if } x \in (2c, -2c + m) \\ \text{incumbent vote} \rightarrow \text{abstain} & \text{if } x \in [-2c + m, 2c + m] \\ \text{always vote for the incumbent} & \text{if } x \in (2c + m, \bar{x}] \end{array} \right. \quad (41)$$

The group that changes its voting behavior is:

$$\bar{A}(m) = [-2c, 2c + m] = [-2c, -2c + m] \cup [2c, 2c + m]. \quad (42)$$

□

Proposition 1. *Suppose that the targeting fails. The populist spends nothing on uniform persuasion:*

$$\text{for any } m \in \mathbb{R}^+, \hat{T}^*(m) = 0. \quad (13)$$

Proof. Denote

$$\underline{T}(m) \quad (43)$$

as the minimum spending that secures an election for the populist when she broadcast a message $m \in \mathbb{R}^+$ to all voters $[\underline{x}, \bar{x}]$.

The optimal uniform spending for $m \in [0, \mu)$ is zero Suppose that the message $m \in [0, \mu)$. Then even if the populist spends $T(\bar{x}, m)$ so that he convinced everyone of the

message m , the populist cannot win the election:

$$\begin{aligned} \text{the populist's vote share} &= F(-2c + m) < F(-2c + \mu) \\ &= 1 - F(\mu + 2c) < 1 - F(m + 2c) = \text{the incumbent's vote share.} \end{aligned} \quad (44)$$

$F(-2c + \mu) = 1 - F(\mu + 2c)$ because $f(\cdot)$ is symmetric around μ . Equation 44 show that $\underline{T}(m) = \infty$. Thus, for any $m \in [0, \mu)$, the populist sets $\hat{T}^*(m) = 0$ and loses the election:

$$\begin{aligned} \text{the populist's vote share} &= F(-2c) = 1 - F(2\mu + 2c) \\ &< F(2c) = \text{the incumbent's vote share.} \end{aligned} \quad (45)$$

The populist's payoff is 0.

The optimal uniform spending for $m \in [2\mu, \infty)$ is zero Suppose that the message $m \in [\mu, \infty)$. Then a populist could win if he spends $T(\bar{x}, m)$ and convince everyone of the message m . The populist could swing voters with

$$x \in \bar{A}(m) = [-2c, -2c + m] \cup [2c, 2c + m]. \quad (46)$$

Further suppose that $m \in [2\mu, \infty)$. To win the election with minimal spending, the populist spends

$$\underline{T}(m) = T(-2c + 2\mu, m) \quad (47)$$

With such a spending, the populist can swing voters with

$$x \in [-2c, -2c + 2\mu]. \quad (48)$$

The persuasion cost for an $m \in [2\mu, \infty)$ is:

$$\int_{\underline{x}}^{\bar{x}} T(-2c + 2\mu, m) f(x) dx = T(-2c + 2\mu, m). \quad (49)$$

Notice that by Assumption 4, such a populist obtains a payoff lower than losing the election:

$$R - T(-2c + 2\mu, m) < R - T(-2c + 2\mu, 2\mu) < 0. \quad (50)$$

Thus, for any $m \in [2\mu, \infty)$, the populist sets $\hat{T}^*(m) = 0$.

The optimal uniform spending for $m \in [\mu, 2\mu]$ is zero Suppose that the message $m \in [\mu, 2\mu)$. To win the election with minimal spending, the populist sets

$$\underline{T}(m) = T(2\mu - (m - 2c), m) \quad (51)$$

and swings voters with

$$x \in [-2c, -2c + m] \cup [2c, 2\mu - (m - 2c)]. \quad (52)$$

$T(2\mu - (m - 2c), m)$ decreases with m :

$$\frac{\partial}{\partial m} T(2\mu - (m - 2c), m) = -T_x + T_m < -T_x + \tilde{t} < 0, \quad (53)$$

with $T_m < \tilde{t}$ by Assumption 5. Thus, with a message $m \in [\mu, 2\mu)$, the populist obtains a payoff:

$$R - T(2\mu - (m - 2c), m) \quad (54)$$

$$< R - T(2\mu - (2\mu - 2c), 2\mu) \quad (55)$$

$$= R - T(2c, 2\mu) \quad (56)$$

$$< R - T(-2c + 2\mu, 2\mu) < 0. \quad (57)$$

Equation 54 to Equation 55 because $T(2\mu - (m - 2c), m)$ decreases with m . Equation 56 to Equation 57 because $2c > -2c + 2\mu \Leftrightarrow c > 2\mu$ (Assumption 2).

Thus, for any $m \in [\mu, 2\mu)$, the populist sets $\hat{T}^*(m) = 0$. \square

Proposition 2. *Suppose that $m \in [2\mu, \infty)$. There exists a unique $\bar{m} \in (2\mu, \infty]$, such that the optimal audience for the message m is*

$$A^*(m) = \begin{cases} \emptyset & \text{for } m \in [\bar{m}, \infty) \\ [-2c, -2c + 2\mu] & \text{for } m \in [2\mu, \bar{m}) \end{cases}. \quad (15)$$

The populist's payoff is:

$$V(m) = \begin{cases} 0 & \text{for } m \in [\bar{m}, \infty) \\ R - \int_{-2c}^{-2c+2\mu} T(x, m) f(x) dx & \text{for } m \in [2\mu, \bar{m}) \end{cases}. \quad (16)$$

Proof. For any $m \in [2\mu, \infty)$, we first derive the minimal persuasion cost that secures the election. This gives us the payoff the populist obtains if he wins the election through targeted

persuasion. We then compare the payoff from winning the election with deliberately losing the election by targeting no voters.

