USC FBE APPLIED ECONOMICS WORKSHOP
presented by: Sevgi Yuksel
Friday,April 15, 2016
1:30 pm - 3:00 pm; Room: **HOH-506**

Media Competition and the Source of Disagreement

Jacopo Perego
New York University

Sevgi Yuksel
UC Santa Barbara

July, 2015

Abstract

We identify a novel channel through which increased competition among information providers decreases the efficiency of electoral outcomes. A number of profit-maximizing firms compete to sell information to a group of Bayesian agents about how two political candidates compare on several issues. Voters can disagree on which issues are important to them (*agenda*) and on how each issue in their agenda should be addressed (*slant*). We show that competition forces firms to differentiate the *type* of information they produce. In particular, differentiation leads to higher provision of information on issues where there is higher disagreement in the electorate. Although voters become individually better informed, the share of votes going to the socially optimal candidate decreases. We also show that this inefficiency is magnified if there is higher polarization in the underlying preferences of the society.

JEL Classification Numbers: D72, L15, L82.

Key Words: Information, Media, Competition, Voting, Spatial models.

We thank Michael Castanheira, Severine Toussaert and seminar participants at SAET Conference in Cambdrige 2015, Stony Brook Conference 2015 and New York University for helpful comments. We thank Guillaume Frechette, Alessandro Lizzeri and Debraj Ray for their guidance and support.
Contacts: jacopo.perego@nyu.edu and sevgi@nyu.edu.

1. Introduction

The past two decades have brought significant changes to the market for news. With the rise in the number of TV channels and the dramatic expansion in Internet access, there is unprecedented number of news sources available to a growing share of the population.[1] How increased competition in the market for news affects voters beliefs and consequently the political process has been a highly debated issue and is of fundamental interest to policy makers.

The traditional perspective on this matter has been that truth prevails in a competitive market for news. According to this view, a "marketplace for ideas" promotes truth, because it brings out a diversity of perspectives which allows people to learn more effectively. In a ruling of the US Supreme Court, it is stated that "the widest possible dissemination of information from diverse and antagonistic sources is essential to the welfare of the public," (Associated Press v. United States, 1945). This viewpoint is still considered as "one of the basic tenets of national communications policy" (Federal Communications Commission (2003)).

Although the primacy of free speech is undisputed among scholars in all fields, it has been noted that forces that characterize competition in the media environment are very different from those described in traditional models. Baker (1978) states this position clearly: "the assumptions on which the classic marketplace of ideas theory rests are almost universally rejected". For these reasons, a more critical view has been instrumental in shaping competition policy in the media market, leading in many instances to antitrust exemptions on ownership regulation, price setting and advertisement, or explicit subsidies for news organizations, precisely to reduce, at least to some extent, the pressures of competition.

Overall, it is today widely accepted that the welfare consequences of media competition are far from obvious. For example, although the last two decades have brought about a significant expansion in the range of news sources available voters, we have also witnessed a dramatic polarization in beliefs of the American public. People that self-identify themselves as Republicans and Democrats are more divided along ideological lines than at any point in the last two decades with growing antipathy in attitudes towards the opposing party. Moreover, these divisions are greatest among those who are the most engaged and active in the political process.[2] Hence, at a first look, it appears as if the increased availability of news outlets in the market for news has not been associated with any type of convergence in political views. If anything, there is growing disagreement.

In this paper, we identify a novel channel through which increased competition in the market for news can intensify disagreement in the society and decrease the efficiency of electoral outcomes. Greater competition among news sources leads to more ideological voting and

[1] Pew Research Center, Nielson report on Media.

[2] According to a 2014 survey by Pew Research Center, "the share of Americans who express consistently conservative or consistently liberal opinions has doubled over the past two decades, from 10% to 21%."

this implies that voting decisions become less and less correlated with the welfare optimal candidate. In our model, news sources are profit-maximizers, uninterested in the outcome of the election, and agents are expected utility-maximizers, sharing common ex-ante beliefs about the candidates. We depart from most of the existing literature in that our model does not feature partisan media or behavioral voters, ingredients that by themselves could introduce inefficiencies to the model. We show that increased competition in the media market forces firms to differentiate their informational products. Such differentiation leads to higher provision of information on issues that are prone to induce disagreement among the electorate. Although voters become individually better informed, in the aggregate the share of votes going to the welfare optimal candidate decreases.

The critical insight in our paper is about how competition affects the *type* of information that is provided to voters in an environment where there is underlying heterogeneity in voter preferences. Voters seek to learn about candidates' characteristics which are fundamentally multidimensional. Some of these characteristics can be referred to as *valence*; on these dimensions voters have identical preferences. Others characteristics are identified as *ideology*; on these voters have conflicting preferences. For example, valence issues might include the technical competence of the candidate or her international credibility, whereas ideological issues might include the candidate's position on financial regulation, affirmative action, marriage, immigration etc. Voters' ideological preferences are heterogeneous in two fundamental ways. Voters can disagree on what's important to them (their *agenda*) and on how each issue in their agenda should be addressed (their *slant*). The interaction between these two types of heterogeneity creates a society where the *ideological distance* between any two voters cannot be reduced to a discrete metric (either two types are far apart or are the same). The idea that voters can have different preferences about different issues is not new and yet has not been emphasized in the literature.[3][4] Accounting for this type of richness creates a *space* on which firms differentiate and specialize with competition, becoming the drivers of the inefficiency we described before.

We assume that each news source faces constraints on how jointly informative it can be on all these dimensions. Either for reasons of costs, limited space or limited time, news sources can only provide a partially informative account about each candidate. Hence, an important choice for firms pertains to how much content is to be allocated to each of these underlying issues (*editorial position*). Constraints on content imply that voters with heterogeneous preferences seek out heterogeneous informational products. For example, a voter who primarily cares about economic issues will be more interested in learning about a candidate's

[3]There are few recent exceptions such as Aragones et al. (2015), Dragu and Fan (2015) and Yuksel (2015). These papers are similar to ours in that they allow voters to have different preferences about what is important to them, but they focus on different aspects of political competition. The first two papers study the behavior of two competing parties trying to strategically "prime" the electorate on the issues that are more convenient to them. The last paper studies polarization of party platforms.

[4]This idea goes back to the seminal paper of Stokes (1963) in which the author recognizes that "what is needed is a language that would express the fact that different weights should be given different dimensions at different times."

position on financial regulation, while another voter who is primarily interested in social issues would like to hear about her position on affirmative action. In our model, competition affects firms' incentives to cater to the distribution of preferences in the population. We solve for the equilibrium in the game where firms choose which information structure to offer to maximize their readership, and voters choose which news source to consume to maximize their expected utility. The equilibrium of this game induces a random mapping from the unknown characteristics of the candidates to the share of votes going to each one of them. Our efficiency benchmark is the decision that a social planner - whose objective is to maximize aggregate voter utility - would take could he observe the unknown quality of the candidate.

In equilibrium, the type of information provided by any news sources is chosen optimally to target a specific group of voters. Hence, how much information is provided on ideology relative to valence by a specific news source depends on how much overlap there is in the type of ideological information that is of interest to the consumers of this news source. Competition leads to segmentation: the share of the market that can be targeted by any firm decreases with the number of firms in the market. This creates incentives for news sources to differentiate even more their product choices, leading to overall more information on ideological issues. It is important to note that competition generates more information on ideological issues in two ways. First, competition leads to differentiation and specialization in the types of ideological information that is provided. Some news sources focus more on information relating to social issues, while others focus more on economic issues. Second, all news sources shift focus from valence to ideological information as the market gets segmented. For example, focus shifts from how technically competent a candidate is, a valence issue, to discussions on her positions on affirmative action, same-sex marriage or financial regulation, etc.

Differentiation of news sources in terms of content allow voters to be better informed about how they should vote. The main point here is that, on an individual level, the impact of increased competition is exactly what one could expect in a competitive market. Competition generates differentiation in the type of information provided, creating a spectrum of options for consumers. This enables voters to find news sources which provide the type of ideological information tailored towards their needs. In a sense, competition in our model endogenously generates what political scientists have often referred to as the "Daily Me": an fictitious newspaper crafted around one's unique tastes.[5] However, the welfare effects of media competition extend beyond information acquisition on an individual level. Agents condition their voting decision on the information they receive from the media. They fail to internalize the effects that their choices have on whole society via the election outcome. Since competition in the media market generates more information about the characteristics of candidates along dimensions with stronger disagreement across the population, voting becomes increasingly ideological. Our main result shows that, as the number of firms in the market increases, the share of votes going to the the highest-valence candidate monotonically

[5] For example, see Sunstein (2001).

declines.

We also show that the inefficiency described above is exacerbated by polarization of political preferences. We measure preference polarization in a simple way. We look at how much weight the electorate puts on ideology relative to valence in forming their political preferences. The equilibrium behavior of firms naturally depends on the preferences of the electorate. As polarization increases, the value of information on ideology increases for all voters. This generates strong incentives for news sources to specialize in ideological information. For any number of firms in the market, as polarization increases, the share of votes going to the highest-valence candidate declines.

Our paper also contains a methodological contribution. We introduce a simple type-space that allows us to keep track in a convenient way of the two-level heterogeneity in voters' preferences (*agenda* and *slant*) described above. By normalizing voters' utility, we map the distribution of voter preferences to a circle where the arc running between any two types of voters measures the correlation in their preferences. For example, voters that are on opposite ends of a circle have ideological views that are perfectly negatively correlated. This implies that they have a similar agenda, but they address these issues in opposite way. Similarly, voters that are located in orthogonal positions have independent ideological preferences. These refer to voters with different agendas. Mapping the population of voters onto a circle provides a framework in which all sources of heterogeneity among voters can be reduced to the correlations in their political views. Ultimately, it allows us to interpret and examine the game among firms as a spatial competition model.

The paper is summarized as follows. In Section 2, we review the related literature. The model is introduced in Section 3. We proceed by characterizing the voters' problem in Section 4 and the game among information providers in Section 5. In Section 6, we present the main results of the paper. Finally, Section 7 discuss robustness and extensions, while Section 8 concludes.

2. Related Literature

There is extensive empirical evidence showing the effects that media can have on political attitudes and electoral outcomes.[6] There is also a growing political and economic literature investigating the potential welfare consequences of media competition. We refer the reader to Gentzkow and Shapiro (2008) and Gentzkow et al. (2014) for a partial review of the literature. The main arguments for how media competition can be welfare increasing rely on how competition can alleviate distortions on the supply side of the market and can be summarized as follows. First, increasing the number of news sources can make it easier for news sources to remain independent when there is threat of government capture (Besley and Prat

[6] Among the others, Stromberg (2004), Gentzkow and Shapiro (2004), Gentzkow (2006), Dellavigna and Kaplan (2007), Gerber et al. (2009), Gentzkow et al. (2011).

(2006)). Second, competition promotes truth by generating a spectrum of news with diverse viewpoints in situations where firms may have incentives other than accurately reporting the truth. In this context, Milgrom and Roberts (1986) and more generally Gentzkow and Kamenica (2015a,b) study impact of competition on information revelation in the context of persuasion games. They identify conditions (on available information structures, distribution of preferences for the firms, etc.) under which equilibrium outcomes are more informative under the competitive regimes.[7] Finally, in addition to these forces, it has been argued that competition among information providers can affect investment in faster, higher quality news.

The impact of competition is less likely to be beneficial when there is demand driven bias in news. Demand driven bias results from the incentives of the news sources to pander to its readers expectations. Mullainathan and Shleifer (2005) analyze a model where readers have a preference for news sources that confirm their prior beliefs. Media outlets confront the same trade-off between catering to the readers priors and providing them with better information.[8][9] As the authors demonstrate, when voters suffer from confirmation-bias, increasing competition can sometimes exacerbate bias in information provision by allowing consumers to self-segregate more effectively in terms of prior biases.[10] Similarly, in the model of Bernhardt et al. (2008), media consumers prefer newspapers that withhold unfavorable information about the party they support. But catering to this preference is socially costly since voters become less informed and elections are less likely to correspond to the efficient outcome. Another possible source of demand-driven distortion is that consumers value politically relevant information less than a social planner would. Similar to the case with confirmation bias, competition can have detrimental effects by allowing self-segregation to news sources which shift focus from "hard news" to "soft news" (entertainment, sports, etc.). Empirically, the effect of this is mixed. Prat and Stromberg (2005) provide evidence that the introduction of private television in Sweden increased political information and political participation relative to a public television monopoly. On the other side, Cage' (2014), using a county-level panel dataset of local newspaper presence and political turnout in France from 1945 to 2012, finds newspaper entry to be associated with a decline in information provision, and to ultimately decrease voter turnout.

[7] Competition does not mitigate all supply-driven bias in the media market. For example, Baron (2006) studies media bias resulting from the ideological bias of reporters/editors and shows that bias can be greater with competition than with a monopoly news organization.

[8] Gentzkow and Shapiro (2006) show that the incentive to conform to the readers prior expectations can arise endogenously when there is uncertainty about the quality of news sources. Media bias is observed in equilibrium as slanting news increases reputation for the news sources even though it makes all market participants worse off.

[9] Gentzkow and Shapiro (2010) construct a new index of media slant that measures the similarity of a news outlets language to that of a congressional Republican or Democrat. They show that for US newspapers most of the slant is driven by the demand side.

[10] Sunstein (2002) has argued that the availability of a vast number of news sources via the Internet can intensify this problem to the extent that news turn into echo-chambers, where citizens only hear news precisely in line with their priors and there is no effective learning.

Our paper differs from these models in that we abstract from these kinds of distortions both in the supply-side and the demand-side. In our model, on the supply side, firms are profit-maximizer, non-biased, non-partisan. On the demand side, voters are Bayesian expected utility maximizers. As a consequence, in contrast to the existing literature, our model predicts competition to bring about a *more* informed electorate.[11] Crucially, the inefficiency identified in our model is not due to a failure in information provision, but stems from how competition causes a shift in the *type* of information that is revealed in the news market.

In this sense, our paper also relates to a literature studying the interaction between ideological and valence issues in political competition. Alesina et al. (1999); Besley and Prat (2006); Lizzeri and Persico (2001) and Fernandez and Levy (2008) study this interaction by investigating how preference heterogeneity affects public good provision. Besley and Prat (2006) examines how increased voter ethnicization, defined as greater voter preference for the party representing her ethnic group, affects legislator quality. Alesina et al. (1999) present a model that links heterogeneity of preferences across ethnic groups in a city to the amount and type of public goods the city supplies. Relatedly, Ashworth and de Mesquita (2009); Carillo and Castanheira (2008); Eyster and Kittsteiner (2007); Groseclose (2001) present models which study the interaction between valence competition and party platforms. The closest to our model, Ashworth and de Mesquita (2009) study a game in which candidates first choose platforms and then invest in costly valences (e.g., engage in campaign spending). The marginal return to valence depends on platform polarization: the closer platforms are, the more valence affects the election outcome. The common insight in these papers is that preference heterogeneity, either intrinsic or accentuated through platform divergence, will hurt valence competition by decreasing its importance in individual voting decisions. Our paper studies how competition in the media market affects this force.

Our paper also relates to spatial competition models pioneered by Hotelling (1929) and Salop (1979). Few papers have incorporated insights from this literature to the media market. In Chan and Suen (2008), voters who are constrained in their information processing abilities choose media outlets to maximize the value of information. They show that voter welfare is typically higher under a duopoly than under a monopoly because in a competitive market, the two firms differentiate and provide two diverse viewpoints leading to more information revelation. Duggan and Martinelli (2011) develop a theory of media slant as a systematic filtering of political news that reduces multidimensional politics to the one-dimensional space perceived by voters. They do not solve for the equilibrium with multiple news sources, but characterize socially optimal slant when there is only one news source. Gul and Pesendorfer (2012) present a mechanism where media competition (via specialization) increases divergence in party platforms. In their model, competition also leads to ideological segmentation of news sources, but this is beneficial for information revelation because voters are assumed to have limited information processing capacity which is offset by the ideological bias of the media sources. Although our paper shares some common elements with these papers, the

[11]Note that this should have testable consequences possibly in terms of turnout, campaign spending, polarization of partisanship, etc.

type of differentiation and segmentation we model in a competitive media market, and the source of inefficiency associated with competition identified in our main result is completely novel. In addition, to our knowledge, we are to first to characterize competitive outcomes in the media market for an arbitrary number of firms allowing us to study the impact of competition for any size of the market.

3. Model

3.1 Candidates and Voters' Heterogeneity

We consider two political candidates, A and B, running for office. Each candidate is born with an ex-ante unknown type. We focus on $\boldsymbol{\theta} := (\theta_v, \theta_{id})$ which expresses the relative comparison of candidate A to candidate B on different issues, θ_v and θ_{id}.

Let T be the set of voters, a compact interval on the real line. Each $t \in T$ denotes the *type* of a voter. Given $\boldsymbol{\theta}$, the utility function $u(\boldsymbol{\theta}, t)$ represents how type t evaluates candidate A relative to B. We assume that u takes the following simple form form:

$$u(\boldsymbol{\theta}, t) := \lambda\theta_v + (1-\lambda) f(\theta_{id}, t).$$

The function f is linear in $\boldsymbol{\theta}$; and $\lambda \in (0,1)$ represents how voters of type t trade off one issue with the other.

The first component, θ_v, enters linearly and is *independent* of the type t. That is, all voters have identical preferences about dimension θ_v. As customary, we refer to this dimension as *valence*, Stokes (1963). The second component θ_{id} enters linearly, through f, but it is type-*dependent*. That is we potentially allow voters to have heterogeneous preferences on dimension θ_{id}. Accordingly, we refer to this dimension as *ideology*, Downs (1957).[12] Notice also that, since we haven't made assumption on f yet, it is with no loss of generality to assume that voters are distributed uniformly on T, $t \sim \mathcal{U}(T)$.

The heterogeneity of voters' preferences clearly depends on the functional form of f. For example, in the extreme case, if f is constant in t, voters' preferences are not heterogeneous at all. Since voters' heterogeneity is going to be one of the crucial ingredient in our model, it is important to find a convenient way to account for such heterogeneity. Fixing $t \in T$, we can define $\rho_t : T \to [-1, 1]$ as $\rho_t(t') := \text{Corr}(f(\theta_{id}, t), f(\theta_{id}, t'))$. The function ρ_t measures how correlated type $t's$ ideological preferences are relative to any other type t'. In a sense, $\rho_t(t')$ is a measure of *ideological distance*. If $\rho_t(t') = 1$, the ideological preferences of type t and t' are completely aligned. In fact, for any realization of θ, types t and t' have identical political views. In contrast, if $\rho_t(t') = -1$, the preferences of type t and t' are perfectly misaligned.

[12] Notice that due to the linearity of the utility function $u(\cdot, t)$, it is without loss of generality to focus on the relative difference between the two candidates, i.e. $\boldsymbol{\theta} := \boldsymbol{\theta}^A - \boldsymbol{\theta}^B$. The vector $\boldsymbol{\theta}$ expresses how candidate A fares relative to candidate B.

Any change in θ_{id} that makes type t better off, it would also make type t' worse off. Finally, if $\rho_t(t') = 0$, the ideological preferences of t and t' are independent. This could correspond to a situation in which the two types employ completely different criteria to compare the two candidates along the dimension θ_{id}. For instance, it could be that type t only cares about economic issues, while t' cares about issues social issues.

The distribution of ρ_t, which we call $P_t \in \Delta([-1, 1])$, measures how "similar" the preferences of t are to those of the rest of the population. P_t captures how likely it for voter t' to find another voter whose ideological preferences are identical to his, orthogonal, completely opposite etc. Similarly, comparing the distributions P_t to $P_{t'}$ is informative about how representative the preferences of type t are of the population relative to t'. Different P_t's produce models in which voters are heterogeneous in fundamentally different ways.

Some of the traditional models of heterogeneous voters preferences can be described in terms of properties on P_t. For example, in models where voters are homogeneous we have that $\text{supp } P_t = \{1\}$ for all $t \in T$. This in fact implies that all voters are in full agreement on how to evaluate the candidates. This is a model of *pure valence.* Alternatively, suppose that for all $t \in T$, $\text{supp } P_t = \{0\}$. In model of this sort, voters have heterogeneous preferences, since their preferences on the ideological dimension are not aligned. However, their ideological tastes would be simple *idiosyncratic noise.* Finally, suppose $\text{supp } P_t = \{-1, 1\}$ for all $t \in T$, This implies that effectively there are only two groups of voters and that there is perfect disagreement between them. This dichotomy in political preferences is peculiar of models that feature *Downsian ideology.*

In this paper, we allow P_t to describe a more comprehensive form of heterogeneity. In our model the distribution P_t will satisfy two attractive properties:

$$\textbf{(Anonymity)}\quad P_t = P \text{ for all } t \in T \qquad \text{and} \qquad \textbf{(Richness)}\quad \text{supp } P = [-1, 1]$$

The first property, *anonymity*, states that the distribution of "ideological distances" is identical for any voter: no ideological position is special. There is no such a thing as a more centrist position or a more extreme one. The second property *richness* directly generalizes the models above. The fact that P_t has full support means that, for each type t, one can find voters that are arbitrary "close" to t - in terms of their ideological preferences - as well as types that are arbitrary "far away."

To produce a model that satisfies the two properties above, we make appropriate assumptions on f and $\boldsymbol{\theta}$. In particular, we assume that $\theta_{id} \in \mathbb{R}^2$. That is, the ideological component of a candidate can indeed be decomposed further into two - more primitive - ideological sub-components, ϑ_1 and ϑ_2, the combination of which generates $f(\theta_{id}, t)$. We assume θ_v, and each component of θ_{id} to be independently distributed according to a normal distribution with mean zero and unit variance.[13] The way ϑ_1 and ϑ_2 are *mixed* depends on the type of the voter. In particular we assume:

[13]It is without loss of generality to assume that the ideological dimensions ϑ_1 and ϑ_2 have mean zero. In reality, it is likely that there are *expected* differences between candidate A and B. In such case, the model

Assumption 1. *For all $t \in T$: $f(\theta_{id}, t) := \vartheta_1 \cos(t) + \vartheta_2 \sin(t)$.*

The interpretation for this specification is straightforward. People are different in two levels. They can disagree on what's important for them (*agenda*) and on how each issue in their agenda should be addressed (*slant*). For example, for some voters, the candidates' position on affirmative action can be extremely important in determining their voting behavior. Others, instead, may have little interest in this issues. Among those who care about affirmative action, there can be voters who are for or against it. Our model allows voters to disagree not only on whether ϑ_1 is good or bad, in absolute terms. But also on how important ϑ_1 is relative to ϑ_2. This is a crucial feature of our model because, as it will become clear later on, it generates the *the space* on which information providers can diversify their products.

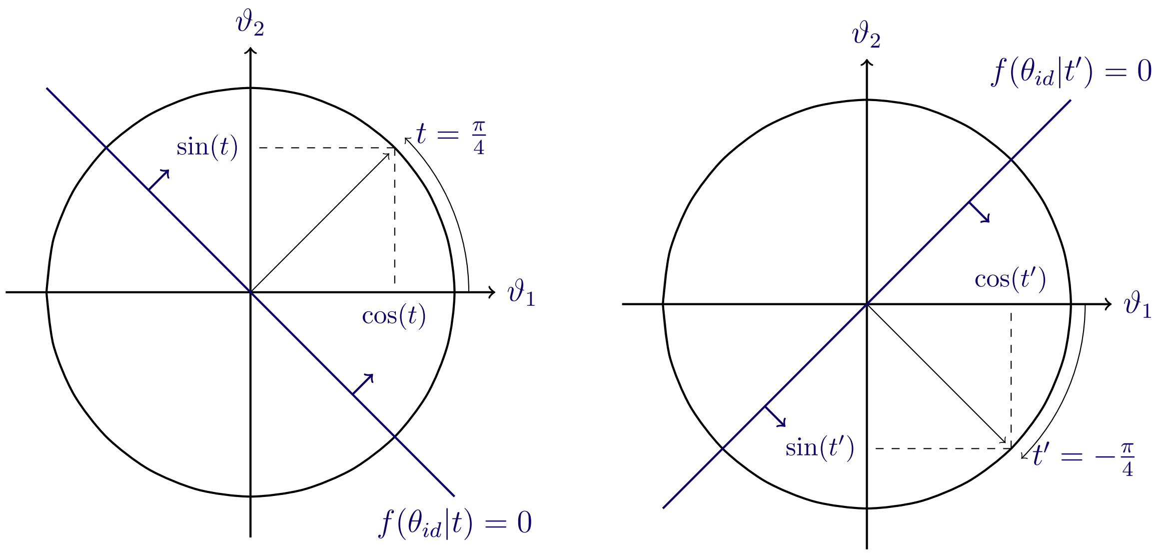

FIGURE 1: The level curves of $f(\theta_{id}, t)$ for two voters: $t = \frac{\pi}{4}$ and $t' = -\frac{\pi}{4}$.

This specification has several nice features. First, it guarantees that the variance of $u(\boldsymbol{\theta}, t)$ is independent of t.[14] This is important because it ensures that all voters *ex ante* value information the same amount. Second, as we show in the appendix, this specification ensures that $\rho_t(t') = \cos(t' - t)$. The correlation in ideological preferences of any two types t and t' can be easily measured as their distance on the circumference of a circle. The farther away they are, the smaller the correlation in their ideological preferences. In line with the richness property, we get a continuum of possible correlations between different types, enriching the heterogeneity among voters preferences. In particular, for every t, there are a pair of types

could still be solved in a very similar fashion by including a type-dependent constant term in $u(\boldsymbol{\theta}, t)$. The random variables ϑ_1 and ϑ_2 would be interpreted as the *residual* uncertainty about how the candidates fare relative to each others.

