Media Competition and the Source of Disagreement

Jacopo Perego
New York University

Sevgi Yuksel
UC Santa Barbara

April, 2017

Abstract

We identify a novel channel through which competition among information providers decreases the efficiency of electoral outcomes. The critical insight we put forward is that the level of competition in the market determines the *type* of information that is provided in equilibrium. In our model, 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 the level of competition in the market determines how much firms differentiate in terms of the type of information they produce. Importantly, differentiation leads to higher provision of information on issues where there is higher disagreement in the electorate. Although voters become individually better informed, voting decisions shift from focusing on valence issues to ideological issues. On aggregate, the share of votes going to the socially optimal candidate decreases. Our model also highlights how competition in the market for news can have negative welfare consequences even in the absence of behavioral agents or partisan media, therefore offering a new, and to some extent more distressing, perspective on the problem.

JEL Classification Numbers: D72, L15, L82.

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

1. Introduction

This paper identifies a new channel through which competition in the media market intensifies disagreement in a society, and thereby decreases the efficiency of electoral outcomes. The critical insight we put forward is that the *type* of information provided to the voters changes as competition among news sources increases. We show that competition pushes media to become more informative on issues where there is greater heterogeneity in voters' preferences. Higher emphasis of these issues allows media to differentiate and specialize in their informational products. In this sense, competition among information providers does not create disagreement, but rather uncovers the underlying heterogeneity in voters' preferences, consequently deepening its negative effect on electoral outcomes.

In our model, a number of profit-maximizing information providers compete to sell information to a group of Bayesian agents, who are seeking to learn about how two political candidates compare to each other. Candidates may differ along several dimensions such as their technical competence, international credibility, or in terms of their position on issues such as affirmative action, immigration, or redistribution. Voters' preferences are heterogeneous in two fundamental ways. Voters can disagree on which dimensions are important to them (their *agenda*) and on how each issue in their agenda should be addressed (their *slant*). We show how this kind of heterogeneity creates the "informational space" on which news sources can differentiate and specialize in terms of their informational products. We show that, as competition intensifies, news sources increasingly emphasize ideological issues, where voters' preferences are more heterogeneous, over valence issues where voter preferences are highly aligned. Our main result shows that, as competition increases, voting becomes more ideological, and the share of votes going to the welfare optimal candidate monotonically declines.

This mechanism illustrates clearly how the news market differs from traditional markets. Markets efficiently respond to demand from individuals, and ours is not an exception. Competition generates differentiation in the type of information provided, creating a spectrum of options for consumers. This is, in fact, beneficial at the individual level. Voters are able to find news sources which provide the type of information tailored towards their needs and hence learn more effectively about

which candidate is optimal for them.[1] Indeed, in this sense, competition generates a more informed electorate. 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 the whole society via the election outcome. The source of the inefficiency lies precisely in the fact that individual preferences are aggregated at the election stage. This stage introduces a wedge between the type of information that is valuable to voters on an individual level and the type of information that is relevant for determining the socially optimal candidate, a feature that is peculiar of this particular market. 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 and decreases the likelihood that any individual vote coincides with the welfare optimal candidate.

Our paper contributes to a growing academic and public debate on how competition in the market for news can affect voters' beliefs and consequently political outcomes. The traditional perspective on this matter has been that truth prevails in a competitive market. A "marketplace for ideas" promotes truth, because it brings out a diversity of perspectives which allows people to learn more effectively.[2] Yet, it has been noted that forces that characterize competition in the media environment are very different from those described in traditional models (Baker (1977)). 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. The debate on the effects of competition has gained significance with recent changes in the news market. With the dramatic expansion in Internet access, there is an unprecedented number of news sources available to a growing share of the

---

[1] In a sense, competition in our model endogenously generates what political scientists have often referred to as the "Daily Me": a fictitious newspaper crafted around one's unique tastes. For example, see Sunstein (2001).

[2] One of the basic tenets of the US communications policy is 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). See, for example, Federal Communications Commission (2003).

population. Despite this expansion, we have also witnessed a dramatic polarization in beliefs of the American public.[3]

In this paper, we show how competition among news sources can lead to further polarization of beliefs. The channel we discuss in this paper is conceptually different from those that have been studied in the literature. It does not rely on either behavioral voters or partisan media, ingredients that, by their own nature, could introduce inefficiencies in the model.[4] In our model, news sources are profit-maximizers, uninterested in the outcome of the election; and agents, who consume these news sources, are expected utility-maximizers, sharing common ex-ante beliefs about the candidates. Instead, we posit that candidates differ over a spectrum of characteristics, and that voters disagree on both the importance and the desirability of these characteristics.[5] Allowing for voters to disagree on *agenda* implies that voters who care about different issues will demand different information structures. While considering such heterogeneity is natural, it presents a challenge in terms of developing a tractable model which can be used to study the effects of competition. We present a model which integrates both types of heterogeneity (*agenda* and *slant*) in a simple way. Specifically, we map the distribution of voter preferences into a circle, and show that the arc running between any two types of voters measures the correlation in their preferences. Mapping the population of voters onto a circle provides us with a convenient framework in which all sources of heterogeneity among voters can be reduced to the correlations in political views. Ultimately, it allows us to consider the game among information providers as a spatial competition model and study comparative statics on the number of news sources in the market.

---

[3]People that self-identify 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. 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%."

[4]See, for example, Mullainathan and Shleifer (2005) and Baron (2006).

[5]This idea goes back to Stokes (1963), in which it is recognized that "what is needed is a language that would express the fact that different weights should be given to different dimensions at different times." In recent surveys on policy priorities, a striking partisan divide emerged over the importance of a number of different political issues, such as the environment, dealing with the poor and needy, strengthening the military, etc. See, for example, Pew Research Center's "Public's Policy Priorities Reflect Changing Conditions At Home and Abroad," January (2015).

Differentiation of news sources is a natural consequence of competition, as it affects firms' incentives to cater to the distribution of preferences in the population. We solve for the equilibrium of the game where firms choose which information structure to offer to maximize their readership, and voters choose which news sources 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 welfare - would take if he could observe exactly how the two candidates compare to each other.

In equilibrium, the type of information provided by any news source is chosen optimally to target a subset of voters. Hence, how much information is provided on each issue by a specific news source depends on how much overlap there is in the preferences of the consumers of this news source on this issue. 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 further in terms of the informational product they offer to the market, leading overall to more information on issues where there is disagreement. 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 are 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.

In this paper, 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.

The paper is structured 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. In Section 7, we generalize our results to cases where voters can consume multiple news sources. Finally, Section 8 provides a discussion of our results in relation to possible extensions, while Section 9 concludes.

2. Related Literature

There is extensive empirical evidence showing that the structure of the news market can have a significant impact on political attitudes and electoral outcomes. There is also a growing political and economic literature investigating specifically the possible welfare consequences of media competition. We refer the reader to Gentzkow and Shapiro (2008) 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 more likely for news sources to remain independent when there is a threat of government capture (Besley and Prat (2006)). Second, competition can lead to the proliferation of a spectrum of news with diverse viewpoints in environments where firms may have incentives other than accurately reporting the truth. In this context, Milgrom and Roberts (1986) and more generally Gentzkow and Kamenica (2016) study the 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 in competitive regimes.[6] Finally, in addition to these forces, it has been argued that competition among information providers can lead to greater investment in faster, higher quality news (Gentzkow and Shapiro (2008)).

---

[6]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.

6

The impact of competition is more likely to be negative when there is demand driven bias in news. Demand driven bias results from the incentives of the news sources to pander to their 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. Following on this idea, Genzkow et al. (2014) provide historical evidence that voters prefer like-minded news and that newspapers strategically use their political orientations to differentiate from competitors to increase readership and advertisement revenue.[7] 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, there is mixed evidence on this. Prat and Stromberg (2005) show that the introduction of private television in Sweden increased political information and political participation relative to a public television monopoly. On the other side, Cagé (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 lead to a decrease in voter turnout.