For any $m \in [2\mu, \infty)$, the minimal cost of targeted persuasion is:

$$T^*(m) \equiv \min_{A(m) \in \bar{A}(m)} \int_{A(m)} T(x, m) f(x) dx \quad (58)$$

such that

$$\int_{(A(m) \cap [-2c, -2c+m]) \cup [\underline{x}, -2c]} f(x) dx \geq \int_{[2c, \bar{x}] \setminus A(m)} f(x) dx. \quad (59)$$

The constraint 59 ensures that the populist will be elected. The left hand side is the vote share received by the populist. Voters with $x \in [\underline{x}, -2c]$ always support the populist. Voters with $x \in [-2c, -2c + m]$ also support the populist upon receiving the message m at an intensity $T(x, m)$. Thus, by sending the message to $A(m)$, the populist gains additional voters with $x \in [-2c, -2c + m] \cap A(m)$. The populist receives support from all voters with:

$$x \in (A(m) \cap [-2c, -2c + m]) \cup [\underline{x}, -2c]. \quad (60)$$

The right hand side of constraint 59 is the vote share received by the incumbent. Voters with $x \in [2c, \bar{x}]$ supports the incumbent had they not received the message m . Upon receiving the message, voters with $x \in A(m)$ withdraw their support of the incumbent. Thus, the incumbent receives support from all voters with:

$$x \in [2c, \bar{x}] \setminus A(m). \quad (61)$$

The optimality of $A^*(m) = [-2c, -2c + 2\mu]$ is intuitive, since these are voters who are the least satisfied with the incumbent in the maximal audience $\bar{A}(m)$, while securing exactly 50% of total votes for the populist. Notice that by symmetry,

$$F(-2c + 2\mu) = 1 - F(2c). \quad (62)$$

$F(-2c + 2\mu)$ is the vote share of the populist who chooses the audience $A^*(m)$, and $F(2c)$ is the vote share of the incumbent.

But the full proof is technical and tedious, because we need to show that

$$A^*(m) \text{ generates the lowest persuasion cost among all Lebesgue measurable subsets of } \bar{A}(m). \quad (63)$$

We will show the case where $[-2c, -2c + m] \cap [2c, 2c + m] = \emptyset$. For the other case where

$[-2c, -2c + m] \cap [2c, 2c + m] \neq \emptyset$, the proof is almost identical and omitted. We will also omit another almost identical step in the proof for Proposition 3.

To show the claim 63, We proceed in two steps.

Step 1. Any audience A' with $\int_{(A' \cap [-2c, -2c+m]) \cup [x, -2c]} f(x)dx > \int_{[2c, \bar{x}] \setminus A'} f(x)dx$ is strictly suboptimal.

Denote the Lebesgue measure of a subset of \mathbb{R} as

$$\nu(\cdot). \tag{64}$$

Notice that $\nu(A') > 0$. There are two cases to consider: $\nu(A' \cap [-2c, -2c + m]) > 0$ and $\nu(A' \cap [2c, 2c + m]) > 0$. We focus on the first case. The proof for the second case is, again, almost identical.

Denote that $D = A' \cap [-2c, -2c + m]$. We have just assumed that $\nu(D) > 0$. We first prove the following lemma.

Lemma Fix any $\epsilon > 0$. For any measurable set $E \subset [-2c, -2c + m]$ with $\nu(E) > 0$, there exists a subset $\Delta \subset E$ such that $\nu(\Delta) > 0$ and $\nu(\Delta) < \epsilon$.

We first partition the set $[-2c, -2c + m]$. Note that for a fixed integer $h \geq \frac{2}{m}$, we can partition the set $[-2c, -2c + m]$ as follows:

$$[-2c, -2c + m] = \cup_{i=1}^{mh} [-2c + \frac{i-1}{h}, -2c + \frac{i}{h}).$$

For each interval $[-2c + \frac{i-1}{h}, -2c + \frac{i}{h})$, the measure is $\frac{2}{h} > 0$. All the partitions are disjoint. Thus, we can apply the additivity of Lebesgue measure, and note that the intersection of two measurable sets is measurable:

$$\begin{aligned} \nu(E) &= \nu\left(E \cap [-2c, -2c + m]\right) \\ &= \nu\left(E \cap \left\{ \cup_{i=1}^{mh} [-2c + \frac{i-1}{h}, -2c + \frac{i}{h}) \right\}\right) \\ &= \nu\left(\cup_{i=1}^{mh} \left\{ E \cap [-2c + \frac{i-1}{h}, -2c + \frac{i}{h}) \right\}\right) \\ &= \sum_{i=1}^{mh} \nu\left(E \cap [-2c + \frac{i-1}{h}, -2c + \frac{i}{h})\right) > 0. \end{aligned}$$

Denote $\Delta_i = E \cap [-2c + \frac{i-1}{h}, -2c + \frac{i}{h})$. Therefore there exists an $i' \in \{1, \dots, mh\}$, such that $\nu(\Delta_{i'}) > 0$. Notice that $\nu(\Delta_{i'}) \leq \nu([-2c + \frac{i-1}{h}, -2c + \frac{i}{h})) = \frac{2}{h}$. Thus, Take $h^* =$

$\max\{\lceil \frac{4}{\epsilon} + 1 \rceil, \lceil \frac{2}{m} + 1 \rceil\}$, $\nu(\Delta_{i'}) \leq \frac{\epsilon}{2}$. Gather the facts that $\nu(\Delta_{i'}) > 0$, $\nu(\Delta_{i'}) < \epsilon$, and $\Delta_{i'} \subset E$, the lemma is proved.

Next, pick

$$0 < \epsilon' < \frac{\int_{(D \cup [\underline{x}, -2c])} f(x) dx - \int_{[2c, \bar{x}] \setminus A'} f(x) dx}{\sup_x f(x)}.$$

Apply the above lemma, we can find a subset of $D = A' \cap [-2c, -2c + m]$, denote as $\Delta_{\epsilon'}$, such that $\nu(\Delta_{\epsilon'}) > 0$ and $\nu(\Delta_{\epsilon'}) < \epsilon'$. Suppose the populist does not send the message to the set $\Delta_{\epsilon'} \subset D$. The votes that the populist receives is:

$$\begin{aligned} & \int_{(D \cup [\underline{x}, -2c]) - \Delta_{\epsilon'}} f(x) dx \\ &= \int_{(D \cup [\underline{x}, -2c])} f(x) dx - \int_{\Delta_{\epsilon'}} f(x) dx \\ &\geq \int_{(D \cup [\underline{x}, -2c])} f(x) dx - \nu(\Delta_{\epsilon'}) \sup_x f(x) \\ &> \int_{(D \cup [\underline{x}, -2c])} f(x) dx - \epsilon' \cdot \sup_x f(x) \end{aligned}$$

The populist still wins the election:

$$\begin{aligned} & \int_{(D \cup [\underline{x}, -2c])} f(x) dx - \epsilon' \cdot \sup_x f(x), \\ &> \int_{(D \cup [\underline{x}, -2c])} f(x) dx - \frac{\int_{(D \cup [\underline{x}, -2c])} f(x) dx - \int_{[2c, \bar{x}] \setminus A'} f(x) dx}{\sup_x f(x)} \cdot \sup_x f(x) \\ &= \int_{[2c, \bar{x}] \setminus A'} f(x) dx. \end{aligned}$$

And the populist generates a strictly positive saving on persuasion cost

$$\begin{aligned} & \int_{(D \cup [\underline{x}, -2c]) - \Delta_{\epsilon'}} T(x, m) f(x) dx \\ &= \int_{(D \cup [\underline{x}, -2c])} T(x, m) f(x) dx - \int_{\Delta_{\epsilon'}} T(x, m) f(x) dx \\ &\leq \int_{(D \cup [\underline{x}, -2c])} T(x, m) f(x) dx - \overbrace{\nu(\Delta_{\epsilon'}) \cdot T(\inf \Delta_{\epsilon'}, m) f(\inf \Delta_{\epsilon'})}^{>0} \end{aligned}$$

$$< \int_{(D \cup [\underline{x}, -2c])} T(x, m) f(x) dx.$$

Step 2 we show that the audience $A^* = [-2c, -2c + 2\mu]$ reaches the lowest persuasion cost among any measurable set A such that $\int_{(A \cap [-2c, -2c+m]) \cup [\underline{x}, -2c]} f(x) dx = \int_{[2c, \bar{x}] \setminus A} f(x) dx$.

By contradiction, suppose that there exists a measurable set $A' \subset \bar{A}$ such that

1. $\nu(A' - A^*) > 0$,
2. $\int_{A'} T(x, m) f(x) dx < \int_{A^*} T(x, m) f(x) dx$,
3. $\int_{(A' \cap [-2c, -2c+m]) \cup [\underline{x}, -2c]} f(x) dx = \int_{[2c, \bar{x}] - A'} f(x) dx$.

The last condition is obtained from Step 1. The last condition implies that:

$$\int_{A'} f(x) dx = \int_{A^*} f(x) dx. \quad (65)$$

Condition 65 implies that:

$$\int_{A^* - A'} f(x) dx = \int_{A' - A^*} f(x) dx \quad (66)$$

To prove 66, by definition,

$$\begin{aligned} \int_{A^*} f(x) dx &= \int_{(A^* \cap A') \cup (A^* - A')} f(x) dx \\ &= \int_{(A^* \cap A')} f(x) dx + \int_{(A^* - A')} f(x) dx. \end{aligned}$$

Similarly, we have:

$$\int_{A'} f(x) dx = \int_{(A^* \cap A')} f(x) dx + \int_{(A' - A^*)} f(x) dx.$$

Recall that $\int_{A'} f(x) dx = \int_{A^*} f(x) dx$ (Equation 65), we obtain Equation 66.

Also, notice that for any $x' \in (A' - A^*)$, $x' \in \bar{A} = [-2c, -2c + m] \cup [2c, 2c + m]$ and $x' \notin [-2c, -2c + 2\mu]$. Thus,

$$\text{for any } x' \in (A' - A^*), x' > -2c + 2\mu. \quad (67)$$

We now compute a lower bound on the cost to persuade the audience A' :

$$\int_{A'} T(x, m) f(x) dx = \int_{(A' \cap A^*) \cup (A' - A^*)} T(x, m) f(x) dx$$

$$\begin{aligned}
&= \int_{A' \cap A^*} T(x, m) f(x) dx + \int_{A' - A^*} T(x, m) f(x) dx. \tag{68} \\
&> \int_{A' \cap A^*} T(x, m) f(x) dx + T(-2c + 2\mu, m) \int_{A' - A^*} f(x) dx \\
&= \int_{A' \cap A^*} T(x, m) f(x) dx + T(-2c + 2\mu, m) \int_{A^* - A'} f(x) dx \\
&> \int_{A' \cap A^*} T(x, m) f(x) dx + \int_{A^* - A'} T(x, m) f(x) dx \\
&= \int_{A^*} T(x, m) f(x) dx.
\end{aligned}$$

A contradiction. This finishes the proof that an optimal audience is $A^*(m) = [-2c, -2c + 2\mu]$. The proof also establishes that any other optimal audience can only differ from $A^*(m) = [-2c, -2c + 2\mu]$ by a measure of zero.

With $A^*(m) = [-2c, -2c + 2\mu]$, we obtain the maximal payoff to a populist who sends out a message $m \geq 2\mu$ to the audience:

$$U(m) \equiv R - \int_{-2c}^{-2c+2\mu} T(x, m) f(x) dx, \tag{69}$$

$U(m)$ is monotonically decreasing in m :

$$U'(m) = - \int_{-2c}^{-2c+2\mu} T_m(x, m) f(x) dx < 0. \tag{70}$$

Also, by Assumption 3,

$$U(2\mu) = R - \int_{-2c}^{-2c+2\mu} T(x, 2\mu) f(x) dx > 0. \tag{71}$$

Thus, there exists a unique $\bar{m} \in (2\mu, \infty]$ such that:

$$\begin{cases} U(m) < 0 & \text{if } m \in (\bar{m}, \infty] \\ U(m) \geq 0 & \text{if } m \in (2\mu, \bar{m}] \end{cases}. \tag{72}$$

If $m \in (\bar{m}, \infty]$, the populist refuses to target his message m to any voters, obtaining a payoff of zero:

$$A^*(m) = \emptyset. \tag{73}$$

If $m \in [2\mu, \bar{m}]$, the populist targets the message m to:

$$A^*(m) = [-2c, -2c + 2\mu]. \quad (74)$$

□

Proposition 3. *Suppose that $m \in [\mu, 2\mu]$. There exists a unique $\underline{m} \in [\mu, 2\mu)$, such that the optimal audience is*

$$A^*(m) = \begin{cases} [-2c, -2c + m] \cup [2c, 2\mu - (m - 2c)] & \text{for } m \in [\underline{m}, 2\mu] \\ \emptyset & \text{for } m \in [\mu, \underline{m}) \end{cases}. \quad (19)$$

The populist's payoff is:

$$V(m) = \begin{cases} R - \int_{-2c}^{-2c+m} T(x, m) f(x) dx - \int_{2c}^{2\mu - (m - 2c)} T(x, m) f(x) dx & \text{for } m \in [\underline{m}, 2\mu] \\ 0 & \text{for } m \in [\mu, \underline{m}) \end{cases}. \quad (20)$$

Proof. The populist solves a problem that is identical to the one in Proposition 2, with the only difference that $m \in [\mu, 2\mu]$.

The minimal persuasion cost that secures the election is:

$$T^*(m) \equiv \min_{A(m) \in \bar{A}(m)} \int_{A(m)} T(x, m) f(x) dx \quad (75)$$

such that

$$\int_{(A(m) \cap [-2c, -2c+m]) \cup [\underline{x}, -2c]} f(x) dx \geq \int_{[2c, \bar{x}] \setminus A(m)} f(x) dx. \quad (76)$$

We now characterize the optimal audience for the problem 75-76. If $m = 2\mu$, Proposition 2 shows that $A^*(m) = [-2c, -2c + 2\mu]$.

If $m < 2\mu$, then just targeting $[-2c, -2c + m]$ is not enough:

$$\text{the populist's vote share} = F(-2c + m) \quad (77)$$

$$< F(-2c + 2\mu) = 1 - F(2c) = \text{the incumbent's vote share}. \quad (78)$$

The populist also needs to target some voters with $x \in [2c, 2c + m]$.²² Denote the targeted

²²Note that when $m \in [\mu, 2\mu]$, the two intervals of the maximal audience $\bar{A}(m) = [-2c, -2c+m] \cup [2c, 2c+m]$ do not overlap: $m \leq 2\mu$ and $\mu < 2c \Rightarrow m < 4c \Rightarrow -2c + m < 2c$.

voter with the maximal satisfaction as \hat{x} . \hat{x} should ensure that:

$$\text{the incumbent's vote share} = \int_{\hat{x}}^{\bar{x}} f(x)dx \quad (79)$$

$$\leq \int_{\underline{x}}^{m-2c} f(x)dx = \text{the populist's vote share}, \quad (80)$$

$$1 - F(\hat{x}) \leq F(m - 2c), \quad (81)$$

$$F(2\mu - \hat{x}) \leq F(m - 2c), \quad (82)$$

$$\hat{x} \geq 2\mu - (m - 2c). \quad (83)$$

Thus, the optimal \hat{x} is:

$$\hat{x}^* = 2\mu - (m - 2c). \quad (84)$$

Notice that $\hat{x}^* = 2\mu - (m - 2c) \in [2c, 2c + m] \Leftrightarrow m \in [\mu, 2\mu]$. So \hat{x}^* is feasible.

Thus, for $m \in [\mu, 2\mu]$, the optimal audience for the populist is:

$$A^*(m) = [-2c, -2c + m] \cup [2c, 2\mu - (m - 2c)]. \quad (85)$$

The minimal persuasion cost is:

$$T^*(m) = \int_{-2c}^{-2c+m} T(x, m)f(x)dx + \int_{2c}^{2\mu-(m-2c)} T(x, m)f(x)dx. \quad (86)$$

The minimal persuasion cost decreases with m :

$$\begin{aligned} T_m^*(m) &= f(-2c + m)T(-2c + m, m) + \int_{-2c}^{-2c+m} T_m(x, m)f(x)dx \\ &\quad - f(2\mu - (m - 2c))T(2\mu - (m - 2c), m) + \int_{2c}^{2\mu-(m-2c)} T_m(x, m)f(x)dx \end{aligned} \quad (87)$$

$$< f(-2c + m)T(-2c + m, m) - f(2\mu - (m - 2c))T(2\mu - (m - 2c), m) + \tilde{t} \quad (88)$$

$$= f(-2c + m)T(-2c + m, m) - f(-2c + m)T(2\mu - (m - 2c), m) + \tilde{t} \quad (89)$$

$$= f(-2c + m) \overbrace{[T(-2c + m, m) - T(2\mu - (m - 2c), m)]}^{< -2T(\mu - m + 2c) < 0} + \tilde{t} < 0 \quad (90)$$

for \tilde{t} sufficiently small.

So the maximal payoff by sending out the populist message, $U(m) = R - T^*(m)$, increases

with m . Also, by Assumption 3,

$$U(2\mu) = R - \int_{-2c}^{-2c+2\mu} T(x, m)f(x)dx - \int_{2c}^{2\mu} T(x, m)f(x)dx \quad (91)$$

$$= R - \int_{-2c}^{-2c+2\mu} T(x, m)f(x)dx > 0. \quad (92)$$

Therefore, there exists a unique $\underline{m} \in [\mu, 2\mu)$ such that:

$$\begin{cases} R - T^*(m) \geq 0 & \text{if } m \in [\underline{m}, 2\mu] \\ R - T^*(m) < 0 & \text{if } m \in [\mu, \underline{m}) \end{cases}. \quad (93)$$

When $m \in [\underline{m}, 2\mu]$, the populist targets the message m to $A^*(m) = [-2c, -2c + m] \cup [2c, 2\mu - (m - 2c)]$. When $m \in [\mu, \underline{m})$, the populist does not send the message to any voter, $A^*(m) = \emptyset$. \square

Proposition 4. *For any $m \in [0, \mu)$, the optimal audience is*

$$A^*(m) = \emptyset. \quad (23)$$

The populist's payoff is:

$$V(m) = 0. \quad (24)$$

Proof. The populist cannot win even if he convinces all voters in the maximal audience $\bar{A}(m) = [-2c, -2c + m] \cup [2c, 2c + m]$. The populist's vote share is

$$F(-2c + m) < F(-2c + \mu), \quad (94)$$

while the incumbent's vote share is:

$$1 - F(2c + m) > 1 - F(2c + \mu) = F(-2c + \mu). \quad (95)$$

Thus,

$$F(-2c + m) < 1 - F(2c + m). \quad (96)$$

Therefore the populist should send the message to no voters, $A^*(m) = \emptyset$. \square

Proposition 5. *The optimal message is:*

$$m^* = 2\mu. \quad (27)$$

With probability π , the message $m^* = 2\mu$ convinces the optimal audience

$$A^*(m^*) = [-2c, -2c + 2\mu]. \quad (28)$$

With probability $1 - \pi$, the message $m^* = 2\mu$ convinces no voters.

Proof. As a preliminary step, notice that formulating a message $m \in [0, \underline{m}] \cup (\bar{m}, \infty)$ yields a payoff of 0, lower than formulating the message 2μ (Assumption 3):

$$0 < \pi[R - \int_{-2c}^{-2c+2\mu} T(x, 2\mu)f(x)dx]. \quad (97)$$

Therefore,

$$m^* = \arg \max_{m \in [\underline{m}, \bar{m}]} \bar{V}(m). \quad (98)$$

First, $\bar{V}'(m) < 0$ for $m \in [2\mu, \bar{m}]$:

$$\bar{V}'(m) = -\pi \left[\int_{-2c}^{-2c+2\mu} T_m(x, m)f(x)dx \right] < 0. \quad (99)$$

Second, $\bar{V}'(m) > 0$ for $m \in [\underline{m}, 2\mu]$:

$$\bar{V}'(m) = -\pi \left\{ T(-2c + m, m)f(-2c + m) + \int_{-2c}^{-2c+m} T_m(x, m)f(x)dx \right. \quad (100)$$

$$\left. - T(2\mu - (m - 2c), m)f(2\mu - (m - 2c)) + \int_{2c}^{2\mu - (m - 2c)} T_m(x, m)f(x)dx \right\} \quad (101)$$

$$< -\pi \left\{ T(-2c + m, m)f(-2c + m) - T(2\mu - (m - 2c), m)f(2\mu - (m - 2c)) + \tilde{t} \right\} \quad (102)$$

$$= -\pi \left\{ T(-2c + m, m)f(-2c + m) - T(2\mu - (m - 2c), m)f(-2c + m) + \tilde{t} \right\} \quad (103)$$

$$= -\pi \left\{ f(-2c + m) \underbrace{[T(-2c + m, m) - T(2\mu - (m - 2c), m)]}_{< -2T(\mu - m + 2c) < 0} + \tilde{t} \right\} > 0. \quad (104)$$

Thus, the opposition's payoff is maximized at

$$m^* = 2\mu. \quad (105)$$

Following Proposition 2, the optimal convinced audience is:

$$A^*(m^*) = [-2c, -2c + 2\mu]. \quad (106)$$

Following Proposition 1, when targeting fails, the opposition spends $\hat{T}(2\mu) = 0$ on the national media and loses the election. \square

Proposition 7. *For any $\alpha_1 > \alpha_2$, there exists a “cutoff” $\mu^* \in (-\infty, \frac{\mu}{2}]$ so that the followings are true.*

1. *If $c < -\mu^*$, $\Pi(\alpha_1) < \Pi(\alpha_2)$. When the voting cost is sufficiently low, a more centrist society further deters digital populism.*
2. *If $c > \mu - \mu^*$, $\Pi(\alpha_1) > \Pi(\alpha_2)$. When the voting cost is sufficiently high, a more centrist society further encourages digital populism.*

Proof. **Lemma 1** There exists $\underline{\mu} < \mu$, such that for all $x \in (\underline{\mu}, \mu]$, and $\alpha_1 > \alpha_2$, $f(x; \alpha_1) > f(x; \alpha_2)$.

By symmetry,

$$\int_{-\infty}^{\mu} f(x; \alpha_1) dx = \int_{-\infty}^{\mu} f(x; \alpha_2) dx = \frac{1}{2}.$$

Claim 1 There exists a set E of measure $\nu(E) > 0$, such that $f(x; \alpha_1) > f(x; \alpha_2)$.

By contradiction, suppose Claim 1 is not true. Further suppose $f(x; \alpha_1) = f(x; \alpha_2)$ for all $x \in (-\infty, \mu] \setminus \Delta$, with $\nu(\Delta) = 0$. Then $f'(x; \alpha_1) = f'(x; \alpha_2)$ for all $x \in (-\infty, \mu] \setminus \Delta$, a contradiction. Then for all $x \in (-\infty, \mu] \setminus \Delta$, with $0 \leq \nu(\Delta) < \infty$, and $\nu((-\infty, \mu] \setminus \Delta) > 0$, $f(x; \alpha_1) < f(x; \alpha_2)$. For $x \in \Delta$, $f(x; \alpha_1) = f(x; \alpha_2)$. Therefore $\int_{-\infty}^{\mu} f(x; \alpha_1) dx > \int_{-\infty}^{\mu} f(x; \alpha_2) dx$. A contradiction.

Claim 2 There exists a connected set E' of measure $\nu(E') > 0$, such that $f(x; \alpha_1) > f(x; \alpha_2)$.

Pick any two $x_1, x_2 \in E$, $x_1 < x_2$, we want to show that for any $x \in (x_1, x_2)$, $g(x) \equiv f(x; \alpha_1) - f(x; \alpha_2) > 0$. By contradiction, suppose that there exists $x_1, x_2 \in E$, such that for an $x' \in (x_1, x_2)$, $g(x') = f(x; \alpha_1) - f(x; \alpha_2) \leq 0$. By continuity, there exists an $\tilde{x} \in (x_1, x')$ such that $g'(\tilde{x}) = \frac{g(x') - g(x_1)}{x' - x_1} < 0$, or $g'(\tilde{x}) = f'(\tilde{x}; \alpha_1) - f'(\tilde{x}; \alpha_2) < 0$, a contradiction. We establish Claim 2 by setting $E' = E \cup \{x \in \mathbb{R} : x = x_1 + x_2, x_1 \in E, x_2 \in E\}$. Notice that because E' is connected with $\nu(E') > 0$, it is an interval. Also, $\inf E' < \mu$.

Claim 3 For all $x \in (\inf E', \mu]$, $f(x; \alpha_1) > f(x; \alpha_2)$.

Note that $g(\inf E') \geq 0$ and $g'(x) > 0$ for $x \in (-\infty, \mu]$. Claim 3 is immediately established.

Let $\underline{\mu} = \inf E$. Lemma 1 is proved.

Lemma 2 There exists $\hat{\mu} < \underline{\mu}$, such that for all $x \in (-\infty, \hat{\mu})$, and $\alpha_1 > \alpha_2$, $f(x; \alpha_1) < f(x; \alpha_2)$.

Claim 1 There exists a set D of measure $\nu(D) > 0$, such that $f(x; \alpha_1) < f(x; \alpha_2)$.

Because $\int_{\underline{\mu}}^{\mu} f(x; \alpha_1) dx > \int_{\underline{\mu}}^{\mu} f(x; \alpha_2) dx$ and $\int_{-\infty}^{\mu} f(x; \alpha_1) dx = \int_{-\infty}^{\mu} f(x; \alpha_2) dx$, there must be a measurable set D with $\nu(D) > 0$, such that $f(x; \alpha_1) < f(x; \alpha_2)$.

Claim 2 There exists a connected set D' of measure $\nu(D') > 0$, such that $f(x; \alpha_1) < f(x; \alpha_2)$.

Pick any two $x_1, x_2 \in D$, $x_1 < x_2$, we want to show that for any $x \in (x_1, x_2)$, $g(x) = f(x; \alpha_1) - f(x; \alpha_2) < 0$. By contradiction, suppose there exists an $x' \in (x_1, x_2)$, such that $g(x') \geq 0$. By continuity, there exists an $\tilde{x} \in (x', x_2)$, such that $g'(\tilde{x}) = \frac{g(x_2) - g(x')}{x_2 - x'} < 0$, or $g'(\tilde{x}) = f'(\tilde{x}; \alpha_1) - f'(\tilde{x}; \alpha_2) < 0$, a contradiction. Define $D' = D \cup \{x \in \mathbb{R} : x = x_1 + x_2, x_1 \in D, x_2 \in D\}$, we have proved Claim 2. Note that D' is an interval with $\nu(D') > 0$.

Claim 3 For all $x \in (-\infty, \sup D']$, $f(x; \alpha_1) > f(x; \alpha_2)$.

This is simply established by the fact that $g(\sup D') \leq 0$ and $g'(x) > 0$ for $x \in (-\infty, \mu]$. Let $\hat{\mu} = \sup D'$, the lemma is established.

Lemma 3 There exists a unique $\mu^* \in (\infty, \frac{\mu}{2})$, such that $g(x) < 0$ for $x \in (-\infty, 2\mu^*)$ and $g(x) > 0$ for $x \in (2\mu^*, \mu]$.

Notice that for the sets D' and E' we have constructed, $\sup D' \leq \inf E'$. If $\sup D' = \inf E'$, the claim is established by setting $2\mu^* = \sup D' = \inf E'$. If $\sup D' < \inf E'$, by continuity and monotonicity of $g(x)$, there exists a unique $2\mu^* \in (\sup D', \inf E')$, such that $g(x) < 0$ for $x \in (-\infty, 2\mu^*)$ and $g(x) > 0$ for $x \in (2\mu^*, \mu]$. This proves Lemma 3.

Now recall that

$$\Pi(\alpha) \equiv \frac{\pi[R - \int_{-2c}^{-2c+2\mu} T(x, 2\mu) f(x; \alpha) dx]}{R}.$$

If $-2c > 2\mu^*$ or $c < -\mu^*$, for all $x \in (-2c, -2c + 2\mu)$, $f(x; \alpha_1) > f(x; \alpha_2)$. Thus, $\Pi(\alpha_1) < \Pi(\alpha_2)$.

If $-2c + 2\mu < 2\mu^*$ or $c > \mu - \mu^*$, for all $x \in (-2c, -2c + 2\mu)$, $f(x; \alpha_1) < f(x; \alpha_2)$. Thus, $\Pi(\alpha_1) > \Pi(\alpha_2)$. \square

B A heterogeneous audience

When the optimal message is interior solution, the optimal audience includes two sets of voters:

$$[-2c, -2c + m^*] \cup [2c, 2\mu - (m^* - 2c)]. \quad (107)$$

The opposition formulate the populist message m^* if and only if:

$$\pi \left[R - \int_{-2c}^{-2c+m^*} T(x, m^*) f(x) dx - \int_{2c}^{2\mu-(m^*-2c)} T(x, m^*) f(x) dx \right] \geq \gamma R. \quad (108)$$

The probability that the opposition runs as a populist is:

$$\Pi(\pi, c) = \frac{\pi}{R} \left[R - \int_{-2c}^{-2c+m^*} T(x, m^*) f(x) dx - \int_{2c}^{2\mu-(m^*-2c)} T(x, m^*) f(x) dx \right]. \quad (109)$$

We can prove the following result:

Proposition 8. *Suppose that Assumption 5 does not hold and $m^* \in (\underline{m}, 2\mu)$.*

1. $\Pi_\pi(\pi, c) > 0$: *the opposition is more likely to run as a populist if the targeting technology improves.*
2. *The sign of $\Pi_c(\pi, c)$ is ambiguous: the effect of voting cost on the probability of opposition running as a populist is ambiguous.*

Proof.

$$\Pi_\pi(\pi, c) = \frac{[R - \int_{-2c}^{-2c+2\mu} T(x, 2\mu) f(x) dx]}{R} > 0. \quad (110)$$

$$\Pi_c(\pi, c) = -\frac{\pi}{R} \cdot \frac{\partial}{\partial c} \int_{-2c}^{-2c+2\mu} T(x, 2\mu) f(x) dx \quad (111)$$

$$= -\frac{\pi}{R} \cdot [T(-2c + 2\mu, 2\mu) f(-2c + 2\mu) \cdot (-2) - T(-2c, 2\mu) f(-2c) \cdot (-2)] \quad (112)$$

$$= \frac{2\pi}{R} \cdot [T(-2c + 2\mu, 2\mu) f(-2c + 2\mu) - T(-2c, 2\mu) f(-2c)] > 0 \quad (113)$$

because $f(-2c + 2\mu) > f(-2c)$ (Assumption 1) and $T(-2c + 2\mu, 2\mu) > T(-2c, 2\mu)$.

$$\Pi_{\pi c}(\pi, c) = \frac{2}{R} \cdot [T(-2c + 2\mu, 2\mu) f(-2c + 2\mu) - T(-2c, 2\mu) f(-2c)] > 0. \quad (114)$$

□

A better targeting technology still facilitates populism. The populist sends a softer message to balance between a more ideal audience and stronger resistance from voters. Yet the populist still wants to send the message to a narrow audience, a situation that is only possible when the targeting technology is effective.

The populist sends a relatively soft message to both marginal populist voters and marginal incumbent supporters. When voting cost increases, the populist persuades fewer marginal

populist voters who are also easier to convince, the same effect as in the main model in Proposition 6. But the populist also needs to persuade marginal incumbent supporters who are more difficult to convince. This additional persuasion produces the ambiguity result.

Proposition 8 further clarifies why voter suppression reduces populism in the main model with Assumption 5. Without the second part of Assumption 5, it is much more expensive to convince any voter of a more inflaming message. Thus, the populist needs to send a softer message to both marginal populist voters and marginal incumbent voters, leading to the ambiguity result in Proposition 8. With Assumption 5, every single voter’s attitude towards inflaming populism is fixed, but the heterogeneity is still enormous across different voters. As a consequence, the populist loses little by sending an extremely inflaming message. But the benefit of the inflaming message is huge: the populist is able to only persuade an audience as ideal as possible; given that different voters react to the populist message so differently, communication with the right audience is all that matters. As a consequence, the populist only target an already angry audience, and a higher voting cost produce an even angrier audience for the populist message. The focus on the angry audience also drives the complementary result in Proposition 6. Thus, Proposition 6 and Proposition 8 together identify the deep source of the key complementary result: the fixity of an individual voter’s attitude towards populism in contrast to the heterogeneity across voters.

C Voter suppression by a populist in power

This short section extends the logic of the main model to study the behavior of a populist in power. We formalize a simple and important insight: because stronger voter suppression *always* reduces persuasion spending (Proposition 6), a digital populist in power faces a strong temptation to suppress voters. Populism indeed will produce some of the most detrimental consequences on voting rights.

Specifically, we endogenize investment in voter suppression as a choice for a populist politician. The model now has two periods. The first period is the same as the main model, with the difference that the utility of the office, now denoted as \tilde{R} , is the populist’s payoff in the second period. The second period remains similar with only two differences. First, the populist politician can only run the re-election campaign as a populist; he cannot run as an establishment candidate. Second, after it is revealed whether media targeting is successful for the re-election campaign, the populist politician may increase the voting cost to $c + e$ by spending $Q(e)$, with standard assumptions that $Q'(e) > 0$ and $Q''(e) > 0$. Then the populist chooses the optimal message and the optimal audience. Also, we assume that $e \leq \bar{e}$, \bar{e} being the maximal increase in voting cost that does not wipe out a facade of democracy. The

populist runs for re-election for the exogenous utility R and then retires.

We also impose another assumption similar to Assumption 4:

Assumption 6. $T(-2(c + \bar{e}) + 2\mu, 2\mu) > R$.

The optimal message and the optimal audience is the same as the main model. The populist incumbent's payoff is:

$$\max_e \left\{ R + \pi \left[-Q(e) - \int_{-2(c+e)}^{-2(c+e)+2\mu} T(x, 2\mu) f(x) dx + R \right] \right\}.$$

This gives us a simple proposition:

Proposition 9. 1. $e^* > 0$: a populist makes a positive investment in voter suppression.
 2. As $Q'(e^*) \rightarrow 0$, $e^* = \bar{e}$: if it is costless to suppress voters, the populist excludes as many voters as possible, including his own base.

The proposition formalizes the notion that in the digital age (when targeting succeeds), populism has an intrinsic impulse to exclude as many voters as possible. Existing literature focuses on how populism threatens liberal institutions in advanced democracies, such as liberal media and checks and balances. It turns out that the problem is far worse: digital populism is a lethal threat to enfranchisement, the very foundation of a minimalist democracy. Indeed, extreme digital populism is indistinguishable from dictatorship. The proposition is directly driven by a main result of our paper, the result that stronger voter suppression *always* reduces persuasion spending (Proposition 6).

In the first period, the opposition solves the same game as the baseline model, with the utility of winning the office as:

$$\tilde{R} = \left\{ R + \pi \left[R - \int_{-2(c+e^*)}^{-2(c+e^*)+2\mu} T(x, 2\mu) f(x) dx - Q(e^*) \right] \right\},$$

Thus, the probability that the opposition runs as a populist is:

$$\tilde{\Pi} \equiv \frac{\pi}{R} \left[\tilde{R} - \int_{-2c}^{-2c+2\mu} T(x, \mu) f(x) dx \right].$$

The extended model affirms the key comparative statics in the main model, as summarized in the following proposition.

Proposition 10. 1. $\tilde{\Pi}_\pi > 0$: the opposition is more likely to run as a populist if the targeted media is more effective.

2. $\tilde{\Pi}_c > 0$: the opposition is more likely to run as a populist with stronger voter suppression.

3. $\tilde{\Pi}_{\pi c} > 0$: the opposition's probability of running as a populist features a complementarity between targeted media and voter suppression.

Proof.

$$\tilde{\Pi} = \frac{\pi}{R} \left[- \int_{-2c}^{-2c+2\mu} T(x, \mu) f(x) dx + R \right] + \frac{\pi^2}{R} \left[-Q(e^*) - \int_{-2(c+e^*)}^{-2(c+e^*)+2\mu} T(x, 2\mu) f(x) dx + R \right].$$

$$\tilde{\Pi}_c = \frac{\pi}{R} \left\{ \frac{\partial}{\partial c} \left[- \int_{-2c}^{-2c+2\mu} T(x, \mu) f(x) dx \right] + \overbrace{\frac{\pi^2}{R} \left[\frac{\partial}{\partial c} \left[- \int_{-2(c+e^*)}^{-2(c+e^*)+2\mu} T(x, 2\mu) f(x) dx \right] \right]}^{\text{Envelope Theorem}} \right\} > 0.$$

$$\tilde{\Pi}_{\pi} = \frac{1}{R} \left[- \int_{-2c}^{-2c+2\mu} T(x, \mu) f(x) dx + R \right] + \frac{2\pi}{R} \left[-Q(e^*) - \int_{-2(c+e^*)}^{-2(c+e^*)+2\mu} T(x, 2\mu) f(x) dx + R \right] > 0,$$

noticing that

$$\begin{aligned} & -Q(e^*) - \int_{-2(c+e^*)}^{-2(c+e^*)+2\mu} T(x, 2\mu) f(x) dx + R \\ & > Q(0) - \int_{-2c}^{-2c+2\mu} T(x, 2\mu) f(x) dx + R \\ & = - \int_{-2c}^{-2c+2\mu} T(x, 2\mu) f(x) dx + R > 0. \end{aligned}$$

$$\tilde{\Pi}_{\pi c} = \frac{1}{R} \left\{ \frac{\partial}{\partial c} \left[- \int_{-2c}^{-2c+2\mu} T(x, \mu) f(x) dx \right] + \overbrace{\frac{2\pi}{R} \left[\frac{\partial}{\partial c} \left[- \int_{-2(c+e^*)}^{-2(c+e^*)+2\mu} T(x, 2\mu) f(x) dx \right] \right]}^{\text{Envelope Theorem}} \right\} > 0.$$

□

D Voter satisfaction and populism

This section analyzes how average voter satisfaction (μ) affects populism. Voter dis-satisfaction, manifested as economic and sociocultural grievances, is usually cited as the major cause of populism. Although the explanation commands strong intuitions, the empirical findings are mixed (Guriev and Papaioannou (2022); Berman (2021)). Our analysis replicates an intuitive mechanism that it is easier to convince angrier voters of a populist message. But we also uncover an additional mechanism in a centrist society, a mechanism that countervails the intuitive mechanism. The next proposition characterizes the populism potential when voters become more satisfied.

Proposition 11. $\Pi_\mu(\pi, c, \mu)$ can be decomposed into two competing effects: in the first effect, μ reduces Π ; in the other effect, μ raises Π .

Proof. Recall that

$$\Pi(\pi, c, \mu) = \frac{\pi[R - \int_{-2c}^{-2c+2\mu} T(x, 2\mu)f(x; \mu)dx]}{R}. \quad (115)$$

Notice that we denote the probability density function as $f(x; \mu)$ to highlight that μ is the key parameter.

$$\Pi_\mu = -\frac{2\pi}{R}[T(-2c+2\mu, 2\mu)f(-2c+2\mu; \mu) + \int_{-2c}^{-2c+2\mu} T(x, 2\mu)f_\mu(x; \mu)dx + \int_{-2c}^{-2c+2\mu} T_m(x, 2\mu)f(x; \mu)dx]. \quad (116)$$

$-\frac{2\pi}{R}T(-2c+2\mu, 2\mu)f(-2c+2\mu; \mu) < 0$ is the effect of the additional marginal audience.

$-\frac{2\pi}{R}\int_{-2c}^{-2c+2\mu} T(x, 2\mu)f_\mu(x; \mu)dx > 0$ is the effect of the shrinking existing audience. $f_\mu(x; \mu) < 0$ because for all $x \in [-2c, -2c+2\mu]$, $x < \mu$ and $f(\cdot)$ is symmetric at μ and uni-modal.

$-\frac{2\pi}{R}\int_{-2c}^{-2c+2\mu} T_m(x, 2\mu)f(x; \mu)dx < 0$ is the effect of a more inflaming message, and it is second order by Assumption 5. \square

A more satisfied citizenry produces two effects.

1. The addition of a marginal audience: the populist needs to persuade a new marginal group of voters, making it more costly to be a populist.
2. The shrinking of the existing audience: the existing audience of the populist message shrinks, making it cheaper to be a populist.

The first effect is straightforward. As the average voter becomes more satisfied with the incumbent (μ increases), the populist needs to expand the optimal audience and persuade

all voters with

$$x \in [-2c, -2c + 2\mu]. \tag{117}$$

The persuasion of the additional audience at $x = -2c + 2\mu$ increases the cost of being a populist.

At the same time, we have a second effect. When the average voter becomes more satisfied with the incumbent, the entire distribution of voter satisfaction shifts rightward. Because the optimal audience is to the left of the average voter, the probability mass of the optimal audience shrinks everywhere. The effect of a shrinking audience is driven by the assumption that generically, the society is predominantly centrist (Assumption 1). Per the logic of this assumption, the higher satisfaction of the average voter reflects the higher satisfaction of the centrist society in general, as well as the shrinking mass of dissatisfied voters. The effect disappears if the voter satisfaction follows a uniform distribution, the most polarized distribution under Assumption 1. In this case, the prediction is un-ambiguous. The model only features the intuitive effect that higher voter satisfaction adds a marginal audience for the populist message.