A Theory on Media Bias and Elections

Sun, J.; Schram, A.; Sloof, R.

DOI:
10.2139/ssrn.3420703

Link to publication

TI 2019-048/I
Tinbergen Institute Discussion Paper

A Theory on Media Bias and Elections

Junze Sun[1,3,4]
Arthur Schram[1,2,3,4]
Randolph Sloof[3,4]

[1] CREED
[2] Department of Economics, EUI
[3] Amsterdam School of Economics
[4] Tinbergen Institute

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

Contact: discussionpapers@tinbergen.nl

More TI discussion papers can be downloaded at https://www.tinbergen.nl

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 598 4580

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900

A Theory on Media Bias and Elections*

Junze Sun† Arthur Schram‡ Randolph Sloof§

July 2, 2019

Abstract

We develop a tractable theory to study the impact of biased media on election outcomes, voter turnout and welfare. News released by media allows voters to infer the relative appeal of the two candidates, and the closeness of elections. In large elections, the former determines the election outcome, whereas the latter drives voter turnout. With a single media outlet, a rise in media bias affects the election outcome in a non-monotonic way, and reduces voter welfare by decreasing the probability of electing the efficient candidate and increasing aggregate turnout costs. Introducing extra media outlets can systematically shift the election outcome and voter turnout in either direction, but it weakly improves voter welfare. The impact of other ways to strengthen media competition – such as increased polarization and prevention of collusion – critically depends on whether media have commitment power; if not, they can worsen information transmission and voter welfare.

JEL Codes: D72, D82, D83

Keywords: media bias, voting, Poisson games, media competition, commitment

---

*We thank Sourav Bhattacharya, Aaron Kamm, Navin Kartik, David Levine, Zara Sharif and Joel Sobel for insightful comments at various stages. We also thank audiences in the 2017 NYU-CESS Political Science Conference, the 2017 ESA World meeting in UCSD, the 2018 CES Annual Conference in Hefei, the 2018 EEA-ESEM Conference in Cologne, the 2018 BiGSEM Workshop on Economic Theory at Bielefeld, the 2018 SAEe at Madrid, the 2018 RES PhD Meeting at London, and also seminar participants from University of Amsterdam, Tinbergen Institute, Erasmus University of Rotterdam, EUI, NYU-Abu Dhabi, and Hong Kong University of Science and Technology (HKUST), for helpful suggestions and comments. All errors are our own.

†CREED, Amsterdam School of Economics and Tinbergen Institute. Email: J.Sun@uva.nl.

‡CREED, Amsterdam School of Economics and Tinbergen Institute, and Department of Economics at EUI, Florence. Email: A.J.H.C.Schram@uva.nl.

§Amsterdam School of Economics and Tinbergen Institute. Email: R.Sloof@uva.nl.

1 Introduction

Mass media are important information sources that voters rely on to make political decisions. However, media may have their own agendas and be biased in news reporting.[1] Over the past few decades, empirical studies have indeed confirmed the pervasiveness of media bias and documented significant impacts of media bias on voting behavior.[2] For example, DellaVigna and Kaplan (2007) find that exposure to Fox News increased the Republican vote share and voter turnout in the 2000 U.S. presidential election. Enikolopov et al. (2011) find that exposure to NTV, the only national TV channel independent of the Russian government in the late 1990s, increased the vote share for major opposition parties and decreased voter turnout in the 1999 parliamentary election. At the same time, empirical evidence shows that media entry may either increase or decrease voter turnout, but it does not systematically affect party vote shares.[3]

Why does media exposure sometimes increase voter turnout while in other cases it does the opposite? Why are media outlets able to affect party vote shares in a particular election, but unable to affect it systematically in the long run? Importantly, even if media cannot systematically affect party vote shares, does this imply that media outlets cannot systematically manipulate election outcomes as well? These are crucial questions for understanding the role of media in modern democracies. Yet, theories providing unified answers are still absent.

In this paper we take a step towards filling this gap by developing a tractable theory to explain how biased media affect election outcomes, voter turnout, and voter welfare.[4]

[1] See Prat and Strömberg (2013) and Gentzkow et al. (2015) for comprehensive overviews of the theoretical discussions on possible sources of media bias.

[2] Groseclose and Milyo (2005), Gentzkow and Shapiro (2010), Larcinese et al. (2011), Puglisi and Snyder (2011), and Puglisi and Snyder (2015a) provide empirical evidence and measures for media bias in the US. See Puglisi and Snyder (2015b) for a comprehensive review. The empirical evidence for impacts of biased media on voting behavior is reviewed in Section 2.

[3] Gentzkow et al. (2011) and Drago et al. (2014) find that newspaper entry increased voter turnout in the U.S. and in Italy, respectively. In contrast, Cagé (2017) find that media entry had a negative impact on voter turnout in France. Gentzkow et al. (2011) find that newspaper entry/exit did not systematically influence party vote shares in the U.S. during 1868–1928.

[4] Our analysis primarily concerns traditional, large-scale media such as newspapers and TV-channels. Obviously, social media play an important role in modern elections (Allcott and Gentzkow, 2017; Smith and Anderson, 2018). Yet, traditional media are still major players. For instance, Kennedy and Prat (2019) examine the media consumption of people across 36 countries and estimate, for a wide variety of news sources, their attention shares and influential power on election outcomes. They conclude that "Of the 88 [most powerful] distinct media organizations, 72 specialize in some form of television programming, 12 are primarily print sources, and only 4 are pure Internet players".

Methodologically, we model the electorate by a canonical pivotal voter framework with a Poisson distributed population (Myerson, 2000) and analytically derive voting equilibria in large elections. Subsequently, we model the interaction between media and voters via a sender-receiver game, and derive communication equilibria under both cheap talk (Crawford and Sobel, 1982) and Bayesian persuasion (Kamenica and Gentzkow, 2011).

We consider an election (under simple majority rule) between two candidates, A and B, who differ along a quality dimension. The quality of a candidate can be interpreted as her integrity, valence, or executive ability, attributes that all voters equally appreciate. Apart from caring about quality, voters also have private ideological preferences over candidates. In this way, it is possible for a voter to prefer the (in his view) ideologically less-favored candidate if her quality is sufficiently superior to that of the other candidate. Voters can vote for either candidate or abstain, but casting a vote is costly; these costs are private information. We assume that the election is symmetric in the sense that a priori (i) the qualities of both candidates are identical, and (ii) voters' ideological preferences do not favor either candidate over the other. Voters know precisely their private ideological preferences, but do not observe candidates' qualities. Media outlets precisely observe candidates' qualities and communicate to voters by disseminating public news.

The first insight from our analyses is that public information about candidates' qualities allows voters not only to assess which candidate is more appealing, but also to gauge the closeness of the election. In large elections, the former determines the expected party vote shares and the election outcome, whereas the latter drives voter turnout. With a higher perceived relative quality, a candidate is able to 1) convince marginal voters to support her, and 2) increase the willingness-to-vote of her own supporters, relative to her opponent's supporters. These increase the candidate's expected vote share and winning chances. On the other hand, with a large perceived quality difference, it becomes very unlikely that the election ends up in a close race since the quality-superior candidate is expected to gain a substantially larger vote share than her opponent. In this situation, voter turnout is low because the chances of casting a pivotal vote are low and voters have little incentive to cast costly votes. Therefore, the perceived quality difference is negatively associated with the closeness of elections. Voter turnout is thus a decreasing function of the perceived quality difference, in line with the "competition effect", which predicts voter turnout to be higher in closer elections (Levine and Palfrey, 2007).

Our second insight is that media outlets can systematically manipulate election outcomes and voter turnout by strategically providing public information, even if voters are Bayesian.

Depending on the information received, voters may either update their posterior beliefs towards candidate A having a superior quality, or the opposite. But their posterior beliefs should – by Bayes' rule – always average back to the prior (i.e., no quality differences). From this perspective, manipulating public information cannot systematically shift rational voters' posterior beliefs. However, perhaps surprisingly, this property does not carry over to election outcomes. This is because the election outcome is generically *nonlinear* in voters' posterior beliefs; it is in some regions much more sensitive to a shift in voters' beliefs than in others. For this reason, the ex-post election outcomes need not average back to the outcome under voters' common prior. This nonlinearity gives room for media outlets to systematically manipulate election outcomes even in a rational voter framework; they may strategically disclose information to shift voters' posterior beliefs in regions with different sensitivities (Kamenica and Gentzkow, 2011). In the same spirit, we also show that voter turnout is generically nonlinear in voters' posterior expectations of candidates' quality difference, which determines the closeness of elections. By affecting the distribution of voters' perceived closeness, media exposure may also systematically influence voter turnout.

Building on these general insights, we deliver three contributions to the literature. First, we derive robust bounds for the extent to which mass media can systematically influence voting behavior and election outcomes. We show that, when the electorate is sufficiently polarized, mass media can hardly affect the expected party vote shares from an ex-ante perspective. Nevertheless, such inability does not prevent mass media from systematically manipulating election outcomes and voter turnout. Perhaps surprisingly, we show that the opposite is true – by strategically releasing public information, mass media can manipulate both the election outcome and voter turnout to an arbitrarily large extent. This result provides a possible rationale for the empirical observation that changes in the media landscape may not systematically affect party vote shares, yet they can either increase or decrease voter turnout (Gentzkow et al., 2011; Drago et al., 2014; Cagé, 2017).

Second, we derive precise comparative statics predictions for how biased media affect election outcomes, voter turnout, and voter welfare. We assume that each media outlet has a commonly known ideological position. An *unbiased* media outlet always prefers the quality-superior candidate. A *biased* media outlet, however, may prefer a certain candidate even if her quality is inferior. When media outlets communicate to voters via cheap talk, we show that in equilibrium their news reporting strategies take a simple cutoff structure: endorsing a candidate only if her relative quality lies above a certain threshold. This threshold is decreasing in an outlet's bias towards that candidate. In this way, media outlets partition the state (i.e., a candidate's relative quality) space into adjacent and disjoint intervals.

We first study the impact of media bias on elections with a single media outlet. We show that, ex-ante, the relationship between media bias and the election outcome is non-monotonic. On the one hand, an outlet with a stronger bias towards (for instance) candidate A is *a priori* more likely to endorse candidate A; this effect *per se* increases A's electoral prospects. On the other hand, this outlet's endorsement for candidate A becomes less credible, and its endorsement for the opposing candidate B becomes more credible, as its bias towards A increases. These effects *per se* reduce A's winning chances. The net effect, therefore, depends on which of these opposite effects dominate. For a small degree of media bias, the first effect dominates and a marginal increase of bias increases candidate A's winning probability. For a sufficiently large degree of bias the opposite applies. The probability of electing the quality-superior candidate, however, unambiguously declines in media bias.

The impact of media bias on voter turnout follows from what an endorsement reveals about the expected closeness of the election. With a strong bias towards candidate A, an outlet's endorsement for A is very uninformative about the quality difference between candidates. Given the symmetric priors, voters will expect the election to end up in a close race, which leads to a high turnout rate. In contrast, this outlet's endorsement for the opposing candidate B is highly informative and signals a substantial quality difference in favor of candidate B. As a consequence, voters will expect the election to result in a landslide victory for B and thus they have little incentive to vote. Therefore, as media bias towards candidate A increases, voter turnout is higher (lower) conditional on an endorsement for candidate A (B). We show, however, that ex-ante the expected voter turnout increases in media bias unambiguously. As a result, an increase in media bias reduces voter welfare not only by decreasing the probability of electing the quality-superior candidate, but also by increasing aggregate turnout costs.

We then study the electoral impact of introducing a second media outlet. Under the cutoff endorsement strategy, having an extra outlet essentially refines the information partition induced by the existing outlet. Hence, more information must be transmitted to voters. We show that, and precisely identify conditions under which, introducing a second outlet can systematically shift the election outcome and voter turnout in either direction. These results may reconcile the mixed empirical evidence on the relationship between media entry and voter turnout (Gentzkow et al., 2011; Drago et al., 2014; Falck et al., 2014; Cagé, 2017). From a welfare perspective, we show that media entry weakly increases voter welfare in sufficiently large elections. This confirms, in an environment in which media outlets communicate to voters via cheap talk, the conventional wisdom that media competition improves information transmission and voter welfare (Besley and Prat, 2006; Gentzkow and Shapiro, 2006, 2008; Prat, 2018).

Finally, we explore the extent to which this conventional wisdom can be generalized, by varying both the communication protocol used by media outlets and the notion of media competition. In all previous analyses we assumed that media outlets communicate to voters via cheap talk; they cannot commit to any specific news reporting strategy. This assumption seems plausible in weak democratic regimes where politicians have strong control (e.g., through censorship or bribing) over the media outlets (McMillan and Zoido, 2004; Besley and Prat, 2006; Prat and Strömberg, 2013; Enikolopov and Petrova, 2015). However, under other circumstances media outlets may possess the possibility of commitment because of substantial costs of switching editorial strategies and reputation concerns in the marketplace (Chan and Suen, 2008; Duggan and Martinelli, 2011; Sobbrio, 2014; Guo and Shmaya, 2019).[5] We study the situation with media commitment using the Bayesian persuasion framework (Gentzkow and Kamenica, 2016, 2017) and find that commitment plays a key role in shaping information transmission in equilibrium. In particular, "conflicting states", where at least two media outlets' partisan preferences disagree, are generically revealed to voters if commitment is possible while they may not be revealed without commitment.

Analogous to Gentzkow and Kamenica (2016) we consider three different notions of media competition: (i) introducing additional media outlets, (ii) increasing media polarization, and (iii) preventing media collusion. In line with their results, all three ways of strengthening media competition improve information transmission and voter welfare in large elections when commitment is possible. However, in the absence of media commitment the opposite may apply under either increased polarization or the prevention of collusion. Our results suggest that increasing media competition may have perverse impacts on both information transmission and voter welfare in the absence of media commitment.

The remainder of this paper is organized as follows. Section 2 reviews the related literature. Section 3 introduces the model. Section 4 derives voting equilibria in large elections when candidates' qualities are common knowledge. It also explores the extent to which manipulating public information can influence voters' behavior and election outcomes.

[5] Chan and Suen (2008), Duggan and Martinelli (2011) and Sobbrio (2014) argue that media may obtain commitment power by hiring editors/journalists with publicly known ideological preferences or editorial rules. These decisions are usually made prior to news reporting and are costly to switch in the short run. Alternatively, Guo and Shmaya (2019) show that a sender can secure her commitment payoff if the punishment for miscalibration (i.e., the discrepancy between the report and the truth) is sufficiently high.

Section 5 presents comparative statics results concerning the influence of biased media on elections. Section 6 discusses how media commitment affects information transmission and voter welfare, as well as its implications for the welfare impacts of increasing media competition. Section 7 concludes and suggests some avenues for future research.

2 Related literature

This paper relates to a large body of empirical literature on the electoral impact of mass media. One branch of this literature exploits reasonably exogenous cross-sectional variations in media exposure to identify media's influence in particular elections (DellaVigna and Kaplan, 2007; Gerber et al., 2009; Chiang and Knight, 2011; Enikolopov et al., 2011; Durante and Knight, 2012; Adena et al., 2015). These studies find that exposure to a biased media outlet can 1) either increase or decrease voter turnout, and 2) shift the party vote share and voters' voting intentions in favor of the endorsed candidate. The strengths of these effects depend on the magnitude of the bias. A second branch exploits local-level discrete changes in the number of media outlets over time to identify the systematic influence of media entry on voter behavior (Gentzkow, 2006; Gentzkow et al., 2011; Drago et al., 2014; Falck et al., 2014; Gavazza et al., 2015; Cagé, 2017; Ellingsen et al., 2018). These studies find mixed evidence on how media entry affects voter turnout. Our theory produces comparative statics predictions that could potentially reconcile such mixed findings in a unified framework. Our paper also relates to theoretical work on the political impact of media bias (Bernhardt et al., 2008; Duggan and Martinelli, 2011; Piolatto and Schuett, 2015). Unlike these studies, our theory builds on a pivotal voter framework and identifies novel channels of media influence. A separate strand of literature studies the impact of media on electoral competition and political accountability in settings where candidates' policy choices are endogenous (Besley and Prat, 2006; Chan and Suen, 2008, 2009; Strömberg, 2004; Snyder and Strömberg, 2010). In contrast, our focus is on the selection of candidates.