Our paper differs from these models in that we abstract from these types of distortions considered in the literature. In our model, on the supply side, firms are profit-maximizers - unbiased with no ideological preferences. On the demand side, voters are Bayesian expected utility maximizers. As a consequence, in contrast to

---

[7]Many scholars have voiced concerns over the potential effects of media competition if voters suffer from confirmation-bias. It has generally been argued that increasing competition can exacerbate bias in information provision by allowing consumers to self-segregate more effectively in terms of priors. 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 sources turn into echo-chambers, where citizens only hear news precisely in line with their priors and there is no effective learning.

the existing literature, our model predicts competition to bring about a *more* informed electorate.[8] 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 ideology and valence in political competition. Besley and Prat (2006), Alesina et al. (1999), Lizzeri and Persico (2001), Lizzeri and Persico (2005), and Fernandez and Levy (2008) study this interaction by investigating how preference heterogeneity affects public good provision. Relatedly, Eyster and Kittsteiner (2007), Groseclose (2001), Carillo and Castanheira (2008), and Ashworth and de Mesquita (2009) present models which study the interaction between valence competition and party platforms.[9] 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 common insight in these papers is that preference heterogeneity, either intrinsic or accentuated through platform divergence, will hurt valence competition by decreasing its importance on individual voting decisions. Our paper studies how competition in the media market affects the interaction between ideology and valence in the absence of any changes in voter preferences or party platforms.

The idea that voters can have different preferences about different issues has been the object of recent studies in the political economy literature. Aragones et al. (2015), Dragu and Fan (2015), and Yuksel (2015) are similar to ours in that they allow voters to have different preferences about what is important to them. However, these papers focus on different aspects of the 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, instead, studies polarization of party platforms. Yet, the critical role that this kind of preference heterogeneity could have on the media market has not been emphasized in the literature.

---

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

[9]More recently, Bandyopadhyay et al. (2015) investigate how high valence candidates could pick extreme party platforms to attract media coverage which consequently helps them signal strength.

8

Finally, 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 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 the 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

9

evaluates candidate A relative to B. We assume that u takes the following simple form:

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

where $\lambda \in (0, 1)$ represents how voters 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).[10]

The kind of heterogeneity in voters' preferences represented in the model will naturally depend on the joint assumptions made on the functional form of f and distribution of θ_{id} and T. Our goal is to capture, with the simplest model, heterogeneity in the population both in terms of *agenda* and *slant*. For example, as two potential extreme cases, we want our model to allow for the ideological preferences of two voters to be fully misaligned, or alternatively to be completely orthogonal. The former would correspond to a situation in which the two types care about the exact same issues (identical *agenda*) but disagree completely on how these issues should be addressed (opposite *slant*). The latter would correspond to a situation where the two types employ completely different criteria to compare the two candidates (non-overlapping *agenda*) along the ideological dimension θ_{id}. For instance, it could be that one type only cares about economic issues, while the other one cares only about social issues. Moreover, we naturally would like our model to also allow for intermediate cases, where two voters can be arbitrary "close" in terms of how correlated their ideological preferences are.

We adopt a model that produces these features in a very simple way. 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

---

[10]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.

with mean zero and unit variance.[11] The way ϑ_1 and ϑ_2 are *mixed* depends on a voter's type and creates the heterogeneity in our model. In particular, we assume for all $t \in T := [-\pi, \pi]$:

$$f(\theta_{id}, t) := \vartheta_1 \cos(t) + \vartheta_2 \sin(t).$$

The interpretation for this specification naturally follows from our motivation. People are different on two levels. They can disagree on which issues are important for them (*agenda*) and on how each of these issues should be addressed (*slant*). Our model allows voters to disagree on whether a high realization of ϑ_1 makes candidate A more desirable relative to candidate B, or vise versa. In addition, it also allows for voters to disagree on how important ϑ_1 is relative to ϑ_2.[12] Conveniently, both types of heterogeneity can be tracked by one's type t. This is a crucial feature of our model because, as it will become clear later on, it generates *the space* on which information providers can diversify their products. Finally, to preserve symmetry, we assume voters to be distributed uniformly on T, that is $t \sim \mathcal{U}(T)$. We discuss robustness to this distributional assumption in Section 8.1.

This specification has several convenient features. First, it ensures that all voters *ex ante* value information the same amount.[13] Second, as we show in the appendix, this specification ensures that 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. Finally, we get a continuum of possible correlations between different types, enriching the heterogeneity among voter preferences that we can account for. In particular, for every t, there are a pair of types $t \pm \frac{\pi}{2}$ who have *orthogonal* preferences to t - meaning that making $t \pm \frac{\pi}{2}$ happier can leave t indifferent - and a

---

[11] 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 candidates A and B. In such case, the model 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 other.

[12] 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 issue. Among those who care about affirmative action, there can be voters who are for or against it.

[13] This is guaranteed by the fact that the variance of $u(\boldsymbol{\theta}, t)$ is independent of t. 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.
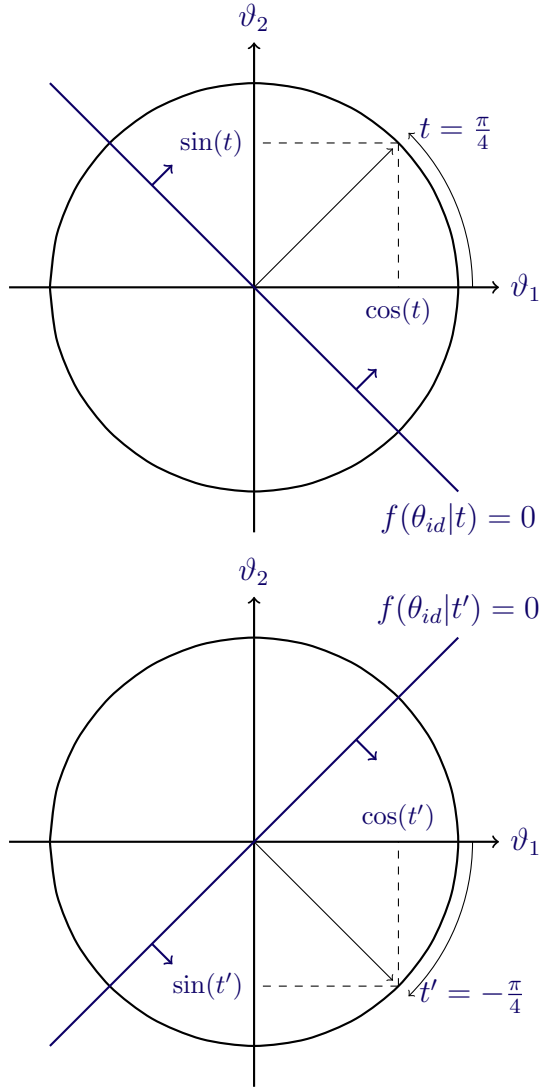
FIGURE 1: The level curves of $f(\theta_{id}, t)$ for two voters: $t = \frac{\pi}{4}$ and $t' = -\frac{\pi}{4}$.

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 on

ideology and for this reason it is relatively more polarized. In Section 8.1, we discuss extensions and robustness of most of these assumptions.

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 is comprised of two independent signals, one for valence and one for ideology. In choosing how accurate these signals are, we assume the news sources 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 source's reporting strategy can be characterized by a *precision* $\tau \in [0, \bar{\tau}]$ and a *position* $x \in T$. This implies that the news source generates two *independent* signals:

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

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 side (in the production of news) and from the demand side (in the consumption of news). Our results until Section 7 do not depend on the specific bound $\bar{\tau}$ for total precision.

3.3 Social Planner

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 maximizes

aggregate voters' welfare.

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

$$r^{SP}(\boldsymbol{\theta}) = 1 \quad \textit{iff} \quad \frac{1}{2\pi} \int_T u(\theta, t) dt > 0$$

Our definition is straightforward. The social planner selects the candidate A if and only if A generates higher welfare for the society than candidate B. 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 have conflicting preferences. 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 voter on any of these dimensions necessarily implies "harming" another.[14] 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.,*

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

However, in light of the restrictions we imposed on the information structures available to the news sources, there is 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

---

[14]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 $\boldsymbol{\theta}$, $f(\theta_{id}, t) = -f(\theta_{id}, t + \pi)$.

about θ_v, by setting $\tau = \bar{\tau}$. The next definition builds on this intuition.

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

$$r^{SP}(\boldsymbol{\theta}) = 1 \quad \textit{iff} \quad \frac{1}{2\pi}\mathbb{E}\Big(\int_T u(\theta, t)dt \,|\, s_v(\boldsymbol{\theta})\Big) > 0$$

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

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

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 type t. This, in turn, will reveal 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 players directly receive utility from voting for a candidate. In this sense, we put aside the issue of why people vote. Indeed, in
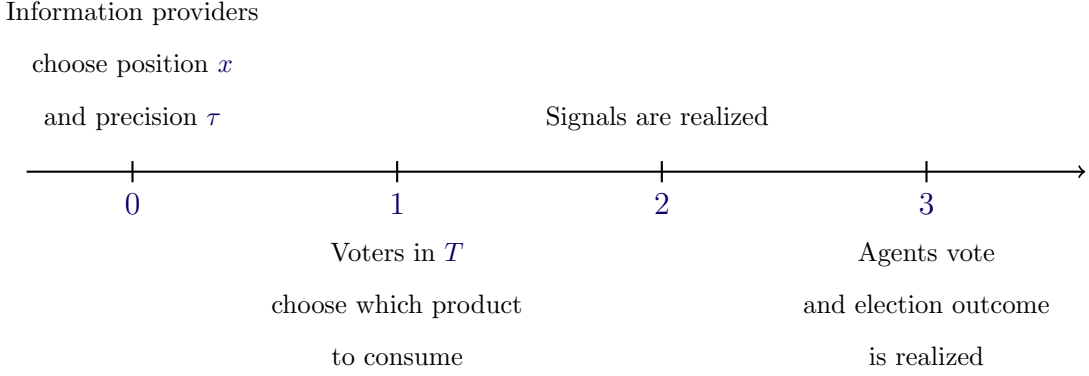
FIGURE 2: Timeline of the game.

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.[15]

Now fix an information structure (τ, x). Given signal realizations s_v and s_id, a voter of type t will compute $\mathbb{E}_{\tau,x}(u(\boldsymbol{\theta}, t)|s_\mathrm{v}, s_\mathrm{id})$, his subjective expected valuation of candidate A given the information received, and will vote for A if and only if this expectation is positive. The *value* of information structure (τ, x), denoted $V(\tau, x|t)$ is therefore defined as:

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

Intuitively, the value of an information structure for a voter of type t represents how much better off that type is expected to be after having received signals from (τ, x) relative to receiving no signals at all. The function $V(\tau, x|t)$ represents a key component in our model as it will determine which information provider a voter of type t will want to acquire information from. Lemma A1 in the Appendix computes analytically $\mathbb{E}_{\tau,x}(u(\boldsymbol{\theta}, t)|s_\mathrm{v}, s_\mathrm{id})$. In the next proposition, instead, we compute and characterize analytically the value of information in this game. To this purpose, let $\sigma^2(\tau, x|t)$ be the variance of the random variable $\mathbb{E}_{\tau,x}(u(\boldsymbol{\theta}, t)|s_\mathrm{v}, s_\mathrm{id})$.

Proposition 2. *The value of information (τ, x) for voter t is strictly increasing in the variance $\sigma^2(\tau, x|t)$ and can be decomposed into two components with the following characteristics:*

---

[15]In Section 8.4, we discuss the robustness of our main results to strategic voting.

- a Public *component that is type-independent and increasing in τ.*

- a Private *component that is type-dependent and decreasing in τ and $|t - x|$ (for $|t - x| < \frac{\pi}{2}$)[16].*

It is intuitive that the value of an information structure is a monotonic transformation of $\sigma^2(\tau, x|t)$. This represents the variance in ex-post political preferences of type t that will be induced by information structure (τ, x). When this variance is higher, it is more likely that the voter's 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 will be the better candidate and, ultimately, the more informative (τ, x) is for such voter.

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 a 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 two 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.

In the Proof of Proposition 2, we show how the variance $\sigma^2(\tau, x|t)$ can be written as

$$\sigma^2(\tau, x|t) = \lambda^2 g(\tau) + (1 - \lambda)^2 \cos^2(t - x) g(\bar{\tau} - \tau),$$

where $g(\tau) = \frac{\tau}{1+\tau}$. From the formula above, we can notice that two information

---

[16]Below, we discuss why it is sufficient to focus on this domain.

structures with the same precision τ, but with different locations x and x', which lie at any two opposite ends of the circle, produce identical values for all voters $t \in T$. This is because the two information structures provide signals on ideology that are perfectly negatively correlated. The first signal is centered around $f(\theta_{id}, x)$, while the second one is centered around $f(\theta_{id}, x + \pi) = -f(\theta_{id}, x)$. Therefore, the informativeness of these two information structures are the same. For this reason, it is without loss of generality for the equilibrium analysis to restrict the location decision x of the information providers to $x \in [-\frac{\pi}{2}, \frac{\pi}{2}] \subset T$.
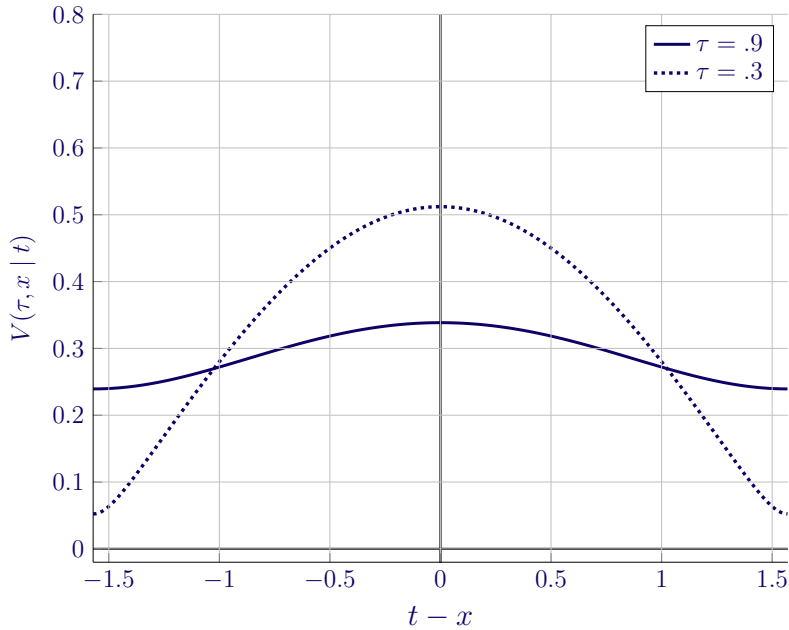

FIGURE 3: The value of information (τ, x) for type t (where $\bar{\tau}$ is set to 1). On the horizontal axis, $t - x$ represents the distance of given type from her information provider.

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 two implicit assumptions 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 com-

18

plexity. More significantly instead, we assume that voters are consuming at most one information structure. We show that our results are robust to relaxing this assumption in Section 7.

5. Information Providers' Problem

In the first stage of this game, a set $N := \{1, \ldots, n\}$ of information providers compete by simultaneously choosing their strategies $(\tau_i, x_i) \in [0, \bar{\tau}] \times [-\frac{\pi}{2}, \frac{\pi}{2}]$. 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 *readership*, the share of voters who choose to acquire information from them. That is, they maximize 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 is positive, is often negligible. More than the price, it is 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 is convenient to visualize the information provider's problem as the choice of a location on a disk. Each firm chooses a position (angle) x and a precision on valence (distance from the center) 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.[17] The second choice, τ, specifies the

---

[17]Note that, fixing τ, setting x to $\frac{\pi}{2}$ and $-\frac{\pi}{2}$ would correspond to providing ideological signals centered around $f(\theta_{id}, \frac{\pi}{2})$ and $f(\theta_{id}, -\frac{\pi}{2})$ which are perfectly negatively correlated. Hence, they would generate the same value for all types. Therefore, 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 this is where we will solve the location

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 disk.
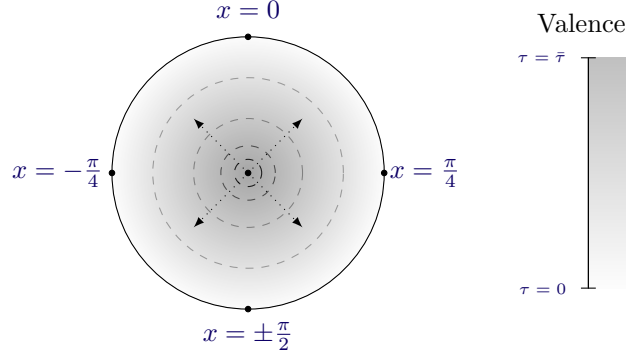

FIGURE 4: Mapping the firm's problem into a circle.

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 $(N, ([0, \bar{\tau}] \times [-\frac{\pi}{2}, \frac{\pi}{2}], \Pi_i)_{i \in N})$ 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 other* $|x_i - x_j| = \frac{\pi}{n}$.

Every symmetric rotation or re-shuffling of the firms' locations $\boldsymbol{x}$ would still be an equilibrium. Thus, provided they exist, 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

problem of the firms.

in these two values completely characterize the impact of competition on welfare. This is the sense in which we discuss uniqueness of symmetric equilibria.

Before we study competition among multiple information providers, let us notice that the monopolist's problem is trivial. For a monopolist, any choice of τ and x would capture the whole market and thus maximize profits. This multiplicity is just an artifact of the absence of prices in the model and therefore disappears as soon as one allows the monopolist to also set a price.[18] Therefore, in the remainder of the paper, we focus on cases where $n \geq 2$, and study comparative statics as n increases.

Proposition 3. *There exists a symmetric Nash equilibrium $(\tau^\star, x^\star)$. For $n \geq 2$, this equilibrium is unique in the class of symmetric equilibria, up to rotations and permutations of the locations.*

An important part of 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 4. *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

---

[18]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. $\tau^\star = \bar{\tau}$. The intuition for this is straightforward. For any other choice of $\tau < \bar{\tau}$, 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.

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 $\bar{\tau} - \tau^\star$ 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 radius 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 $(\bar{\tau} - \tau^\star)$ 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.
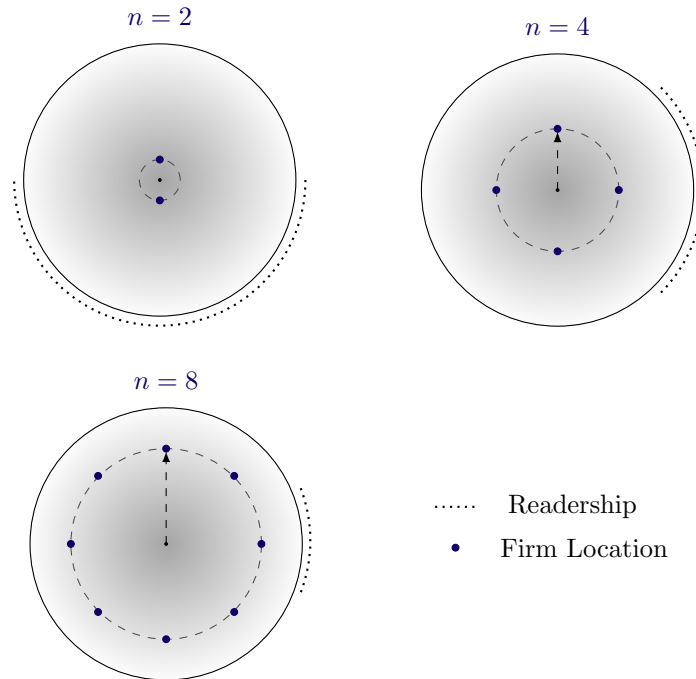

FIGURE 5: The representation of the symmetric equilibrium for several values of n.

22

6. Elections

In this section, we study how changes in 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 dimensions.

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 the realizations of the signals s_v and s_{id}, her expected utility $\mathbb{E}_{\tau,x}(u(\boldsymbol{\theta}, t)|s_v, s_{id})$ 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 is easy to derive the expression for μ:

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

From Proposition 4, 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 1. *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 2. *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 this impacts voters individually. Proposition 5 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 5. *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 5 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 $\boldsymbol{\theta}$,
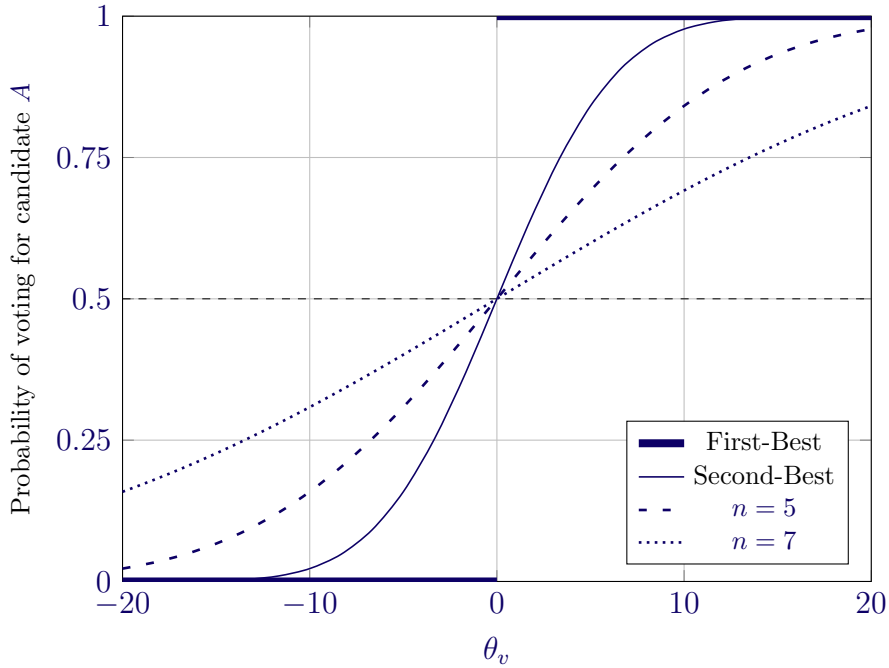
24

FIGURE 6: The probability of voting for A conditional on θ_v for an arbitrary voter t.

Voting for A is socially optimal if only if $\theta_v > 0$.

voters acquiring information from the same or different information sources receive mutually independent signals $(s_{\mathrm{v}}, s_{\mathrm{id}})$. 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 affecting ex-post beliefs.[19]

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*

---

[19]More formally, a voter who acquires information (τ, x) receives two signals: $s_{\mathrm{v}} = \theta_v + \varepsilon_{t,v}$ and $s_{\mathrm{id}} = f(\theta_{id}, x) + \varepsilon_{t,id}$. The error terms are s.t. $\varepsilon_{t,v} \sim \mathcal{N}(0, \tau^{-1})$ and $\varepsilon_{t,id} \sim \mathcal{N}(0, (\bar{\tau} - \tau)^{-1})$. 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 forms of correlation in the error terms, e.g. when *different* types acquiring information from the *same* source receive the *same* signal.

optimal candidate decreases.

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 corresponding from the stochastic behavior of *noise voters*.[20] Second, the distribution of votes affects outcomes directly in proportional electoral systems, or more generally whenever a mixture of majoritarian *and* proportional systems is used. Lastly, the distribution of votes can affect voters' welfare indirectly, by putting pressure on the winning candidate to compromise with the other side.

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 much voters care about valence relative to 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.*

Theorem 2 shows that the inefficiency associated with competition is exacerbated by

---

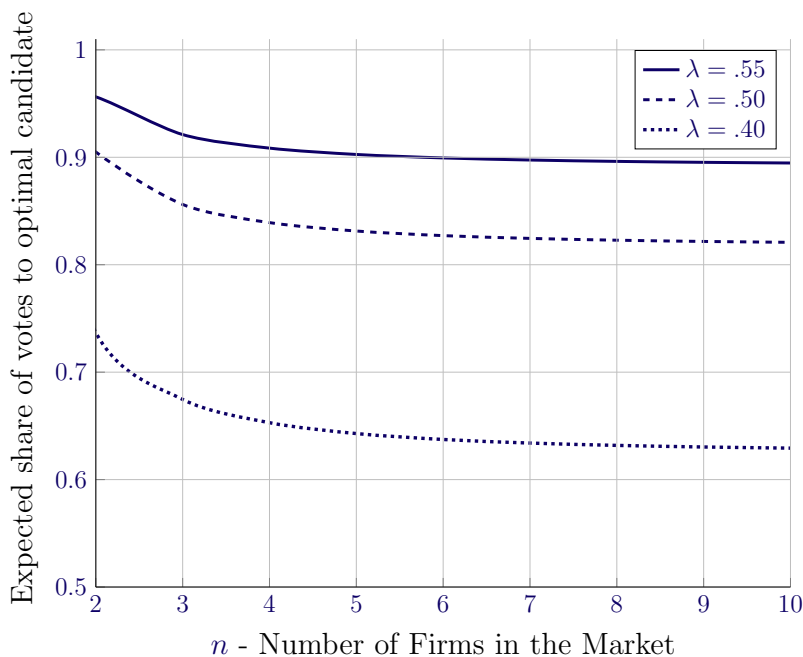[20]See Baron (1994) and Grossman and Helpman (1996).

FIGURE 7: A graphical representation of the main comparative statics: n and λ

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. Consuming Multiple News Sources

An important assumption we maintained in the paper so far is that each voter chooses one and only one information provider. With this assumption, we showed that each voter chooses to consume the news source that is closest to her, namely the news source whose informational product is most correlated with her own preferences. This significantly reduced the complexity of the game played by the news sources. It allowed us to demonstrate the main forces driving our results in a simple and transparent way. We studied 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 coun-

teracting force. Assume that voters can freely access all signals produced by the n information providers.[21] As n increases, voters would have access to an increasing number of signals on valence and ideology and, without additional structure, it is not possible to determine how this would affect voting behavior. Yet, assuming that voters can consume *every* signal produced by the market, irrespective 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, cognitive constraints, opportunity costs, etc.

In this section, we show that our main result generalizes to allowing voters to consume multiple news sources. We assume there is a cap $\kappa \in \mathbb{N}$ on the number of news sources a voter can consume, and we study how competition affects electoral outcomes when this constraint is binding, namely when $n \geq \kappa$.

Theorem 3. *Let $2\kappa\bar{\tau} < 1$ and $n \geq \kappa$. A symmetric equilibrium always exists. Moreover, as competition increases, the expected share of votes going to the socially optimal candidate decreases.*

Once again, firms compete for readership. When voters can consume multiple news sources, the challenge is to define readership for each news source carefully. Theorem 3 puts constraints on how much information voters are able to extract from κ news sources. This provides a sufficient condition under which voters always pick the news sources that are closest to them. That is, the optimal learning strategy entails choosing the κ news sources that are individually ranked highest (for type t) in terms of the value associated with the information structure they provide.[22]

Once this is established, the game played among the news sources can be mapped back to the $\kappa = 1$ case we've solved before with adjustments on how market share is defined. For example, if $n = 8$ and $\kappa = 2$ as shown in Figure 8, each firm will cater to a quarter of the market with neighboring firms serving overlapping shares of the population. However, once these adjustments are made, forces underlying

---

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

[22]Without any constraints on how much information can be transmitted with κ news sources, we can encounter situations where voters optimally choose to consume a different set of news sources. In these examples, voters care about how symmetrically distributed the news sources are around t more than how informative the news sources are individually.

the structure of the symmetric equilibria are identical to the $\kappa = 1$ case. Each firm will choose their reporting strategy to maximize the value of the signals they provide for their most extreme readers - the threshold types that are indifferent between consuming this news source and another. The only difference will be that each news source will effectively be competing over these threshold types with news sources that are κ to the right and to the left.
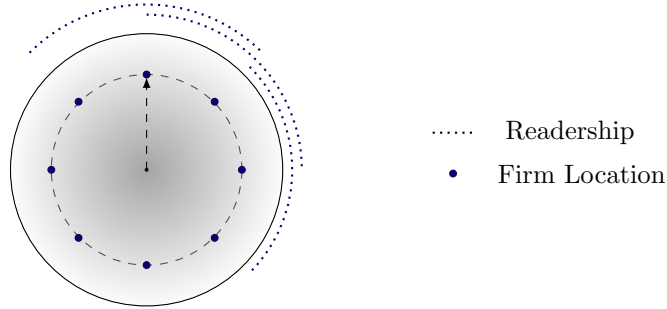

FIGURE 8: The representation of a symmetric equilibrium for $n = 8$ and $\kappa = 2$. Overlapping readership is marked for the three adjacent firms located in the first quadrant.

The competitive tensions that this situation generates are very similar to the $\kappa = 1$ case. In fact, as before, firms will choose precision of their signal on ideology relative to valence depending on how correlated the preferences of their readers are on ideology. For any κ, as n increases the market will be segmented into smaller and smaller groups with more correlated preferences. Consequently, news sources will shift focus to ideological issues, foregoing those customers that are "far away" from their location, thus creating more value for those that are close by.

8. Discussion

8.1 Voters' heterogeneity and the distribution of preferences

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, θ_{id}, and not on va-

lence, θ_v. Furthermore, since $t \sim \mathcal{U}(T)$, the distribution of preferences is symmetric on all issues labeled as ideology, making them essentially *zero-sum* issues in aggregate for the society. It is precisely for this reason that the planner's solution does not depend on the comparison of the candidates on ideology. Hence, any information provided to the voters on these issues generates higher disagreement, leading to a decrease in the share of votes going to the welfare-optimal candidate. Our assumptions on the structure and distribution of preferences are, of course, *only* a crude description of the environment we are studying and intended to capture its most general characteristics. They serve the purpose of making the illustration of the mechanism behind our main result more transparent and our analysis more tractable.

In reality, the *true* preferences of voters are likely to be highly multi-dimensional, and voters, on their own, do not differentiate between issues as valence and ideology. While their preferences are vastly heterogeneous, to a certain extent, they are also correlated. We think of valence as capturing the principal component of such heterogeneity, a statistical dimension along which voters' preferences are maximally correlated and with respect to which the residual heterogeneity, which we call ideology, is indeed *zero-sum.* Reformulating the preference space in these terms is essentially without loss of generality. From this point of view, the welfare-optimal solution of our model is, by construction, associated with valence, as it maximizes that particular mixture of political issues, whose exact composition can indeed be quite complex, along which preferences are maximally aligned.

A second aspect of voters' heterogeneity that we have simplified is the relative weight that each individual puts on valence. This was done by assuming that λ is type-independent; that is, we focused on a population of voters for which individual preferences are equally correlated with valence. Given the valence-ideology separation we discussed above, it is still possible that, in the real-world, different voters put different weights on valence.[23] Relaxing this assumption, letting the preferences of some voters be more correlated than others with valence would make the analytical solution for the firms' equilibrium reporting strategies significantly more cumbersome, without altering the main equilibrium tension of our model: namely,

---

[23]In the framework of our model, the distribution of types would correspond to a distribution on a disk, rather then on a circle.

that as n increases, firms struggle to differentiate their information structures by shifting precision away from valence in favor of ideology (Proposition 4).

This brings us to a fundamental question: What assumptions in our model are truly responsible for such a relative shift of precision from valence to ideology? This result is a direct consequence of the fact that voters disagree on what issues are important to them, namely that there is heterogeneity in *agenda*. Allowing agents to disagree on agenda implies that voters who care about different issues will demand different information structures. In setting up our model, we have implicitly assumed that heterogeneity in *agenda* plays a more important role on issues where there is also significant heterogeneity in *slant*, namely ideological issues. We find this assumption natural and compelling. Our goal is to capture a world in which information on ideology necessarily needs to be *targeted*. Interpretation of our model along these lines highlights why higher levels of competition among information providers necessarily shifts focus from valence to ideology. Competition leads to differentiation and specialization, which brings out components of voters' preferences that are heterogeneous. From this perspective, it is important to notice that, in our model, information providers are not *creating* disagreement but, rather, *uncovering* the primitive heterogeneity that already exists in voters' preferences.

8.2 Competition and the provision of public goods

Our main result also relates to the traditional literature on public goods provision. In the context of our model, however, the distinction between what is public or private is more abstract and depends on the *content* of the information provided by the news sources. One classic result in the public good literature is to show how competitive markets favor private goods over public ones, leading to the under-provision of the latter. In our context, a similar force is at work: the heterogeneity in voters' preferences gives information providers, who face a shrinking market share, incentives to design informational products that are increasingly more *private*, as they cater more exclusively to a smaller portion of the electorate (Proposition 2). As discussed in Section 8.1, the subdivision of the preference space into a public and a private dimension is more general than the specific model introduced in this paper. Therefore, we would expect similar mechanisms, characterized by differentiation of information providers through the oversupply of private information, to have

significant consequences in other contexts as well where information is disseminated to a population of agents with heterogeneous preferences.

Yet, our model also highlights that the market for political news is unique in that information is consequential to how people vote. Since agents do not take this externality into account, changes in the types of information provided to the voters, purely due to competitive forces, can have significant negative effects on aggregate voting patterns.

8.3 Introducing a government-funded news source

Our main results highlight how competition in the media market can deepen disagreement in a society by shifting focus from valence issues to ideological issues. Our model demonstrates that this can be a natural consequence of the contrast between the type of information that is relevant for voters on an individual level vs. the type of information that is relevant for determining the socially optimal candidate. Profit maximizing firms (news sources in our model) shift their informational products to cater more to individual demand as the market gets more and more segmented.

It is interesting to consider what role government-funded news sources that are not affected by competitive forces can play in such an environment. After all, despite the dramatic increase in the number of news sources available to voters, government-funded news sources remain in most countries.[24] Here, we discuss a simple extension of our model with the addition of a news source that reports *only* on valence. It is clear that the presence of such a news source can partially counteract the focus on ideological issues driven by other news sources in the competitive market. Naturally, the magnitude of this effect will depend on further assumptions we make on how and when voters consume this public news source. For example, if the public news source simply provides an alternative to the privately operated news sources in the market, we would need to investigate the conditions under which (at least a share of the) voters choose this news source over the alternatives. In this context, our model would predict that the attractiveness of the public information provider decreases as competition increases, since voters are able to find private providers supplying information that better match their individual preferences.

---

[24]BBC, PBS/NPR, Deutsche Welle, CBC, VOA are some prominent examples.

On the other hand, it could also be the case that a publicly funded news source broadcasts information in ways that are more easily accessible to the voters. In this case, we would expect the publicly funded news source to be consumed by all voters. Note that this would effectively be equivalent to changing voters' prior on valence before the game described in our model takes place. Privately operated news sources in equilibrium would adjust their reporting strategy to take into account information on valence already provided by the public source. In response, we would expect news sources in the competitive market to shift focus even further to ideological issues. Nonetheless, for any level of competition, the consumption of the public news source should increase how informed voters are on valence issues relative to ideology. In conclusion, although the presence of a public news source can partially alleviate the emphasis on ideology driven by a competitive market as suggested above, it is important to point out that our main comparative result on the effects of competition will still go through.

While we do not solve these extensions formally, there is clear intuition on how a public news source, by providing information on valence, can play an important role counteracting market forces that emphasize ideological issues.[25] Furthermore, our results also suggest that the role played by public news sources can change with the level of competition in the market. Public news sources have historically been founded on principles that emphasize "universal geographic accessibility," "attention to minorities," "contribution to national identity and sense of community," and "distance from vested interests."[26] As acquisition of political news shifts online and the number of news sources simultaneously available to voters dramatically increases, there is arguably less concern on some of the issues addressed above. Nonetheless, our model demonstrates that, as the level of competition in the market increases, public news sources can have a critical role to play in shaping public opinion by refocusing public discourse on issues which are of relevance to the population as a whole and on which there is general agreement.

---

[25]Government funded news sources can also be manipulated and censored more easily. In this discussion, we assumed the public source to be unbiased. We refer the reader to Besley and Prat (2006) for a study of competition and media capture.

[26]These highly referenced principles were first stated by the Broadcasting Research Unit in Britain in 1985.

8.4 Strategic Voting

While solving the voter's problem, we 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 levels. 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. Such effects crucially hinge on the mechanism that maps vote 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. This allows each voter to have an effect on the final outcome regardless of the voting behavior of others. It is easy to see that, in this case, strategic voting moves closer to sincere 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. 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.

9. 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 a whole.

References

ALESINA, A., R. BAQIR, AND W. EASTERLY (1999): "Public goods and ethnic divisions," *The Quarterly Journal of Economics*, 111, 1243–1284.

ARAGONES, E., M. CASTANHEIRA, AND M. GIANI (2015): "Electoral Competition through Issue Selection," *American Journal of Political Science*, 59, 71–90.

ASHWORTH, S. AND E. DE MESQUITA (2009): "Elections with Platform and Valence Competition," *Games and Economic Behavior*, 67, 191–216.

BAKER, E. C. (1977): "Scope of the First Amendment Freedom of Speech," *UCLA Law Review*, 25, 964.

BANDYOPADHYAY, S., K. CHATTERJEE, AND J. ROY (2015): "Manufacturing extremism: political consequences of profit-seeking media," Working Paper.

BARON, D. (1994): "Electoral competition with informed and uninformed voters," *American Political Science Review*, 88, 33–47.

——— (2006): "Persistent 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*, 92, 1092–1104.

BESLEY, T. AND A. PRAT (2006): "Handcuffs for the grabbing hand? Media capture and government accountability," *American Economic Review.*, 96, 720–736.

CAGÉ, 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*, 118, 845–874.

CHAN, J. AND W. SUEN (2008): "A spatial theory of news consumption and electoral competiton," *Review of Economic Studies*, 75, 699–728.

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*, 78, 640–666.

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*, 92, 925–943.

GENTZKOW, M. AND E. KAMENICA (2016): "Competition in Persuasion," *Review of Economic Studies, forthcoming.*

GENTZKOW, M. AND J. SHAPIRO (2008): "Competition and truth in the market for news," *Journal of Economic Perspective*, 22, 133–154.

GENZKOW, M., J. M. SHAPIRO, AND M. SINKINSON (2014): "Competition and Ideological Diversity: Historical Evidence from US Newspapers," *American Economic Review*, 104, 3073–3114.

GROSECLOSE, T. (2001): "A Model of Candidate Location When One Candidate Has a Valence Advantage," *American Journal of Political Science*, 45, 862–886.

GROSSMAN, G. AND E. HELPMAN (1996): "Electoral competition and special interest politics," *Review of Economic Studies*, 63, 265–286.

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, 225–239.

——— (2005): "A drawback of electoral competition," *Journal of the European Economic Association*, 3, 1318–1348.

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*, 95.

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*, 57, 368–377.

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

——— (2002): "The Law of Group Polarization," *The Journal of Political Philosophy*, 10, 175–195.

YUKSEL, S. (2015): "Specialized Learning and Political Polarization," *Working Paper*.

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 $\theta_v > 0$. $\qquad\square$

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

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

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

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

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

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

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

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

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

Proof of Proposition 2. We begin by computing the value of an information structure characterized by (τ,x) for type t. We will show that

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

where $\sigma^2(\tau,x|t) = \lambda^2 g(\tau) + (1-\lambda)^2\cos^2(t-x)g(\bar\tau-\tau)$ and $g(\tau) = \frac{\tau}{1+\tau}$. Recall that if $X \sim \mathcal{N}(0,\sigma^2)$, then $\mathbb{E}(|X|) = \sigma\sqrt{2/\pi}$. In our case, since – unconditionally – $s_v \sim \mathcal{N}(0,1+\tau^{-1})$ and $s_{id} \sim \mathcal{N}(0,1+(\bar\tau-\tau)^{-1})$ we have

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

39

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

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

Therefore, this shows how $V(\tau, x|t)$ is a monotonic transformation of σ^2 and how the latter can be separated into two main components: the public part, $\lambda^2 g(\tau)$, increasing in τ, and the private one, $(1-\lambda)^2 \cos^2(t-x)g(\bar{\tau} - \tau)$ decreasing in τ and in $|t - x|$. $\qquad\square$

Proof of Proposition 3. 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
$$\frac{\partial}{\partial \tau_i}\left(V(\tau_i, x_i|t_r) - V(\tau^\star, x_{i+1}^\star|t_r)\right) = 0$$

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

$$\frac{\partial t_r}{\partial \tau_i} = \frac{\lambda^2 g'(\tau_i) - (1-\lambda)^2 \cos^2(t_r - x_i)g'(\bar{\tau} - \tau_i)}{(1-\lambda)^2\left(g(\bar{\tau} - \tau_i)\sin 2(t_r - x_i) + g(\bar{\tau} - \tau^\star)\sin 2(x_{i+1}^\star - t_r)\right)}$$

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

$$\frac{\partial t_l}{\partial \tau_i} = -\frac{\lambda^2 g'(\tau_i) - (1-\lambda)^2 \cos^2(x_i - t_l)g'(\bar{\tau} - \tau_i)}{(1-\lambda)^2\left(g(\bar{\tau} - \tau_i)\sin 2(x_i - t_l) + g(\bar{\tau} - \tau^\star)\sin 2(t_l - x_{i-1}^\star)\right)}$$

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 $\tau^\star$, $t_l - x_{i-1}^\star = x_{i+1}^\star - t_r$ and $t_r - x_i = x_i - t_l$. Thus we get

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

or

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

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

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

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

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 $x_i^\star = 0$.*

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

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

which gives us the following expression

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

where

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

In a similar fashion we can get

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

where

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

41

When $\tau_i = \tau^\star$ for all $i \in N$, it is 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 $f\pi/n$. $\qquad\square$

This concludes the Proof of Proposition 3. $\qquad\square$

Proof of Proposition 4: In the Proof of Lemma A2 we implicitly derived a

solution for $\tau^\star$ in terms of the parameter n:

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

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 $\frac{g'(\tau^\star)}{g'(\bar\tau - \tau^\star)}$ is strictly decreasing in τ. Thus an increase in n can be compensated only by a decrease in $\tau^\star$. $\qquad\square$

Proof of Lemma 1. Conditional on $\boldsymbol\theta$, the expected utility of type t consuming infrormation structure (τ, x) is distributed $\mathcal{N}(\mu, \nu^2)$, with

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

and

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

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

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

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, $a = \frac{(1-\lambda)g(\bar\tau - \tau)\cos(t-x)}{\nu}f(\theta_{id}, t)$, $b = \frac{\lambda g(\tau)}{\nu}$ 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 4, 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

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

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

$$-\frac{2}{1+\bar\tau - \tau}\left(\lambda^2\frac{\tau}{1+\tau} + (1-\lambda)^2\cos^2(t-x)\frac{\bar\tau - \tau}{(1+\bar\tau - \tau)^2}\right) - (\bar\tau - \tau)\left(\lambda^2\frac{1}{(1+\tau)^2} - (1-\lambda)^2\frac{\cos^2(t-x)(1-\bar\tau}{(1+\bar\tau - \tau)^3}\right)$$

44

This is equivalent to

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

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 $f(\theta_{id},t)$. $\qquad\square$.

Proof of Lemma 2. 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

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

and get the following expression:

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

where

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

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. $\qquad\square$

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 $\alpha=\lambda\frac{\tau}{1+\tau}$, $\alpha'=\lambda\frac{1}{(1+\tau)^2}$, $\beta=\lambda^2\frac{\tau}{(1+\tau)^2}+(1-\lambda)^2\cos^2(t-x)\frac{\bar\tau-\tau}{1+\bar\tau-\tau}$ and $\beta'=\lambda^2\frac{1-\tau}{(1+\tau)^3}-(1-\lambda)^2\cos^2(t-x)\frac{1}{(1+\bar\tau-\tau)^2}$. Thus, $2\alpha'\beta-\alpha\beta'$ is equal to

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

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

The first term is positive since $2\tau - \tau(1-\tau) > 0$. The second term is also positive. It is straightforward to see that $A(\tau, x|t)|$ is increasing in $|t-q|$. Thanks to these two facts, it is straightforward to see $A(\tau, x|t) \in [0,1]$. It is also trivial to see that $A(\tau, x|t)$ is decreasing in $|t-x|$. $\qquad\square$

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

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

$\qquad\square$

Proof of Proposition 5. 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 A6 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 4 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.

$$\int_{\mathbb{R}}\sum_{i\in N}\int_{x_i^\star-\frac{\pi}{2n}}^{x_i^\star+\frac{\pi}{2n}}\Pr\Big(\mathbb{E}_{\tau^\star,x_i^\star}(u(\theta_v,\theta_{id},t)|s_v,s_{id})>0\Big)dH(t)d\Phi(\theta_{id})$$

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

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

$$=\sum_{i\in N}\int_{x_i^\star-\frac{\pi}{2n}}^{x_i^\star+\frac{\pi}{2n}}\Phi\Big(A(\tau^\star,x_i^\star|t)\theta_v\Big)dH(t)$$

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

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

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

since $\Phi(A(\tau^\star,0|\frac{\pi}{2n})\theta_v)=\Phi(A(\tau^\star,0|-\frac{\pi}{2n})\theta_v)$. It is 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. $\qquad\square$

Proof of Theorem 2. This result follows from the proof of Theorem 1 and from Lemma A7. $\qquad\square$

Proof of Theorem 3. We'll make use of the following lemmas.

Lemma A8. *For any two news sources on the same side of t, if the agent is consuming the farthest one, then he must be consuming the one closer as well.*

Proof. Fix the learning strategy used by an agent of type t. Since we are focusing on symmetric equilibria where all news sources provide the same precision of signals on valence vs. ideology, we can focus on learning from the signals associated with ideology. Any learning strategy consists of two parts. Set of κ chosen news sources and a vector $\mathbf{v} = (v_1, v_2, ... v_\kappa)$ which specifies how the signals from news sources are used in calculating the expected $f(\theta_{id}, t)$. Namely, $\mathbb{E}(f(\theta_{id}, t) | \mathbf{s}) = \sum v_i s_i$ where linearity follows from the fact that the signals are normally distributed. The agent will choose the learning strategy that minimizes $\mathbb{E}(\mathbb{E}(f(\theta_{id}, t) | \mathbf{s}) - f(\theta_{id}, t))^2$. This can be rewritten as $\delta \cdot \delta + \sum v_i^2 \sigma^2$, where $\delta = (\delta_1, \delta_2)$, $\delta_1 = \cos(t) - \sum v_i \cos(t_i)$ and $\delta_2 = \sin(t) - \sum v_i \sin(t_i)$ (t_i signifies the type targeted by news source i) and σ^2 is the variance of the noise associated with the ideology signal from any news source. To facilitate the exposition of the proof, without any loss of generality, we fix $t = \pi/2$.

Claim 1. $\frac{\delta_2}{1 - \delta_2} > \frac{\sigma^2}{\kappa}$

Proof. Look at the best case where there are κ news sources that are all perfectly targeting $t = \pi/2$ (which cannot happen for finite n). It is easy to see that the optimal strategy would be to choose $\mathbf{v} = (v, v, ..v)$ to minimize $(1 - \kappa v)^2 - \kappa v^2 \sigma^2$. Solving this problem gives us $v = \frac{1}{\kappa + \sigma^2}$ which implies $\delta_2 = 1 - \kappa v = \frac{\sigma^2}{\kappa + \sigma^2}$. Hence at this bound $\frac{\delta_2}{1 - \delta_2} = \frac{\sigma^2}{\kappa}$. $\square$

Claim 2. $|\delta_1| < \tan(\alpha/2)(1 - \delta_2)$ *where* $\alpha = \frac{\pi}{\kappa}$

Proof. Let θ be the angle between vectors $(\cos(t), \sin(t))$ and $(\sum v_i \cos(t_i), \sum v_i \sin(t_i))$. Since the news sources are symmetrically distributed, this angle cannot exceed $\alpha/2$. If it did, then without changing $\mathbf{v}$, the agent could shift the set of new sources consumed all to the left (or right) by one, and this would decrease θ by α. This can done till the resulting vector is at least $\alpha/2$ either to the left or right of $(\cos(t), \sin(t))$. At the maximal angle, using $t = \frac{\pi}{2}$, $|\delta_1| = \tan(\alpha/2)(1 - \delta_2)$. $\square$

Now we come back to the proof of Lemma A8. Assume for contradiction that the agent is not consuming the closest news source. We show that the agent will be better off using the same $\mathbf{v}$ but replacing the farthest news source with the closest one. Call the old news source o, and the new one n and v be the component of $\mathbf{v}$ that is associated with the old news source. Let $\mathbf{d} = (\cos(t_n) - \cos(t_o), \sin(t_n) - \sin(t_o))$.

Since $\mathbf{v}$ remains unchanged, it will only change the first part of $\delta \cdot \delta + \sum v_i^2 \sigma^2$. Specifically, the change will be $\Delta = (\delta - v\mathbf{d}) \cdot (\delta - v\mathbf{d}) - \delta \cdot \delta$ which we will show to be negative. Assume $\delta_1 > 0$ and $t_n = \frac{\pi}{2}$ and $t_o = \frac{\pi}{2} + \alpha$. It can be shown that all other cases will follow a similar structure to what's below and will hold true if this case holds true. Note that for such t_n and t_0, $\mathbf{d} = (\sin(\alpha), 1 - \cos(\alpha))$.

$$
\begin{aligned}
\Delta &= v^2(\mathbf{d} \cdot \mathbf{d}) - 2v(\delta \cdot \mathbf{d}) \\
&= v^2[d_1^2 + d_2^2] - 2v[d_1\delta_1 + d_2\delta_2] \\
&< v^2[d_1^2 + d_2^2] + 2v[d_1\tan(\alpha/2)(1 - \delta_2) - d_2\delta_2] \\
&< v^2[\sin(\alpha)^2 + (1 - \cos(\alpha))^2] + 2v[\sin(\alpha)\tan(\alpha/2)(1 - \delta_2) - (1 - \cos(\alpha))\delta_2] \\
&= 2v^2(1 - \cos(\alpha)) + 2v[\sin(\alpha)\tan(\alpha/2)(1 - \delta_2) - (1 - \cos(\alpha))\delta_2] \\
&= 2\Big(v\sin(\alpha)\tan(\alpha/2)(1 - \delta_2) - (\delta_2 v - v^2)(1 - \cos(\alpha))\Big) \\
&= 2\Big(2v\sin(\alpha/2)\cos(\alpha/2)\tfrac{\sin(\alpha/2)}{\cos(\alpha/2)}(1 - \delta_2) - (\delta_2 v - v^2)(2\sin(\alpha/2)^2)\Big) \\
&= 4(1 - \delta_2)v(\sin(\alpha/2))^2[1 - \tfrac{\delta_2}{1-\delta_2} + \tfrac{v}{1-\delta_2}]
\end{aligned}
$$
(1)

For the third line we used Claim 2. In the fourth line, we focused on the extreme $\mathbf{d}$ stated above. Δ is negative whenever $1 - \frac{\delta_2}{1-\delta_2} + \frac{v}{1-\delta_2} < 0$. Note that $\frac{v}{1-\delta_2}$ can at most be one. Thus, $\frac{\delta_2}{1-\delta_2} > 2$ is sufficient to guarantee the result. By Claim 1, $\frac{\delta_2}{1-\delta_2} > \frac{\sigma^2}{\kappa}$ which is assumed to be larger than 2 by the theorem. $\qquad \square$

Lemma A9. *In the optimal learning strategy, an agent consumes the closest news sources.*

Proof. In Lemma A8, we already showed that there cannot be a gap in the news sources consumed to the right and to the left. Now assume for contradiction that the set of new sources chosen is not actually the set closest to the agent. Let $t_m = \frac{1}{\kappa}\sum = t_i$ be the mean type of the chosen set of new sources. Without loss of generality, assume that t_m is to the right of t. By assumption $t - t_m = \frac{\pi}{2n} + \theta$ for some $\theta > 0$. This also suggests that t is closer to the mid point of an alternative set of new sources that have been shifted to the left. First we show that the original set of news sources provides a more effective learning strategy for all types between t and t_m. We can always take the most right and left news sources, and we can replace $\tilde{v}_1 = (1-\rho)v_1 + \rho\frac{v_1+v_\kappa}{2}$ and $\tilde{v}_\kappa = (1-\rho)v_\kappa + \rho\frac{v_1+v_\kappa}{2}$. We can do this iteratively for all other news sources as well. As $\rho \to 1$, the estimated type shifts towards t_m

and the variance goes down. