OXFORD

Microtargeting, voters' unawareness, and democracy[†]

Freek van Gils[1], Wieland Müller[2,3,*], Jens Prüfer[3,4]

[1]The Netherlands Authority for Consumers and Markets, The Hague, The Netherlands
[2]University of Vienna & VCEE, Vienna, Austria
[3]Tilburg University & TILEC, Tilburg, The Netherlands
[4]University of East Anglia and Centre for Competition Policy, Norwich, Norfolk, United Kingdom

[*]Corresponding author: E-mail: wieland.mueller@univie.ac.at
[†]This article previously circulated under the title "Big Data and Democracy."

ABSTRACT

Recent technological developments have raised concerns about threats to democracy because of their potential to distort election outcomes: (a) data-driven voter research enabling political microtargeting and (b) growing news consumption via social media and news aggregators that obfuscate the origin of news items, leading to voters' unawareness about a news sender's identity. We provide a theoretical framework in which we can analyze the effects that microtargeting by political interest groups and unawareness have on election outcomes in comparison to "conventional" news reporting. We show which voter groups suffer from which technological development (a) or (b). While both microtargeting and unawareness have negative effects on voter welfare, we show that only unawareness can flip an election. Our model framework allows the theory-based discussion of policy proposals, such as to ban microtargeting or to require news platforms to signal the political orientation of a news item's originator (*JEL* C72, D72, D82, D83).

KEYWORDS: disinformation, interest groups, news platforms, microtargeting, voter awareness

1. INTRODUCTION

Democracy comes with many virtues. To hold governments accountable in representative democracies, voters depend on political information provision. The level and credibility of voters' information affects their trust in political leaders, institutions, election outcomes, and, hence, in the functioning of democracy itself (Van der Meer 2017). As a principal source of political information, the media is of critical importance to democratic societies.

However, recent technological developments affecting information available *about* voters, means to provide information *to* voters, and the nature of information acquisition *by* voters have raised concerns about threats to democracy because of their alleged potential to distort election outcomes (Kavanagh et al. 2019). These developments concern (a) data-driven

voter research and the possibility of political microtargeting and (b) news consumption of growing numbers of people using social media and news aggregators that obfuscate the origin of news, leading to voter unawareness about the news sender's identity.

Platforms collect vast amounts of data on users' preferences and characteristics by tracking them on and outside of the platform and by acquiring third-party data. Platforms can infer a range of attributes from these data, most notably users' political views (Kosinski et al. 2013). Some platforms also offer microtargeted advertising services, which can be used by political interest groups to tailor news to the preferences and characteristics of individual voters. Microtargeting allows interest groups (or advertisers) to differentiate their news reports, which may contain disinformation, to influence voters' beliefs in their favor in each subgroup of the electorate.[1]

Today more than half of digital news consumers use an algorithm-driven platform such as social media, search engines, and news aggregators as their main way to obtain news (Newman et al. 2020). Users may find it difficult to distinguish among the multitude of news senders, which, arguably, leads to voters' unawareness of a sender's identity. Indeed, platform users demonstrate a lack of recognition of outlet identity and are less able to attribute news items to the outlets that reported them if they saw the news on a platform than if they accessed it directly.[2]

A variety of actors spread untrustworthy content on platforms with the aim of promoting their own political goals (Tucker et al. 2018). Disinformation produced by highly partisan websites (Faris et al. 2017), false news websites (Allcott and Gentzkow 2017), and foreign governments (Maréchal, 2017) has been disseminated on news platforms, which are also reported to advance blurring of the line between fact and opinion (Kavanagh and Rich 2018). If the electorate is led astray by disinformation, it is not clear that election outcomes reflect voters' true preferences.[3] This is already problematic *per se*. Even more concerning, it calls into question the legitimacy of elections and, hence, may undermine citizens' trust in democracy.[4] Our article suggests explanations on how this is possible and studies that are most affected.

Despite the widespread attention to disinformation, microtargeting, and potential user unawareness on news platforms in public debate, the academic literature is arguably lagging behind. There is only one empirical research paper that shows a causal link from politically motivated social media use in an election (the 2016 US Presidential election) to voting behavior (Liberini et al. 2020). There is no political theory about the influence of political interest groups on voting behavior that captures the specificities of news platforms.[5]

---

[1]   Facebook's Custom Audience is a prominent example of a microtargeted advertising service. According to investigative journalism outlet ProPublica, Facebook offers a list of 29,000 user categories that ad buyers can use to determine their target audience (https://www.propublica.org/article/facebook-doesnt-tell-users-everything-it-really-knows-about-them).

[2]   Kalogeropoulos and Newman (2017) report that, when a news item was accessed directly on the original website, users' recall rate of the originator's identity in their study was 81%, as compared to 47% if the news item was accessed via social media and to 37% if accessed via a search engine. Kang et al. (2011) find that a news portal website user's assessment of the credibility of a news story tends to be primarily influenced by the identity of the portal and less by the original source of the news if the user reports to be not too "involved" with the news story. Even if a user notices the source of a political message, it might be hard to find out the ideological leaning of the sender. Despite implemented transparency initiatives, political campaign groups with undisclosed funding entities ran more political ads on Facebook than any of the registered parties in the two months before the 2019 United Kingdom general election (https://www.ft.com/content/f42f9aa2-16ba-11ea-8d73-6303645ac406).

[3]   We follow the definition of Tucker et al. (2018): "Disinformation […] is intended to be a broad category describing the types of information that one could encounter online that could possibly lead to misperceptions about the actual state of the world." For instance, by selectively reporting one-sided information (truthfully) an interest group produces disinformation but not fake news.

[4]   Sixty-four percent of US adults say fabricated news stories cause a great deal of confusion about the basic facts of current issues and events (https://www.pewresearch.org/fact-tank/2017/10/04/key-trends-in-social-and-digital-news-media/). Sixty-eight percent of US adults say false news undermines their confidence in the government (https://www.journalism.org/2019/06/05/many-americans-say-made-up-news-is-a-critical-problem-that-needs-to-be-fixed/).

[5]   We use the term *interest group* as reference to all ideologically motivated suppliers of political content, that is, it includes traditional newspapers or TV channels but also the websites and social media accounts of parties, political organizations, and individuals. Whereas many of these organizations also have other communication channels, where voters are aware of a

We address this gap in the literature and ask what role microtargeting technologies and voter unawareness about the political position of a news sending interest group play for the potential to manipulate elections. Can rational voters' ex-post beliefs be affected by the news they receive via news platforms from ideological interest groups, such that voters make "wrong" voting decisions in equilibrium? If so, what is the bigger problem, microtargeting or voters' unawareness about interest groups' political positions, and why?

1.1 Overview of our theory

We incorporate the two technological key features of today's news platforms into a game-theoretical model: platforms may (a) enable *microtargeted* matching of news to users based on users' preferences and characteristics and (b) *impede users' awareness* of the original interest group that reports the news. Our model comprises two kinds of active players, an interest group and voters. The interest group reports political news that is disseminated via a news platform and voters consume news and can elect political parties. A binary state of the world, which is drawn from a commonly known probability distribution, objectively favors either a left-wing or right-wing policy. The interest group knows the state of the world but voters do not. On a classical left-right political spectrum, nature determines the positions of the interest group and two political parties (who are committed to implementing commonly known policies if elected). Voters are uniformly distributed over the political spectrum and have a privately known cost of voting.

The timing of actions is that, first, the interest group sends a message to voters about whether the state of the world is either favoring a left- or a right-leaning policy. Voters receive the message, update their beliefs about the state of the world, and then cast their vote for one of the parties, or abstain. While voters maximize expressive utility from voting, the interest group minimizes the weighted mismatch costs between its own preferred policy position and the ones of the two parties, depending on the realized state of the world.

We consider four different games, which are determined by varying the two essential features discussed above. First, the interest group must either send the same message to all voters (*public*) or can let the message depend on individual voters' ideological position (*microtargeting*). Second, when updating beliefs, voters can either be *aware* or *unaware* of the political position of the interest group.

For each of the four games, we characterize the Perfect Bayesian Equilibrium with the highest voter welfare. We show that in all games some voters and interest group types always prefer the same party over the other one independent of the state of the world (which we call "radicals"), whereas others change their party preferences in line with the state of the world (called "moderates"). Because radical interest groups ignore their information about the state of the world and always try to send messages that support their preferred party, in games with awareness about the sender's position, all voters ignore messages from radical groups. In contrast, we show that the messages sent by moderate interest groups can be truthful in equilibrium: those groups have an incentive to inform moderate voters truthfully as their goals are aligned. If moderate groups are constrained by public news dissemination, they inevitably also inform radical voters about the truth. However, with microtargeting that *disciplining effect* disappears and radical voters do not receive valuable information about the state of the world anymore because all types of interest groups have an incentive to manipulate their beliefs. Hence, radical voters, who have very strong party preferences but still benefit from truthful news, now rationally ignore all news and suffer most from microtargeting.

sender's identity, we focus on communication via platforms, which disguises the original sender. *News platforms* include but are not limited to social media, news aggregators, and search engines.

In contrast, the switch from awareness to unawareness hurts in particular "moderate" voters. In equilibrium, these voters know that a moderate interest group would inform them correctly but a radical interest group would always send them the same uninformative message. Thus, with awareness, they can either completely rely on the message received or completely discard it. Without awareness, moderate voters have to guess, which means that they will sometimes discard a truthful message from a moderate interest group and sometimes believe an uninformative message sent by a radical interest group. Both changes hurt the payoffs of moderate voters.

A voter-welfare ranking among our games produces policy-relevant results: the *public game with voter awareness* is ranked highest, whereas the *microtargeting game with voter unawareness* is ranked lowest. The other two games occupy intermediate ranks, depending on parameter values. Additionally, we show that voter unawareness is a necessary condition for *election flipping* (to change the election winner). Microtargeting alone cannot distort election outcomes qualitatively. Studying competition among interest groups, we show that voter welfare increases with increasing competition.

These results allow the theory-based discussion of policy proposals. One proposal is that news platforms could be compelled to *implement technologies by which users can identify a message's original sender*. This should help users to also infer the sender's political position. In our model's language, this provision would help to establish *awareness* among voters and thereby decrease the risk of flipping election outcomes even in the presence of *microtargeting*.[6]

2. LITERATURE REVIEW

This is the first article analyzing the effects of interest groups' information provision via news platforms on voting behavior. It contributes to the literature on interest groups' influence on policy outcomes through information transmission to voters.[7] Yu (2005) models how two competing interest groups influence both an incumbent government and homogeneous voters, who are exposed to the same message. Voters' posterior beliefs are an exogenous function of a prior belief and the number of messages received from both interest groups. In contrast, we consider heterogenous voters, study the possibility of microtargeted communication and model voters as Bayesian updaters, taking into account all aspects of the game.

Like our article, Shapiro (2016) studies the effects of false claims made by interest groups on voting behavior. Whereas we study direct communication from an interest group to voters, in Shapiro (2016), an interest group can only reach voters through a journalist's news coverage.[8] In Shapiro (2016), disclosure of the information sender's political position helps voters because it takes away the journalist's reputational incentive to report ambiguous news when facts are not in line with his predisposition. In our framework, voters benefit from disclosure even without reputational concerns.

Our article also adds to theoretical research on supply-driven media bias and political outcomes. There, media organizations may manipulate news content to advance the ideological

---

[6] Facebook recently implemented a requirement in its Custom Audience service that "In the drop-down menu of each ad, the "Why am I seeing this?" section will show people the *source of the information (advertiser or partner)* […]" (our emphasize). See https://www.facebook.com/business/news/introducing-new-requirements-for-custom-audience-targeting. However, it remains unclear how much this feature is used and to what extent it diminishes users' unawareness.

[7] Most of this literature studies direct lobbying of politicians by interest groups, which we do not study. See Grossman and Helpman (2001) and Van Winden (2004) for broad discussions.

[8] Related, Sobbrio (2011) considers policy-motivated media outlets; Petrova (2012) analyzes the link between advertising profitability and media bias.

agenda of journalists (Baron 2006) or editors (Sobbrio 2014) or yield to pressure from governments (Besley and Prat 2006). A typical supply-driven media bias model contains a media outlet that commits to a (potentially biased) reporting strategy to maximize a payoff function that captures both a profit motive and a political or commercial motive.[9] In contrast, our model abstracts from profit motives and the reporting strategy is unobservable to voters.

Alonso and Padró I Miquel (2023) consider a model in which two special interest groups with diametrically opposed interests can spend resources to capture media sources. Once captured, media sources can disseminate any message (from a continuous message space) independent of the underlying binary state of the world. In equilibrium, the levels of capture and lying by interest groups lead to polarization regarding news as more extreme messages are sent more often, but rational news consumers (i.e., citizens) cannot be deceived completely as they become skeptical. However, in equilibrium, social learning is weakened as informative messages are jammed, so that overall informativeness is reduced. While we mainly analyze a model featuring just one interest group, in an extension, we show that allowing for competition among interest groups increases voter welfare in all of our four games. Note also that, in contrast to our model, there is no voting in Alonso and Padró I Miquel (2023).

We contribute to an emerging, and mostly empirical, literature on the political effects of social media, reviewed by Zhuravskaya et al. (2020) and Tucker et al. (2018).[10] Our theory complements Liberini et al. (2020), who study Facebook advertisement price variation for different audiences and ask to what extent political campaigns during the 2016 US Presidential elections used social media to microtarget voters. They find that the Republican campaign used extensive Facebook ads and microtargeting and conclude that microtargeted ad campaigns had significant effects on voting behavior.

In an early contribution to the empirical literature on political microtargeting, Hillygus and Shields (2014) use a large variety of data sources to argue that political candidates exploit data-mining technologies and enormous voter databases to identify and target voters by raising "wedge issues," on which voters share a candidate's opinion. Contrasting this view, Hersh (2015) claims that even highly advanced political campaigns often lack accurate information about voter preferences and shows that the political campaigns in the United States in the period 2008–2012 mostly relied on a limited set of public voter-data records, even when alternative data sources were available.

Theoretical work on microtargeted advertising in political campaigns includes Schipper and Woo (2019) and Hoffmann et al. (2020). Schipper and Woo (2019) develop a model in which two political candidates with fixed policy positions on multiple political issues communicate (some) information about their political stance to voters, who vote for the candidate that they perceive as closest to their preferred policy position. The authors show that, with sophisticated voters, microtargeting abilities of candidates are sufficient to get an election outcome that is equivalent to the outcome under complete information. A voter realizes that a political candidate only shares information with her when it makes him look more attractive to her, which gives rise to an unraveling of all relevant information. Without microtargeting abilities, a voter might think that information is withheld to her because it would make the candidate look bad to other voters (but not to her), which implies less information revelation. Within our framework, microtargeting abilities can only hurt and never help

---

[9] Gentzkow et al. (2015) survey this literature.
[10] Germano and Sobbrio (2020) theoretically study opinion formation through the usage of algorithm-driven platforms such as search engines and highlight the platforms' role in the diffusion of misinformation but do not address the implications for political outcomes.

information transmission to voters. This stark difference in findings results from the different ways we model communication. Schipper and Woo (2019) study truthful but potentially imprecise communication, whereas communication is cheap talk in our model. Hoffmann et al. (2020) apply their selective disclosure model to study microtargeted political campaigning and find that both campaigners and voters can benefit from selective disclosure based on voter data. Unlike in our model, information disclosure is necessarily truthful and the incentives of the campaigners are independent of the state of the world.

Levy and Razin's (2019) study (online) echo chambers, which result from the choice of news consumers to cluster with like-minded others in combination with a number of behavioral biases. In our model, voters do not choose their news source (the interest group). Instead, we assume that voters are randomly matched to an interest group to capture the gatekeeper role of the platform's algorithm.

Lastly, our model builds on cheap talk models, in which a sender observes a payoff-relevant state and sends a costless and non-verifiable message to an uninformed receiver, who then chooses an action that determines the payoffs of both sender and receiver together.[11] Li and Madarász (2008) exclusively consider sender types with state-dependent preferences, whereas we also consider sender types with state-independent preferences (radical interest groups) over voters' actions. Disclosure of interests may harm receivers in Li and Madarász (2008) but unambiguously benefits them in our model.

Farrell and Gibbons (1989) and Goltsman and Pavlov (2011) study cheap talk communication with multiple audiences, which resembles our microtargeting games. They show (in a model without uncertainty about the sender's type) that private communication is beneficial for receivers under some biased configurations and hurts them under other biased configurations.[12] Like them, we find that a restriction to send the same message to multiple audiences can discipline the sender to transmit information to audiences that otherwise would receive uninformative communications.

3. A FORMAL MODEL OF POLITICAL NEWS DISSEMINATION

We first describe the model and discuss its key assumptions thereafter. The game features two kinds of players, an interest group and voters, who act in a world in which a news platform disseminates political news from interest groups to voters and voters can elect political parties. There are two political parties, indexed by $j \in \{L, R\}$.

3.1 Building blocks of the model
3.1.1 Voters' preferences and actions

It is common knowledge that party L and party R are committed to implement policies x_L and x_R, respectively, if elected. These policies are elements of a left-right political spectrum (more on this below) and are exogenously given. Voters have preferences over policies and receive expressive utility from voting. Voters are characterized by their position y on the uniformly distributed left-right spectrum $[-b, b]$: voter $-b$ has the most left-wing and voter b has the most right-wing ideological position.[13] A voter's ideal policy x_v depends on her

---

[11] There is a long tradition of applying cheap talk models to explain political phenomena in economics and political science. See for instance Austen-Smith and Banks (2002).
[12] Galeotti et al. (2013) study public and private cheap talk communication in an environment with multiple players who are both senders and receivers of messages. Our model is closer to Farrell and Gibbons (1989) and Goltsman and Pavlov (2011).
[13] Our results are robust to other distributions of voter ideologies, as long as they have strictly positive density on the interval $[-b, b]$.

ideological position y and the state of the world $\theta \in \{-1, 1\}$, which, however, voters cannot observe:

$$x_v(y, \theta) = y + \theta. \tag{1}$$

Hence, all voters prefer a relatively more left-wing policy if $\theta = -1$ and a more right-wing policy if $\theta = 1$. Each voter chooses a voting action $a \in \{L, R, 0\}$, where $a = L$ if she votes for party L, $a = R$ if she votes for party R, and $a = 0$ if she abstains from voting. A voter's utility from voting, in the absence of voting costs, is specified as follows:

$$U(a, y, \theta) = \begin{cases} g - t\left(x_a - x_v(y, \theta)\right)^2 & \text{if } a \in \{L, R\} \\ 0 & \text{if } a = 0. \end{cases} \tag{2}$$

The parameter $g > 0$ represents the psychological gain from voting and is needed to endogenize abstention. The utility received from voting for party j is decreasing in the mismatch cost $t(x_j - x_v)^2$ that the voter incurs if her ideal policy differs from the party j's policy position. Abstaining yields 0 utility. Each voter y incurs a cost of voting, $c_y \in [0, \bar{c}]$, which is independent of y and θ and is an *i.i.d.* draw from a uniform distribution over $[0, \bar{c}]$. Net voter utility is equal to $U(a, y, \theta) - c_y$ if $a \in \{L, R\}$ and 0 otherwise. Voter welfare is defined as:

$$W(a, c, \theta) = \int_{-b}^{b} U(a, y, \theta) dy - \int_{-b}^{b} \mathbb{1}_{\{a_y \in \{L,R\}\}} c_y dy. \tag{3}$$

3.1.2 Voters' beliefs

Voters have common prior beliefs $p = Pr(\theta = -1)$ about the state of the world, where $0 < p < 1$. Before voting, voters receive via a news platform a single cheap-talk news item $m \in \{-1, 1\}$ concerning the state of the world. Denote by $\mu(m) = \mu(\theta = -1|m)$, the probability (*posterior belief*) that a voter assigns to the event $\theta = -1$ after observing news item m. Denote by $\mathbb{E}[U(a, y, \theta)|\mu(m)]$ a voter's expected utility from voting if she has an ideological position y and posterior belief $\mu(m)$.

3.1.3 Interest group

There is a single ideologically motivated interest group that accurately observes the state of the world θ and sends a news item m about θ.[14] The interest group has ideological position z, which is drawn from a uniform distribution over $[-h, h]$ and is unobserved by voters. Corresponding to voters' preferences, the ideal policy position of the interest group, x_n, is determined by its ideological position z and the state of the world θ:

$$x_n(z, \theta) = z + \theta. \tag{4}$$

The interest group reports cheap talk news item $m \in \{-1, 1\}$ and earns the payoff

$$\Pi(a, z, \theta) = -\nu_L(a)\left(x_L - x_n(z, \theta)\right)^2 - \left(1 - \nu_L(a)\right)\left(x_R - x_n(z, \theta)\right)^2, \tag{5}$$

---

[14] We study competition among interest groups below.

where

$$\nu_L(a) = \frac{\int_{-b}^{b} \mathbb{1}_{\{a_y=L\}} dy}{\int_{-b}^{b} \mathbb{1}_{\{a_y=L\}} dy + \int_{-b}^{b} \mathbb{1}_{\{a_y=R\}} dy} \tag{6}$$

denotes party L's vote share, and $(x_L - x_n(z, \theta))^2$ is the interest group's mismatch cost. Analogous for party R.

3.1.4 Four games

We study the interaction of two news dissemination technologies and two awareness states of the voters about the ideological position of the interest group z, resulting in four different games. In the two *public games*, the interest group is restricted to producing a single news item $m \in \{-1, 1\}$ for the entire electorate. In the two *microtargeting games*, the interest group reports a news item $m_y \in \{-1, 1\}$ for each voter y, unobserved by others. In all games, the interest group may misrepresent the true state of the world. Voters have no way to learn about θ apart from observing m.

All aspects of the game, including the distribution of interest group ideologies and the news dissemination technology but not the realizations of θ, z, and c_y, are common knowledge.

3.1.5 Timing

The timing of each game is as follows:

Stage 0: Nature determines θ according to p, draws $c_y \sim U[0, \bar{c}]$ for each voter y, and draws $z \sim U[-h, h]$ for the interest group. Each voter y privately learns c_y and in games with voter awareness also z. The interest group observes z and θ.

Stage 1: The **interest group** chooses $\mathbf{m} \in \{-1, 1\}$ in the **public games** and $\mathbf{m_y} \in \{-1, 1\}$ for each voter y in the **microtargeting games**.

Stage 2: **Each voter y** observes news item m if a public news dissemination technology is in place and privately observes m_y in games with microtargeting. Each voter **updates belief** $\boldsymbol{\mu(m)}$, and **chooses voting action a** $\in \{\mathbf{L, R}, 0\}$. All payoffs are realized.

3.1.6 Equilibrium concept

Our solution concept is Perfect Bayesian Nash Equilibrium (PBE). A PBE of the game consists of a reporting strategy m^* of an interest group and a voting strategy a^* and a belief μ^* of a voter, which maximize a player's expected payoff, given her beliefs about other players. The games we analyze have multiple PBEs. Therefore, we focus on the *Voter Welfare-Maximizing Perfect Bayesian Equilibrium (VWMPBE)*, that is, the equilibrium with the highest voter welfare. As we will show for all games, this equilibrium coincides with the equilibrium in which most voters receive news that is informative to them. A formal definition of the equilibrium concept and a set of technical assumptions that make sure the solutions are well-behaved are in the Supplementary Appendix.

3.2. Model discussion
3.2.1 Voters' utility

The probability that a single vote is decisive is low in large elections. Hence, rational voters are unlikely to turn out to vote if they are solely interested in the election outcome (Downs 1957). Turning out to vote is not paradoxical if voters derive direct *expressive utility*

from voting.[15] Following Chan and Suen (2009), the utility that a voter derives from voting for a political party depends on the party's policy platform, the voter's individual-specific taste ("ideology"), and an unobserved state of the world. A mathematical equivalent to expressive voting is modeling one representative voter, who is pivotal for election outcomes, by definition. Close to our formulation of voters' utility, in Binswanger and Prüfer (2012) the voter's optimal voting action also depends on a politician's policy platform and an unknown state of the world.

3.2.2 Interest group's payoffs and information

Whereas *individual* voters have a negligible impact on the election outcome, an interest group could be decisive for the election outcome by influencing the voting behavior of *multiple* voters. For this reason, we let the payoffs of an interest group depend on the realized election outcome. This modeling decision does not drive our results. Alternatively, we could have assumed that an interest group enjoys expressive utility from a vote for its favored party and expressive disutility from a vote for the opposing party. This would not have altered our findings. As we assume the same structure of the ideal policy position of voters (1) and interest groups (4), the latter can be thought of as being managed by members of the electorate. The weighted mismatch cost of interest groups (5) captures that parties' political influence, for example, the number of seats held in parliament or the amount of campaign contributions received, usually depends on their vote shares. We also assume that interest groups are only motivated by political interests and not by a preference for truth telling. We will nevertheless show that, depending on the game, it is a result, not an assumption, that moderate interest groups report truthfully in specific, clearly delineated cases. Shapiro (2016) and Kartal and Tremewan (2018) offer discussions and justification for the assumption that interest groups, having access to expert knowledge and resources, are perfectly informed about the true state of the world.

3.2.3 A single interest group

With one interest group, the models' mechanisms can be easily understood. Results for the case of competing messages received by a voter are shown after the baseline model.

3.2.4 Perfect rationality

We model perfectly rational voters, who understand the incentives of interest groups to misreport news. This is a strong assumption as many voters have cognitive limitations and imperfect foresight. However, the significance of our results is only strengthened if we can show that and how even rational voters can be manipulated in equilibrium and make voting decisions that are against their own interests. Then, voters with naive beliefs about political messages could, arguably, be manipulated even easier by demagogues. Complementarily, if only a share $\alpha \in (0, 1)$ of voters are rational (Bayesian updaters) and $1 - \alpha$ voters do not update their beliefs (and if α is not correlated with y), our results hold for the rational voters and, hence, qualitatively for the entire electorate.

3.2.5 Common beliefs

We assume common prior beliefs for tractability. Replacing the common prior belief p by an individual belief p_y for each voter y, would not affect our results qualitatively. Posterior beliefs $\mu(m)$ can differ across voters, which is a crucial feature in our microtargeting games.

---

[15] Expressive voters are frequently modeled (e.g., Schuessler (2000); Glaeser et al. 2005). There is ample empirical support for the expressive voting theory (e.g., Pons and Tricaud 2018). See Tyran and Wagner (2019) for a survey on expressive voting experiments in the laboratory.

3.2.6 Exogenous awareness

In practice, interest groups have various communication channels to send messages to voters. In this article, however, we take the overarching empirical relevance of news platforms for political communication as given and study the consequences of this news consumption pattern for voters' political beliefs, where awareness about a sender's identity is decreased exogenously (Kalogeropoulos and Newman 2017). The case of endogenous awareness is briefly discussed in footnote 21.

4. EQUILIBRIUM ANALYSIS

In all four games, the critical task of each voter y is to update her prior belief p about the state of the world θ to posterior belief $\mu(m)$ after receiving message m. Then casting a vote for the party that is perceived to be located closest to y's ideal policy $x_v(y, \theta)$ is straightforward. The difficulty of voter y is that m is potentially valuable because the interest group has perfect knowledge about θ but the interest group's payoff increases if it can make the voter vote for *its* preferred party. Hence, the voter should not trust the message blindly.

4.1 Public game with voter awareness

Here the voter knows the political position z and, hence, the objectives of the interest group. However, the voter does not know θ. The interest group, in turn, knows all aspects of the game apart from the realization of an individual voter's voting cost c_y. This ignorance does not affect its decision, though.

Denote the ideology of a voter with belief μ', who is *indifferent* between voting for parties L and R, by $\hat{y}_{\{\mu'\}}$. Similarly, denote the ideology of an indifferent interest group by $\hat{z}(\cdot)$. We show in the Supplementary Appendix that the following proposition holds.

Proposition 1. The following strategy profiles and beliefs constitute the voter welfare-maximizing Perfect Bayesian Equilibrium of the public game with voter awareness:

$$a^*(y, c_y, m, z) = \begin{cases} L & \text{if } y < \hat{y}_{\{\mu^*(m,z)\}} \text{ and } c_y < \mathbb{E}[U(a = L, y, \theta)|\mu^*(m,z)] \\ R & \text{if } y > \hat{y}_{\{\mu^*(m,z)\}} \text{ and } c_y < \mathbb{E}[U(a = R, y, \theta)|\mu^*(m,z)] \\ 0 & \text{otherwise,} \end{cases} \quad (7)$$

$$m^*(z, \theta) = \begin{cases} -1 & \text{if } z \leq \hat{z}(\theta) \\ 1 & \text{if } z > \hat{z}(\theta), \end{cases} \quad (8)$$

$$\mu^*(m = -1, z) = \begin{cases} p & \text{if } z \leq \hat{z}(\theta = 1) \text{ or } z > \hat{z}(\theta = -1) \\ 1 & \text{if } \hat{z}(\theta = 1) < z \leq \hat{z}(\theta = -1), \end{cases}$$

$$ \quad (9)$$

$$\mu^*(m = 1, z) = \begin{cases} p & \text{if } z \leq \hat{z}(\theta = 1) \text{ or } z > \hat{z}(\theta = -1) \\ 0 & \text{if } \hat{z}(\theta = 1) < z \leq \hat{z}(\theta = -1). \end{cases}$$

Intuitively, the decision rule (7) shows that a voter only votes for her preferred party if the expected utility from voting exceeds her voting cost. An interest group with rather left-wing (right-wing) ideology prefers voters to cast their vote for the left (right) party. However, "rather left-wing" depends on the state of the world, as depicted in Figure 1. If the interest group ideology z is "moderate," that is, if $\hat{z}(\theta = 1) < z \leq \hat{z}(\theta = -1)$, then it depends on
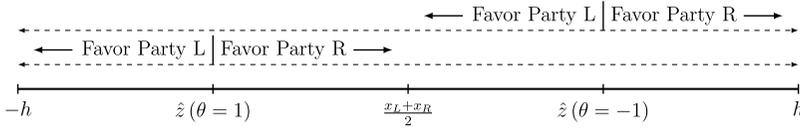
Figure 1. Location of the indifferent *interest group* conditional on the state of the world.
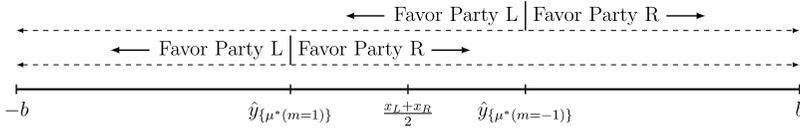

Figure 2. Location of the indifferent *voter* conditional on the public news item received.

the state of the world which party the interest group favors. If interest group ideology is "radical," that is, if $z \leq \hat{z}(\theta = 1)$ or $z > \hat{z}(\theta = -1)$, then the interest group's favorite party is state-independent.

Equation (8) states that the interest group reports $m = -1$ if it favors party L and $m = 1$ if it favors party R. Hence (endogenously emerging), moderate interest groups truthfully report about the state of the world, whereas radical interest groups always report $m = -1$ (if left-wing radical) or $m = 1$ (if right-wing radical), whether it is truthful, or not.

The rational voters in this game understand the interest group's strategy. Therefore, if the interest group's ideology is moderate, voters assign a higher probability to $\theta = -1$ if they receive the message $m = -1$ rather than $m = 1$. Therefore, the indifferent voter is located further to the right of the political spectrum for $m = -1$ than for $m = 1$. This is visualized in Figure 2, which also shows that, just as interest groups, voters endogenously come in two variants: "moderate" voters with state-dependent party preferences (with $\hat{y}_{\{\mu^*(m=1)\}} < y < \hat{y}_{\{\mu^*(m=-1)\}}$) and "radical" voters (with y outside of these bounds) whose party preferences do not depend on their beliefs about θ.

As voters can observe z perfectly in this game, they can update their belief about the state of the world conditional on z. Equation (9) states that voters trust the messages of moderate interest groups fully, whereas they refrain from updating their prior beliefs after receiving news from a radical interest group.

4.2. Public game with voter unawareness

Now assume that voters are unaware of the political position z of the message sender.

Proposition 2. The following strategy profiles and beliefs constitute the voter welfare-maximizing Perfect Bayesian Equilibrium of the public game with voter unawareness:

$$
a^*(y, c_y, m) = \begin{cases} L & \text{if } y < \hat{y}_{\{\mu^*(m)\}} \text{ and } c_y < \mathbb{E}[U(a = L, y, \theta)|\mu^*(m)] \\ R & \text{if } y > \hat{y}_{\{\mu^*(m)\}} \text{ and } c_y < \mathbb{E}[U(a = R, y, \theta)|\mu^*(m)] \\ 0 & \text{otherwise}, \end{cases} \tag{10}
$$

$$
m^*(z, \theta) = \begin{cases} -1 & \text{if } z \leq \hat{z}(\theta) \\ 1 & \text{if } z > \hat{z}(\theta), \end{cases} \tag{11}
$$

$$\mu^*(m = -1) = \begin{cases} \dfrac{p(2(h+1) + x_L + x_R)}{2(h + 2p - 1) + x_L + x_R}, \end{cases}$$

$$\mu^*(m = 1) = \begin{cases} \dfrac{p(2(h-1) - x_L - x_R)}{2(h - 2p + 1) - x_L - x_R}. \end{cases} \tag{12}$$

Proposition 2 shows that the incentives of interest groups (11) are similar to the public game with voter awareness: moderate interest groups report truthfully, whereas radical groups report their preferred message state-independent. Voters' best response to this reporting strategy is therefore also unchanged (10). See Figures 1 and 2.

What differs now is that voters do not know whether the sender of the message is a moderate or radical interest group. Therefore, voters have to form a belief that depends on the relative shares of left-wing radical, right-wing radical, and moderate interest groups (12). Consequently, they trust all news a bit—and hence news reporting is payoff-relevant for all voters. This implies that, as compared to the public game with awareness, in equilibrium, voters may trust the disinformation of a radical interest group or discount the truthful information from a moderate interest group.

4.3. Microtargeting with voter unawareness

Now the platform still distributes news to voters who are unaware of the news source's political position. However, the platform knows a lot about each voter and enables the interest group to microtarget its message to every voter's individual characteristic (i.e., to her political position y).

Proposition 3. The following strategy profiles and beliefs constitute the voter welfare-maximizing Perfect Bayesian Equilibrium of the microtargeting game with voter unawareness:

$$a^*(y, c_y, m_y) = \begin{cases} L & \text{if } y < \hat{y}_{\{\mu^*(m_y)\}} \text{ and } c_y < \mathbb{E}[U(a = L, y, \theta)|\mu^*(m_y)] \\ R & \text{if } y > \hat{y}_{\{\mu^*(m_y)\}} \text{ and } c_y < \mathbb{E}[U(a = R, y, \theta)|\mu^*(m_y)] \\ 0 & \text{otherwise,} \end{cases} \tag{13}$$

$$m^*(z, \theta) = \begin{cases} m_y = -1 & \text{for } Y_1^U < y \le Y_2^U & \text{if } z \le \hat{z}(\theta) \\ m_y = -1 \text{ (or 1) with prob. } p \text{ (or } 1-p) & \text{for } y \le Y_1^U \text{ and } y \ge Y_2^U & \\[2mm] m_y = 1 & \text{for } Y_1^U < y \le Y_2^U & \text{if } z > \hat{z}(\theta), \\ m_y = -1 \text{ (or 1) with prob. } p \text{ (or } 1-p) & \text{for } y \le Y_1^U \text{ and } y > Y_2^U & \end{cases} \tag{14}$$

where $Y_1^U = \min\{x_L, \hat{y}_{\{\mu_1\}}\}$ and $Y_2^U = \max\{x_R, \hat{y}_{\{\mu_2\}}\}$,

$$\mu^*(m_y = -1) = \begin{cases} \mu_2 = \dfrac{p\left(2(h+1) + x_L + x_R\right)}{2(h + 2p - 1) + x_L + x_R} & \text{for } Y_1^U < y \le Y_2^U \\ p & \text{for } y \le Y_1^U \text{ and } y > Y_2^U, \end{cases}$$

$$\mu^*(m_y = 1) = \begin{cases} \mu_1 = \dfrac{p\left(2(h-1) - x_L - x_R\right)}{2(h - 2p + 1) - x_L - x_R} & \text{for } Y_1^U < y \le Y_2^U \\ p & \text{for } y \le Y_1^U \text{ and } y > Y_2^U. \end{cases} \tag{15}$$
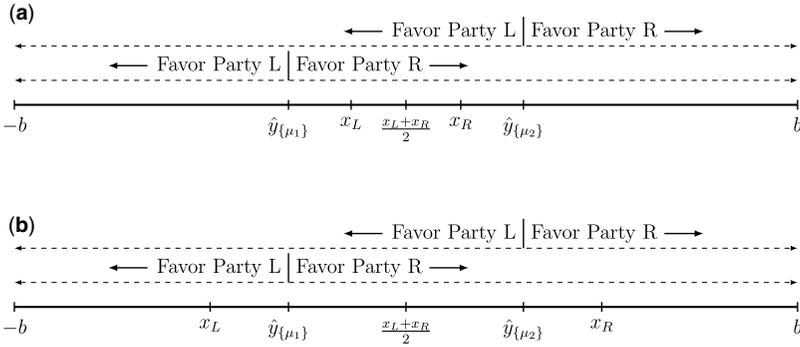
(a)

← Favor Party L │ Favor Party R →

← Favor Party L │ Favor Party R →

$-b$ $\hat{y}_{\{\mu_1\}}$ x_L $\frac{x_L+x_R}{2}$ x_R $\hat{y}_{\{\mu_2\}}$ b

(b)

← Favor Party L │ Favor Party R →

← Favor Party L │ Favor Party R →

$-b$ x_L $\hat{y}_{\{\mu_1\}}$ $\frac{x_L+x_R}{2}$ $\hat{y}_{\{\mu_2\}}$ x_R b

Figure 3 Location of the indifferent voter conditional on the microtargeted news item received. (a) $x_L > \hat{y}_{\{\mu_1\}}$ and $x_R < \hat{y}_{\{\mu_2\}}$. (b) $x_L \le \hat{y}_{\{\mu_1\}}$ and $x_R \ge \hat{y}_{\{\mu_2\}}$.

Proposition 3 shows that voters' best voting action (13) has the same structure as in the public games (see equations (7) and (10)). However, the structure of the interest group's reporting strategy differs: while in the public games, messages depend only on a group's own political position z (see equations (8) and (11)), now they also depend on voter characteristics y equation (14).

Nevertheless, the interest group is still more likely to favor party L if $\theta = -1$ than if $\theta = 1$, as depicted in Figure 1. Therefore, moderate voters assign a higher probability to $\theta = -1$ if they receive the message $m_y = -1$ than if they receive $m_y = 1$ (15). Notably, which voters are "moderate" or "radical" slightly differs from the public games. Figure 3(a) captures the case where $x_L > \hat{y}_{\{\mu_1\}}$ and $x_R < \hat{y}_{\{\mu_2\}}$. Here, voters with positions $y \in [\hat{y}_{\{\mu_1\}}, \hat{y}_{\{\mu_2\}}]$ are moderate, that is, they react to the message received by updating beliefs. In contrast, Figure 3(b) captures the case where $x_L \le \hat{y}_{\{\mu_1\}}$ and $x_R \ge \hat{y}_{\{\mu_2\}}$. Here, voters with positions $y \in [x_L, x_R]$ are "moderate."[16]

The news strategy equation (14) shows that moderate voters get truthful news from moderate interest groups. Their problem is that they cannot identify a news sender's type. Therefore, just as in the public game with unawareness, moderate voters react to news a bit (where the updating probability depends on the share of moderate vs. radical left and radical right interest groups; see μ_1 and μ_2 in Equation (15)). However, "radical" voters realize that interest groups have an overwhelming incentive to disinform them in order to make them vote for the interest group's preferred party—and hence rationally ignore the content of all news.

Summarizing, microtargeting makes radical voters dismiss all incoming news. Moderate voters do take news into account to some extent, which makes news payoff-relevant to them.

4.4. Microtargeting with voter awareness
Finally, assume the interest group can microtarget voters but voters are aware of the political position z of the interest group.

Proposition 4. The following strategy profiles and belief constitute the voter welfare-maximizing Perfect Bayesian Equilibrium of the microtargeting game with voter awareness:

---

[16]   In both cases, moderate voters are those voters who become more likely to vote for Party L (Party R) or less likely to vote for Party R (Party L) if they receive the message $m = -1$ ($m = 1$).

$$a^*(y, c_y, m_y, z) = \begin{cases} L & \text{if } y < \hat{y}_{\{\mu^*(m_y)\}} \text{ and } c_y < \mathbb{E}[U(a = L, y, \theta)|\mu^*(m_y)] \\ R & \text{if } y > \hat{y}_{\{\mu^*(m_y)\}} \text{ and } c_y < \mathbb{E}[U(a = R, y, \theta)|\mu^*(m_y)] \quad (16) \\ 0 & \text{otherwise,} \end{cases}$$

$$m^*(z, \theta) = \begin{cases} m_y = -1 & \text{for } Y_1^A < y \leq Y_2^A \quad \text{if } z \leq \hat{z}(\theta) \\ m_y = -1 \text{ (or 1)} \left(\text{with prob. } p \text{ (or } 1-p)\right) & \text{for } y \leq Y_1^A \text{ and } y > Y_2^A \\ m_y = 1 & \text{for } Y_1^A < y \leq Y_2^A \quad \text{if } z > \hat{z}(\theta), \\ m_y = -1 \text{ (or 1)} \left(\text{with prob. } p \text{ (or } 1-p)\right) & \text{for } y \leq Y_1^A \text{ and } y > Y_2^A \end{cases}$$

$$(17)$$

where $Y_1^A = min\{x_L, \hat{y}_{\{\mu=0\}}\}$ and $Y_2^A = max\{x_R, \hat{y}_{\{\mu=1\}}\}$,

$$\mu^*(m_y = -1, z) = \begin{cases} 1 & \text{if } \hat{z}(\theta = 1) < z \leq \hat{z}(\theta = -1) \text{ and } Y_1^A < y \leq Y_2^A \\ p & \text{otherwise,} \end{cases}$$

$$(18)$$

$$\mu^*(m_y = 1, z) = \begin{cases} 0 & \text{if } \hat{z}(\theta = 1) < z \leq \hat{z}(\theta = -1) \text{ and } Y_1^A < y \leq Y_2^A \\ p & \text{otherwise.} \end{cases}$$

Proposition 4 shows that the structure of the interest group's reporting is similar to Proposition 3; see equation (17): moderate interest group's message contains correct information for moderate voters and uninformative news for radical voters; a message from a radical interest group is always uninformative. As opposed to the microtargeting game with voter unawareness, voters know who is sending the news they receive. Therefore, moderate voters rationally ignore messages by radical groups and fully trust messages from moderate groups. Radical voters still do not trust any messages.

4.5. The persuasion, (de-)mobilization, and disciplining effects

Based on Propositions 1–4, with voters' unawareness, the interest group can have a *persuasion effect* on moderate voters and a *(de)mobilization effect* on radical voters. The first channel affects the *incentive-compatibility constraint* of voters, persuading them to vote for the interest group's preferred party instead of the opponent. The second channel affects the *participation constraint* of voters, either mobilizing radicals favoring the interest group's preferred party to participate in the election or demobilizing radicals supporting the opponent by convincing them to abstain.[17]

Both effects are illustrated in Figure 4 for an interest group supporting party L in a game in which $\hat{y}_{\{\mu_1\}} < x_L < x_R < \hat{y}_{\{\mu_2\}}$: "+" indicates a wanted effect for the interest group; "−" indicates an unwanted side-effect of sending $m = -1$. The interest group can persuade moderate voters ($\hat{y}_{\{\mu_1\}} < y < \hat{y}_{\{\mu_2\}}$) to favor party L over party R in both the public game (panel

---

[17] Empirically, Liberini et al. (2020) show that the advertisement strategies employed in the 2016 US Presidential elections to (a) persuade swing voters to vote for candidate Trump and (b) to make Republican voters vote, differed substantially. In our theory, these two different effects arise endogenously despite the simple uni-dimensional message space, $\{-1, 1\}$.
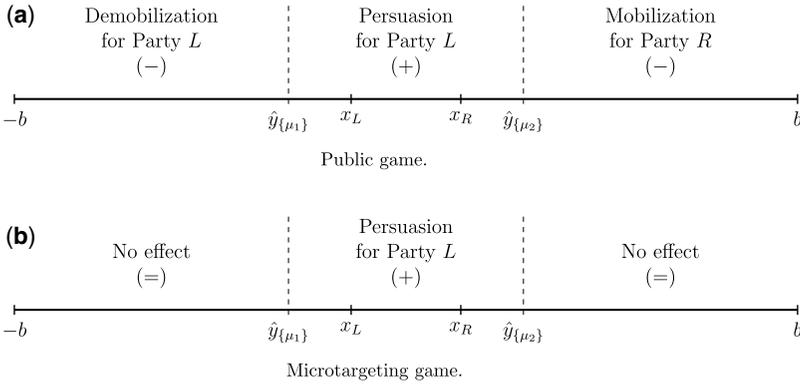
Figure 4 Influence on voting behavior by an interest group that favors party L. (a) Public game. (b) Microtargeting game.

4(a)) and the microtargeting game (panel 4(b)). However, in the public game sending $m = -1$ has an unwanted side-effect on radical voters,[18] which disappears if messages can be personalized to voter types.[19]

Moreover, Propositions 1–4 show that a moderate interest group has an incentive to communicate truthfully to moderate voters because their voting preferences are aligned. This is different w.r.t. radical voters. However, if the interest group can only send one uniform message to all voters, there is a *disciplining effect of the public news dissemination technology*: moderate interest groups report truthfully to *all* voters. Therefore, voters fully trust the message of a moderate interest group. In contrast, a radical interest group always reports the same news. Consequently, if voters can identify a radical interest group, they ignore its message and do not update p.

On the voters' side, there are three groups that are affected differently by a change in their awareness of the interest group's ideological position (see Figure 5): *stable moderate voters* (with ideological position $Y_1^U < y \leq Y_2^U$) and *stable radical voters* ($y \leq Y_1^A$ and $y > Y_2^A$) are moderate or radical, respectively, in all games. However, *unstable moderate voters* ($Y_1^A < y \leq Y_1^U$ and $Y_2^U < y \leq Y_2^A$) are moderate in games with awareness but radical in games with unawareness.[20] Then they always prefer their favorite party over the opponent. With awareness, in contrast, news from an identified, moderate interest group can persuade them because it is truthful, without a shadow of doubt.

5. WELFARE, ELECTION FLIPPING, AND COMPETITION
5.1 Voter welfare

A news report that is (partly) trusted serves two functions for voters. First, it informs them whether it is worthwhile to turn out to vote, or not. Second, it helps moderate voters to find out which party's policy position is closest to their ideal position. A higher share of trustworthy news and a higher number of voters receiving such news increase total voter welfare.

---

[18] Sending $m = -1$ demobilizes radical left voters, who think party L is not left enough, and mobilizes radical right voters for party R, who now believe that R is acceptable and do not abstain.
[19] In the example, the left-favoring interest group would send $m = -1$ to moderate voters and would like to send $m = 1$ to all radicals. The latter, however, is no part of an equilibrium because voters could then infer θ from m_y. Only a mixed strategy solves this dilemma; see Equations (14) and (17).
[20] Unstable moderate voters only exist if $x_L > y_{\{\mu=0\}}$ or $x_R < y_{\{\mu=1\}}$.
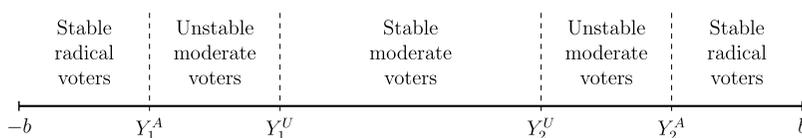
Figure 5. Overview voter groups.

Proposition 5. (Voter Welfare) In the voter welfare-maximizing equilibria of the four games analyzed, voter welfare compares as follows:

1) *Total voter welfare is strictly higher in public games than in microtargeting games.*
2) *Total voter welfare is strictly higher in games with awareness than in games with unawareness.*

Proposition 5 yields the following Corollary:

Corollary 1. (*Voters' Ranking of Games*)
   Total voter welfare across the four games ranks as follows:
   1. Public news with Awareness
   2./3. Microtargeting with Awareness
   2./3. Public news with Unawarenes
   4. Microtargeting with Unawareness

The intuition of Corollary 1, which is a central result of this article, is straightforward. Public news dissemination and awareness of the interest group's political position maximize voter welfare. Public news helps *radical* voters benefit from the discipline effect, whereas awareness enables *all* voters to recognize if an interest group is *moderate* (and hence trustworthy) or *radical* (and hence untrustworthy). On the flip-side, voters fare worst under microtargeting and unawareness, where both of these effects do not exist. The ranking of the two intermediate regimes depends on parameter realizations.

5.2 Flipping an election with disinformation

While we have shown that manipulating voters is possible in equilibrium, a critical question remains whether this effect could be large enough to flip an election through disinformation if the winning party would have lost the election (a) in the absence of news and (b) if voters had complete information about the state of the world.

   The degree to which election flipping is possible depends on the probability that voters are exposed to a malevolent interest group with an incentive to flip an election and on the ability of this interest group to actually flip an election. While the probability of exposure to a malevolent interest group is the same in all games, the ability of an interest group to flip an election differs across games. We construct a single measure, the *election flipping potential*, to compare an interest group's ability to influence the outcome of an election in its favor across the games.

Definition 1 (Election Flipping Potential). The election flipping potential is defined as the maximum (minimum) prior belief $\hat{p}$ up to which a right-wing (left-wing) interest group can get party R (party L) elected.

Proposition 6. (Election Flipping)

1) *With voter unawareness, an interest group's election flipping potential is larger in the microtargeting game than in the public game.*
2) *Voter unawareness is a necessary condition for election flipping.*

Proof: see Supplementary Appendix. The key insight of Proposition 6 is part (2): as long as voters are aware of an interest group's political position, election flipping is impossible. Without such awareness, elections can be flipped, in principle.[21]

5.3 Interest group competition

Assume there are $K \geq 1$ interest groups, each sending one message to voters. Each interest group $k \in \{1, \ldots, K\}$ has an ideological position, z_k, which is *i.i.d.* from a uniform distribution over $[-h, h]$. Denote by q the number of messages $m = -1$ and by s the number of messages $m = 1$ that a voter receives. Hence, $K = q + s$. The remaining elements of the model are unchanged. We prove in the Supplementary Appendix:

Proposition 7. (Interest Group Competition) Interest group competition increases voter welfare in all games.

In games with voter unawareness, posterior beliefs of moderate voters are increasing in q, decreasing in s and concave in both q and s. Hence, a marginal news report with the same message still affects beliefs but does so at a decreasing rate. Crucially, conflicting messages do not cancel each other out (meaning that voters' posterior beliefs are not equal to their prior beliefs) unless party positions are exactly symmetric. For instance, suppose that $x_L + x_R < 0$, implying that there are more radical right interest groups than radical left interest groups. Receiving a message $m = -1$ (which favors party L) is now more informative about the true state than $m = 1$, which implies that a (moderate) voter weighs the former messages more strongly than the latter. Moreover, moderate voters' beliefs converge to the truth if K increases.[22] Finally, the ratio of radical to moderate voters weakly decreases if K increases. The intuition is that one news item might not be enough to persuade voters to change their party preference but multiple news items with the same message can.

For games with voter awareness, equations (9) and (18) show that a single news item from a moderate interest group is sufficient to resolve all uncertainty about the state of the world. Hence, receiving an additional news item from a moderate interest group has no added value (and radical interest groups' messages are ignored, anyways). However, the ex-ante probability that a voter is exposed to news from a moderate interest group increases in K. Therefore, voters also benefit from interest group competition in games with voter awareness.

---

[21] Our findings imply that a radical interest group is best off in an environment with voter unawareness. With unawareness, moderate voters believe that every message is somewhat credible, even if it is sent by an unreliable radical interest group. Hence, the radical interest group can deceive some moderate voters to vote for its preferred party. The moderate interest group type, however, is better off with voter awareness, which prevents moderate voters from discounting its message. In an environment in which voter (un)awareness could be endogenously determined by the interest groups (e.g., by disclosure of their types), moderate interest groups would like to distinguish themselves from radical interest groups. If a moderate interest group can perfectly disclose its type, a radical interest group would no longer be able to influence voting and election flipping would not occur anymore. If disclosure is imperfect (i.e., if there is a positive probability that voters do not observe the interest group type), voters still (somewhat) believe a message from a radical interest group but less so than without any disclosure. Hence, election flipping would still be possible but less likely than without any disclosure.

[22] For instance, consider a game with voter unawareness and parameter values $p = 0.5$, $x_L = -1$, $x_R = 1$ and $h = 3$. Suppose $\theta = -1$. In expectation, a moderate voter holds belief $\mathbb{E}[\mu(\cdot)|K = 1] \approx 0.56$ if she receives one message, $\mathbb{E}[\mu(\cdot)|K = 5] \approx 0.76$ if she receives five messages, and $\mathbb{E}[\mu(\cdot)|K = 25] \approx 1$ if she receives 25 messages.

6. DISCUSSION AND CONCLUSION

Our analysis has shown that microtargeting hurts especially radical voters (because the discipline effect of public news dissemination falls away) and that unawareness about a sender's political position prevents all voters from recognizing whether an interest group is *moderate* (and hence trustworthy) or *radical* (and hence untrustworthy). These key insights can inform policymaking.

To reduce microtargeting, a committee of the UK's House of Commons suggested "a minimum limit for the number of voters sent individual political messages […] at a national level" (House of Commons 2019, paragraph 142). Consequently, voters could not be targeted individually or in small groups. However, in games with microtargeting, we show that even with our most limited message space of $\{-1, 1\}$ it is possible to manipulate an election. Customization of the message content at the individual level is not necessary as long as every voter can be attributed to a certain group with homogeneous characteristics and this group is not too small. Therefore, according to our results, it would be possible to both comply with the House of Commons' proposal and to effectively deceive voters and manipulate an election.

Arguably today's news platforms are best characterized by the *microtargeting game with voter unawareness*, which yields the worst welfare outcome of the four games analyzed.[23] Hence, a better intervention seems to promote policies reestablishing awareness about interest group's political positions, such as, compelling platforms to implement technologies by which users can easily recognize the identity of a specific message's original sender. It should help voters to assess the political position of a news item's originator, to update beliefs about the credibility of news received and, hence, to make voting decisions that are more aligned with their own interests.[24] This would particularly benefit moderate voters.

A complementary policy proposal supported by our analysis of competing interest groups is to increase competition among senders of political news, for instance, by increasing media plurality, supporting local news stations, and preventing large conglomerates from monopolizing news dissemination.

Our model relies on stylized assumptions. Confirmation bias is said to be a relevant phenomenon among voters (Plous 1993), which might be included in our framework in future research. Similarly, deviations from the randomized draws and uniform distribution of messages, especially on the news platform's side (platform bias) and on the voters' side (endogenous news consumption) are promising extensions.[25] This requires further empirical work to assess how far these assumptions are acceptable simplifications of reality. Among the most pressing questions are: To which extent do voters fall prey to disinformation about politically relevant events depending on the information they have about interest group's political positions? To which extent can political interest groups actually make use of this weakness and manipulate election outcomes? To get clean results, such empirical testing could first be conducted in lab experiments. Then, to verify the external validity of lab results, they should be tested in the field. A lot of future work is waiting.

---

[23] Moreover, in the Supplementary Appendix, we derive an empirical hypothesis: if unawareness of voters about interest groups' political positions grows over time, for example, through the proliferation of social media in political news reporting, we expect fewer moderate and more radical voters.

[24] For an example that moves in the direction of such a policy, see footnote 6.

[25] Both characteristics were suggested by Piolatto and Schuett (2015).

SUPPLEMENTARY MATERIAL

Supplementary material is available at *Journal of Law, Economics, & Organization* online.

REFERENCES

Allcott, Hunt, and Matthew Gentzkow. 2017. "Social Media and Fake News in the 2016 Election," 31 *Journal of Economic Perspectives* 211–36.

Alonso, Ricardo and Gerard Padró I. Miquel 2023. "Competitive Capture of Public Opinion." NBER Working Paper 31414, NBER.

Austen-Smith, David, and Jeffrey S. Banks 2002. "Costly Signaling and Cheap Talk in Models of Political Influence," 18 *European Journal of Political Economy* 263–80.

Baron, David P. 2006. "Persistent Media Bias," 90 *Journal of Public Economics* 1–36.

Besley, Timothy, and Andrea Prat. 2006. "Handcuffs for the Grabbing Hand? Media Capture and Government Accountability," 96 *American Economic Review* 720–36.

Binswanger, Johannes and Jens Prüfer. 2012. "Democracy, Populism, and (un)Bounded Rationality," 28 *European Journal of Political Economy* 358–72.

Chan, Jimmy, and Wing Suen. 2009. "Media as Watchdogs: The Role of News Media in Electoral Competition," 53 *European Economic Review* 799–814.

Downs, Anthony. 1957. "An Economic Theory of Political Action in a Democracy," 65 *Journal of Political Economy* 135–50.

Faris, Robert, Hal Roberts, Bruce Etling, Nikki Bourassa, Ethan Zuckerman, and Yochai Benkler. 2017. "Partisanship, Propaganda, and Disinformation: Online Media and the 2016 US Presidential Election." 6 *Berkman Klein Center Research Publication*, 2017–6.

Farrell, Joseph and Robert Gibbons. 1989. "Cheap Talk with Two Audiences," 79 *American Economic Review* 1214–23.

Galeotti, Andrea, Christian Ghiglino, and Francesco Squintani. 2013. "Strategic Information Transmission Networks," 148 *Journal of Economic Theory* 1751–69.

Gentzkow, Matthew, Jesse M. Shapiro, and Daniel F. Stone. 2015. "Media Bias in the Marketplace: Theory," in Simon Anderson, Joel Waldfogel, and David Stromberg, eds., *Handbook of Media Economics*, Vol. 1B, 623–45. Amsterdam, Netherlands: Elsevier.

Germano, Fabrizio, and Francesco Sobbrio. 2020. "Opinion Dynamics via Search Engines (and Other Algorithmic Gatekeepers)," 187 *Journal of Public Economics* 104188.

Glaeser, Edward L., Giacomo A.M. Ponzetto, and Jesse M. Shapiro. 2005. "Strategic Extremism: Why Republicans and Democrats Divide on Religious Values," 120 *The Quarterly Journal of Economics* 1283–330.

Goltsman, Maria, and Gregory Pavlov. 2011. "How to Talk to Multiple Audiences," 72 *Games and Economic Behavior* 100–22.

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

Hersh, Eitan D. 2015. *Hacking the Electorate: How Campaigns Perceive Voters*. Cambridge: Cambridge University Press.

Hillygus, D. Sunshine, and Todd G. Shields. 2014. *The Persuadable Voter*. Princeton, NJ: Princeton University Press.

Hoffmann, Florian, Roman Inderst, and Marco Ottaviani. 2020. "Persuasion through Selective Disclosure: Implications for Marketing, Campaigning, and Privacy Regulation," 66 *Management Science* 4958–79.

House of Commons. 2019. "Disinformation and 'Fake News': Interim Report," *Digital, Culture, Media and Sport Committee*. United Kingdom Parliament. Authority of the House of Commons.

Kalogeropoulos, Antonis and Nic Newman. 2017. *I Saw the News on Facebook*. England: Reuters Institute for the Study of Journalism.

Kang, Hyunjin, Keunmin Bae, Shaoke Zhang, and S Shyam Sundar. 2011. "Source Cues in Online News: Is the Proximate Source More Powerful than Distal Sources," 88 *Journalism & Mass Communication Quarterly* 719–36.

Kartal, Melis, and James Tremewan. 2018. "An Offer You Can Refuse: The Effect of Transparency with Endogenous Conflict of Interest" 161 *Journal of Public Economics* 44–55.

Kavanagh, Jennifer, William Marcellino, Jonathan S. Blake, Shawn Smith, Steven Davenport, and Mahlet Gizaw. 2019. *News in a Digital Age: Comparing the Presentation of News Information over Time and across Media Platforms*. Santa Monica, CA: RAND Corporation.

Kavanagh, Jennifer, and Michael D. Rich 2018. *Truth Decay: An Initial Exploration of the Diminishing Role of Facts and Analysis in American Public Life*. Santa Monica, CA: Rand Corporation.

Kosinski, Michal, David Stillwell, and Thore Graepel. 2013. "Private Traits and Attributes Are Predictable from Digital Records of Human Behavior," 110 *Proceedings of the National Academy of Sciences* 5802–5.

Levy, Gilat, and Ronny Razin. 2019. "Echo Chambers and Their Effects on Economic and Political Outcomes," 11 *Annual Review of Economics* 303–28.

Li, Ming, and Kristóf Madarász. 2008. "When Mandatory Disclosure Hurts: Expert Advice and Conflicting Interests," 139 *Journal of Economic Theory* 47–74.

Liberini, Federica, Michela Redoano, Antonio Russo, Angel Cuevas, and Ruben Cuevas. 2020. "Politics in the Facebook Era. Evidence from the 2016 US Presidential Elections." *CESifo Working Paper No. 8235.*

Maréchal, Nathalie. 2017. "Networked Authoritarianism and the Geopolitics of Information: Understanding Russian Internet Policy," 5 *Media and Communication* 29–41.

Newman, Nic, Richard Fletcher, Anne Schulz, S imge Andi, and Rasmus Kleis Nielsen. 2020. *Reuters Institute Digital News Report 2020*. England: Reuters Institute for the Study of Journalism.

Petrova, Maria. 2012. "Mass Media and Special Interest Groups." *Journal of Economic Behavior & Organization* 84 (1):17–38.

Piolatto, Amedeo, and Florian Schuett. 2015. "Media Competition and Electoral Politics," 130 *Journal of Public Economics* 80–93.

Plous, Scott. 1993. *The Psychology of Judgment and Decision Making*. New York, NY: Mcgraw-Hill Book Company.

Pons, Vincent and Clémence Tricaud. 2018. "Expressive Voting and Its Cost: Evidence from Runoffs with Two or Three Candidates," 86 *Econometrica* 1621–49.

Schipper, Burkhard C., and Hee Yeul Yoo, 2019. "Political Awareness, Microtargeting of Voters, and Negative Electoral Campaigning," 14 *Quarterly Journal of Political Science* 41–88.

Schuessler, Alexander A. 2000. *A Logic of Expressive Choice*. Princeton, NJ: Princeton University Press.

Shapiro, Jesse M. 2016. "Special Interests and the Media: Theory and an Application to Climate Change," 144 *Journal of Public Economics* 91–108.

Sobbrio, Francesco. 2011. "Indirect Lobbying and Media Bias," 6 *Quarterly Journal of Political Science* 3–4.

Sobbrio, Francesco. 2014. "Citizen-Editors' Endogenous Information Acquisition and News Accuracy," 113 *Journal of Public Economics* 43–53.

Tucker, Joshua, Andrew Guess, Pablo Barberá, Cristian Vaccari, Alexandra Siegel, Sergey Sanovich, Denis Stukal, and Brendan Nyhan. 2018. "Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature," Available at SSRN: https://ssrn.com/abstract=3144139.

Tyran, Jean-Robert and Alexander K. Wagner. 2019. "Experimental Evidence on Expressive Voting," in Roger D. Congleton, Bernard N. Grofman, and Stefan Voigt, eds., *The Oxford Handbook of Public Choice*, Vol. 2, 928–40. Oxford: Oxford University Press.

Van der Meer, Tom W. G. 2017. "Political Trust and the 'Crisis of Democracy'," in William R. Thompson, ed., *Oxford Research Encyclopedia of Politics*. Oxford: Oxford University Press. https://doi.org/10.1093/acrefore/9780190228637.013.77.

Van Winden, Frans. 2004. "Interest Group Behavior and Influence," in Charles Rowley and Friedrich Schneider, eds., *The Encyclopedia of Public Choice*, 118–29. Berlin, Germany: Springer Science & Business Media.

Yu, Zhihao. 2005. "Environmental Protection: A Theory of Direct and Indirect Competition for Political Influence," 72 *The Review of Economic Studies* 269–86.

Zhuravskaya, Ekaterina, Maria Petrova, and Ruben Enikolopov. 2020. "Political Effects of the Internet and Social Media," 12 *Annual Review of Economics* 415–38.