This paper also contributes to a strand of literature that studies the impact of media competition on voter welfare. On the one hand, media competition can improve political accountability by increasing the costs of media capture (Besley and Prat, 2006). Competition may also urge media to provide information that aligns better with readers' interests (Gentzkow and Shapiro, 2006; Chan and Suen, 2008). These insights support the view that media competition is welfare-improving. On the other hand, the literature also identifies channels through which media competition can deteriorate voter welfare. For instance, competition can drive profit-maximizing media to invest fewer resources in the provision of political news or topics of common interests (Cagé, 2017; Perego and Yuksel, 2018; Chen and Suen, 2018).[6] Our paper contributes to this debate by highlighting the role of media's commitment ability, which has been consistently overlooked, in determining the welfare impact of media competition.

Our theoretical framework combines the pivotal voter model (Palfrey and Rosenthal, 1985; Ledyard, 1984) with Poisson games (Myerson, 1998, 2000). We contribute to this literature by analytically deriving voting equilibria in large elections with a Poisson distributed population. We demonstrate that the asymptotic voting equilibria can be (i) tractably computed with respect to generic distributions of voters' ideological preferences, and (ii) expressed conveniently as functions of a few variables with clear economic interpretations.

In addition, our analysis of media outlets' information transmission behavior without commitment relates to the literature on multi-sender cheap talk games (Gilligan and Krehbiel, 1989; Krishna and Morgan, 2001a,b; Battaglini, 2002; Chan and Suen, 2009). This literature primarily focuses on the single-peaked payoff structure studied in Crawford and Sobel (1982) and aims at identifying conditions under which fully revealing equilibria exist. We depart by studying an alternative setting in which senders' payoffs are monotonic in the receiver's action, conditional on the realized state. We explicitly construct the most informative equilibria with an arbitrary number of senders.[7]

Finally, our equilibrium analysis with media having commitment power relates to the recent literature on Bayesian persuasion (Rayo and Segal, 2010; Kamenica and Gentzkow, 2011; Gentzkow and Kamenica, 2016, 2017). Our analysis with multiple media outlets is a special case of Gentzkow and Kamenica (2016). Because we focus on a specific payoff structure, however, we derive sharper results on how media bias shapes equilibrium information transmission behavior. Our analysis with a single outlet relates to several papers. Alonso and Câmara (2016) study how a politician can strategically design a policy experiment to persuade voters under super-majority rule and compulsory voting. We instead focus on elections under simple majority rule and assume voting to be voluntary and costly. Kolotilin et al. (2017) and Ginzburg (2019) consider persuasion problems in an environment similar to ours. We use their insights in deriving the media outlet's equilibrium behavior.

[6] See Gentzkow and Shapiro (2008) for a comprehensive review of this debate.

[7] Rubanov (2014) and Lu (2017) study robust informative equilibria in the presence of multiple senders under Crawford and Sobel type preferences. Ravid and Doron (2018) study a cheap talk game with one sender who has a "transparent motive"; the sender's preferences over the receiver's actions are state-independent. This nests a special case of our model, where voters only hear from an extremely biased media outlet.

3 The model

Candidates. We consider an election with two candidates, A and B. A commonly valued state k is drawn from a commonly known prior $F(\cdot)$ on $[-1,1]$. One can interpret k as the relative quality of candidate A. If $k=0$, the qualities of both candidates are identical. If instead $k>(<)0$, then candidate A has a higher (lower) quality than candidate B. Unless explicitly stated otherwise, k is assumed to be unobservable to voters.

Electorate. The electorate consists of a set of voters, N. Following Myerson (1998, 2000), we assume that the electorate size $|N|$ follows a Poisson distribution with mean $n>0$.[8] Each voter can choose an action from $\{A,B,O\}$, representing voting for candidate A, candidate B, and abstaining, respectively. Any voter $i \in N$ is characterized by a pair $(v_i, c_i) \in [-\delta,\delta]\times[0,C]$, with $\delta>1$ and $C>0$. v_i represents voter i's private ideological preference. c_i represents her private voting costs, which are incurred only if she casts a vote. We normalize voter i's payoff, excluding voting costs, to 0 if candidate B is elected and let voter i's payoff be $k+v_i$ if A is elected. Therefore, voter i prefers candidate A to B if and only if $k+v_i \geq 0$, and $|k+v_i|$ measures voter i's stake at the election. Formally, voter i's utility function is given by[9]

$$u(k,v_i,c_i,a_i,\Omega) = (k+v_i)\cdot \mathbb{1}_{\Omega=A} - c_i \cdot \mathbb{1}_{a_i \neq O} \tag{1}$$

where $\Omega \in \{A,B\}$ indicates the winning candidate, $a_i \in \{A,B,O\}$ is voter i's action, and $\mathbb{1}_E$ is an indicator function that equals 1 (0) if event E is true (false). Both v_i and c_i are independently and identically distributed across voters. v_i follows distribution $G(\cdot)$ on $[-\delta,\delta]$, and c_i is uniformly distributed on $[0,C]$. Both $G(\cdot)$ and C are common knowledge. We impose Assumption 1 in the remainder of this paper to obtain our main results.

Assumption 1. *[Symmetric election] Both $F(\cdot)$ and $G(\cdot)$ are symmetric distributions and admit positive density functions $f(\cdot)$ and $g(\cdot)$ on $[-1,1]$ and $[-\delta,\delta]$, respectively.*

Assumption 1 implies that a priori each candidate is equally likely to have a higher quality, and that the candidate with a superior quality is efficient in the sense of maximizing the electorate's ex-ante Utilitarian welfare.

[8] As argued by Myerson (1998), in large elections it seems reasonable to assume that voters do not know the exact size of the electorate.

[9] For our analysis to apply, we only need to assume that voters' utility function is monotonically increasing in both k and v, and is additively separable with respect to both arguments. This setup nests the spatial voting model with commonly valued valence (Enelow and Hinich, 1984), which is widely applied in studies concerning media's electoral influences (e.g., Chiang and Knight (2011) and Durante and Knight (2012)).

Media outlets and media bias. There is a finite set of media outlet(s), M. Each outlet $m \in M$ is precisely informed of k and communicates to voters by sending a public message $s_m \in S$. S is a common and generic message space and will be explicitly defined in subsequent sections. Any media outlet's utility function has a similar structure as voters', except that it cannot vote:

$$V(k, \chi_m, \Omega) = (k + \chi_m) \cdot \mathbb{1}_{\Omega=A} \tag{2}$$

This implies that outlet m strictly prefers candidate A (B) if and only if $k > (<) - \chi_m$. If $\chi_m = 0$, m is said to be *unbiased* since it always prefers the candidate with a higher quality. If $\chi_m > (<)0$, m is said to be *A(B)-biased* as it may prefer candidate A(B) to be elected even if it has a lower quality. For this reason, we refer to χ_m as the *media bias*. We distinguish between two types of communication protocols, cheap talk and Bayesian persuasion. In the former case media outlets cannot commit to any reporting strategy before observing the realized state k, while in the latter case they can. We primarily focus on the cheap talk in deriving the comparative statics of media influence on elections, and compare equilibrium information transmission behavior under cheap talk and Bayesian persuasion in Section 6.

Timing and equilibrium concept. The timing of the game is as follows:

1. (Bayesian persuasion only) Each outlet $m \in M$ commits to a reporting strategy σ_m : $[-1,1] \mapsto \Delta(S)$.

2. Nature draws the realizations of k, N and the profile of voter types $\{(v_i, c_i)\}_{i \in N}$.

3. Observing k, outlets $m \in M$ simultaneously send public messages $s_m \in S$.

4. Observing the message profile $\{s_m\}_{m \in M}$, voters simultaneously make their voting decisions (vote for A, for B, or abstain).

5. The winning candidate is determined by simple majority rule, with ties broken by a fair coin toss. All payoffs then realize.

A media outlet's reporting strategy maps the realized state to a probability distribution on the message space, $\sigma_m : [-1,1] \mapsto \Delta(S)$. A voter's voting strategy maps the observed message profile and the realized private type to a probability distribution of possible actions $q : S^{|M|} \times [-\delta, \delta] \times [0, C] \mapsto \Delta(\{A, B, O\})$.[10] Because this is a dynamic game of incomplete information, we derive the Perfect Bayesian Equilibrium (PBE). In a PBE, voters' strategies

[10] Under Bayesian persuasion, the voting strategy is also a function of media outlets' strategy profile.

are best responses to their posterior beliefs conditional on the observed message profile. Media outlets' strategies are best responses to voters' strategies. Voters' posterior beliefs are formed by Bayes' rule whenever possible. Without loss of generality, we focus on type-symmetric PBE where voters of the same type (v_i, c_i) adopt the same voting strategy in equilibrium.[11]

Although our analysis applies under much wider scenarios, we will use Example 1 below to graphically illustrate our main results.

Example 1. $n = 200$, $k \in U[-1,1]$, $v_i \sim U[-1.6,1.6]$ *and* $c_i \sim U[0,0.2]$ *for all* $i \in N$.

4 Preliminaries: Voting equilibria in large elections

In this section we derive voting equilibria in large elections and thereby also illustrate two general insights from our pivotal voter framework: (i) public information about k not only allows voters to assess which candidate is more appealing, but also how close the election is likely to be, and (ii) through the strategic provision of information media may systematically manipulate election outcomes, even when voters are fully rational. In subsequent sections we build on the formal results reported here to study the comparative statics of the impact of media bias on elections. Readers solely interested in the latter comparative statics could just skim through this section.

In Section 4.1, we derive the equilibrium aggregate voting behavior and election outcomes in large elections (i.e., $n \to \infty$) assuming that k is common knowledge. This serves as a basis for Section 4.2, where we assume k is unobservable to voters and quantify the extent to which these aggregate outcomes can be systematically influenced by mass media. Detailed derivations and proofs are relegated to Online Appendices A and B, respectively.

4.1 Voting equilibria when k is common knowledge

If k is common knowledge, it is clear from voters' utility function (1) that any voter i with $v_i > (<) -k$ strictly prefers candidate A (B) to be elected, and will vote for candidate A (B) if she votes. Hence, the only strategic decision is whether or not to cast a vote. To make this decision a voter compares the expected benefits to the costs of casting a vote. The optimal turnout strategy must then take a cutoff form; vote if and only if the voting costs

[11] Myerson (1998) (page 391) argues that all equilibria are type-symmetric under population uncertainty.

c_i are below the expected benefits (Ledyard, 1984; Palfrey and Rosenthal, 1985; Myerson, 2000). The expected benefits are the product of (i) the individual stake at the election, $|k+v_i|$, and (ii) the probability of casting a pivotal vote. Hence, voter i votes only if:

$$c_i \leq |k+v_i| \cdot PivA(k,n), \text{ if } v_i > -k$$
$$c_i \leq |k+v_i| \cdot PivB(k,n), \text{ if } v_i < -k$$

where $PivA(k,n)$ and $PivB(k,n)$ denote the probabilities that a single vote for candidate A respectively B is pivotal, conditional on k and n. These pivotal probabilities are endogenously determined in equilibrium. Consistent with common intuition, in equilibrium both pivotal probabilities converge to zero as $n \to \infty$. For sufficiently large n and any given k and v_i, the expected benefits of voting must then decrease in n and lie below C. The voter's turnout probability then equals $\frac{|k+v_i| \cdot PivA(k,n)}{C}$ for $v_i > -k$ and $\frac{|k+v_i| \cdot PivB(k,n)}{C}$ for $v_i < -k$, since $c_i \sim U[0,C]$.

Let $q_A(k,n)$ and $q_B(k,n)$ be the probability that a randomly sampled voter votes for A and B, respectively. $q_A(k,n)$ and $q_B(k,n)$ can be obtained by integrating the turnout probabilities of "A-supporters" (voters with $v_i > -k$) and "B-supporters" (voters with $v_i < -k$), respectively. For sufficiently large n we have

$$q_A(k,n) = \int_{-k}^{\delta} \frac{|k+v| \cdot PivA(k,n)}{C} dG(v) \equiv \frac{\alpha(k)}{C} \cdot PivA(k,n) \tag{3}$$
$$q_B(k,n) = \int_{-\delta}^{-k} \frac{|k+v| \cdot PivB(k,n)}{C} dG(v) \equiv \frac{\beta(k)}{C} \cdot PivB(k,n) \tag{4}$$

where $\alpha(k) \equiv \int_{-k}^{\delta} |k+v| dG(v) = k + \int_{-\delta}^{-k} G(k)dk$ and $\beta(k) \equiv \int_{-\delta}^{-k} |k+v| dG(v) = \int_{-\delta}^{-k} G(k)dk$. Both $\alpha(k)$ and $\beta(k)$ have straightforward economic interpretations. $\alpha(k)$ represents the *expected stake* of A-supporters and $\beta(k)$ represents the *expected stake* of B-supporters. It is clear from these expressions that $\alpha(k)$ increases in k whereas $\beta(k)$ decreases in k. The influence of a rise in k on $\alpha(k)$ and $\beta(k)$ can be decomposed into three effects. First, it convinces the marginal B-supporters to switch to A. Second, it strengthens the expected stakes of the infra-marginal A-supporters. Third, it weakens the expected stakes of the infra-marginal B-supporters. Under Assumption 1 it holds that $\alpha(k) - \beta(k) = k$. Hence, k directly measures the difference in expected stakes between A- and B-supporters.

Because $\delta > 1$ and $G(\cdot)$ has full support on $[-\delta, \delta]$ (cf. Assumption 1), it holds that $\beta(1) = \int_{-\delta}^{-1} G(v)dv > 0$. Define $\nu \equiv \beta(1)$, which equals the expected stakes of *extreme*

B-supporters with $v_i < -1$, in the state most favorable to candidate A (i.e., $k = 1$). By symmetry of distribution $G(\cdot)$, it holds that $\alpha(-1) = \beta(1) = \nu$. Therefore, ν reflects the minimal aggregate expected stakes of ideologically extreme voters and can be interpreted as a measure for the degree of *polarization*. As will become clear below, the degree of polarization ν plays an important role in determining voting equilibria in large elections.

As elaborated in our Online Appendix A, both $PivA(k,n)$ and $PivB(k,n)$ are functions of $q_A(k,n)$ and $q_B(k,n)$. Therefore, the system of equations (3) and (4) forms a self-mapping and a voting equilibrium can be characterized by its fixed point, denoted by $(q_A^*(k,n), q_B^*(k,n))$. In our Online Appendix A we show that $q_A^*(k,n)$ and $q_B^*(k,n)$ in large elections can be analytically approximated as follows[12]

$$q_A^*(k,n) \approx \begin{cases} \left(\frac{\alpha(0)}{2\sqrt{\pi C}}\right)^{\frac{2}{3}} \cdot \frac{1}{\sqrt[3]{n}}, & \text{if } k = 0 \\ \left(\frac{\mu(k)}{1-\mu(k)}\right)^2 \cdot \frac{\ln n}{n}, & \text{if } k \neq 0 \end{cases} \tag{5}$$

$$q_B^*(k,n) \approx \begin{cases} \left(\frac{\alpha(0)}{2\sqrt{\pi C}}\right)^{\frac{2}{3}} \cdot \frac{1}{\sqrt[3]{n}}, & \text{if } k = 0 \\ \left(\frac{1}{1-\mu(k)}\right)^2 \cdot \frac{\ln n}{n}, & \text{if } k \neq 0 \end{cases} \tag{6}$$

where $\mu(k) \equiv \sqrt[3]{\frac{\alpha(k)}{\beta(k)}}$ equals the cubic root of the expected stakes ratio. Let $VS(k,n) \equiv \frac{q_A^*(k,n)}{q_A^*(k,n)+q_B^*(k,n)}$ and $T(k,n) \equiv q_A^*(k,n) + q_B^*(k,n)$ be the equilibrium expected vote share of candidate A and the expected voter turnout, conditional on k and n, respectively. We also use approximations for (5) and (6) to compute candidate A's equilibrium winning probability, denoted by $\pi(k,n)$. Finally, let $W(k,n)$ be the expected utility of a randomly selected voter in equilibrium. Theorem 1 characterizes equilibrium voting behavior, election outcomes and voter welfare in large elections.

Theorem 1. *Suppose Assumption 1 holds. Then for all $k \in [-1,1]$ it holds that:*

1. $VS(k,n) \approx \frac{\mu^2(k)}{1+\mu^2(k)}$.

2. $\lim\limits_{n\to\infty} \pi(k,n) = \begin{cases} 1, & \textit{if } k > 0 \\ \frac{1}{2}, & \textit{if } k = 0 \\ 0, & \textit{if } k < 0 \end{cases}$.

[12] Throughout the paper, the expression $x_n \approx y_n$ denotes $\lim\limits_{n\to\infty} \frac{x_n}{y_n} = 1$.

3. $T(k,n) \approx \begin{cases} 2\left(\frac{\alpha(0)}{2\sqrt{\pi}C}\right)^{\frac{2}{3}} \frac{1}{\sqrt[3]{n}}, & \text{if } k=0 \\ \gamma(k)\frac{\ln n}{n}, & \text{if } k \neq 0 \end{cases}$, *where* $\gamma(k) \equiv \frac{1+\mu^2(k)}{(1-\mu(k))^2}$ *for* $k \neq 0$.

4. $\lim_{n\to\infty} W(k,n) = \max\{k,0\}$.

Proof. See Online Appendix B.1. □

Theorem 1 is graphically illustrated by Figure 1a to 1d, under the model parameters from Example 1. Figure 1a (black solid line) depicts candidate A's equilibrium vote share as a function of k. As discussed above, with a higher k candidate A is able to (i) convince marginal voters to switch to her, and (ii) increase (reduce) the expected stakes of her own (opponent's) supporters. For these reasons, candidate A's expected vote share increases in k (Theorem 1.1).

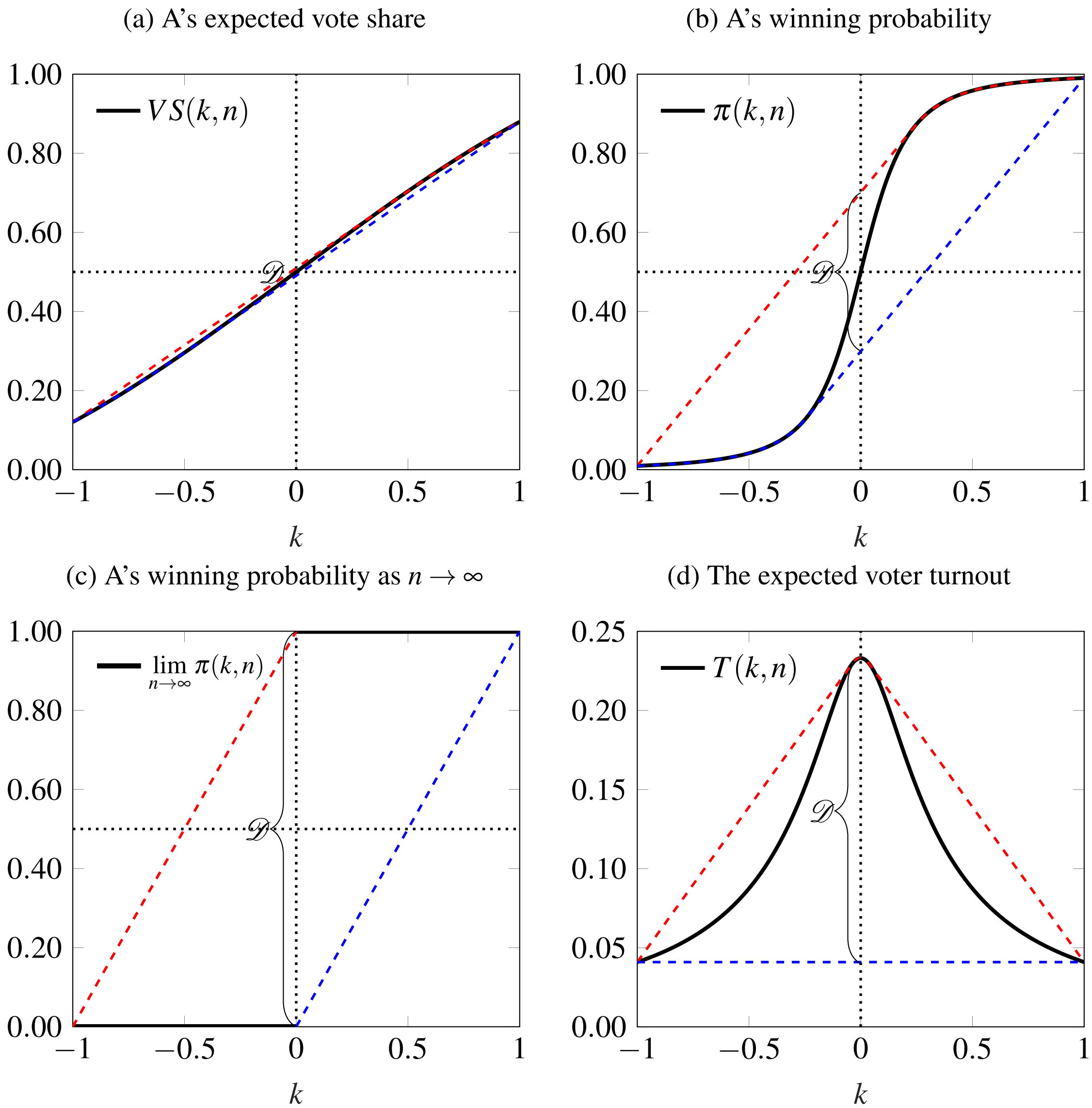
Figure 1b (black solid line) depicts candidate A's equilibrium winning probability as a function of k. It is evident graphically that A's equilibrium winning probability has an "*S-shape*" property; it increases in k convexly for $k < 0$ and concavely for $k > 0$.[13] This implies that the election outcome responds more sensitively to marginal variations in k when the quality difference $|k|$ is closer to 0. Under simple majority rule, candidate A only requires a vote share exceeding 50% to assure a victory. Therefore, thanks to the law of large numbers, A's winning probability converges to a step function as $n \to \infty$: $\lim_{n\to\infty} \pi(k,n) = 1(0)$ if and only if $k > (<)0$ (Theorem 1.2; cf. Figure 1c).

Figure 1d (black solid line) depicts the expected voter turnout in equilibrium, $T(k,n)$, as a function of k. $T(k,n)$ is symmetric around $k = 0$ and decreases in the quality difference $|k|$ (Theorem 1.3). If $k = 0$, the election is expected to end up in a close race since the two candidates are expected to get equal vote shares. As a result, the chances of being pivotal are highest and voters have the strongest incentive to cast costly votes. Voter turnout is therefore highest. If instead $|k|$ is large, the election is expected to end in a landslide; the quality-superior candidate is expected to get a substantially larger vote share than her opponent. In this case the chances of being pivotal are low and, as a result, voter turnout will be low. The negative relationship between the closeness of an election and turnout is known as the "competition effect" (Levine and Palfrey, 2007).[14] Since the quality difference $|k|$ is negatively associated with the closeness of the election, a higher $|k|$ lowers voter turnout.

[13] Lemma 9 in Online Appendix A.3 provides analytical approximations for the curvature of $\pi(k,n)$ as $n \to \infty$. In Section 5 we provide a sufficient condition (Assumption 2) under which the S-shape property of $\pi(k,n)$ holds in large elections.

[14] In fact, our results are more general than the original concept of the competition effect proposed by Palfrey and Rosenthal (1985) and Levine and Palfrey (2007). In their models the individual stakes are fixed

Figure 1: Equilibrium voting behavior and election outcomes as functions of k



Note: In Panel (a), $VS(k,n)$ denotes the equilibrium vote share of candidate A conditional on k, n. In Panel (b) and (c), $\pi(k,n)$ denotes the winning probability of candidate A, conditional on k, n. In Panel (d), $T(k,n)$ denotes the expected voter turnout conditional on k,n. The red (blue) curves in all panels are the upper-concave (lower-convex) envelopes of the depicted outcome variable. The difference between the upper-concave and lower-convex envelopes, evaluated at the prior mean $k=0$, is denoted by $\mathscr{D}$. $\mathscr{D}$ provides a robust bound of the extent to which an outcome variable can be systematically influenced by manipulating public information. This is discussed in Section 4.2. Model parameters are taken from Example 1.

Lemma 10 in Online Appendix A.3 precisely characterizes the curvature of $T(k,n)$ as $n \to \infty$. It shows that the asymptotic voter turnout is generically a *convex* function of $|k|$ if the degree of polarization ν exceeds a certain bound.[15] This convexity indicates that the competition effect exhibits a diminishing marginal impact; a marginal increase in quality difference has a larger negative impact on turnout when the quality difference is smaller. As explained below, this convexity property plays an important role in deriving unambiguous comparative static predictions on how biased media affect voter turnout.

Finally, Theorem 1.4 follows from the fact that with k common knowledge, the election outcome is *asymptotically efficient*; the quality-superior candidate is elected with probability approaching 1 as $n \to \infty$. Recall that k reflects the relative quality of candidate A, i.e., the quality of candidate B is normalized to zero. Moreover, Theorem 1.3 suggests that in large elections voter turnout vanishes to zero and thus so do the expected voting costs. Together, this explains the expression for the asymptotic welfare in Theorem 1.4.

4.2 Robust bounds for media influence

In this section, we assume that k is unobservable to voters and propose two robust bounds to gauge the extent to which voter behavior and the election outcomes can be systematically influenced by mass media. We show that, by manipulating public information, mass media outlets can substantially influence the ex-ante election outcome and voter turnout, even in cases where systematically affecting party vote shares is nearly impossible.

Following the information design literature (e.g., Kamenica and Gentzkow (2011) and Taneva (2019)), we define an *information structure* as a pair $\Xi = (X, \sigma)$, which consists of a signal realization space X and a signal distribution $\sigma : [-1,1] \mapsto \Delta(X)$.[16] In our model, $X \equiv S^{|M|}$ is the set of all possible message profiles jointly sent by all media outlets, and $\sigma \equiv \prod_{m \in M} \sigma_m$ is a profile of reporting strategies that maps each state k to a distribution of message profiles. Because voters' utility function is linear in k (cf. (1)), any public message profile $x \in X$ affects voters' behavior only through its impact on $E[k|x, \Xi]$, the expectation of k conditional on x under information structure Ξ. Consequently, voters behave

and homogeneous; each voter receives a fixed reward if her preferred candidate is elected. So the closeness of an election depends only on the relative number of voters supporting each candidate. In our model the closeness of an election is not only determined by this extensive margin, but also by the average stakes of voters supporting each candidate (an intensive margin), similar to Krishna and Morgan (2011, 2015).

[15] In Figure 1d, voter turnout decreases concavely in $|k|$ for $|k|$ close to 0. This is an artifact of the finite electorate size ($n = 200$). As $n \to \infty$, the concave region vanishes.

[16] In the information design literature, a signal distribution is a function mapping from the state space to the set of distributions on the message space. In our model, the state space is the domain of k, i.e., $[-1,1]$.

as if $E[k|x,\Xi]$ is common knowledge and the voting equilibria follow immediately from Theorem 1 by replacing k with $E[k|x,\Xi]$. Therefore, it is convenient to characterize an information structure Ξ by the distribution of posterior means, denoted by τ_Ξ, it induces. Let Γ_F denote the set of all distributions of posterior means that can be induced by some information structure Ξ, given prior $F(\cdot)$.

Let $\xi(k,n)$ denote the outcome variable of interest, where $\xi \in \{VS, \pi, T\}$. Given any prior $F(\cdot)$, the difference between the upper and lower bounds of expectations of $\xi(k,n)$ precisely measures the extent to which $\xi(k,n)$ can be systematically influenced by manipulating public information, under prior $F(\cdot)$. In general, these bounds depend strongly on the specific functional form of $F(\cdot)$, which is often unknown. We thus consider the following two alternative bounds that do not rely on the exact specification of $F(\cdot)$ as long as Assumption 1 holds:

$$\mathscr{D}(\xi,\Gamma) \equiv \sup_{\tau\in\Gamma} E_\tau[\xi(k,n)] - \inf_{\tau\in\Gamma} E_\tau[\xi(k,n)] \tag{7}$$

$$\widetilde{\mathscr{D}}(\xi,\Gamma) \equiv \frac{\mathscr{D}(\xi,\Gamma)}{\max_{k\in[-1,1]}\xi(k,n) - \min_{k\in[-1,1]}\xi(k,n)} \tag{8}$$

where Γ denotes the set of all distributions of posterior means that average back to 0, the prior mean under $F(\cdot)$.[17]

In words, $\sup_{\tau\in\Gamma} E_\tau[\xi(k,n)]$ is the precise upper bound of the ex-ante expectation of $\xi(k,n)$ that can be induced by a distribution of posterior means that average back to 0, and $\inf_{\tau\in\Gamma} E_\tau[\xi(k,n)]$ the lower bound. Hence, $\mathscr{D}(\xi,\Gamma)$ provides a robust bound to gauge the extent to which the ex-ante expectation of $\xi(k,n)$ can be influenced by manipulating voters' information. $\widetilde{\mathscr{D}}(\xi,\Gamma)$ measures the extent to which $\xi(k,n)$ can be systematically influenced, relative to the maximal range it can possibly vary. It rescales $\mathscr{D}(\xi,\Gamma)$ to a score ranging between 0 to 1.

We can geometrically derive and illustrate $\mathscr{D}(\xi,\Gamma)$ using convex and concave envelopes of function $\xi(k,n)$, as demonstrated in Figures 1a to 1d. The red dashed curves in these figures depict $\xi_+(k,n)$, the upper concave envelope of $\xi(k,n)$, while the blue dashed ones depict $\xi_-(k,n)$, the lower convex envelope of $\xi(k,n)$.[18] By construction, $\mathscr{D}(\xi,\Gamma)$ equals

[17] Note that Γ_F is a proper subset of Γ, because Γ contains distributions of posterior expectations that cannot be induced by any information structure given prior $F(\cdot)$.

[18] Formally, let $co(\xi(\cdot,n))$ denote the convex hull of the graph of $\xi(k,n)$: $\{(k,\xi(k,n))\}_{k\in[-1,1]}$. Then $\xi_+(k,n) \equiv \sup\{z|(k,z) \in co(\xi(\cdot,n))\}$ is the *upper-concave envelope* of function $\xi(k,n)$, and $\xi_-(k,n) \equiv \inf\{z|(k,z) \in co(\xi(\cdot,n))\}$ is the *lower-convex envelope* of function $\xi(k,n)$.

the vertical distance between the blue and the red curve evaluated at the prior mean $k = 0$; in these figures this distance is denoted by $\mathscr{D}$. The other bound $\widetilde{\mathscr{D}}(\xi,\Gamma)$ can be obtained through dividing $\mathscr{D}(\xi,\Gamma)$ by the range of $\xi(k,n)$, which can also be straightforwardly identified from Figures 1a to 1d.

It is evident from Figure 1a that, under Example 1, both $\mathscr{D}(VS,\Gamma)$ and $\widetilde{\mathscr{D}}(VS,\Gamma)$ are almost negligible in magnitude: $VS_+(0,n) < 0.505$ and $VS_-(0,n) > 0.495$, such that $\mathscr{D}(VS,\Gamma) < 0.01$ and $\widetilde{\mathscr{D}}(VS,\Gamma) < 0.0125$. In other words, no matter how one manipulates voters' information, *a priori* the expected vote share for any candidate can be systematically shifted by no more than 1%-point, which corresponds to 1.25% of the range of $VS(k,n)$. This is because, under the model parameters from Example 1, $VS(k,n)$ is almost linear in k; variations in the expected vote share are almost proportional to variations in voters' posterior expectations of k. The latter, however, cannot be systematically affected because voters are Bayesian.

The situation is entirely different for the election outcome and voter turnout, as depicted in Figure 1b and 1d, respectively. In Figure 1b, $\pi_+(0,n) = 0.7$ and $\pi_-(0,n) = 0.3$, so that $\mathscr{D}(\pi,\Gamma) = 0.4$ and $\widetilde{\mathscr{D}}(\pi,\Gamma) = 0.407$. Therefore, manipulating public information can systematically influence A's winning probability by up to 40%-points, which corresponds to around 40.7% of the range of $\pi(k,n)$. In Figure 1d, $T_+(0,n) = 0.23$ and $T_-(0,n) = 0.04$, so that $\mathscr{D}(T,\Gamma) = 0.19$ and $\widetilde{\mathscr{D}}(T,\Gamma) = 1$. Manipulating public information may thus systematically affect the expected turnout rate by up to 18%-points, which corresponds to 100% of the range of $T(k,n)$. Taken together, the aforementioned observations suggest that, by manipulating public information, one can substantially influence the ex-ante election outcome and voter turnout, even when systematically influencing party vote shares is virtually impossible. This is a consequence of the fact that both $\pi(k,n)$ and $T(k,n)$ are much more sensitive to shifts of k when k is close to 0 than when k is away from 0. This gives media the possibilities to strategically disclose information and shifting voters' posterior expectations of k in regions with different sensitivities, and thereby systematically influence election outcomes and voter turnout.

Going beyond Example 1, Theorem 2 characterizes asymptotic properties of $\mathscr{D}(\xi,\Gamma)$ and $\widetilde{\mathscr{D}}(\xi,\Gamma)$ for symmetric elections (i.e., under Assumption 1).

Theorem 2. *Suppose Assumption 1 holds, then*

1. $\lim_{\nu\to\infty}\lim_{n\to\infty}\mathscr{D}(VS,\Gamma) = \lim_{\nu\to\infty}\lim_{n\to\infty}\widetilde{\mathscr{D}}(VS,\Gamma) = 0.$

2. $\lim_{n\to\infty}\mathscr{D}(\pi,\Gamma) = \lim_{n\to\infty}\widetilde{\mathscr{D}}(\pi,\Gamma) = 1.$

3. $\lim_{n\to\infty} \mathscr{D}(T,\Gamma) = 0$ *and* $\lim_{n\to\infty} \widetilde{\mathscr{D}}(T,\Gamma) = 1$

Proof. See Online Appendix B.2. □

Theorem 2 has three implications for large and symmetric elections. First, the extent to which the expected party vote shares can be systematically affected depends on the degree of polarization of the electorate, ν; with a strongly polarized electorate, the expected party vote share can hardly be systematically influenced according to both $\mathscr{D}(VS,\Gamma)$ and $\widetilde{\mathscr{D}}(VS,\Gamma)$. Second, in sharp contrast with party vote shares, the election outcome can be arbitrarily affected by strategically providing public information. Third, though manipulating public information can hardly affect the expected voter turnout in absolute levels (according to $\mathscr{D}(T,\Gamma)$) in large elections, its influence relative to the maximal range of turnout becomes arbitrarily large (according to $\widetilde{\mathscr{D}}(T,\Gamma)$).

We conclude this section by demonstrating how a media outlet can arbitrarily manipulate the election outcome in large and symmetric elections. As illustrated in Figure 1c, when $n \to \infty$, $\mathscr{D}(\pi,\Gamma)$ converges to one and $\pi(k,n)$ converges to a step function with threshold at the prior mean $k = 0$, where voters are (in aggregate) indifferent between the two candidates. Therefore, the election outcome is extremely sensitive (insensitive) to shifts in voters' beliefs when $k = (\neq)0$, as $n \to \infty$. Exploiting this feature, a media outlet can guarantee an arbitrarily large winning probability of candidate A by endorsing A if $k \in [-1+\varepsilon, 1]$ and endorsing B if $k \in [-1, -1+\varepsilon)$, with $\varepsilon > 0$ arbitrarily close to zero. By doing so, an endorsement for A provides very weak evidence for the superiority of A and only slightly increases the vote share of A. Yet it is sufficient to tip the balance and assure victory of A as $n \to \infty$. Though an endorsement for B provides extraordinarily strong evidence for the superiority of B and almost assures a certain victory of B, it is sent with vanishing probability as $\varepsilon \to 0$. Therefore, as $\varepsilon \to 0$, an endorsement for A is sent almost surely ex-ante, which in turns guarantees a certain victory of A as $n \to \infty$.

5 The influence of biased media on elections

This section studies the influence of biased media on voting behavior and election outcomes under a cutoff information environment explained in Section 5.1. Section 5.2 explores the electoral impacts of media bias, in an environment with a single media outlet (i.e., $|M| = 1$). In Section 5.3 we consider an increase in $|M|$ from 1 to 2 to explore the electoral impacts of introducing a second media outlet.

Before proceeding, we make two additional technical assumptions 2 and 3 on the distribution of voters' ideological preference $G(\cdot)$ to simplify our comparative statics analyses. Assumption 2 guarantees, through Lemma 9 in Online Appendix A.3, that in large elections candidate A's equilibrium winning probability, $\pi(k,n)$, is an S-shape function of k as demonstrated by Figure 1b. Despite its complexity, Assumption 2 holds for a wide range of distribution functions $G(\cdot)$.[19] Assumption 3 specifies a strictly positive lower bound for the degree of polarization in the electorate. This assumption guarantees, through Lemma 10 in Online Appendix A.3, that in large elections the expected voter turnout will be a convex function of the quality difference between candidates, $|k|$. This assumption also seems plausible, for example, under the current political landscape in the U.S., where the degree of polarization has increased sharply in recent decades (Pew Research Center, 2017; Martin and Yurukoglu, 2017; Gentzkow et al., 2019). As will be made clear below, Assumption 2 and 3 enables us to derive unambiguous comparative static impacts of biased media on election outcomes and voter turnout, respectively. Throughout Sections 5 and 6, we will maintain Assumptions 1 to 3. Importantly, our illustrating Example 1 satisfies all Assumptions 1 to 3.

Assumption 2. *[S-shape of $\pi(k,n)$]* $\left(\frac{1}{2}\frac{\mu-1}{\mu+1}\frac{\mu'}{\mu}+\frac{\mu\frac{\alpha'}{\alpha}-\frac{\beta'}{\beta}}{\mu-1}\right)\left(\frac{1}{\sqrt[3]{\alpha(k)}-\sqrt[3]{\beta(k)}}\right)^3$ *decreases in k on* $(0,1]$.[20]

Assumption 3. *[Moderate polarization]* $\nu \equiv \int_{-\delta}^{-1} G(v)dv > \nu^*$*, where ν^* is implicitly characterized by (A.27) in Online Appendix A.3.*[21]

5.1 The cutoff information environment

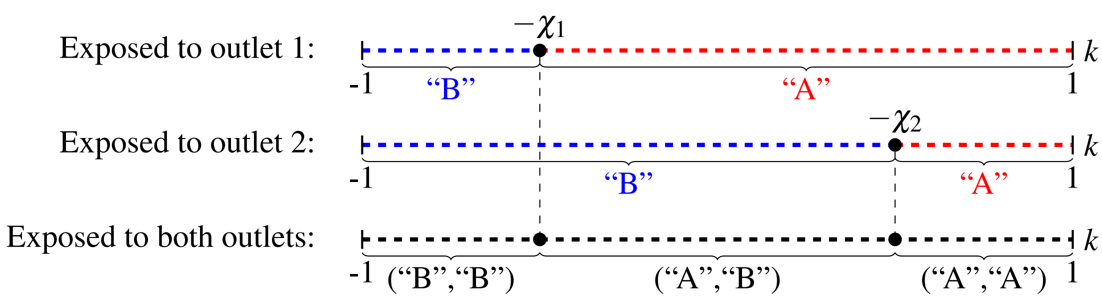
By construction (cf. (2)), the preference of media outlet $m \in M$ takes a cutoff form: it prefers candidate A if and only if $k > -\chi_m$. It is thus natural to conjecture that media outlets use a simple cutoff endorsement strategy with a binary message space $S = \{\text{“A”}, \text{“B”}\}$:

$$\sigma_m(k) = \begin{cases} \text{“A”}, & \text{if } k > -\chi_m \\ \text{“B”}, & \text{if } k \leq -\chi_m \end{cases} \tag{9}$$

[19] We have not been able to find any $G(\cdot)$ that violates Assumption 2, when Assumptions 1 and 3 are met.
[20] In this expression α', β' and μ' denote the first order derivatives of $\alpha(k)$, $\beta(k)$ and $\mu(k)$, respectively.
[21] The numerical value of ν^* lies between 0.044 and 0.045. In our illustrating Example 1, $G(\cdot)$ is a uniform distribution on $[-\delta, \delta]$ with $\delta = 1.6$, and it holds that $\nu = \int_{-1.6}^{-1} G(v)dv = 0.056 > \nu^*$.

That is, each outlet $m \in M$ sends message "A" ("B") if $k > (\leq) -\chi_m$.[22] Message "A" ("B") can be interpreted as an *endorsement* for candidate A (B). Let $\{\chi^{(m)}\}_{m\in M}$ be a descending permutation of media biases: $\chi^{(1)} \geq \chi^{(2)} \geq \cdots \geq \chi^{(|M|)}$. Under strategy (9), the $|M|$ media outlets jointly induce a "cutoff information environment", which partitions the state space into adjacent and disjoint intervals with cutoffs $-\chi^{(1)}, -\chi^{(2)}, ..., -\chi^{(|M|)}$. Notation wise, we denote the information partition induced by the $|M|$ outlets by the set of cutoffs points: $\mathscr{P}_C(\chi_1, \cdots, \chi_{|M|}) = \{-\chi_1, \cdots, -\chi_{|M|}\}$.[23] Figure 2 illustrates strategy (9) and the information partition it induces in an example where $|M| = 2$, $\chi_1 = 0.5$ and $\chi_2 = -0.5$.

Figure 2: The Cutoff Endorsement Strategy and Information Environment



Note: The figure illustrates the case in which there are two media outlets with biases $\chi_1 = 0.5$ and $\chi_2 = -0.5$, respectively. The red (blue) segments denote states where media outlet $m \in \{1,2\}$ will send message "A" ("B"). The pair (s_1, s_2) in the bottom line denotes the message combination the two media outlets send in each realized state. Since voters read from both outlets, voters' posterior knowledge is a coarse interval partition of the state space.

We focus on this cutoff information environment for two reasons. First, as shown in Section 6 this environment endogenously arises as the most informative coordination-proof equilibrium outcome if media outlets communicate to the electorate via cheap talk. Second, this environment is widely applied in both theoretical and empirical papers studying the electoral impacts of mass media (Chan and Suen, 2008; Chiang and Knight, 2011; Durante and Knight, 2012; Puglisi and Snyder, 2015a).

5.2 The influence of media bias

Let $|M| = 1$ and χ be the bias of this single media outlet. Without loss of generality we assume $\chi \in [0,1]$, i.e., the media is either unbiased ($\chi = 0$) or A-biased ($\chi > 0$). We study

[22] Under Assumption 1, $k = -\chi_m$ is a zero probability event and thus ties can be broken arbitrarily.

[23] The subscript C denotes a cutoff partition made up of adjacent and disjoint intervals, and it distinguishes this particular category from other partitions discussed in Section 6 below.

both the *interim* and the *ex-ante* impacts of increased media bias on the voting equilibrium; the former impact is conditional on the realized message, while the latter is unconditional. Denote voters' rational posterior expectations of k as $k_A(\chi) \equiv E_F[k|k > -\chi]$ conditional on message "A", and $k_B(\chi) \equiv E_F[k|k \leq -\chi]$ conditional on message "B". Since voters' utility function is linear in k, media bias χ affects voting equilibria only through its impact on $k_A(\chi)$, $k_B(\chi)$ and the likelihood of sending each message. In this regard, the following three effects are crucial for understanding the electoral impact of media bias.[24]

Effect I A more biased media outlet (larger χ) is *a priori* more likely to send message "A" (with probability $1-F(-\chi)$) and induce posterior expectation $k_A(\chi)$.

Effect II Both $k_A(\chi)$ and $|k_A(\chi)|$ decrease in χ; message "A" becomes less credible and the election is more likely to end up in a close race under message "A", as χ increases.

Effect III $k_B(\chi)$ decreases while $|k_B(\chi)|$ increases in χ; message "B" becomes more credible and the election is less likely to end up in a close race under message "B", as χ increases.

Proposition 1 summarizes the interim and ex-ante impacts of increased media bias on the expected party vote shares and the election outcome.

Proposition 1. *Suppose $|M| = 1$. The following properties hold:*

1. $\forall \chi' > \chi \geq 0$, *for both $s \in \{A, B\}$ and sufficiently large n it holds that $VS(k_s(\chi'), n) > VS(k_s(\chi), n)$ and $\pi(k_s(\chi'), n) > \pi(k_s(\chi), n)$.*
2. *For a sufficiently large n, $E_{\mathscr{P}_{C(\chi)}}[\pi(k,n)]$ varies non-monotonically with χ on $[0,1]$:*[25]
   (a) $E_{\mathscr{P}_{C(\chi)}}[\pi(k,n)] = \frac{1}{2}$ *for* $\chi \in \{0,1\}$, *and* $E_{\mathscr{P}_{C(\chi)}}[\pi(k,n)] > \frac{1}{2}$ *for* $\chi \in (0,1)$.
   (b) $E_{\mathscr{P}_{C(\chi)}}[\pi(k,n)]$ *increases (decreases) in χ if χ is sufficiently close to* 0 (1).

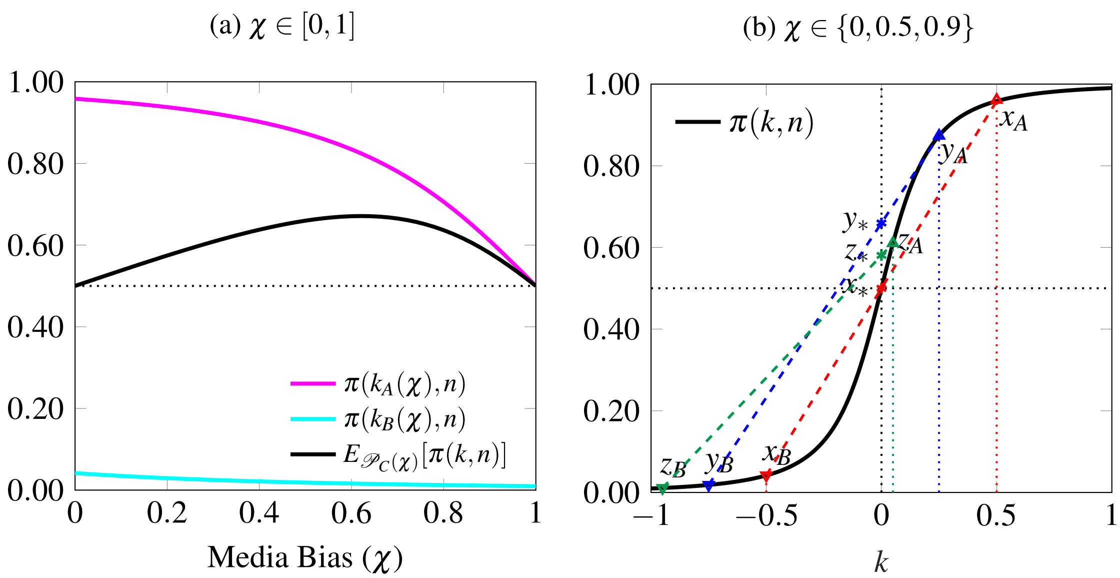
Proof. See Online Appendix C.1. □

Figure 3a illustrates both the interim and the ex-ante impacts of increasing media bias on the election outcome, based on the model parameters from Example 1. As Effect II and III indicate, both $k_A(\chi)$ and $k_B(\chi)$ decrease in χ. Therefore, both candidate A's expected vote

[24] These effects are formally stated and proved in Lemma 12 of Online Appendix C.1.

[25] Throughout the paper, we use notation $E_{\mathscr{P}}[z(k)]$ to denote the expectation of any real valued function $z(k)$ under the partition information environment characterized by $\mathscr{P}$.

Figure 3: The interim and ex-ante impact of media bias on the election outcome



Note: Panel (a) depicts the interim and the ex-ante impacts of media bias for all $\chi \in [0,1]$. Panel (b) geometrically demonstrates these impacts for $\chi \in \{0,0.5,0.9\}$. The function $\pi(k,n)$ depicted in Panel (b) is identical to $\pi(k,n)$ in Figure 1b. In Panel (b), the selected levels of media bias increase from red to blue to green: $\chi = 0$ (unbiased, nodes x), $\chi = 0.5$ (weakly biased, nodes y), and $\chi = 0.9$ (strongly biased, nodes z). Nodes w_A, w_B and w_*, where $w \in \{x,y,z\}$, represent candidate A's winning probability conditional on message "A", message "B" and unconditionally, respectively. The dashed line segments represent the sets of all convex combinations of w_A and w_B, for $w \in \{x,y,z\}$. Model parameters are taken from Example 1.

share and winning probability decrease in media bias after either message (Proposition 1.1), since $VS_A(k,n)$ and $\pi(k,n)$ are decreasing functions of k. Effect I, however, provides a force in the opposite direction; a more A-biased outlet is *a priori* more likely to send message "A" and induce posterior expectation $k_A(\chi)$ instead of $k_B(\chi)$. Since $k_A(\chi) > k_B(\chi)$, and hence $\pi(k_A(\chi),n) > \pi(k_B(\chi),n)$, Effect I *per se* improves candidate A's electoral prospects. The overall net effect depends on which of the opposing forces is strongest. For a sufficiently small bias (χ close to 0), Effect I dominates so that a marginal increase in media bias increases candidate A's winning chances. For a sufficiently large degree of media bias (χ close to 1), however, the opposite applies. This explains the non-monotonic relationship between media bias and A's winning probability (Proposition 1.2).

The interim and ex-ante impacts of media bias on election outcomes can also be illustrated through an elegant geometric approach presented in Figure 3b. With k uniformly distributed on $[-1,1]$ (cf. Example 1), it holds that $k_A(\chi) = \frac{1-\chi}{2}$ and $k_B(\chi) = \frac{-1-\chi}{2}$. In Figure 3b we consider three levels of media bias $\chi \in \{0,0.5,0.9\}$. First suppose $\chi = 0$, that is, the media outlet is unbiased. Then $k_A(0) = 0.5$ and $k_B(0) = -0.5$, and hence the electorate selects candidate A with probability $\pi(0.5,n)$ ($\pi(-0.5,n)$) conditional on message "A" ("B"), represented by the red node x_A (x_B) in Figure 3b. Ex-ante, candidate A's winning probability is a weighted average of $\pi(0.5,n)$ and $\pi(-0.5,n)$, and must lie on the red dashed line segment connecting x_A and x_B. By the law of iterated expectations, the posterior means must average back to the prior mean, which is 0. Therefore, candidate A's ex-ante winning probability can be geometrically represented by the red node x_*, the intersection of line segment $x_A x_B$ with the vertical line $k = 0$.

A similar exercise can be done for a moderately biased outlet with $\chi = 0.5$, yielding interim election probabilities represented by the blue nodes y_A and y_B and an ex-ante winning probability given by y_*. Likewise, for a strongly biased media with $\chi = 0.9$ the election outcomes conditional on message "A" and "B", along with the ex-ante impact, are represented by the green nodes z_A, z_B and z_*, respectively. The interim impacts of an increase in media bias can now be straightforwardly represented by the movement from x_s to y_s to z_s, for $s \in \{A,B\}$. This reveals that candidate A's winning probability decreases as media bias increases conditional on both message "A" and "B", as predicted by Proposition 1.1. The ex-ante impact of increasing media bias is given by the movement from x_* to y_* to z_*. With $x_* < z_* < y_*$, the geometric analysis clearly illustrates the non-monotonic ex-ante impact of media bias on the election outcome, derived in Proposition 1.2.

We next turn to the interim and ex-ante impacts of media bias on voter turnout.

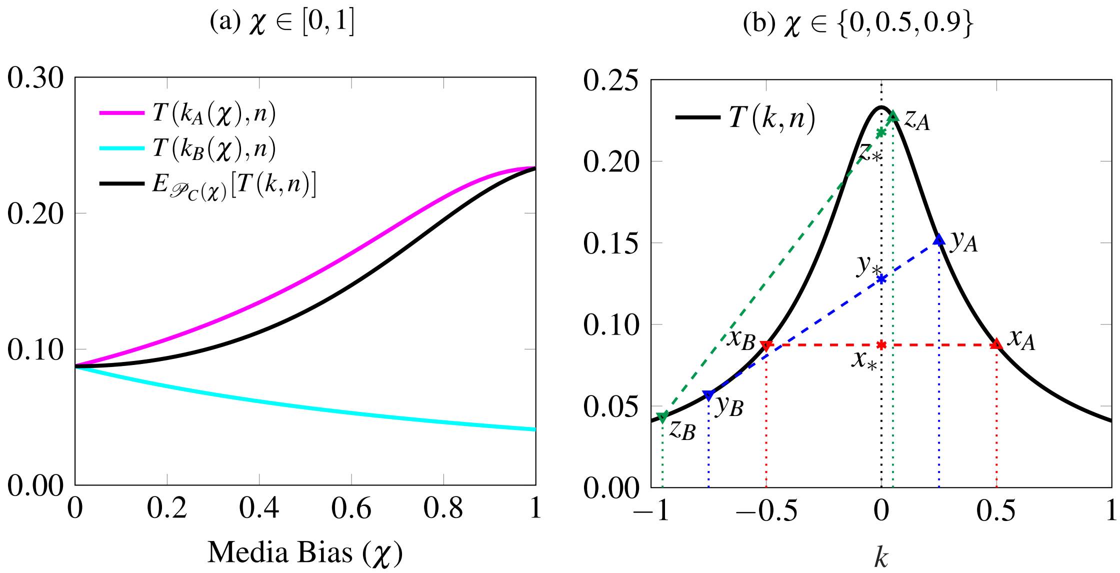
Proposition 2. *Suppose* $|M| = 1$. $\forall \chi' > \chi \geq 0$ *and for sufficiently large n:*

1. $T(k_s(\chi'),n) > (<) T(k_s(\chi),n)$ *for* $s = A(B)$.
2. $E_{\mathscr{P}_C(\chi')}[T(k,n)] > E_{\mathscr{P}_C(\chi)}[T(k,n)]$.

Proof. See Online Appendix C.1. □

Figure 4a illustrates Proposition 2. The interim impact of media bias on voter turnout depends on the message sent by the outlet (Proposition 2.1). Given message "A", the expected voter turnout is higher as media bias increases. This holds because, by Effect II, $|k_A(\chi)|$ decreases in χ and the election is more likely to end up in a close race, driving up voter turnout. Based on a similar intuition (cf. Effect III), conditional on message "B" the expected voter turnout is lower as media bias increases. From the ex-ante perspective voter turnout increases with media bias unambiguously (Proposition 2.2) for two reasons. First, with $|k_A(\chi)|$ decreasing and $|k_B(\chi)|$ increasing in χ, the distribution of posterior expectations of $|k|$ becomes more spread out as χ increases. This would lead to a higher

Figure 4: The interim and ex-ante impact of media bias on voter turnout



Note: Panel (a) depicts the interim and the ex-ante impacts of media bias for all $\chi \in [0,1]$. Panel (b) geometrically demonstrates these impacts for $\chi \in \{0, 0.5, 0.9\}$. The function $T(k,n)$ depicted in Panel (b) is identical to $T(k,n)$ in Figure 1d. In Panel (b), the selected levels of media bias increase from red to blue to green: $\chi = 0$ (unbiased, nodes x), $\chi = 0.5$ (weakly biased, nodes y), and $\chi = 0.9$ (strongly biased, nodes z). Nodes w_A, w_B and w_*, where $w \in \{x,y,z\}$, represent the expected voter turnout conditional on message "A", message "B" and unconditionally, respectively. The dashed line segments represent the sets of all convex combinations of w_A and w_B, for $w \in \{x,y,z\}$. Model parameters are taken from Example 1.

expected voter turnout if the average posterior mean of $|k|$ remains constant, due to the convexity of $T(|k|,n)$ in large elections. Second, due to Effect I the average posterior mean of $|k|$ puts more weight on $|k_A(\chi)|$ as media bias increases, since message "A" is more likely to be sent ex-ante. As a result, the average posterior mean of $|k|$ decreases in media bias.[26] This further increases voter turnout because $T(|k|,n)$ decreases in $|k|$.

Figure 4b takes the same geometric approach as in Figure 3b. Similar to the previous analysis for election outcomes, the interim impacts conditional on message "A" and "B", as well as the ex-ante impact of media bias on voter turnout, are geometrically represented by nodes w_A, w_B and w_*, respectively, for $w \in \{x,y,z\}$ (again corresponding to $\chi \in \{0, 0.5, 0.9\}$). The interim impact conditional on message "A" ("B") again follows from the movement from x_A to y_A to z_A (x_B to y_B to z_B). Ex-ante voter turnout increases in media bias since $x_* < y_* < z_*$.

Finally, we study how media bias affects the probability of electing the first-best (i.e., quality-superior) candidate and voter welfare, as $n \to \infty$. Unlike previous propositions,

[26] This is formally proved in Lemma 12 of Online Appendix C.1.

which focus on $|M| = 1$ and information partition $\mathscr{P}_C(\chi)$, Proposition 3 quantifies these outcomes for generic information partitions $\mathscr{P} \subset [-1,1]$.[27]

Proposition 3. *For any partition $\mathscr{P} \subseteq [-1,1]$, let $x^* \equiv \min\{|x| | x \in \mathscr{P}\}$. It holds that*

1. *the probability of electing the first-best candidate converges to $\frac{1}{2} + F(-x^*)$ as $n \to \infty$.*
2. $\lim_{n\to\infty} E_{\mathscr{P}}[W(k,n)] = \int_{x^*}^{1} k dF(k)$.

Proof. See Online Appendix C.1. □

With $|M| = 1$ and $\chi \geq 0$, $\mathscr{P}_C(\chi) = \{-\chi\}$. It follows directly from Proposition 3 that, as $n \to \infty$, the probability of electing the first-best candidate and the expected voter welfare $E_{\mathscr{P}_C}[W(k,n)]$ converge to $\frac{1}{2} + F(-\chi)$ and $\int_{\chi}^{1} k dF(k)$, respectively. Both are decreasing in media bias, χ. Moreover, by Proposition 2, increasing media bias also systematically increases voter turnout and thus the aggregate voting costs. This effect further reduces voters' welfare. Proposition 3 suggests that, when the media outlet is unbiased (i.e., $\chi = 0$), the election outcome is *asymptotically efficient* in the sense that the asymptotic voter welfare, $\lim_{n\to\infty} E_{\mathscr{P}}[W(k,n)]$, is maximized. This holds because, with an unbiased media outlet, voters learn precisely ex-post whether k is above or below 0, and hence elect the quality-superior candidate with probability approaching 1 as $n \to \infty$. Conversely, if the media outlet is extremely biased (i.e., $\chi = 1$), voters remain uninformed ex-post because the media outlet always endorses candidate A regardless of the realized state. In this case, the election outcome is completely independent of the realized state. Therefore, both the probability of electing the first-best candidate and the asymptotic voter welfare are lowest when $\chi = 1$.

5.3 The influence of introducing a second media outlet

In this section we vary $|M|$ from 1 to 2 to study the electoral influence of introducing an extra media outlet. Let the biases of outlet 1 and 2 be χ_1 and χ_2, respectively. Without loss of generality, we assume $\chi_1 \in [0,1]$. We say outlets 1 and 2 are *like-minded* (*opposite-minded*) if χ_1 and χ_2 have the same (opposite) sign. Moreover, we say outlet 2 has a *stronger (weaker)* bias than outlet 1 if $|\chi_2| > (<) |\chi_1|$. Recall from Figure 2 that outlet 1 partitions the posterior state space into two intervals $[-1, -\chi_1]$ and $(-\chi_1, 1]$. Introducing

[27] We define a generic information partition $\mathscr{P}$ as follows: $x \in \mathscr{P}$ if and only if voters precisely learn ex-post whether $k \leq x$ or $k > x$. With this definition, $\mathscr{P}$ can also contain continuous intervals. For any interval $[a,b] \subset [-1,1]$ with $a < b$, if $[a,b] \subset \mathscr{P}$ then any $k \in [a,b]$ is precisely revealed under partition $\mathscr{P}$.

outlet 2 affects voters' information environment by inducing a finer partition on $[-1,-\chi_1]$ ($(-\chi_1,1]$), if $\chi_2 > (<) \chi_1$. Proposition 4 characterizes how introducing a second media outlet systematically affects the election outcome.

Proposition 4. *For all $\chi_1 \in [0,1]$, $\chi_2 \neq \chi_1$, and for sufficiently large n, $E_{\mathscr{P}_C(\chi_1,\chi_2)}[\pi(k,n)] > (<) E_{\mathscr{P}_C(\chi_1)}[\pi(k,n)]$ if and only if $\chi_2 > (<) \chi_1$.*

Proof. See Online Appendix C.2. □

In words, Proposition 4 states that in sufficiently large elections, introducing a like-minded outlet with a stronger bias increases the ex-ante winning probability of the candidate favored by the incumbent outlet. Conversely, if outlet 2 is opposite-minded, or like-minded with a weaker bias, then introducing outlet 2 decreases the winning chances of that candidate. Loosely put, in large elections introducing outlet 2 systematically shifts the election outcome in the direction of outlet 2's bias (relative to outlet 1's).

Proposition 4 is primarily driven by the S-shape of $\pi(k,n)$: it is convex on $[-1,0)$ and concave on $(0,1]$. Recall that outlet 1 partitions the state space into two intervals $[-1,-\chi_1]$ and $(-\chi_1,1]$. Suppose $\chi_1 > 0$ (i.e., A-biased), then $\pi(k,n)$ is convex on $[-1,-\chi_1]$ due to its S-shape. If outlet 2 is like-minded with a stronger bias (i.e., $\chi_2 > \chi_1$), then it refines the information partition on $[-1,-\chi_1]$ and hence induces a mean-preserving spread of voters' posterior expectations of k on the same interval, which necessarily increases the expectation of any convex function. If instead $\chi_2 < \chi_1$, then outlet 2 refines the information partition on $(-\chi_1,1]$, where $\pi(k,n)$ is neither convex nor concave in k. In this case, the argument for why introducing outlet 2 systematically decreases A's winning probability is more subtle (see Online Appendix C.2 for details).

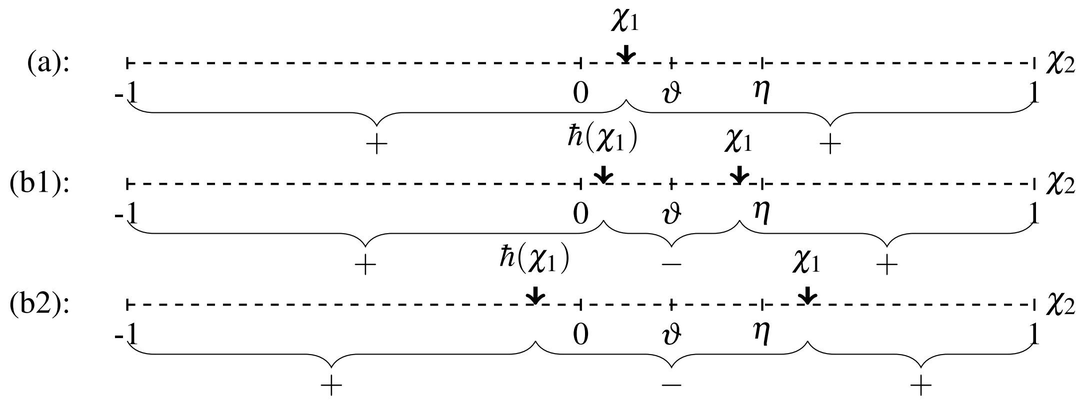
Our next proposition concerns how introducing a second media outlet systematically affects voter turnout.

Proposition 5. *For sufficiently large n, there exists $\vartheta \in (0,1)$ and a decreasing function $\hbar(\chi)$ satisfying (i) $-\chi < \hbar(\chi) < \chi$ for all $\chi > \vartheta$ and (ii) $\hbar(\eta) = 0$ for some $\eta \in (\vartheta,1)$, such that:*

1. *if $0 \leq \chi_1 \leq \vartheta$ and $\chi_2 \neq \chi_1$, then $E_{\mathscr{P}_C(\chi_1,\chi_2)}[T(k,n)] > E_{\mathscr{P}_C(\chi_1)}[T(k,n)]$.*
2. *if $\chi_1 > \vartheta$, then $E_{\mathscr{P}_C(\chi_1,\chi_2)}[T(k,n)] < E_{\mathscr{P}_C(\chi_1)}[T(k,n)]$ if $\chi_2 \in (\hbar(\chi_1),\chi_1)$ and the opposite holds otherwise.*

Proof. See Online Appendix C.2. □

Figure 5: A graphical illustration of Proposition 5



Note: The intervals in curly braces with a "+" ("−") sign denote the sets of χ_2 under which introducing a second media outlet increases (decreases) ex-ante voter turnout, for a given $\chi_1 \geq 0$.

Figure 5 graphically illustrates Proposition 5. Panel (a) corresponds to Proposition 5.1 stating that if the bias of outlet 1 is sufficiently low (i.e., $\chi_1 < \vartheta$), then introducing a second outlet systematically increases voter turnout regardless of χ_2. In contrast, if outlet 1 is at least moderately biased (i.e., $\chi_1 > \vartheta$), then introducing outlet 2 systematically decreases voter turnout if outlet 2 has a weaker bias sufficiently close to the bias of outlet 1 (see panel (b1)). For a rather strongly biased outlet 1 (i.e., $\chi_1 > \eta$), this less biased outlet 2 could even be opposite minded (see panel (b2)). Note that in the latter case introducing an unbiased outlet can systematically decrease voter turnout.

Proposition 5 is driven by the curvature of $T(k,n)$. If outlet 2 is like-minded and has a stronger bias than outlet 1 (i.e., $\chi_2 > \chi_1 \geq 0$), then outlet 2 induces a finer information partition on $[-1,-\chi_1]$ and necessarily increases the expectation of $T(k,n)$, which is convex on $[-1,-\chi_1]$. If instead outlet 2 is opposite-minded (i.e., $\chi_2 < 0$), or like-minded with a weaker bias (i.e., $0 < \chi_2 < \chi_1$), then outlet 2 induces a finer partition on $[-\chi_1, 1]$, where voter turnout $T(k,n)$ is neither convex nor concave in k. In this case, introducing outlet 2 can either systematically increase or decrease voter turnout, depending on the biases of both outlets. In this way, Proposition 5 provides a possible explanation for the mixed empirical evidence on the relationship between media entry and voter turnout, in a unified framework.

The literature suggests some alternative mechanisms through which media entry may affect voter turnout. First, media entry may increase voter turnout by simply notifying voters about an upcoming election (Gentzkow et al., 2011). Second, media entry may decrease voter turnout through a substitution effect; the entry of commercial TV channels and Internet may shift voters' interests from politics to entertainment, decreasing their exposure to political information and thereby voter turnout (Gentzkow, 2006; Falck et al., 2014; Ellingsen et al., 2018). Third, competition in the marketplace may affect media outlets' provision of political information and, in turn, affect voter turnout (Cagé, 2017). All these mechanisms rely on the assumption that informed voters are more likely to turnout than their uninformed counterparts. This assumption finds support from the literature on individual decision making (Matsusaka, 1995), the *swing voter's curse* (Feddersen and Pesendorfer, 1996, 1999; Battaglini et al., 2008, 2010) and government's responsiveness to the demand of informed voters (Strömberg, 2004; Gavazza et al., 2015). Nevertheless, none of these mechanisms can explain, in a unified framework, why media entry can sometimes increase and in other cases decrease voter turnout. To the best of our knowledge, Piolatto and Schuett (2015) is the only other theoretical paper able to reconcile the mixed empirical findings. They do so in an ethical voter framework and the mechanism therein is different from ours.[28]

Given that introducing an additional media outlets can systematically drive both the election outcome and voter turnout in either direction, how does it affect voter welfare? One might expect that introducing extra outlets necessarily makes voters better off, because it improves information transmission and thus allows rational voters to make better decisions. Proposition 6 confirms this conventional wisdom for large elections under the cutoff information environment.

Proposition 6. *For all* $\chi_1 \in [0,1]$*, it holds that*

1. $\lim_{n\to\infty} E_{\mathscr{P}(\chi_1,\chi_2)}[W(k,n)] \geq \lim_{n\to\infty} E_{\mathscr{P}(\chi_1)}[W(k,n)]$ *for all* $\chi_2 \in [-1,1]$.
2. $\lim_{n\to\infty} E_{\mathscr{P}(\chi_1,\chi_2)}[W(k,n)] > \lim_{n\to\infty} E_{\mathscr{P}(\chi_1)}[W(k,n)]$ *if and only if* $|\chi_1| > |\chi_2| \geq 0$.

Proposition 6 is an immediate corollary of Proposition 3; with $|M| = 2$ and $\mathscr{P}_C(\chi_1,\chi_2) = \{-\chi_1,-\chi_2\}$, the probability of electing the quality-superior candidate and voter welfare converges to $\frac{1}{2}+F(-x^*)$ and $\int_{x^*}^{1} kdF(k)$, respectively, where $x^* = \min\{|\chi_1|,|\chi_2|\}$. As $x^* \leq |\chi_1|$, introducing an extra outlet never deteriorates asymptotic voter welfare (Proposition 6.1). In fact, introducing an outlet strictly increases asymptotic voter welfare if and only if $|\chi_2| < |\chi_1|$, i.e., the second outlet has a strictly weaker bias than the existing one (Proposition 6.2).

[28] In Piolatto and Schuett (2015), media entry affects voter turnout through influencing independent and partisan voters' decisions to acquire information. They show that information increases the turnout rate of independent voters but decreases the turnout rate of partisan voters in expectation. The net effect is thus ambiguous and critically dependent on the degree of polarization of the electorate. In our model, however, media's impacts on voter turnout is essentially driven by their influences on voters' perceptions about the closeness of elections.

More generally, in the presence of an arbitrary number of media outlets, asymptotic voter welfare is determined solely by the *least biased* outlet, whose χ_m is closest to 0.

6 Media competition and voter welfare

In Section 5 we analyzed media's influence on voting equilibria under the assumption that all media outlets use the cutoff endorsement strategy (9) and we found that the introduction of an additional media outlet necessarily improves voter welfare in large elections. In this section we explore how robust this finding is to different communication protocols and different ways of increasing media competition.

For this purpose we first endogenize media outlets' reporting strategies and characterize equilibrium information transmission under two different communication protocols: cheap talk (CT) and Bayesian persuasion (BP). In the former setup, media outlets cannot commit to any reporting strategy before observing the realized state whereas in the latter setup they can. The comparison between the two protocols thus informs us how media commitment affects information transmission and voter welfare. Indeed, we find that media commitment plays a key role in determining the amount of information that can be transmitted in equilibrium (cf. Section 6.1). We then turn to the question to what extent the conventional wisdom – media competition improves information revelation and voter welfare – can be generalized to different communication protocols. This is explored in Section 6.2.

6.1 Preliminaries: Information transmission under CT and BP

We use a sender-receiver game framework to study the information transmission between media outlets and the electorate. Since media outlets send public information, all voters form a common posterior conditional on the message profile x produced by all media outlets. From the media outlets' perspective, the entire electorate is then equivalent to a single receiver who rationally updates her belief to obtain posterior $E[k|x,\Xi]$ and subsequently selects candidate A with probability $\pi(E[k|x,\Xi],n)$. Recall that we denote media m's reporting strategy by $\sigma_m(k)$, which is a mapping from the state space ($[-1,1]$) to the message space S. In this section, we assume $S=[-1,1]$ so that a message can be literally interpreted as a report of the state. As in all cheap talk games, there are multiple equilibria. For tractability we impose the following *monotonicity constraint* on the strategies $\sigma_m(\cdot)$ for all $m \in M$:

$$\forall k, k' \in [-1,1], k > k' \implies \sigma_m(k) \geq \sigma_m(k') \tag{10}$$

In words, constraint (10) requires that each outlet $m \in M$ must announce weakly higher messages in higher states. The monotonicity constraint implies that (i) a strictly higher message must reveal strictly higher states, and (ii) if a message s is sent in two different states, then the same s must also be sent in all states in between. Taken together, any reporting strategy $\sigma_m(\cdot)$ must induce a *partition*, denoted by $\mathscr{P}^m$, of the state space. As in Section 5.1, we characterize $\mathscr{P}^m$ by the set of cutoff points; $x \in \mathscr{P}^m$ if and only if $\sigma_m(k) > (\leq)\sigma_m(x)$ implies $k > (\leq)x$ for all $k, x \in [-1,1]$. In other words, $x \in \mathscr{P}^m$ if and only if the electorate is precisely informed of whether the realized k is strictly above, or weakly below x, by reading from outlet m.[29] The information environment created by all media outlets $m \in M$ can be represented by the union of all partitions $\mathscr{P} = \cup_{m \in M} \mathscr{P}^m$. Any $\mathscr{P}$ partitions the state space $[-1,1]$ into adjacent and disjoint *separating* or *pooling* intervals. The realized state k is precisely revealed in a separating interval, while unrevealed in a pooling interval.

We focus on the most informative equilibria, i.e., equilibria that induce the finest partition $\mathscr{P}$, and do not distinguish between equilibria inducing an identical $\mathscr{P}$.[30] When $|M| \geq 2$, we also derive the most informative *coordination-proof* equilibria, in which no subset of media outlets can strictly benefit by jointly deviating from their equilibrium strategy profile. Theorem 3 precisely characterizes the most informative partition in equilibrium under CT.

Theorem 3. *Suppose media outlets communicate via cheap talk and $\chi_1 \geq \chi_2 \geq \cdots \geq \chi_{|M|}$.*

1. *(Discipline equilibrium) The most informative partition in equilibrium equals*

$$\mathscr{P}_D(\chi_1, \cdots, \chi_{|M|}) = \begin{cases} \{-\chi_1\}, & \text{if } |M| = 1 \\ [-1, -\chi_1] \cup [-\chi_2, 1], & \text{if } |M| = 2 \\ [-1, 1], & \text{if } |M| \geq 3 \end{cases}$$

[29] To simplify the presentation and proofs, we assume any cutoff point x is downward pooled. Under Assumption 1, $k = x$ is a zero-probability event and the allocation of x in the partition is inconsequential.

[30] We focus on the most informative equilibria for two reasons. First, this is a standard equilibrium selection practice in the cheap talk literature. Second, by Proposition 3 the asymptotic voter welfare is highest in the most informative equilibrium. Therefore, by focusing on the most informative equilibria, we provide an upper bound for asymptotic voter welfare in equilibrium.

2. *(Cutoff equilibrium) The most informative partition among coordination-proof equilibria equals* $\mathscr{P}_C(\chi_1, \cdots, \chi_{|M|}) = \{-\chi_1, -\chi_2, \cdots, -\chi_{|M|}\}$.

Proof. See Online Appendix D.1. □

To illustrate Theorem 3, we start with $|M| = 1$. In this case, the most informative equilibrium yields a unique information partition $\{-\chi_1\}$, which consists of two pooling intervals $[-1, -\chi_1]$ and $(-\chi_1, 1]$. This is equivalent to the information partition created by the cutoff endorsement strategy (9) introduced in Section 5.1 (cf. Figure 2). Since the outlet prefers candidate A(B) to be elected if $k > (\leq) -\chi_1$ and A's winning probability increases in voters' posterior expectations of k, it is optimal for this outlet to send the highest (lowest) message when $k > (\leq) -\chi_1$. Therefore, whether k lies above or below $-\chi_1$ must be precisely communicated, and no further information (except for the zero-probability event $k = -\chi_1$) can be credibly revealed. This yields the aforementioned pooling intervals.

When there are two media outlets with biases $\chi_1 \geq \chi_2$, the most informative CT equilibrium, referred to as the Discipline equilibrium (Theorem 3.1), yields an information partition $\mathscr{P}_D(\chi_1, \chi_2) = [-1, -\chi_1] \cup [-\chi_2, 1]$. This partition consists of two separating intervals $[-1, -\chi_1]$ and $(-\chi_2, 1]$, and a pooling interval $(-\chi_1, -\chi_2]$ in between. In line with the literature on multiple-sender cheap talk games, the receiver (electorate) can discipline senders (media outlets) to reveal more information by cross-checking their messages and punish senders by a strictly undesirable action if their messages do not match (Gilligan and Krehbiel, 1989; Krishna and Morgan, 2001a,b). Specifically, the electorate can discipline media outlets to reveal all "aligned states" (i.e., $k \in [-1, -\chi_1) \cup (-\chi_2, 1]$), conditional on which both outlets prefer the same candidate to be elected, by the following cross-checking strategy: if $s_1 = s_2 = s$ all voters believe $k = s$ for sure, otherwise they form posterior expectation $E_F[k|k \in (\chi_1, -\chi_2)]$. With this cross-checking strategy, both outlets are worse off in aligned states if their messages do not match, as sending inconsistent messages would only decrease the winning chances of their preferred candidate. In "conflicting states", where the two outlets support different candidates (i.e, $k \in (-\chi_1, -\chi_2)$), such a common punishment does not exist and they cannot be disciplined to reveal any information. With more than two media outlets, full information revelation is generically possible under the Discipline equilibrium. This is because for all realized states k, there always exist (at least) two media outlets whose interests are aligned. These interest alignments can be exploited to construct a fully revealing equilibrium under appropriate cross-checking strategies.

The construction of Discipline equilibria depends critically on the premise that senders cannot coordinate their messages. To see this, note that when $k > -\chi_2$ and coordination is possible, both outlets would be better off by coordinating on $s_1 = s_2 = 1$ (rather than on the equilibrium strategy $s_1 = s_2 = k$) to maximize A's winning chances, given the aforementioned cross-checking strategy of the electorate. Therefore, if we require media outlets' reporting strategy profile to be coordination-proof, then aligned states can no longer be credibly communicated in equilibrium (Theorem 3.2). In fact, the most informative coordination-proof equilibrium generates the cutoff information partition described in Section 5.1 and is thus referred to as the Cutoff equilibrium.

Theorem 4 characterizes the most informative equilibria in the BP setup, under an additional technical assumption defined in Online Appendix D.2 (Assumption 4 therein).[31]

Theorem 4. *Suppose media outlets can commit to any reporting strategy profile and $\chi_1 \geq \chi_2 \geq \cdots \geq \chi_{|M|}$. Under Assumption 4 defined in Online Appendix D.2 it holds that:*

1. *The most informative partition in equilibrium equals*

$$\mathscr{P}(\chi_1, \cdots, \chi_{|M|}) = \begin{cases} [a(\chi_1;n), b(\chi_1;n)], & \text{if } |M| = 1 \\ [-1,1], & \text{if } |M| > 1 \end{cases}$$

2. *(BP equilibrium) The coordination-proof equilibrium is unique and yields information partition $\mathscr{P}_{BP}(\chi_1, \cdots, \chi_{|M|}) = [a(\chi_1;n), b(\chi_{|M|};n)]$.*

Functions $a(x;n)$ and $b(x;n)$ are implicitly defined in Online Appendix D.2 and satisfy $\forall x \in [-1,1]$: (i) $-1 \leq a(x;n) < -x < b(x;n) \leq 1$; (ii) both $a(x;n)$ and $b(x;n)$ weakly decrease in x, and (iii) $\lim_{n\to\infty} a(x;n) = -x$ for $x < 0$ and $\lim_{n\to\infty} b(x;n) = -x$ for $x > 0$.

Proof. See Online Appendix D.2. □

Under BP, an information partition can be supported in equilibrium if and only if no senders can profit from unilaterally committing to any finer partition (Gentzkow and Kamenica, 2016). If $|M| = 1$, the equilibrium partition is unique and equals $[a(\chi_1;n), b(\chi_1;n)]$; it consists of two pooling intervals $[-1, a(\chi_1;n)]$ and $(b(\chi_1;n), 1]$, and a non-empty separating interval $(a(\chi_1;n), b(\chi_1;n)]$ which contains $-\chi_1$. In words, the optimal reporting

[31] Loosely speaking, this additional assumption puts regularity restrictions on the shape of the media's expected payoff function, under which the optimal reporting strategy necessarily takes a simple "pooling at the tails" structure; if two different states are precisely revealed, then all the states in between are revealed as well. Pooling, if it occurs, can thus only happen at the tails of the state space. Again, our illustrating Example 1 satisfies Assumption 4.

strategy precisely reveals states with low payoff relevance (i.e., k sufficiently close to $-\chi_1$), and pools all remaining high states ($k > b(\chi_1;n)$) into a single message (say, "A") and low states ($k \leq a(\chi_1;n)$) into another message (say, "B").[32] In deciding whether to separate the boundary state (say, $b(\chi_1;n)$) from the pooling message (say, "A"), the outlet balances two opposite forces. On the one hand, separating $b(\chi_1;n)$ from "A" lowers candidate A's winning probability when the marginal state $b(\chi_1;n)$ realizes. On the other hand, this separation strengthens the credibility of message "A" and thus increases A's electoral prospects in all states $k > b(\chi_1;n)$. For k sufficiently close to $-\chi_1$ the latter effect always dominates so $b(\chi_1;n) > -\chi_1$ generically holds. This is because the outlet is almost indifferent between candidates for k close to $-\chi_1$ and the marginal losses from the former effect is negligible compared to the marginal gains from the latter. The choice of the other boundary state $a(\chi_1;n)$ follows the analogous principle.

If $|M| \geq 2$, then fully revealing equilibria generically exist under BP without coordination-proofness (Theorem 4.1). Nevertheless, these equilibria are implemented in weakly dominated strategies.[33] For this reason, the literature on Bayesian persuasion with multiple senders typically focuses on the least informative equilibria (Gentzkow and Kamenica, 2016, 2017), which coincide with the unique coordination-proof equilibrium under BP, referred to above as the BP equilibrium (Theorem 4.2). The critical impact of introducing extra media outlets on the BP equilibrium is that all "conflicting states" must be fully revealed. Because $\pi(k,n)$ is generically nonlinear in k, in any non-trivial pooling interval containing conflicting states, inducing a finer information partition can always systematically push the election outcome in favor of one candidate. Because at least two media outlets have opposite partisan preferences in those conflicting states, there must exist at least one media outlet finding unilaterally committing to a finer information partition profitable.[34]

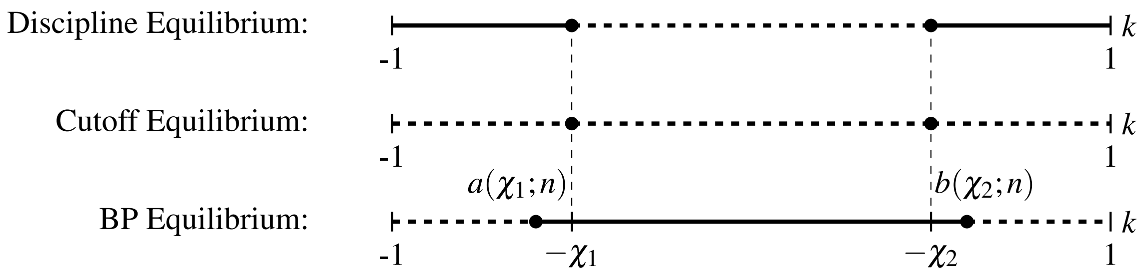
For ease of comparison, we graphically illustrate the information partitions with $|M| = 2$ under the Discipline, Cutoff and BP equilibria in Figure 6. Figure 6 highlights two sharp discrepancies between the equilibrium information partitions under CT and BP setups when $|M| = 2$. First, in aligned states where all media outlets' partisan preferences agree (e.g., intervals $[-1,-\chi_1]$ and $(-\chi_2,1]$), the information partition exhibits "semi-separating" in

[32] It is possible for an optimal reporting strategy to have $a(\chi_1;n) = -1$ or $b(\chi_1;n) = 1$. In these cases, the outlet precisely reveals all states $k \in [-1,-\chi_1]$ or $k \in (-\chi_1, 1]$, respectively. However, $a(\chi_1;n) = -1$ and $b(\chi_1;n) = 1$ never simultaneously hold, that is, full information revelation is impossible under $|M| = 1$.

[33] To see this, suppose outlet 1 unilaterally commits to full revelation. It is then weakly optimal for all other outlets to do so as well because they cannot unilaterally influence voters' information.

[34] See also Gentzkow and Kamenica (2016, 2017) for similar insights.

Figure 6: Equilibrium information partitions under CT and BP when $|M| = 2$



Note: The solid line segments and black nodes include all elements in each equilibrium information partition. Therefore, the solid (dashed) line segments also represent separating (pooling) intervals.

the BP equilibrium. In the two classes of cheap talk equilibria we investigate, however, these states are either fully separated (the Discipline equilibrium) or pooled (the Cutoff equilibrium). Second, the "conflicting states" (i.e., interval $(-\chi_1, -\chi_2])$ are fully revealed in all BP equilibria. In sharp contrast, however, these conflicting states can never be revealed in any cheap talk equilibrium when $|M| = 2$. In what follows, we use these insights to analyze how information transmission and voter welfare in large elections are affected as we increase media competition.

6.2 Competition, information transmission and voter welfare

Following Gentzkow and Kamenica (2016), we consider the following three notions of increasing media competition: (i) introducing extra media outlets, (ii) increasing media polarization, and (iii) preventing media collusion. To compare the informativeness of two partitions $\mathscr{P}_1$ and $\mathscr{P}_2$, we say $\mathscr{P}_1$ is *more informative* than $\mathscr{P}_2$ if $\mathscr{P}_1$ induces a strictly finer information partition than $\mathscr{P}_2$ (i.e., $\mathscr{P}_2 \subset \mathscr{P}_1$). To compare asymptotic voter welfare, we say $\mathscr{P}_1$ is *welfare superior (inferior)* to $\mathscr{P}_2$ if $\lim_{n\to\infty} E_{\mathscr{P}_1}[W(k,n)] > (<) \lim_{n\to\infty} E_{\mathscr{P}_2}[W(k,n)]$. Recall from Proposition 3 that $\lim_{n\to\infty} E_{\mathscr{P}}[W(k,n)] = \int_{x^*}^{1} k dF(k)$ with $x^* = \min\{|x| | x \in \mathscr{P}\}$. Asymptotic voter welfare is maximized if and only if $0 \in \mathscr{P}$; voters learn precisely whether k lies above or below 0 ex-post. For this reason, we say a partition $\mathscr{P}$ is *asymptotically efficient* if $0 \in \mathscr{P}$.

6.2.1 Introducing extra media outlets

We first show that both CT and BP models agree that introducing extra media outlets weakly improves information transmission and voter welfare. However, these models disagree sharply on what type of media outlet should be introduced to maximize these benefits in large elections. We illustrate this insight by increasing $|M|$ from 1 to 2. Let the biases of the incumbent and entrant outlets be χ_1 and χ_2, respectively. We assume $\chi_1 > 0$ and say the two outlets are like-minded (opposite-minded) if χ_1 and χ_2 have the same (opposite) sign. Define $\chi_+ \equiv \max\{\chi_1, \chi_2\}$ and $\chi_- \equiv \min\{\chi_1, \chi_2\}$.[35]

In a Discipline equilibrium it holds that $\mathscr{P}_D(\chi_1, \chi_2) = [-1, -\chi_+] \cup [-\chi_-, 1]$ (cf. the top panel of Figure 6). As is evident graphically, $\mathscr{P}_D(\chi_1, \chi_2)$ is asymptotic efficient (i.e., $0 \in \mathscr{P}_D(\chi_1, \chi_2)$) if and only if the two outlets are like-minded. Moreover, if $\chi_2 = \chi_1$, then $\mathscr{P}_D(\chi_1, \chi_2) = [-1, 1]$ and the Discipline equilibrium is fully revealing.

In a Cutoff equilibrium we have $\mathscr{P}_C(\chi_1, \chi_2) = \{-\chi_1, -\chi_2\}$ (cf. the central panel of Figure 6). Hence, $\mathscr{P}_C(\chi_1, \chi_2)$ is asymptotically efficient if and only if $\chi_m = 0$ for some $m = 1, 2$, i.e., at least one outlet is unbiased.

Finally, in a BP equilibrium we have $\mathscr{P}_{BP}(\chi_1, \chi_2) = [a(\chi_+; n), b(\chi_-; n)] \supset [-\chi_+, -\chi_-]$ (cf. the bottom panel of Figure 6). It is evident graphically that, if the two outlets are opposite-minded, $\mathscr{P}_{BP}(\chi_1, \chi_2)$ must be asymptotically efficient since $0 \in [-\chi_+, -\chi_-]$. Moreover, $\mathscr{P}_{BP}(\chi_1, \chi_2)$ must be fully revealing if $\chi_1 = 1$ and $\chi_2 = -1$, i.e., the two outlets are opposite-minded with extremely strong biases. Importantly, however, if the two outlets are both biased and like-minded (i.e., $\chi_1 \geq \chi_2 > 0$) then $\mathscr{P}_{BP}(\chi_1, \chi_2)$ is no longer asymptotic efficient. This is because, by Theorem 4, $\lim_{n\to\infty} b(\chi_m; n) = -\chi_m < 0$ since $\chi_m > 0$ for both $m \in \{1, 2\}$. Therefore, $b(\chi_-; n) < 0$ and $0 \notin \mathscr{P}_{BP}(\chi_1, \chi_2)$ for sufficiently large n. We summarize these observations in Proposition 7.

Proposition 7. *Let $|M| = 2$ and $\chi_1 > 0$, it holds that:*

1. *$\mathscr{P}_D(\chi_1, \chi_2)$ is asymptotically efficient if and only if $\chi_2 \geq 0$, and it is fully revealing if $\chi_2 = \chi_1$.*
2. *$\mathscr{P}_C(\chi_1, \chi_2)$ is asymptotically efficient if and only if $\chi_2 = 0$.*
3. *$\mathscr{P}_{BP}(\chi_1, \chi_2)$ is asymptotically efficient if and only if $\chi_2 \leq 0$, and it is fully revealing if $\chi_1 = 1$ and $\chi_2 = -1$.*

Proposition 7 highlights the critical role of media commitment in identifying the type of entrant that maximizes asymptotic voter welfare and information transmission. Under CT (i.e., without media commitment), asymptotic efficiency is achieved only if the entrant

[35] We do not consider $\chi_1 = 0$ because in this case, by Proposition 3, the election will be asymptotically efficient regardless of the biases of other media outlets.

is unbiased (Cutoff equilibrium) or like-minded (Discipline equilibrium). Under BP (i.e., with media commitment), however, asymptotic efficiency can be achieved with an opposite-minded entrant only. Regarding information transmission, introducing an outlet that is *like-minded with an identical bias* (*opposite-minded with an extremely strong bias*) maximizes information revelation in a Discipline (BP) equilibrium.

6.2.2 Increasing media polarization

We focus on $|M| = 2$ and ask how information transmission and asymptotic voter welfare are affected by increasing the misalignment of interests between the two media outlets. This informs us how voters are affected by an increase in the polarization of media environment. We study the impact of increasing media polarization in a context with two opposite-minded and equally biased outlets; $|M| = 2$, $\chi_1 = \chi$ and $\chi_2 = -\chi$ with $\chi > 0$. We say *media polarization increases* if χ increases.

By Theorem 3, Discipline and Cutoff equilibria yield partitions $\mathscr{P}_D(\chi,-\chi) = [-1,-\chi] \cup [\chi,1]$ and $\mathscr{P}_C(\chi,-\chi) = \{-\chi,\chi\}$, respectively. By Theorem 4, the BP equilibrium produces partition $\mathscr{P}_{BP}(\chi,-\chi) = [a(\chi;n), b(-\chi;n)] \supset [-\chi,\chi]$. It is evident from Figure 6 that, as media polarization increases, $\mathscr{P}_D(\chi,-\chi)$ becomes less informative whereas $\mathscr{P}_{BP}(\chi,-\chi)$ becomes more informative. In both the Discipline and Cutoff equilibria, asymptotic voter welfare equals $\int_\chi^1 k dF(k)$, which declines as media polarization increases. In contrast, the BP equilibrium is always asymptotically efficient, as $0 \in \mathscr{P}_C(\chi,-\chi)$ holds for all $\chi \geq 0$. We summarize these observations in Proposition 8.

Proposition 8. *Let $|M| = 2$ and for all $1 \geq \chi' > \chi > 0$, it holds that:*

1. *$\mathscr{P}_D(\chi',-\chi')$ is less informative than and welfare inferior to $\mathscr{P}_D(\chi,-\chi)$.*
2. *$\mathscr{P}_C(\chi',-\chi')$ is welfare inferior to $\mathscr{P}_C(\chi,-\chi)$.*
3. *$\mathscr{P}_{BP}(\chi',-\chi')$ is more informative than $\mathscr{P}_{BP}(\chi,-\chi)$. Moreover, $\mathscr{P}_{BP}(\chi,-\chi)$ is asymptotically efficient for all $\chi \in (0,1]$.*

6.2.3 Preventing media collusion

Finally, we characterize how media collusion (e.g., media merge) affects information transmission and voter welfare in large elections. Suppose $|M| = 2$ and $\chi_1 > \chi_2$. We assume that, if the two outlets collude, they jointly maximize their average payoff:

$$\bar{V}(k,\chi_1,\chi_2,\Omega) = \frac{V(k,\chi_1,\Omega)+V(k,\chi_2,\Omega)}{2} = \left(k+\frac{\chi_1+\chi_2}{2}\right)\cdot \mathbb{1}_{\Omega=A} \tag{11}$$

In this way, media collusion essentially yields a single "colluded outlet" with an average bias $\chi \equiv \frac{\chi_1+\chi_2}{2}$. Under CT, by Theorem 3, the information partition produced by this colluded outlet equals $\mathscr{P}_D(\frac{\chi_1+\chi_2}{2}) = \mathscr{P}_C(\frac{\chi_1+\chi_2}{2}) = \{\frac{\chi_1+\chi_2}{2}\}$. Hence, in both the Discipline and Cutoff equilibria, media collusion results in an asymptotic voter welfare equal to $\int_{|\frac{\chi_1+\chi_2}{2}|}^1 kdF(k)$.

Suppose instead that collusion is not possible and the two outlets are like-minded. It follows from Section 6.2.1 that the Discipline equilibrium is asymptotically efficient and hence welfare-superior to all other equilibria. In the Cutoff equilibrium asymptotic voter welfare equals $\int_{\min\{|\chi_1|,|\chi_2|\}}^1 kdF(k)$, which is solely determined by the least biased outlet. Because χ_1 and χ_2 have the same sign, it holds that $\min\{|\chi_1|,|\chi_2|\} < |\frac{\chi_1+\chi_2}{2}|$ and hence collusion of two like-minded outlets unambiguously reduces asymptotic voter welfare in either Discipline or Cutoff equilibria. If instead the two outlets are opposite-minded, then (without collusion) asymptotic voter welfare in both Discipline and Cutoff equilibria equals $\int_{\min\{|\chi_1|,|\chi_2|\}}^1 kdF(k)$. Therefore, media collusion improves voter welfare in large elections if and only if $|\frac{\chi_1+\chi_2}{2}| < \min\{|\chi_1|,|\chi_2|\}$, i.e., the colluded outlet is less biased than either of the original outlets. This is possible with two opposite-minded outlets with sufficiently close biases; for example, $\chi_1 = -\chi_2$. Hence, media collusion may either increase or decrease voter welfare under CT.

Finally, in a BP equilibrium, Theorem 4 implies that the colluded outlet generates information partition $\mathscr{P}_{BP}(\frac{\chi_1+\chi_2}{2}) = \left[a(\frac{\chi_1+\chi_2}{2};n), b(\frac{\chi_1+\chi_2}{2};n)\right]$. Without media collusion, $\mathscr{P}_{BP}(\chi_1,\chi_2) = [a(\chi_1;n), b(\chi_2;n)]$. Since $\chi_1 \geq \frac{\chi_1+\chi_2}{2} \geq \chi_2$ and both $a(\chi;n)$ and $b(\chi;n)$ are weakly decreasing in χ, it holds that $\mathscr{P}_{BP}(\frac{\chi_1+\chi_2}{2}) \subset \mathscr{P}_{BP}(\chi_1,\chi_2)$. Media collusion thus yields a strictly less informative and weakly welfare-inferior information partition, compared to that without collusion. We summarize these observations in Proposition 9.

Proposition 9. *Let $|M| = 2$ and $\chi_1 > \chi_2$, it holds that:*

1. *in both Discipline and Cutoff equilibria, media collusion increases asymptotic voter welfare if and only if $|\frac{\chi_1+\chi_2}{2}| < \min\{|\chi_1|,|\chi_2|\}$.*

2. *the BP equilibrium is less informative, and the asymptotic voter welfare weakly decreases if media outlets collude.*

Consistent with Gentzkow and Kamenica (2016, 2017), we find that media collusion can never improve information transmission and voter welfare under BP. Under CT, however, media collusion can be welfare improving if the two media outlets are opposite-minded with similar magnitudes of biases. The collusion of two like-minded media outlets, however, unambiguously reduces voter welfare.

7 Conclusion

We develop a general and tractable theoretical framework to analyze the influence of biased media on voting behavior and election outcomes. A key observation is that news released by mass media allows voters not only to infer the relative appeal of candidates in terms of their qualities, but also to gauge the closeness of elections. In large elections, the former information determines party vote shares and the election outcome, whereas the latter information drives voter turnout through the "competition effect" (Levine and Palfrey, 2007). Our analyses deliver three contributions.

First, we derive robust bounds for the extent to which public mass media can systematically influence voting behavior and election outcomes. We show that, when the electorate is sufficiently polarized, mass media can hardly affect the expected party vote shares from the ex-ante perspective. Nevertheless, such inability does not prevent mass media from systematically manipulating election outcomes and voter turnout. In fact, we show that the opposite is true – by strategically releasing public information, mass media can potentially manipulate both the election outcome and voter turnout to an arbitrarily large extent.

Second, we derive precise comparative static predictions regarding the impact of biased media on election outcomes, voter turnout and welfare. We first study a scenario with a single media outlet and obtain three findings. First, the relationship between media bias and the election outcome is non-monotonic; ex-ante, a candidate's winning probability will first increase and then decrease as the media outlet becomes more biased towards her. Second, from the ex-ante perspective, voter turnout increases in media bias unambiguously. Third, an increased media bias reduces voter welfare by both decreasing the probability of electing the quality-superior candidate and increasing aggregate turnout costs. We also show that introducing an extra media outlet can systematically shift the election outcome and voter turnout in either direction, yet it unambiguously improves voter welfare. These results provide a potential rationale for the mixed empirical evidence on the relationship between media entry and voter turnout.

Third, we study whether increased media competition necessarily improves voter welfare. We explore this issue under three different notions of increased competition: introducing additional media outlets, increasing media polarization, and preventing media collusion. We show that when media outlets can a priori commit to any specific reporting strategy, all three ways of increasing media competition weakly improve voter welfare unambiguously. Without such commitment ability, however, the last two ways of increasing media competition may instead deteriorate voter welfare. We therefore contribute to the literature by highlighting the crucial role of media's ability to commit determining the welfare impact of media competition.

We suggest three avenues for future research. First, instead of reading from all media outlets, in real life voters might be highly selective in media exposure (Gentzkow and Shapiro, 2010, 2011; Durante and Knight, 2012). Exploring how media consumption correlates with voters' characteristics and whether large elections can aggregate dispersed voter information is undoubtedly a salient issue. Second, voters may also have private information beyond the public news provided by mass media (Liu, 2019). Our approach does not directly extend to these contexts since voters no longer have common posterior beliefs, due to idiosyncratic private information. Third, our Poisson game setup precludes communication among voters and turnout mobilization through social norms (e.g., peer pressure), which are empirically relevant in determining political participation (Grosser and Schram, 2006; DellaVigna et al., 2016). The interactions between media outlets' electoral impact with various communication networks and social norms remain insufficiently understood. These questions might be explored by embedding media outlets in a pivotal voter framework allowing for correlated equilibria (Pogorelskiy, 2014), or in a group-based framework allowing for explicitly modeling of social norms (Levine and Mattozzi, 2016).

References

Adena, M., Enikolopov, R., Petrova, M., Santarosa, V., and Zhuravskaya, E. (2015). Radio and the Rise of the Nazis in Prewar Germany. *Quarterly Journal of Economics*, 130(4):1885–1939.

Allcott, H. and Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2):211–36.

Alonso, R. and Câmara, O. (2016). Persuading voters. *American Economic Review*, 106(11):3590–3605.

Battaglini, M. (2002). Multiple referrals and multidimensional cheap talk. *Econometrica*, 70(4):1379–1401.

Battaglini, M., Morton, R., and Palfrey, T. (2008). Information and Pivotal Voter Models in Large Laboratory Elections. *American Economic Review*, 98(2):194–200.

Battaglini, M., Morton, R. B., and Palfrey, T. R. (2010). The swing voter's curse in the laboratory. *Review of Economic Studies*, 77(1):61–89.

Bernhardt, D., Krasa, S., and Polborn, M. (2008). Political polarization and the electoral effects of media bias. *Journal of Public Economics*, 92(5):1092–1104.

Besley, T. and Prat, A. (2006). Handcuffs for the grabbing hand? The role of the media in political accountability. *American Economic Review*, 96(3):720–736.

Cagé, J. (2017). Media Competition, Information Provision and Political Participation: Evidence from French Local Newspapers and Elections, 1944-2014. CEPR Discussion Paper No. DP12198. Available at SSRN: https://ssrn.com/abstract=3016784.

Chan, J. and Suen, W. (2008). A spatial theory of news consumption and electoral competition. *Review of Economic Studies*, 75(3):699–728.

Chan, J. and Suen, W. (2009). Media as watchdogs: The role of news media in electoral competition. *European Economic Review*, 53(7):799–814.

Chen, H. and Suen, W. (2018). Competition for attention in the news media market. Working paper, University of Hong Kong.

Chiang, C.-F. and Knight, B. (2011). Media bias and influence: Evidence from newspaper endorsements. *Review of Economic Studies*, 78(3):795–820.

Crawford, V. P. and Sobel, J. (1982). Strategic information transmission. *Econometrica*, 50(6):1431–1451.

DellaVigna, S. and Kaplan, E. (2007). The Fox News effect: Media bias and voting. *Quarterly Journal of Economics*, 122(3):1187–1234.

DellaVigna, S., List, J. A., Malmendier, U., and Rao, G. (2016). Voting to tell others. *Review of Economic Studies*, 84(1):143–181.

Drago, F., Nannicini, T., and Sobbrio, F. (2014). Meet the press: How voters and politicians respond to newspaper entry and exit. *American Economic Journal: Applied Economics*, 6(3):159–188.

Duggan, J. and Martinelli, C. (2011). A spatial theory of media slant and voter choice. *Review of Economic Studies*, 78(2):640–666.

Durante, R. and Knight, B. (2012). Partisan control, media bias, and viewer responses: Evidence from Berlusconi's italy. *Journal of European Economic Association*, 10(3):451–481.

Ellingsen, S., Hernæs, Ø., et al. (2018). The impact of commercial television on turnout and public policy: Evidence from Norwegian local politics. *Journal of Public Economics*, 159:1–15.

Enelow, J. M. and Hinich, M. J. (1984). *The spatial theory of voting: An introduction*. Boston: Cambridge Univ. Press.

Enikolopov, R. and Petrova, M. (2015). Media capture: Empirical evidence. In *Handbook of media Economics*, volume 1, pages 687–700. Elsevier.

Enikolopov, R., Petrova, M., and Zhuravskaya, E. (2011). Media and political persuasion: Evidence from Russia. *American Economic Review*, 101(7):3253–3285.

Falck, O., Gold, R., and Heblich, S. (2014). E-lections: Voting Behavior and the Internet. *American Economic Review*, 104(7):2238–2265.

Feddersen, T. J. and Pesendorfer, W. (1996). The Swing Voter's Curse. *American Economic Review*, 86(3):408–424.

Feddersen, T. J. and Pesendorfer, W. (1999). Abstention in elections with asymmetric information and diverse preferences. *American Political Science Review*, 93(2):381–398.

Gavazza, A., Nardotto, M., and Valletti, T. M. (2015). Internet and politics: Evidence from UK local elections and local government policies. CEPR Discussion Paper No.DP10991. Available at SSRN: https://ssrn.com/abstract=3039013.

Gentzkow, M. (2006). Television and voter turnout. *Quarterly Journal of Economics*, 121(3):931–972.

Gentzkow, M. and Kamenica, E. (2016). Competition in persuasion. *Review of Economic Studies*, 84(1):300–322.

Gentzkow, M. and Kamenica, E. (2017). Bayesian persuasion with multiple senders and rich signal spaces. *Games and Economic Behavior*, 104:411–429.

Gentzkow, M. and Shapiro, J. M. (2006). Media bias and reputation. *Journal of Political Economy*, 114(2):280–316.

Gentzkow, M. and Shapiro, J. M. (2008). Competition and Truth in the Market for News. *Journal of Economic Perspectives*, 22(2):133–154.

Gentzkow, M. and Shapiro, J. M. (2010). What drives media slant? Evidence from us daily newspapers. *Econometrica*, 78(1):35–71.

Gentzkow, M. and Shapiro, J. M. (2011). Ideological segregation online and offline. *Quarterly Journal of Economics*, 126(4):1799–1839.

Gentzkow, M., Shapiro, J. M., and Sinkinson, M. (2011). The effect of newspaper entry and exit on electoral politics. *American Economic Review*, 101(7):2980–3018.

Gentzkow, M., Shapiro, J. M., and Stone, D. F. (2015). Media bias in the marketplace: Theory. In *Handbook of media economics*, volume 1, pages 623–645. Elsevier.

Gentzkow, M., Shapiro, J. M., and Taddy, M. (2019). Measuring Group Differences in High-Dimensional Choices: Method and Application to Congressional Speech. *Econometrica*. Forthcoming.

Gerber, A. S., Karlan, D., and Bergan, D. (2009). Does the media matter? A field experiment measuring the effect of newspapers on voting behavior and political opinions. *American Economic Journal: Applied Economics*, 1(2):35–52.

Gilligan, T. W. and Krehbiel, K. (1989). Asymmetric information and legislative rules with a heterogeneous committee. *American Journal of Political Science*, pages 459–490.

Ginzburg, B. (2019). Optimal information censorship. *Journal of Economic Behavior and Organization*, 163:377–385.

Groseclose, T. and Milyo, J. (2005). A measure of media bias. *Quarterly Journal of Economics*, 120(4):1191–1237.

Grosser, J. and Schram, A. (2006). Neighborhood information exchange and voter participation: An experimental study. *American Political Science Review*, 100(2):235–248.

Guo, Y. and Shmaya, E. (2019). Costly miscalibration. Working paper, Northwestern University.

Kamenica, E. and Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review*, 101(6):2590–2615.

Kennedy, P. J. and Prat, A. (2019). Where Do People Get Their News? *Economic Policy*.

Kolotilin, A., Mylovanov, T., Zapechelnyuk, A., and Li, M. (2017). Persuasion of a privately informed receiver. *Econometrica*, 85(6):1949–1964.

Krishna, V. and Morgan, J. (2001a). Asymmetric information and legislative rules: Some amendments. *American Political Science Review*, 95(2):435–452.

Krishna, V. and Morgan, J. (2001b). A model of expertise. *Quarterly Journal of Economics*, 116(2):747–775.

Krishna, V. and Morgan, J. (2011). Overcoming ideological bias in elections. *Journal of Political Economy*, 119(2):183–211.

Krishna, V. and Morgan, J. (2015). Majority Rule and Utilitarian Welfare. *American Economic Journal: Microeconomics*, 7(4):339–375.

Larcinese, V., Puglisi, R., and Snyder, J. M. (2011). Partisan bias in economic news: Evidence on the agenda-setting behavior of US newspapers. *Journal of Public Economics*, 95(9):1178–1189.

Ledyard, J. O. (1984). The pure theory of large two-candidate elections. *Public choice*, 44(1):7–41.

Levine, D. K. and Mattozzi, A. (2016). Voter Turnout with Peer Punishment. Working paper, European University Institute.

Levine, D. K. and Palfrey, T. R. (2007). The paradox of voter participation? A laboratory study. *American Political Science Review*, 101(1):143–158.

Liu, S. (2019). Voting with Public Information. *Games and Economic Behavior*, 113:694–719.

Lu, S. E. (2017). Coordination-free equilibria in cheap talk games. *Journal of Economic Theory*, 168:177–208.

Martin, G. J. and Yurukoglu, A. (2017). Bias in Cable news: Persuasion and polarization. *American Economic Review*, 107(9):2565–99.

Matsusaka, J. G. (1995). Explaining voter turnout patterns: An information theory. *Public Choice*, 84(1-2):91–117.

McMillan, J. and Zoido, P. (2004). How to subvert democracy: Montesinos in Peru. *Journal of Economic Perspectives*, 18(4):69–92.

Myerson, R. B. (1998). Population uncertainty and Poisson games. *International Journal of Game Theory*, 27(3):375–392.

Myerson, R. B. (2000). Large Poisson games. *Journal of Economic Theory*, 94(1):7–45.

Palfrey, T. R. and Rosenthal, H. (1985). Voter participation and strategic uncertainty. *American Political Science Review*, 79(1):62–78.

Perego, J. and Yuksel, S. (2018). Media competition and social disagreement. Working paper.

Pew Research Center (2017). Political Polarization, 1994-2017. Washington, DC: Pew Research Center. http://www.people-press.org/interactives/political-polarization-1994-2017/.

Piolatto, A. and Schuett, F. (2015). Media competition and electoral politics. *Journal of Public Economics*, 130:80–93.

Pogorelskiy, K. (2014). Correlated equilibria in voter turnout games. APSA 2014 Annual Meeting Paper. Available at SSRN: https://ssrn.com/abstract=2451740.

Prat, A. (2018). Media power. *Journal of Political Economy*, 126(4):1747–1783.

Prat, A. and Strömberg, D. (2013). The political economy of mass media. *Advances in Economics and Econometrics*, 2:135.

Puglisi, R. and Snyder, J. M. (2011). Newspaper coverage of political scandals. *Journal of Politics*, 73(3):931–950.

Puglisi, R. and Snyder, J. M. (2015a). The balanced US press. *Journal of European Economic Association*, 13(2):240–264.

Puglisi, R. and Snyder, J. M. (2015b). Empirical studies of media bias. In *Handbook of media economics*, volume 1, pages 647–667. Elsevier.

Ravid, E. and Doron, L. (2018). Cheap Talk with Transparent Motives. Working paper. Available at SSRN: https://ssrn.com/abstract=2941601 or http://dx.doi.org/10.2139/ssrn.2941601.

Rayo, L. and Segal, I. (2010). Optimal information disclosure. *Journal of Political Economy*, 118(5):949–987.

Rubanov, O. (2014). Asymptotic Full Revelation in Cheap Talk with Large Number of Senders. Working paper. Available at SSRN: https://ssrn.com/abstract=2479293 or http://dx.doi.org/10.2139/ssrn.2479293.

Smith, A. and Anderson, M. (2018). Social Media Use in 2018. Pew Internet and Technology, March 1, 2018. Accessed at http://www.pewinternet.org/2018/03/01/social-media-use-in-2018/ on September 11, 2018.

Snyder, J. M. and Strömberg, D. (2010). Press coverage and political accountability. *Journal of Political Economy*, 118(2):355–408.

Sobbrio, F. (2014). Citizen-editors' endogenous information acquisition and news accuracy. *Journal of Public Economics*, 113:43–53.

Strömberg, D. (2004). Radio's impact on public spending. *Quarterly Journal of Economics*, 119(1):189–221.

Taneva, I. A. (2019). Information Design. *American Economic Journal: Microeconomics*. Forthcoming.