[14]This follows from $f(\theta_{id}, t) \sim \mathcal{N}(0, 1)$ for any t, which is a consequence of $\cos^2 t + \sin^2 t = 1$ for any t.

$t \pm \frac{\pi}{2}$ who have *orthogonal* preferences to t - meaning that making $t \pm \frac{\pi}{2}$ happier can leave t indifferent - and a type $t \pm \pi$ that has *opposite* preferences to t - meaning that making $t \pm \pi$ happier necessarily makes t unhappy.

Summing up, the utility specification that we will use in the rest of the paper is

$$u(\boldsymbol{\theta}, t) := \lambda\theta_v + (1-\lambda)\Big(\vartheta_1 \cos(t) + \vartheta_2 \sin(t)\Big).$$

We will refer to θ_v as the valence dimension and to ϑ_1 and ϑ_2 as the ideological dimensions. For all voters, the valence dimension receives weight λ. We think of $1-\lambda$ as a simple reduced-form parameter that measures how ideologically *polarized* a society is. If λ is high, the society puts little weight on ideology, hence is relatively homogeneous. Vice versa, when λ is low, the society puts a higher weight to ideology and for this reason it is relatively more polarized.

3.2 Information Providers

There are n information providers, or news sources, who are competing for readership. Voters pick among these news sources which are offering information structures. An information structure associated with a news source comprises of two independent signals, one for valence and one for ideology. In choosing how accurate these signals are, we assume the news sources to face a trade-off: they cannot increase the informativeness of one signal without reducing the informativeness of the other.

When sending the signal for ideology the information provider needs to decide what mixture of news about ϑ_1 and ϑ_2 to provide. For example, one news source could decide to focus uniquely on ϑ_1, whereas another could do the same with ϑ_2. Effectively, by choosing what to talk about, each news source *targets* specific types of voters.

In summary, each news sources chooses a *precision* $\tau \in [0,1]$ and a *position* $x \in T$. This implies that the news source generates two *independent* signals:

$$s_v \sim \mathcal{N}\Big(\theta_v, \tau^{-1}\Big) \quad \text{and} \quad s_{\text{id}} \sim \mathcal{N}\Big(f(\theta_{id}, x), (1-\tau)^{-1}\Big).$$

The interpretation for these restrictions is that news sources are somewhat limited in the amount of information they can communicate to the readers. Such limitations can be justified in several ways: the most natural being restrictions on space or time, both from the supply (in the production of news) and from the demand side (in the consumption of news). Our results do not depend on specific bounds for total precision, and we normalize it at $\tau \leq 1$ for notational convenience.

3.3 Social Planner

In our model, voters do not observe the realization of $\boldsymbol{\theta}$, and gather information from competing information providers to form their political views. We contemplate two different efficiency benchmarks: first-best and second-best. In the former, the social planner is perfectly informed, i.e. she knows $\boldsymbol{\theta}$ and selects the candidate that maximize aggregate voters' welfare.

Definition 1. *(First-Best) The social planner decision's rule is a function $r^{SP} : \Theta \rightarrow \{0, 1\}$ defined as follows:*

$$r^{SP}(\boldsymbol{\theta}) = \begin{cases} 1 & \text{if } \frac{1}{2\pi}\int_T u(\theta, t)dt > 0 \\ 0 & \text{else.} \end{cases}$$

The social planner selects the candidate A if and only if the total welfare generated by A, namely $\frac{1}{2\pi}\int_T u(\boldsymbol{\theta}, t)dt$, is strictly positive. Proposition 1 below captures the fundamental tension between ideology and valence. Notice that the solution to the social planner's problem is independent of $\theta_{id} = (\vartheta_1, \vartheta_2)$, the dimensions along which voters disagree. The socially optimal solution is for candidate A to be selected if and only if candidate A compares better than candidate B on the first dimension, θ_v, i.e. the valence dimension. This is because, the voters have symmetrically heterogeneous tastes on ideology. "Favoring" one voters necessarily implies "harming" another. The easiest way to see this is to consider *polar types*: $(t, t+\pi)$, types with ideological views that are perfectly negatively correlated. For any realization of $\boldsymbol{\theta}$, $f(\theta_{id}, t) = -f(\theta_{id}, t + \pi)$. This implies that ideological preferences cannot be part of a welfare calculation among such types.

Proposition 1. *(Efficiency) The social planner's solution is to select candidate A iff candidate A compares better than B on the valence dimension, i.e.,*

$$r^{SP}(\boldsymbol{\theta}) = 1 \quad \text{iff} \quad \theta_v > 0$$

However, in light of the restrictions we imposed on the information structures available to the news sources, there is an another, equally interesting, efficiency benchmark that we can consider. Under this benchmark, the social planner cannot observe $\boldsymbol{\theta}$, but she has access to the same *technology* that is available to the information providers. Since θ_v is still the only socially valuable dimension, the optimal choice for the social planner is to produce a signal $s_v(\boldsymbol{\theta})$ that is maximally informative about θ_v, by setting $\tau = 1$. The next Definition builds on this intuition.

Definition 2. *(Second-Best) The constrained-efficient decision rule is a random variable $r^{SB} : \Theta \rightarrow \{0, 1\}$ defined as follows:*

$$r^{SB}(\boldsymbol{\theta}) = \begin{cases} 1 & \text{if } \frac{1}{2\pi}\int_T u(s_v(\boldsymbol{\theta}), t)dt > 0 \\ 0 & \text{else.} \end{cases}$$

with $s_v(\boldsymbol{\theta}) \sim \mathcal{N}(\theta_v, 1)$.

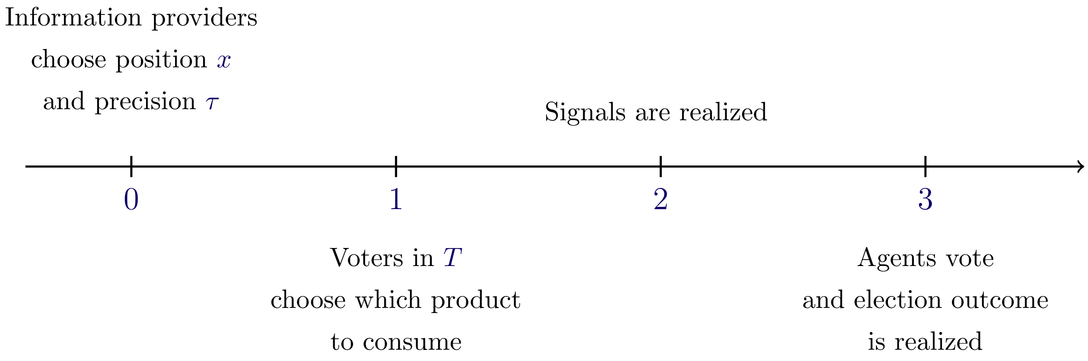

FIGURE 2: Timeline of the game.

The main difference from Definition 1 is that the total welfare is now computed using $s_{\mathrm{v}}(\boldsymbol{\theta})$ instead of θ. Ultimately, that the social planner selects candidate A if and only if $\boldsymbol{\theta}$ is such that $s_{\mathrm{v}}(\theta) > 0$.

4. Voter's Problem

We proceed by characterizing the equilibrium of the game between the voters and the news sources, and studying how it changes with the number of competing media firms n. Formally, this is a complete information dynamic game, and the equilibrium concept we utilize is the one of sub-game perfection. To compute it, we proceed by backward induction. In this Section, we determine the optimal voting strategy of an agent $t \in T$ who is consuming information structure (τ, x), from which she has received signals s_v and s_{id}. This will allow us to determine the *value* of information structure (τ, x) for voter of type t, and in turns voters' preferences over the set of all information structures. In the next Section, we will solve for the equilibrium in the simultaneous move game among information providers.

There are three actions available to each agent: vote for candidate A, abstain from the election, or vote for candidate B, respectively given by $a \in \{1, 0, -1\}$. We assume that the utility that type t derives by choosing to cast vote a for candidate $\boldsymbol{\theta}$ is:

$$\tilde{u}(a, \boldsymbol{\theta}, t) := a\ u(\boldsymbol{\theta}, t).$$

This specification implies that citizens directly receive utility from voting for a candidate. In this sense, we put aside the issue of why people vote. Indeed, in a model with a continuum of voters, no individual has an impact on the election outcome. A direct utility from honest voting (perhaps rising from a sense of civic responsibility) is the most straightforward and possibly most realistic assumption in this context.

Given signals realizations s_{v} and s_{id} from the information structure (τ, x), voter of type t

forms the following expectation: $\mathbb{E}_{\tau,x}(u(\boldsymbol{\theta},t|s_{\text{v}},s_{\text{id}}))$.[15] Here, each voter updates his beliefs about how the candidates compare given the information provided to her by the news source. It is easy to see that the voter casts her vote for candidate A if $\mathbb{E}_{\tau,x}(u(\theta,t)|s_{\text{v}},s_{\text{id}}) > 0$, B if $\mathbb{E}_{\tau,x}(u(\boldsymbol{\theta},t|s_{\text{v}},s_{\text{id}})) < 0$, and abstains otherwise.

Using this, we can easily compute the *value* of information structure (τ,x) for type t. In intuitive terms, the value of an information structure is how better off a voter can expect to be when conditioning her voting behavior on the information she receives, relative to receiving no information and staying ignorant. Formally, the value of information (τ,x) for type t is

$$V(\tau,x|t) := \mathbb{E}\Big(\max_a \mathbb{E}_{\tau,x}\big(\tilde{u}(a,\boldsymbol{\theta},t)|s_{\text{v}},s_{\text{id}}\big)\Big)$$

In the following Lemma we derive the expression of $V(\tau,x|t)$, which will play a key role in the rest of the paper.

Lemma 1. *The value of* (τ,x) *for type* t *is*

$$V(\tau,x|t) = \sigma(\tau,x|t)\sqrt{2/\pi},$$

where $\sigma^2(\tau,x|t) = \lambda^2 g(\tau) + (1-\lambda)^2\cos^2(t-x)g(1-\tau)$ *and* $g(\tau) = \frac{\tau}{1+\tau}$.

There are few important observations with respect to Lemma 1 that are useful to make at this point. The value of an information structure is a monotonic transformation of $\sigma^2(\tau,x|t)$, the variance of $\mathbb{E}_{\tau,x}(u(\boldsymbol{\theta},t)|s_{\text{v}},s_{\text{id}})$. This represents the variance in political preferences of type t that will be induced by information structure (τ,x) or, loosely speaking, how much her political preferences will be shifted towards one or the other candidate conditional on the signals received from the news source. The stronger these shifts, the smaller the uncertainty on who is the preferred candidate and, ultimately, the more informative is (τ,x).

Let's decompose the expression of $\sigma^2(\tau,x|t)$. Notice first that all voters t derive the same value from the signal s_{v} independently of the position of the news source. This value is represented by the term $\lambda^2 g(\tau)$, and is an increasing and concave function of the precision of the signal on valence s_{v}. On the other hand, the value of signal s_{id}, represented by the term $(1-\lambda)^2\cos^2(t-x)g(1-\tau)$, in addition depends both on the type of the voter and on the position of the news source. The "distance" between type t and position x, $\cos(t-x)$, determines how valuable signal s_{id} is for type t when it is generated by a news source with position x. The term $\cos(t-x)$ is in fact the correlation coefficient between the random variables $f(\theta_{id},t)$ and s_{id}. When $\cos(t-x)$ is close to zero, s_{id} is almost orthogonal to $f(\theta_{id},t)$. The variability of the former explains little or nothing about the variability of the latter. In such case, signal s_{id} has little value for type t. In the opposite case, i.e. when $\cos(t-x)$ is far away from zero, s_{id} and $f(\theta_{id},t)$ are strongly correlated either positively or negatively. In both cases, the variability of s_{id} explains a good part of the variability of $f(\theta_{id},t)$. For this reason, $\cos^2(t-x)$ is high and the value produced by s_{id} for type t is high. Notice also that

[15] In the appendix we show that $\mathbb{E}_{\tau,x}(u(\boldsymbol{\theta},t|s_{\text{v}},s_{\text{id}})) = \lambda g(\tau)s_{\text{v}} + (1-\lambda)\cos(t-x)g(1-\tau)s_{\text{id}}$. (Lemma A1)

the *value* of a news source is the same for *polar types*: voters with ideological views that are perfectly negatively correlated. These types simply use ideological signals in opposing ways: if it makes one type more likely to vote for candidate A, it makes the other more likely to vote for candidate B. Since there are only two candidates, the informativeness of the signal remains the same.

The fact that $V(\tau, x|t)$ has one component that is type-independent and another that is type-dependent generates, from the point of view of the information providers, a trade-off between the two signals s_{v} and s_{id}. This trade-off is similar to the one between a public and local good. Increasing informativeness on valence is similar to a public good - it increases the value of the news source for all voters. On the other hand, increasing informativeness on ideology is similar to providing a local good - it increases the value of the news source only for voters who have ideological preferences correlated with the signal provided. This trade-off is represented in Figure 3 in which we plot the value of several information structures as a function of t, x and τ. When τ is high, the information structure is *generalist*, highly informative on valence and with little information on ideology. The value associated with this news source is not particularly high, even for the voters that are perfectly targeted ($t = x$), but remains steadily high even for voters that are "far away" from x. On the other hand, when τ is low, the information structure is *specialist*, highly informative on ideology and not so much on valence. The value associated with this structure is high for the types that are close to x, but drops significantly for voters whose ideological preferences are farther away.

Now that we computed the value of an information structure, it is straightforward to derive agent t's ranking of the set of possible information structures. We have that type t prefers information (τ, x) over (τ', x') if and only if $V(\tau, x|t) \geq V(\tau', x'|t)$. When choosing which information to consume, voter t simply selects the news source whose associated product (x, τ) is the one producing the highest value $V(x, \tau|t)$.

There are a two implicit assumption that we are making at this point. First we assume that there is no cost associated with acquiring information. This assumption is immaterial and could be relaxed at the expense of additional analytical complexity. More significantly instead, we assume that voters can consuming at most one information structure. We postpone discussion of the robustness of our results in light of this assumptions to Section 7.

5. Information Providers' Problem

In the first stage of this game, a set $N := \{1, \dots, n\}$ of information providers compete by simultaneously choosing their strategies $(\tau_i, x_i) \in [0, 1] \times T$. In this section, we characterize the equilibrium as a function of n. The number of firms represents a measure of the level of competition in the news market.

Information providers maximize the share of voters who choose to acquire information from

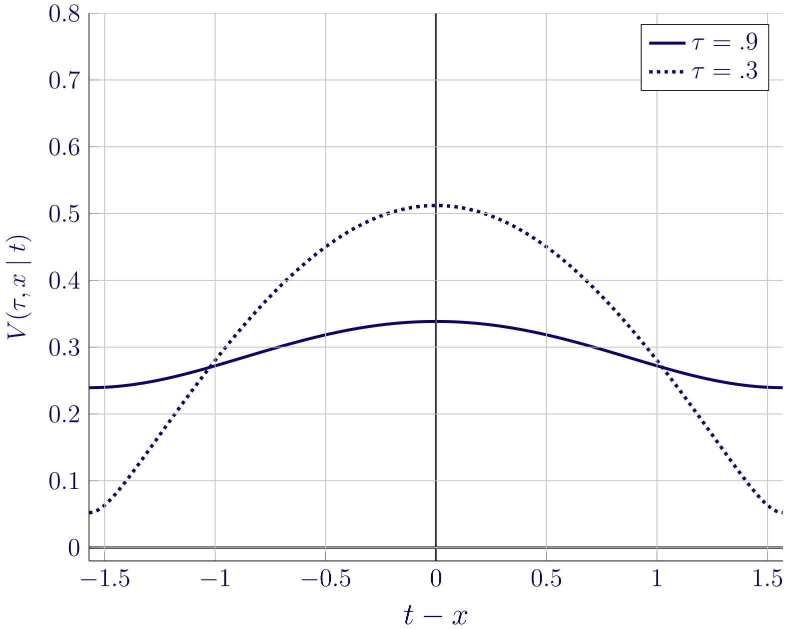

FIGURE 3: The value of information (τ, x) for type t. On the horizontal axis, $t - x$ represents the distance of given type from her information provider.

them. That is, they maximize *readership* or market capture. We don't allow firms to compete on prices to keep the model simple. Although restrictive, we believe this assumption is sensible for at least three reasons. First, price competition in the market for news is generally highly regulated (Newspaper Preservation Act, 1970). Most of the revenues nowadays come from advertisement, that mainly depend on readership. Second, the price for political news, even when it's positive, is often negligible and more than the price, it's the content that differentiates one news source from another. Lastly, price competition would set an even stronger case for product differentiation, which is the main driver of our result. We show - even in the absence of price competition - that incentives for differentiation are strong enough to have negative welfare implications.

It's convenient to think of the information provider's problem as the choice of a location in a circle. Each firm chooses a position (angle) x and a precision on valence (depth) τ. The former choice specifies what kind of ideology mix the signal s_{id} is informative about. For example, setting $x = 0$ implies that the firm only reports about subdimension ϑ_1 , while setting $x = \pm\frac{\pi}{2}$ implies that the firm is informative only about subdimension ϑ_2. Setting $x = \frac{\pi}{4}$ or $x = -\frac{\pi}{4}$ corresponds to putting equal weight on subdimensions ϑ_1 and ϑ_2. However, in the former case, the resulting signal is increasing in ϑ_1 and decreasing in ϑ_2, where as in the latter case, the resulting signal is decreasing in ϑ_1 and increasing in ϑ_2. The second choice, τ, specifies the precision of the signal on valence, s_{v}. As discussed in Section 4, the informativeness of s_{v} affects the value of the information structure for the population as a whole. This can visually be represented by how close this information structure is to the center of the circle.

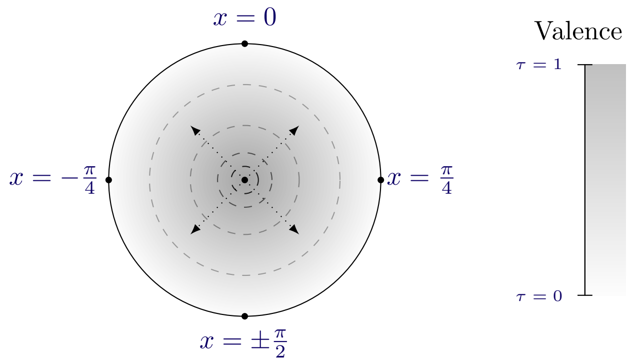

FIGURE 4: Mapping the firm's problem into a circle.

A choice of (τ, x) will generate a value $V(\tau, x|t)$ for every $t \in T$. From Lemma 1, where we characterized the expression of the value, we notice that two information structures with the same τ that are located at opposite ends of the circle produce identical values, for all voters $t \in T$. This is because they provide ideological signals that are perfectly negatively correlated. The first signal is centered around $f(\theta_{id}, x)$, while the second one is centered around $f(\theta_{id}, x + \pi) = -f(\theta_{id}, x)$. Thus the informativeness of the signals are the same. For this reason, it is without loss of generality in the equilibrium analysis to restrict the location decision of the information providers to half a circle $x \in [-\frac{\pi}{2}, \frac{\pi}{2}] \subset T$. However, for the same reason, since $\frac{\pi}{2}$ and $-\frac{\pi}{2}$ are themselves at opposite ends of the circle, they generate the same value for all types. Hence, effectively, we can think of these two ends as a single point and connect together the ends of the half-circle, by forming a new circle with circumference π. This new circle is depicted in Figure 4 and it is where we will solve the location problem of the firms.

In a market with n information providers who play a profile of strategies $(\boldsymbol{\tau}, \boldsymbol{x})$, define the equilibrium function $d_{\boldsymbol{\tau},\boldsymbol{x}} : T \to N$ to be $d_{\boldsymbol{\tau},\boldsymbol{x}}(t) = \arg\max_{i \in N} V(\tau_i, x_i|t)$. That is, $d_{\boldsymbol{\tau},\boldsymbol{x}}(t) \in N$ is the news source that type t optimally chooses to acquire information from. Thus, firm i maximizes the following objective:

$$\Pi_i(\boldsymbol{\tau}, \boldsymbol{x}) := \frac{1}{2\pi} \int_T \mathbb{1}(d_{\boldsymbol{\tau},\boldsymbol{x}}(t) = i) dt.$$

We focus on Nash equilibria of the complete information game $(N, ([0, 1] \times [-\frac{\pi}{2}, \frac{\pi}{2}], \Pi_i)_{i \in N})$ that satisfy the following symmetry property:

Definition 3. *A Nash equilibrium* $(\boldsymbol{\tau}, \boldsymbol{x})$ *is **symmetric** if* $\tau_i = \tau^\star$ *for all* $i \in N$ *and information providers are located equidistantly, that is, for every* $i, j \in N$ *which are immediate neighbors of each others* $|x_i - x_j| = \frac{\pi}{n}$.

Every symmetric rotation or re-shuffling of the firms' locations $\boldsymbol{x}$ would still be an equilibrium. Thus, provided they exists, there is multiplicity of these equilibria, indeed a continuum. However, the equilibrium value of $\boldsymbol{\tau}$ and the distance between any two neighboring firms will be pinned down uniquely. As we show below, changes in these two values completely characterize the impact of competition on welfare. This is the sense in which we will talk about unique symmetric equilibria.

Before we study competition among multiple information provides, we discuss the monopolist's problem. Note that the value associated with any information structure is weakly positive. Since a voter that does not consume information gets a "value" of zero, any choice of τ and x will produce an information structure with a positive value. This, however, trivializes the monopolist problem and makes it uninteresting.

Observation 1. *(Monopolist) Any choice of* τ *and* x *is a solution to the monopolist problem.*

This multiplicity is an artifact of the absence of prices in the model and therefore disappears as soon as one allows the monopolist to also set the price. To see this notice that when the monopolist charges a price, only voters for whom the value of the information structure is higher than the price will acquire the monopolist's information structure. It is easy to show that in any solution to this problem where the monopolist captures the whole market, it must be providing only information on valence, i.e. $\tau^\star = 1$. The intuition for this is straightforward. For any other choice of $\tau < 1$, there will be a segment of the population for whom the signal on ideology carries little or no information, driving down the price. The monopolist, by shifting precision from ideology to valence, can improve the value of the news source for these voters, and consequently increase prices.

Due to the multiplicity of equilibria in the monopolist problem, we focus on markets where there are more than two firms.

Proposition 2. *There exists a symmetric Nash equilibrium* $(\boldsymbol{\tau}^\star, \boldsymbol{x}^\star)$*. For* $n > 2$*, this equilibrium is unique in the class of symmetric equilibria, up to rotations and permutations of the locations.*

An important part our analysis is to understand how the equilibrium level of $\tau^\star$, which measures how informative firms are on the valence dimension, changes as the number of firms in the market increases. The next proposition shows that this effect is negative, meaning that as the market becomes more competitive, the equilibrium precision on valence, the only socially relevant dimension, decreases.

Proposition 3. *As competition increases, news sources become less informative on valence, i.e.* $\tau^\star$ *decreases.*

In equilibrium, the type of information provided by any news source is chosen optimally to target a specific group of voters. Hence, how much information is provided on ideology relative to valence by a specific news source depends on how much overlap there is in the type of ideological information that is of interest to the consumers of that news source. News sources balance informativeness on valence and ideology to win over voter types that are indifferent in terms of which news source to consume. In other words, the information structure is chosen to maximize the value for these threshold types. Competition leads to segmentation: the share of the market that can be targeted by any firm decreases with n. This implies that threshold types move closer in terms of their ideological distance. This creates incentives for news sources to differentiate their product choices, leading to more information on ideological issues.

We can also see this graphically as depicted in Figure 5. In a symmetric equilibrium, all firms set the same $\tau^\star$ and locate equidistantly in terms of position x. The value of $1-\tau^\star$ denotes the precision on ideology; hence, it can be considered as a measure of how specialized news sources are. Visually, this is captured by the size of the circle on which firms locate. As n increases, firms are forced to locate closer to one another. But, in equilibrium, focus on ideology $(1-\tau^\star)$ also increases with n. This corresponds to moving farther away from the center of a circle. In a sense, by increasing the size of the circle, firms are able to ease competition.

6. Elections

In this section, we study how changes it the market for news affect election outcomes. We start by studying how voting decisions depend on the information structure consumed. In this model, voters' preferences have two components: valence, θ_v, and ideology, θ_{id}. Conditional on $\boldsymbol{\theta}$, we aim to study how increasing competition affects the dependency of voters' behavior on each of these dimension.

Assume that in equilibrium voter $t \in T$ acquires information structure (τ, x). Such a voter will vote in favor of candidate A if and only if, given her signals realizations s_{v} and s_{id}, her expected utility $\mathbb{E}_{\tau,x}(u(\boldsymbol{\theta},t)|s_{\mathrm{v}}, s_{\mathrm{id}})$ is positive. Conditional on $\boldsymbol{\theta}$, this expected utilty is normally distributed with mean μ and variance ν^2. That is, conditional on $\boldsymbol{\theta}$, the probability that type t votes for candidate A is $\Phi(\mu/\nu)$, where Φ is the cumulative density function of a standard normal. Since μ and ν^2 explicitly depend on τ and x, this provides a complete characterization of how voting behavior depends on the information structure that is consumed by each type. From Lemma A1 (in the Appendix), it's easy to derive the expression for μ:

$$\mu := \lambda g(\tau)\theta_v + (1-\lambda) g(1-\tau)\cos(t-x) f(\theta_{id}, t).$$

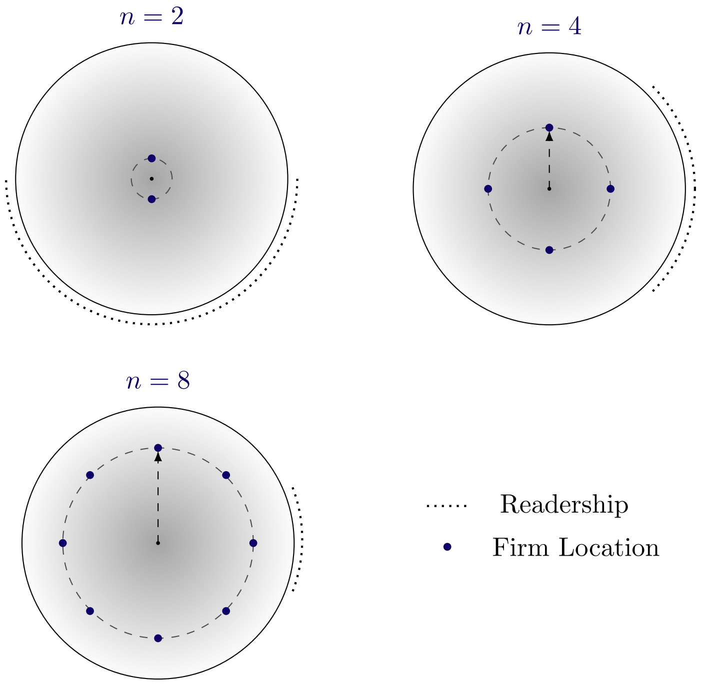

FIGURE 5: The representation of the symmetric equilibrium for several values of n.

From Proposition 3, we know that, as n increases, the equilibrium precision of valence τ decreases and the equilibrium distance between each type and her information provider, $\cos(t-x)$, increases. Hence, in the expression for μ, the *weight* that type t puts on dimension θ_v decreases with n, whereas the *weight* on dimension θ_{id} increases. As a result, the voting behavior of type t becomes increasingly correlated with $f(\theta_{id}, t)$ and increasingly uncorrelated with θ_v. Both of these forces imply stronger ideological voting. In fact, if voters cared only about ideology ($\lambda = 0$), they would want their vote to be perfectly correlated with $f(\theta_{id}, t)$. For such a voter, A would be the preferred candidate if and only if $f(\theta_{id}, t)$ is positive. This is the main intuition behind the proof of the next result:

Lemma 2. *As n increases, the voting behavior of voter t becomes increasingly ideological, that is increasingly correlated with $f(\theta_{id}, t)$.*

Although the preferences of the voters remain unchanged, the effective importance of ideology relative to valence in voting behavior increases. This is not without consequences. In fact, as stated in Proposition 1, the socially optimal candidate is determined by $\theta_v > 0$, i.e. comparison on the valence dimension. Thus, as competition increases, the probability that one's vote coincides with the socially optimal candidate decreases. Figure 6 illustrates how the probability of voting in line with the social planner (both using the first-best and the second-best benchmark) changes with n.

Lemma 3. *As n increases, the voting behavior of type t becomes increasingly uncorrelated with the choice of the social planner.*

Before assessing the aggregate effects of increased competition on electoral outcomes, one might wonder how it impacts voters individually. The next Proposition shows that on an individual level, the impact of increased competition in the news market is exactly what one would expect in a market with profit-maximizing firms and rational consumers. Competition generates differentiation in the space of products, creating a larger spectrum of options for voters. This enables voters to select news sources that provide the type of information that is better tailored towards their needs.

Proposition 4. *As n increases, voters become individually more informed. That is, for all $t \in T$, the value associated with the closest news source increases.*

Proposition 4 points out that the inefficiency identified in this paper is not due to some form of market failure. On the contrary, competition enables voters to learn more effectively. The fact that voters are individually better informed naturally implies that the probability they vote for the candidate that is *ex post* better *for them* increases. The key point is that there is a disconnect between what is individually optimal and what is socially optimal. Improving information acquisition on an individual level doesn't necessarily lead to *better* election outcomes. What exactly voters become informed about is critical for this analysis.

Now, we focus on aggregate voting behavior. To do that we need to specify how the noise associated with learning varies across voters and information sources. We adopt a very simple structure: Signals s_{v} and s_{id} are *conditionally independent* across firms and voters. This assumption requires that any correlation in the signals received by two different voters can be described by the precision and position of the firms from which they are receiving information. This implies that conditional on $\boldsymbol{\theta}$, voters acquiring information from the same or different information sources receive mutually independent signals $(s_{\text{v}}, s_{\text{id}})$. A natural interpretation for this assumption is that the noise associated with learning is idiosyncratic originating from either external or internal biases on an individual level. As the informativeness of a news source increases, these idiosyncratic factors play less of a role in the affecting ex-post beliefs.[16]

With this assumption, we can now study the aggregate impact of competition on election outcomes.

Theorem 1. *As competition increases, the share of votes received by the socially optimal candidate decreases.*

[16]More formally, a voter who acquires information (τ, x) receives two signals: $s_{\text{v}} = \theta_v + \varepsilon_{t,v}$ and $s_{\text{id}} = f(\theta_{id}, x) + \varepsilon_{t,id}$. The error terms are s.t. $\varepsilon_{t,v} \sim \mathcal{N}(0, \tau^{-1})$ and $\varepsilon_{t,id} \sim \mathcal{N}(0, (1-\tau)^{-1})$. We assume $\varepsilon_{t,v}$ and $\varepsilon_{t,id}$ are not correlated. In particular, in the rest of the paper, we assume that *different* types that consume the *same* information structure receive conditionally *independent* signals. However, our results go through even when we allow for more sophisticated form of correlation in the error terms, e.g. when *different* types acquiring information from the *same* source receive the *same* signal.

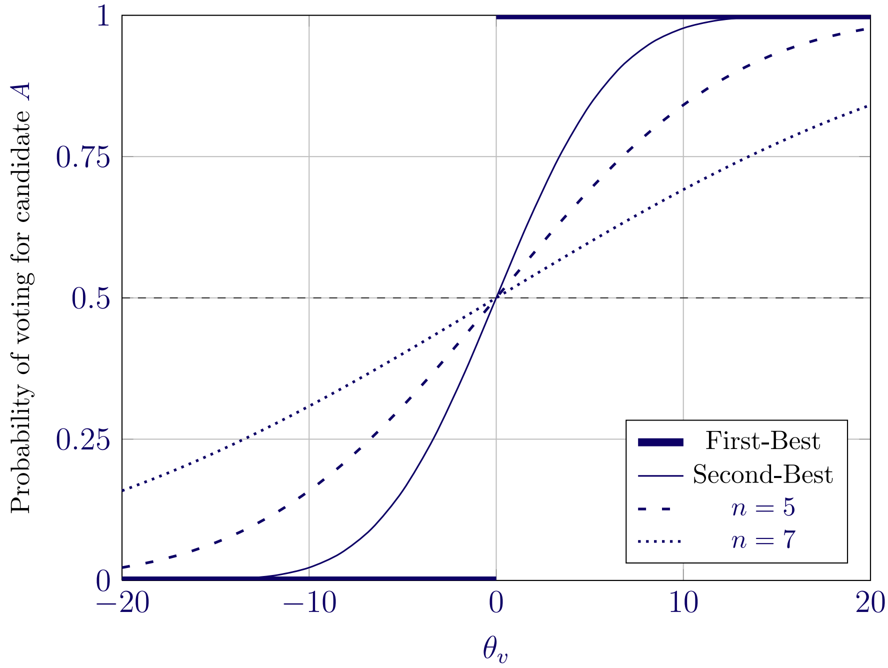

FIGURE 6: The probability of voting for A conditional on θ_v for an arbitrary voter t. Voting for A is socially optimall if only if $\theta_v > 0$.

Theorem 1 demonstrates that the inefficiency identified on an individual level also translates into vote shares. Competition increases ideological voting which intensifies disagreement in voting decisions. The electorate gets divided along ideological lines, and the vote advantage enjoyed by the socially optimal candidate decreases. We illustrate this effect in Figure 7.

Clearly, there are many important settings in which the *distribution* of votes - and not only who wins the majority - has an impact on voters welfare. Consider two scenarios such that in both of them candidate A is the socially optimal candidate, but in one she is expected to receive 51% of the voters, whereas in the latter she is expected to receive 99% of votes. These scenarios can imply profoundly different outcomes. We provide some examples here. First, even under the majority rule, the distribution of votes matters if votes shares are subject to aggregate shocks. Higher vote shares translate into a higher probability of election, as higher shares of votes won by a candidate are less likely to be overturned by a "negative" realization of the aggregate shock. Aggregate shocks can be interpreted as temporary shifts in voter preferences or as the result of the random behavior of *noise voters*.[17] Second, the distribution of votes matters in proportional electoral systems or more generally whenever legislators are elected using a mixture of majoritarian *and* proportional. Finally, the distribution of votes can affect voters' welfare indirectly, by putting pressure on the winning party to balance among the interests of different groups.

Our last comparative static is with respect to the parameter λ. As we argued before, λ provides a simple measure of the degree of homogeneity of the society. Since λ captures how

[17]See Baron (1994) and Grossman and Helpman (1996).

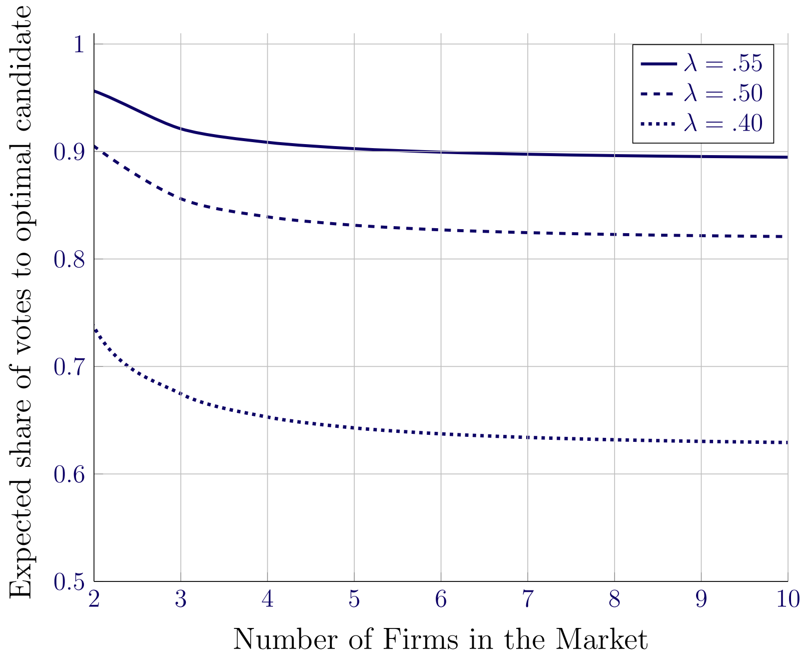

FIGURE 7: The main comparative statics: n and λ

much voters care about valence relative the ideology, $1 - \lambda$ can be thought of as a reduced-form parameter that measures 'polarization' in the political preferences of the electorate. We find that in more polarized societies the inefficiency created by competition is even larger.

Theorem 2. *For all n, increasing preferences' polarization, i.e. lowering λ, decreases the share of votes received by the socially optimal candidate decreases.*

Theorem 2 show that the inefficiency associated with competition is exacerbated by polarization in the distribution of political preferences across voters. As polarization increases, demand for information on ideology increases. In a competitive market, firms respond to this demand by shifting precision from valence to ideology. Once again, we illustrate this effect in Figure 7.

7. Discussion

In this section, we discuss the robustness of our main results in relation to some of the assumptions we have made throughout the paper.

7.1 Strategic Voting

While solving the voter's problem, we have assumed that voters vote sincerely. This is a common assumption in the literature and possibly the most realistic description of voters' behavior. It allows us to work with a continuum of voters and to abstract away from the specifics of the electoral rule. Yet, strategic voting could potentially affect our results at multiple different levels, and thus it is worth considering. First, it could affect the way voters vote given the information they acquired. As a result of this, it could affect how voters value information, and hence which news source they decide to acquire. Finally, given all the above, it could affect the information provision stage in which firms compete with each other. The literature on strategic voting assumes that voters are motivated by instrumental considerations of how their voting behavior can affect the electoral outcome. That is, each voter contemplates the effects of her vote only insofar as it changes the final outcome. Such effects crucially hinge on the mechanism that maps votes shares into electoral outcome. In this paper, as discussed in Section 6, we focused on settings where the *distribution* of votes - and not only who wins the majority - has an impact on voters welfare. In such cases, *regardless* of others' behavior, each voter *always* affects the final outcome. This limits what voters can learn from others by conditioning on events where their vote will be affecting their final utility. It is easy to see that, in this case, strategic voting coincides with sincere voting. Thus, the main results in our paper directly extend to strategic voting.

Interestingly, some of the main forces uncovered in this paper, in particular the fact that competition leads to higher ideological voting, appear to be present even if we also consider the majoritarian electoral rule in combination with strategic voting. Here we provide intuition for this. Under the majority rule, voters affect final outcomes only when they happen to be pivotal; namely, when half of the population votes for A and the other half votes for B. Due to the symmetry in ideological preferences, conditioning on such an event would be more informative about valence. As a result, a strategic voter would put more weight on the signal on ideology relative to the signal on valence in determining how the candidates compare. This reinforces the demand for ideological information, and consequently incentives for product differentiation as a result of competition.

7.2 Diversification on Valence

A key insight in our paper concerns the effect of competition on the *type* of information that is provided to voters. Crucial to this result is the structure of the underlying heterogeneity in preferences. In our model, disagreement among voters, both in terms of *slant* and *agenda*, is present only on ideology, and not on valence. In fact, it is almost the definition of valence itself that precludes voters to disagree on *slant*; but nothing precludes valence to be multi-dimensional and voters to disagree on what sub-dimensions of valence are more important to them.[18] In our model, this kind of disagreement is shut down since valence is assumed

[18] For example, a voter could care more about the technical competence of a candidate, whereas another could care more about her international reputation.

to be one-dimensional. This is, of course, a simplifying assumption that makes our analysis more tractable and the illustration of our main result more salient.

In reality, the "true" preferences of voters are likely to be highly multi-dimensional both in valence and ideology. We can imagine competition to bring about informational products that are in fact diversified on both valence and ideology. In general, our model predicts competition to cause firms to provide more information on issues where consumers are heterogeneous. Hence, an important question is whether or not these issues correspond to ideology. We take this connection to be very likely. There are multiple reasons for which information on ideology needs to be more targeted than information on valence. Concerning the heterogeneity on *agenda*, it is natural to expect more disagreement on ideology relative to valence. For example, some voters could believe that religious issues are of prime importance, while others may not find them important at all. On the other hand, it is unlikely for voters to disagree as strongly on how much the technical competence of a candidate matters relative to his experience. But even if voters' *agendas* about valence issues were no less different than those one ideological issues, there are still reasons to expect information providers to differentiate more on the latter. In fact, heterogeneity on valence is fully explained by differences in voters' *agenda*.This is not the case with ideology. There is additional heterogeneity on *slant*, namely how people would like issues to be addressed. Since firms consider both types of heterogeneity in choosing their informational products, ideology allows for further differentiation.

7.3 Consuming Multiple News

Another assumption we maintained in the paper is that each voter has to choose one and only one information provider. This assumption reduces the complexity of the game played by the news sources and it allows us to solve for the equilibrium in a relatively simple way. Our main results focus on how competition affects individual learning, and consequently electoral outcomes, by increasing the spectrum of information structures available to the voters. Allowing for voters to consume multiple news sources introduces a possibly counteracting force. Assume that voters can freely access all signals produced by the n information providers.[19] As n increases, voters would have access to an increasing number of signals on valence and ideology. Without further specification, it is not possible to determine how this would affect voting behavior. Notice that this would critically depend on whether or not voters are learning more about ideology or valence as the number of information sources available to them increase.

Yet, assuming that voters can consume every signal produced by the market, irrespectively of n, is possibly even more extreme than assuming they can only acquire one. More realistically, agents have limitations on how many signals they can process due to time constraints, rational inattention, cognitive ability, opportunity cost etc. The main intuition behind our

[19]Note that there are also concerns in terms of how competition among information providers should be defined when *all* voters consume *all* products.

result goes through even if we allow voters to consume multiple signals, as long as there is a cap $\bar{\kappa}$ on the number of signals a voter can acquire. For simplicity, our paper sets $\bar{\kappa} = 1$. More generally, if $\bar{\kappa} \in \mathbb{N}$, as soon as $n > \bar{\kappa}$, voters will pick the best $\bar{\kappa}$ signals that they can find on the market. Once again, firms will compete for being selected by the highest number of voters. The competitive tensions that this situation generates are very similar to the $\bar{\kappa} = 1$ case. In fact, firms that are "underchosen" will have incentive to increase the precision of the signal on ideology, foregoing those customers that are "far away" from the their location, thus creating more value for those that are close by.

8. Conclusions

Our paper illustrates a novel channel through which competition among information providers can affect the distribution of political views and produce negative welfare consequences for the society. In our model, competition pushes firms to differentiate their informational products. Differentiation forces firms to provide more information about issues on which there is greater disagreement among voters. Since voters use this information to learn about political candidates, competition creates an electorate which effectively puts higher weight on ideological issues relative to valence issues. Our main result shows that competition generates more ideological voting which leads to a decline in the share of votes going to the socially optimal candidate. This illustrates clearly how the market for news differs from traditional markets. Markets respond to demand from individuals. The resulting differentiation is in fact optimal at the individual level: voters are able to learn more effectively whenever there are more sources competing with each other. In this sense, competition does not create, but simply uncovers the underlying heterogeneity in voters' preferences. The source of the inefficiency lies in the fact that there is a discrepancy between the type of information that is valuable to individuals and the one that is valuable for the society as whole.

A. Proofs.

Proof of Proposition 1. Notice that $\int_T u(\theta|t)dt = \lambda\theta_v + (1-\lambda)(\vartheta_1 \int_T \cos t dt + \vartheta_2 \int_T \sin t dt)$. But $\int_T \sin t dt = \int_T \cos t dt = 0$. Hence, $\int_T u(\theta|t)dt = \lambda\theta_v$, which is positive if and only if $\theta_v > 0$. □

Lemma A1. $\mathbb{E}_{\tau,x}(u(\boldsymbol{\theta}, t|s_v, s_{id})) = \lambda g(\tau)s_v + (1-\lambda)\cos(t-x)g(1-\tau)s_{id}$.

Proof of Lemma A1. Given our assumptions on the distributions of θ, s_{v} and s_{id}, we have that

$$\mathbb{E}_{\tau,x}\Big(u(\theta,t)\Big|s_{\text{v}}, s_{\text{id}}\Big) = \mathbb{E}_{\tau,x}\Big(\lambda\theta_v + (1-\lambda)\Big(\vartheta_1\cos(t) + \vartheta_2\sin(t)\Big)\Big|s_{\text{v}}, s_{\text{id}}\Big) =$$

$$= \lambda\mathbb{E}_{\tau}(\theta_v|s_{\text{v}}) + (1-\lambda)\cos(t)\mathbb{E}_{\tau,x}(\vartheta_1|s_{\text{id}}) + (1-\lambda)\sin(t)\mathbb{E}_{\tau,x}(\vartheta_2|s_{\text{id}}).$$

From the properties of conditional expectation of multivariate normal distributions we have

$$\mathbb{E}_{\tau}(\theta_v|s_{\text{v}}) = \frac{\tau}{1+\tau}s_{\text{v}};$$

$$\mathbb{E}_{\tau,x}(\vartheta_1|s_{\text{id}}) = \cos(x)\frac{1-\tau}{2-\tau}s_{\text{id}};$$

$$\mathbb{E}_{\tau,x}(\vartheta_2|s_{\text{id}}) = \sin(x)\frac{1-\tau}{2-\tau}s_{\text{id}}.$$

Letting $g(\tau) = \frac{1}{1+\tau}$ and noticing that $\cos(t)\cos(x) + \sin(t)\sin(x) = \cos(t-x)$ prove the result. □

Proof of Lemma 1. Recall that if $X \sim \mathcal{N}(0,\sigma^2)$, then $\mathbb{E}(|X|) = \sigma\sqrt{2/\pi}$. In our case, since -unconditionally- $s_{\text{v}} \sim \mathcal{N}(0, 1+\tau^{-1})$ and $s_{\text{id}} \sim \mathcal{N}(0, 1+(1-\tau)^{-1})$ we have

$$\mathbb{E}_{\tau,x}\Big(u(\theta,t)\Big|s_{\text{v}}, s_{\text{id}}\Big) \sim \mathcal{N}(0,\sigma^2)$$

where $\sigma^2 = \lambda^2 u(\tau) + (1-\lambda)^2\cos^2(t-x)u(1-\tau)$. Voter t will vote $a = 1$ if and only if $\mathbb{E}_{\tau,x}(u(\theta,t)|s_{\text{v}}, s_{\text{id}}) > 0$. Thus,

$$V(\tau, x|t) = \mathbb{E}\Big(\max_a \mathbb{E}_{\tau,x}(au(\theta,t)|s_{\text{v}}, s_{\text{id}}\Big) = \mathbb{E}\Big(\Big|\mathbb{E}_{\tau,x}(u(\theta,t)|s_{\text{v}}, s_{\text{id}})\Big|\Big) = \sigma\sqrt{2/\pi}.$$

□

Proof of Proposition 2. Let $n > 2$ and let's focus on the behavior of media i. Fix a symmetric strategy of i's opponents. Every $-i$ chooses the same $\tau^\star$. Every $-i$ whose immediate neighbors is not i is $\frac{\pi}{2n}$-away from both its neighbors. We want to prove that i has no profitable deviation off the symmetric strategy, that is every deviation leads to a profit which is lower or equal than π/n. We divide this proof in 4 Lemmas. First we fix the location $x_i^\star$ and we show that there are no profitable deviation on precision τ_i.

Lemma A2. *Fixing everyone's location, $\tau^\star$ exists and i has no incentive to unilaterally deviate on τ_i.*

Proof Lemma A2. Without loss of generality let $x_i = 0$. By definition of t_r we have that

$$\frac{\partial}{\partial \tau_i}\Big(V(\tau_i, x_i|t_r) - V(\tau^\star, x^\star_{i+1}|t_r)\Big) = 0$$

This equilibrium condition allows us to retrieve an expression for $\frac{\partial t_r}{\partial \tau_i}$. In fact

$$\frac{\partial t_r}{\partial \tau_i} = \frac{\lambda^2 g'(\tau_i) - (1-\lambda)^2 \cos^2(t_r - x_i) g'(1-\tau_i)}{(1-\lambda)^2\big(g(1-\tau_i)\sin 2(t_r - x_i) + g(1-\tau^\star)\sin 2(x^\star_{i+1} - t_r)\big)}$$

Similarly for we can use $V(\tau_i, x_i|t_l) - V(\tau^\star, x^\star_{i-1}|t_l) = 0$ to derive

$$\frac{\partial t_l}{\partial \tau_i} = -\frac{\lambda^2 g'(\tau_i) - (1-\lambda)^2 \cos^2(x_i - t_l) g'(1-\tau_i)}{(1-\lambda)^2\big(g(1-\tau_i)\sin 2(x_i - t_l) + g(1-\tau^\star)\sin 2(t_l - x^\star_{i-1})\big)}$$

The first order condition on media profits tells us that in equilibrium

$$\frac{\partial t_r}{\partial \tau_i} = \frac{\partial t_l}{\partial \tau_i}.$$

Notice that since $x_{i+1} - x_i = x_i - x_{i-1}$ and both $i-1$ and $i+1$ are playing a symmetric strategy $\tau^\star$, $t_l - x^\star_{i-1} = x^\star_{i+1} - t_r$ and $t_r - x_i = x_i - t_l$. Thus we get

$$\lambda^2 g'(\tau_i) = (1-\lambda)^2 \cos^2(t_r - x_i) g'(1-\tau_i)$$

or

$$\frac{\lambda^2}{(1-\lambda)^2}\frac{g'(\tau_i)}{g'(1-\tau_i)} = \cos^2(t_r - x_i)$$

This condition could have been derived also by setting $\frac{\partial t_r}{\partial \tau_i} = 0$. The left-hand side is decreasing in τ_i, while the right-hand side is increasing. In equilibrium we have

$$\frac{\lambda^2}{(1-\lambda)^2}\frac{g'(\tau^\star)}{g'(1-\tau^\star)} = \cos^2\left(\frac{\pi}{2n}\right)$$

which implicitly defines $\tau^\star$ as a function of n. □

Second we fix precision $\tau_i = \tau^\star$ and we show that there are no profitable deviations on x_i.

Lemma A3. *Fixing precision $\tau^\star$ and $-i$ location, media i has no incentive to deviate away from $x^\star_i = 0$.*

Proof of Lemma A3. By the equilibrium condition $\frac{\partial}{\partial x_i}V(\tau_i, x_i|t_r) = \frac{\partial}{\partial x_i}V(\tau^\star, x^\star_{i+1}|t_r)$ we can derive expressions for $\frac{\partial t_r}{\partial x_i}$. Indeed, we get

$$(1-\lambda)^2 g(1-\tau_i)\sin 2(t_r - x_i)\big(1 - \frac{\partial t_r}{x_i}\big) = (1-\lambda)^2 g(1-\tau_{i+1})\sin 2(x_{i+1} - t_r)\frac{\partial t_r}{x_i}$$

which gives us the following expression

$$\frac{\partial t_r}{\partial x_i} = \frac{g(1-\tau_i)}{g(1-\tau_i) + \psi_r g(1-\tau_{i+1})},$$

where

$$\psi_r := \frac{\sin 2(x_{i+1} - t_r)}{\sin 2(t_r - x_i)}$$

In a similar fashion we can get

$$\frac{\partial t_l}{\partial x_i} = \frac{g(1-\tau_i)}{g(1-\tau_i) + \psi_l g(1-\tau_{i-1})},$$

where

$$\psi_l := \frac{\sin 2(t_l - x_{i-1})}{\sin 2(x_i - t_l)}.$$

When $\tau_i = \tau^\star$ for all $i \in N$, it's easy to see that the thresholds t_r and t_l are at the midpoints of the media location, i.e. $t_r = (x_i + x_{i+1})/2$ and $t_r = (x_i + x_{i-1})/2$. This implies that $\psi_r = \psi_l = 1$. This implies that

$$\frac{\partial t_r}{\partial x_i} - \frac{\partial t_l}{\partial x_i} = 0.$$

Thus, firm i does not strictly gain by locating itself away from $x_i^\star$. □

It remains to show that there is no joint deviation in τ_i *and* x_i that could make firm i better off. We do this in the next two Lemmas. In the first we consider a joint deviation that both increase the location x_i and the precision τ_i.

Lemma A4. *For all $\tau_i > \tau^\star$ and all locations x_i, firm i's profit are smaller than π/n.*

Proof Lemma A4. Fix $\tau_i > \tau^\star$ and $x_i > x_i^\star$ (the case in which $x_i < x_i^\star$ is symmetric). Consider the type $\tilde{t} := (x_i + x_{i+1}^\star)/2$ which is midway between x_i and $x_{i+1}^\star$. We want to show that $\tilde{t}$ does prefer $i+1$ to i. Notice that since $x_i > x_i^\star$ and, by Definition 4.1, $x_{i+1}^\star - x_i^\star = \pi/2n$, we have that $\tilde{t} - x_i = x_{i+1}^\star - \tilde{t} < \pi/2n$. By construction, $\tau^\star$ is the optimal level of valence for a type t who is $\pi/2n$-away from the information provider. All types that are closer than $\pi/2n$ would prefer less valence. Thus, $\tilde{t}$ strictly prefers firm $i+1$ since, compared with firm i, it offers a lower level of valence, $\tau^\star < \tau_i$. We conclude that $x_{i+1}^\star - t_r > t_r - x_i > 0$, hence $\psi_r > 1$. Now let's consider t_l. If it is such that $t_l - x_{i-1} > x_i - t_l$ then firm i's profits are necessarily less than $\pi/2n$. Thus, the only case we need to consider is the one in which $t_l - x_{i-1} < x_i - t_l$. In this case, $\psi_l < 1$. Summing up, we have that $\psi_r > 1$ and $\psi_l < 1$, implying that $\frac{\partial t_r}{\partial x_i} - \frac{\partial t_l}{\partial x_i} < 0$. Since $x_i > x_i^\star$ was arbitrary and since we know from Lemma 1 that at $x_i^\star$, if $\tau_i > \tau^\star$ the profits of firm i are less than $\pi/2n$, we can conclude the proof. □

It remains to consider a joint deviation that increases the location x_i and decreases the precision τ_i.

Lemma A5. *For all $\tau_i < \tau^\star$ and all locations x_i, firm i's profit are smaller than π/n.*

Proof Lemma A5. Fix $\tau_i < \tau^\star$ and $x_i > x_i^\star$ (the case in which $x_i < x_i^\star$ is symmetric). There are two subcases to consider here. Either the left threshold type t_l is indifferent between firm i and $i-1$ (as it was in the previous Lemmas), or is indifferent between firm i and $i+1$. This second case is possible because firm i now providing more ideology (lower τ_i) than its neighbors. On the other side, the right threshold t_r will always correspond to a type who is indifferent between firm i and $i+1$.

Subcase 1: Lets's assume t_l is indifferent between i and $i-1$. A similar argument to Lemma 3 above will show that $t_l - x_{i-1}^\star > x_i - t_l > 0$. In fact the midpoint $\tilde{t} := (x_i + x_{i-1}^\star)/2$ is now more than $\frac{\pi}{2n}$-away from both x_i and $x_{i-1}^\star$. Thus she would prefer more valence than $\tau^\star$. Since $\tau_i < \tau^\star$, type $\tilde{t}$ prefers x_{i-1}. This shows $t_l - x_{i-1}^\star > x_i - t_l > 0$. This implies that $\psi_l > 1$. Now we look at t_r. Once again, either *(a)* firm i is conquering more than half of the market, i.e. $x_{i+1} - t_r < t_r - x_i$ or *(b)* firm $i+1$ does, i.e. $x_{i+1} - t_r > t_r - x_i$. If *(b)* is the case, then firm i's profits are necessarily less than π/n and we are done. Thus, we only need to consider case *(a)*. In such case, $\psi_r < 1$ (it can actually be even negative in this case if t_r is to the right of x_{i+1}). This gives us that $\frac{\partial t_r}{\partial x_i} - \frac{\partial t_l}{\partial x_i} > 0$. Since $x_i \in [x_i^\star, x_{i+1}^\star]$ was arbitrary, we proved that the derivative of profits is strictly increasing in such region. Thus, firm i will keep increasing x_i, getting closer and closer to x_{i+1}. Eventually, firm x_i will locate in the same spot of x_{i+1}, but with a lower τ_i. Thus the threshold type t_l will be no longer indifferent between firm i and $i-1$, but rather with firm i and $i+1$. This is Subcase 2, which we analyze next.

Subcase 2: Let's assume t_l is indifferent between i and $i+1$. It must be that t_l is closer to $i+1$ than $i-1$. If not, t_l should prefer $i-1$ to $i+1$, a contradiction. Now consider $\tilde{t} = \frac{x_{i+1}^\star + x_{i+2}^\star}{2}$, which is the midpoint between firm $i+1$ and $i+2$. Notice that since $x_i \in [x_i^\star, x_{i+1}^\star]$, $\tilde{t} - x_{i+1}^\star \geq \tilde{t} - x_i$. Since firm i, relative to firm $i+1$, is offering lower valence τ_i and it is weakly farther away to $\tilde{t}$, then such type will prefer firm $i+1$ to i. Since by construction $\tilde{t} - t_l \leq \pi/n$, firm i 's profit are lower than π/n. ☐

This concludes the Proof of Proposition 2. ☐

Proof of Proposition 3: In the Proof of Lemma A2 we implicitly derived a solution for $\tau^\star$ in terms of the parameter n:

$$\frac{\lambda^2}{(1-\lambda)^2}\frac{g'(\tau^\star)}{g'(1-\tau^\star)} = \cos^2\left(\frac{\pi}{2n}\right)$$

Notice that the right-hand side is increasing in n. On the left-hand side, we can substitute the definition of $g(\tau) = \frac{\tau}{1+\tau}$ to find that

$$\frac{g'(\tau^\star)}{g'(1-\tau^\star)} = \left(\frac{2-\tau}{1+\tau}\right)^2$$

is strictly decreasing in τ. Thus an increase in n can be compensated only by a decrease in $\tau^\star$. ☐

Proof of Lemma 2. Conditional on $\boldsymbol{\theta}$, the expected utility of type t consuming infromation structure (τ, x) is distributed $\mathcal{N}(\mu, \nu^2)$, with

$$\mu := \lambda g(\tau)\theta_v + (1-\lambda)g(1-\tau)\cos(t-x)f(\theta_1, t).$$

and

$$\nu^2 := \lambda^2 \frac{\tau}{(1+\tau)^2} + (1-\lambda)^2 \cos^2(t-x)\frac{1-\tau}{(2-\tau)^2}.$$

Thus voter t votes for candidate A with probability $\Phi(\mu/\nu)$, where Φ is the cdf of a standard normal. On the other side, a completely ideological ($\lambda = 0$) and pefectly informed type t would vote for A if and only if $f(\theta_{id}, t) \geq 0$. Conditional on θ_{id}, the probability of voting for candidate A is $\Phi(B(\tau, x|t)\theta_{id})$ where

$$B(\tau, x|t) := \frac{(1-\lambda)g(1-\tau)\cos(t-x)}{\sqrt{\lambda^2 g(\tau) + (1-\lambda)^2\cos^2(t-x)\frac{1-\tau}{(2-\tau)^2}}}.$$

In fact, recall that $\theta_v \sim \mathcal{N}(0,1)$ and it is indepenedent from $f(\theta_{id}, t)$. With this in mind, we can apply the identity

$$\int_{\mathbb{R}} \Phi(a + bx)d\Phi(x) = \Phi\left(\frac{a}{\sqrt{1+b^2}}\right),$$

which applies to normal distributions. In our case, $a = \frac{(1-\lambda)g(1-\tau)\cos(t-x)}{\nu} f(\theta_{id}, t)$, $b = \frac{\lambda g(\tau)}{\nu}$ and $x = \theta_v$. This gives us the expression of $B(\tau, x|t)$ above.

We show next that $B(\tau, x|t)$ is increasing in n. From Proposition 3, when n increases τ decreases and $\cos|t-x|$ increases. It is straightforward to see $B(\tau, x|t)$ is increasing in $\cos|t-x|$. We turn to proving that $\frac{\partial A}{\partial \tau}$ is negative. Denote $B(\tau, x|t) := \alpha/\sqrt{\beta}$. Deriving with respect to τ we have

$$\frac{\partial A}{\partial \tau} = \frac{\partial}{\partial \tau}\frac{\alpha}{\sqrt{\beta}} = \frac{1}{\beta}\Big(\alpha'\sqrt{\beta} - \frac{1}{2\sqrt{\beta}}\alpha\beta'\Big) = \frac{1}{2\sqrt{\beta}\beta}(2\alpha'\beta - \alpha\beta').$$

It is enough to show that $2\alpha'\beta - \alpha\beta'$ is negative. This expression is equal to

$$-\frac{2}{2-\tau}\Big(\lambda^2\frac{\tau}{1+\tau} + (1-\lambda)^2\cos^2(t-x)\frac{1-\tau}{(2-\tau)^2}\Big) - (1-\tau)\Big(\lambda^2\frac{1}{(1+\tau)^2} - (1-\lambda)^2\cos^2(t-x)\frac{\tau}{(2-\tau)^3}\Big).$$

This is equivalent to

$$-\frac{\lambda^2}{1+\tau}\Big(\frac{2\tau}{2-\tau} + \frac{1-\tau}{1+\tau}\Big) + \frac{(1-\lambda)^2\cos^2(t-x)(1-\tau)}{(2-\tau)^3}\Big(\frac{2}{2-\tau} - \tau\Big) < 0,$$

which is easy to see is negative. This proves $B(\tau, x|t)$ is increasing in n and, *a fortiori*, that as n increases the voting behavior of type t becomes more correlated with $f(\theta_{id}, t)$. $\square$.

Proof of Lemma 3. Recall that $f(\theta_{id}, t) \sim \mathcal{N}(0,1)$ for all t, thus we can compute the probability that type t votes in favor of candidate A conditional only on θ_0. To do that we integrate $f(\theta_{id}, t)$ out. To do so we use the following property of normal distributions:

$$\int_{\mathbb{R}} \Phi(a+bx)d\Phi(x) = \Phi\left(\frac{a}{\sqrt{1+b^2}}\right).$$

We can apply the identity above to

$$\Pr\Big(\mathbb{E}_{\tau^\star, x_i^\star}(u(\theta_0, \theta_{id}, t)|s_{\text{v}}, s_{\text{id}}) > 0\Big) = \int_{\mathbb{R}} \Phi\Big(\frac{\lambda g(\tau)}{\nu}\theta_0 + \frac{(1-\lambda)g(1-\tau)\cos(t-x)}{\nu} f(\theta_{id}, t)\Big) d\Phi(f(\theta_{id}, t)),$$

and get the following expression:

$$\Pr\Big(\mathbb{E}_{\tau^\star, x_i^\star}(u(\theta_0, \theta_{id}, t)|s_{\text{v}}, s_{\text{id}}) > 0\Big) = \Phi\Big(A(\tau, x|t)\theta_0\Big),$$

where

$$A(\tau, x|t) := \frac{\lambda g(\tau)}{\sqrt{\nu^2 + (1-\lambda)^2 g(1-\tau)^2\cos^2(t-x)}}.$$

Hence, the probability that type t's voting behavior matches the first best depends on value of $A(\tau, x|t)$. In Lemma A6 we show that $A(\tau, x|t)$ is increasing in τ and decreasing in $|t-x|$, hence decreasing in n. This proves the claim. □

Lemma A6. *For every $t \in T$, the coefficient $A(\tau, x|t) \in [0,1]$ is increasing in τ and decreasing in $|t-x|$.*

Proof of Lemma A6. We compute the derivate $\frac{\partial A}{\partial \tau}$. Denote $A(\tau, x|t) := \alpha/\sqrt{\beta}$. Deriving with respect to τ we have

$$\frac{\partial A}{\partial \tau} = \frac{\partial}{\partial \tau}\frac{\alpha}{\sqrt{\beta}} = \frac{1}{\beta}\Big(\alpha'\sqrt{\beta} - \frac{1}{2\sqrt{\beta}}\alpha\beta'\Big) = \frac{1}{2\sqrt{\beta}\beta}(2\alpha'\beta - \alpha\beta').$$

It is enough to show that $2\alpha'\beta - \alpha\beta'$ is positive. We have that $\alpha = \lambda\frac{\tau}{1+\tau}$, $\alpha' = \lambda\frac{1}{(1+\tau)^2}$, $\beta = \lambda^2\frac{\tau}{(1+\tau)^2} + (1-\lambda)^2\cos^2(t-x)\frac{1-\tau}{2-\tau}$ and $\beta' = \lambda^2\frac{1-\tau}{(1+\tau)^3} + (1-\lambda)^2\cos^2(t-x)\frac{1}{(2-\tau)^2}$. Thus, $2\alpha'\beta - \alpha\beta'$ is equal to

$$2\lambda\frac{1}{(1+y)^2}\Big(\lambda^2\frac{\tau}{(1+\tau)^2} + (1-\lambda)^2\cos^2(t-x)\frac{1-\tau}{2-\tau}\Big) - \lambda\frac{\tau}{1+\tau}\Big(\lambda^2\frac{1-\tau}{(1+\tau)^3} + (1-\lambda)^2\cos^2(t-x)\frac{1}{(2-\tau)^2}\Big) =$$

$$= \frac{\lambda^2}{(1+\tau)^3}\Big(2\tau - \tau(1-\tau)\Big) + (1-\lambda)^2\cos^2(t-x)\frac{1}{2-\tau}\Big(\frac{2(1-\tau)}{1+\tau} - \frac{\tau}{2-\tau}\Big).$$

The first term is positive since $2\tau - \tau(1-\tau) > 0$. The second term is positive too since

$$\frac{2(1-\tau)}{1+\tau} > \frac{\tau}{2-\tau}$$

in the interval $\tau \in [0,1]$. It is straihgtforward to see that $A(\tau, x|t)|$ is increasing in $|t-q$. Thanks to these two facts, it's straightforward to see $A(\tau, x|t) \in [0,1]$. It's also trivial to see that $A(\tau, x|t)$ is decreasing in $|t-x|$. □

Lemma A7. *For every $t \in T$, the coefficient $A(\tau, x|t)$ is increasing in λ.*

Proof of Lemma A7. We compute the derivative with respect to λ to find that find that

$$\begin{aligned}
&\frac{\partial}{\partial\lambda}A(\tau,x|t) = \\
&= 2g(\tau)\Big[\lambda^2\frac{g(\tau)}{1+r} + (1-\lambda)^2\cos^2(t-x)g(1-\tau)\Big] - \lambda g(\tau)\Big[2\lambda\frac{g(\tau)}{1+r} - 2(1-\lambda)\cos^2(t-x)g(1-\tau)\Big] = \\
&= \lambda^2\frac{g(\tau)^2}{1+r} + (1-\lambda)^2\cos^2(t-x)g(1-\tau)g(\tau) - \lambda^2\frac{g(\tau)^2}{1+r} + \lambda(1-\lambda)\cos^2(t-x)g(1-\tau)g(\tau) = \\
&= (1-\lambda)\cos^2(t-x)g(1-\tau)g(\tau) > 0.
\end{aligned}$$

□

Proof of Proposition 4. As we showed in the proof of Lemma A2, each information provider picks τ to maximize the value its information structure for threshold types, i.e. $\frac{\partial t_r}{\partial \tau_i} = 0$. From Lemma 1 it is easy to see that, fixing τ, the value of an information structure (x, τ) for type t is decreasing in $|x - t|$. This implies that all voters consuming (x, τ) value this information more than the threshold types, and prefer lower τ. We have shown in Proposition 3 that as n increases, τ decreases and the expected distance between a voter and the closest information structure decreases. Both of these imply value of the closest news source to increase.

Proof of Theorem 1. First we compute the expected share of votes that a candidate with valence θ_v gets in a symmetric equilibrium when there are n competing firms.

$$\begin{aligned}
&\int_{\mathbb{R}}\sum_{i\in N}\int_{x_i^\star-\frac{\pi}{2n}}^{x_i^\star+\frac{\pi}{2n}} \Pr\Big(\mathbb{E}_{\tau^\star,x_i^\star}(u(\theta_v,\theta_{id},t)|s_{\text{v}},s_{\text{id}})>0\Big)dH(t)d\Phi(\theta_{id}) \\
&= \sum_{i\in N}\int_{\mathbb{R}}\int_{x_i^\star-\frac{\pi}{2n}}^{x_i^\star+\frac{\pi}{2n}} \Pr\Big(\mathbb{E}_{\tau^\star,x_i^\star}(u(\theta_v,\theta_{id},t)|s_{\text{v}},s_{\text{id}})>0\Big)dH(t)d\Phi(\theta_{id}) \quad \text{(Independence across firms)} \\
&= \sum_{i\in N}\int_{x_i^\star-\frac{\pi}{2n}}^{x_i^\star+\frac{\pi}{2n}}\int_{\mathbb{R}} \Pr\Big(\mathbb{E}_{\tau^\star,x_i^\star}(u(\theta_v,\theta_{id},t)|s_{\text{v}},s_{\text{id}})>0\Big)d\Phi(\theta_{id})dH(t) \quad \text{(Independence across types)} \\
&= \sum_{i\in N}\int_{x_i^\star-\frac{\pi}{2n}}^{x_i^\star+\frac{\pi}{2n}} \Phi\Big(A(\tau^\star,x_i^\star|t)\theta_v\Big)dH(t) \\
&= \frac{n}{\pi}\int_{x_i^\star-\frac{\pi}{2n}}^{x_i^\star+\frac{\pi}{2n}} \Phi\Big(A(\tau^\star,x_i^\star|t)\theta_v\Big)dt
\end{aligned}$$

Fix any $\theta_v > 0$. We want to show that $\frac{n}{\pi}\int_{x_i^\star-\frac{\pi}{2n}}^{x_i^\star+\frac{\pi}{2n}} \Phi(A(\tau^\star, x_i^\star|t)\theta_v)dt$ is decreasing in n. That is, the share of favorable vote for a a socially optimal candidate is decreasing in n. To make our point we assume we can differentiate in n and show that

$$\frac{\partial}{\partial n}n\int_{\frac{\pi}{2n}}^{\frac{\pi}{2n}} \Phi\Big(A(\tau^\star,0|t)\theta_v\Big)dt < 0.$$

Using Leibniz integral rule we have

$$\frac{\partial}{\partial n} n \int_{\frac{\pi}{2n}}^{\frac{\pi}{2n}} \Phi\Big(A(\tau^\star, 0|t)\theta_v\Big) dt = \int_{\frac{\pi}{2n}}^{\frac{\pi}{2n}} \Big[\Phi(A(\tau^\star, 0|t)\theta_v) + n\frac{\partial}{\partial n}\Phi(A(\tau^\star, 0|t)\theta_v)\Big]\, dt,$$

since $\Phi(A(\tau^\star, 0|\frac{\pi}{2n})\theta_v) = \Phi(A(\tau^\star, 0| - \frac{\pi}{2n})\theta_v)$. It's enough to show that for all t,

$$\frac{\partial}{\partial n} \ln \Phi(A(\tau^\star, 0|t) < -\frac{1}{n}.$$

To see this we notice two facts. First, from Lemma A6, as n increases, $\Phi(A(\tau^\star, 0|t)$ decreases, and so does $\ln \Phi(A(\tau^\star, 0|t)$. Thus, the derivative is negative. Second, $\Phi(A(\tau^\star, 0|t)$ is always smaller than 1. The derivative of the log in such interval has magnitude bigger than 1. □

Proof of Theorem 2. This result follows from the proof of Theorem 1 and from Lemma A7. □

References

Alesina, A., R. Baqir, and W. Easterly (1999): "Public goods and ethnic divisions," *The Quarterly Journal of Economics*, 111.

Aragones, E., M. Castanheira, and M. Giani (2015): "Electoral Competition through Issue Selection," *American Journal of Political Science*.

Ashworth, S. and E. de Mesquita (2009): "Elections with Platform and Valence Competition," *Games and Economic Behavior*.

Baker, E. C. (1978): "Scope of the First Amendment Freedom of Speech," *UCLA Law Review*.

Baron, D. (1994): "Electoral competition with informed and uninformed voters," *American Political Science Review*.

——— (2006): "Persisten media bias," *Journal of Public Economics*, 90, 1–36.

Bernhardt, D., S. Krasa, and M. Polborn (2008): "Political Polarization and the Electoral Effects of Media Bias," *Journal of Public Economics*.

Besley, T. and A. Prat (2006): "Handcuffs for the grabbing hand? Media capture and government accountability," *American Economic Review*.

Cage', J. (2014): "Media competition, information provision and political participa- tion," *Working Paper*.

Carillo, J. and M. Castanheira (2008): "Information and Strategic Political Polarization," *Economic Journal*.

Chan, J. and W. Suen (2008): "A spatial theory of news consumption and electoral competiton," *Review of Economic Studies*.

Dellavigna, S. and E. Kaplan (2007): "The Fox News Effect: Media Bias and Voting," *Quarterly Journal of Economics*, 122.

Downs, A. (1957): *An Economic Theory of Democracy*, New York: Harper.

Dragu, T. and X. Fan (2015): "An Agenda-Setting Theory of Electoral Competition," *Working Paper*.

Duggan, J. and C. Martinelli (2011): "A Spatial Theory of Media Slant and Voter Choice," *Review of Economic Studies*.

Eyster, E. and T. Kittsteiner (2007): "Party platforms in electoral competition with heterogeneous constituencies," *Theoretical Economics*, 2, 41–70.

Federal Communications Commission (2003): *Report and Order and Notice of Proposed Rulemaking.*, FCC 03-127. Washington, DC: Federal Communications Commission.

Fernandez, R. and G. Levy (2008): "Diversity and redistribution," *Journal of Public Economics*.

Gentzkow, M. (2006): "Television and Voter Turnout," *Quarterly Journal of Economics*, 121.

Gentzkow, M. and E. Kamenica (2015a): "Competition in Persuasion," *Working paper*.

——— (2015b): "Information environments and the impact of competition on information provision," *Working Paper*.

Gentzkow, M. and J. Shapiro (2004): "Media, Education, and Anti-Americanism in the Muslim World," *Journal of Economic Perspectives*, 18.

——— (2006): "Media bias and reputation," *Journal of Political Economy*.

——— (2008): "Competition and truth in the market for news," *Journal of Economic Perspective*.

——— (2010): "What drives media slant? Evidence from U.S. daily newspapers," *Econometrica*.

Gentzkow, M., J. Shapiro, and M. Sinkinson (2011): "The effect of Newspaper entry and exit on electoral politics," *American Economic Review*.

Gentzkow, M., J. Shapiro, and D. Stone (2014): "Media Bias in the Marketplace: Theory," *NBER Working Papers*.

Gerber, A., D. Karlan, and D. Bergan (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, 35–52.

Groseclose, T. (2001): "A Model of Candidate Location When One Candidate Has a Valence Advantage," *American Journal of Political Science*.

Grossman, G. and E. Helpman (1996): "Electoral competition and special interest politics," *Review of Economic Studies*.

Gul, F. and W. Pesendorfer (2012): "Media and Policy," *Working Paper*.

Hotelling, H. (1929): "Stability in Competition," *The Economic Journal*.

Lizzeri, A. and N. Persico (2001): "The provision of public goods under alternative electoral incentives," *American Economic Review*, 91.

Milgrom, P. and J. Roberts (1986): "Relying on the information of interested parties." *RAND journal of economics*, 17, 18–22.

Mullainathan, S. and A. Shleifer (2005): "The market for news," *American Economic Review*.

Prat, A. and D. Stromberg (2005): “Commercial Television and Voter Information,” *CEPR Discussion Paper.*

Salop, S. (1979): “Monopolistic Competition with Outside Goods,” *The Bell Journal of Economics.*

Stokes, D. (1963): “Spatial Models of Party Competition,” *American Political Review.*

Stromberg, D. (2004): “Radio’s Impact on Public Spending,” *Quarterly Journal of Economics*, 119, 189–221.

Sunstein, C. (2001): *Republic.com*, Princeton University Press.

——— (2002): “The Law of Group Polarization,” *The Journal of Political Philosophy.*

Yuksel, S. (2015): “Specialized Learning and Political Polarization,” *Working Paper.*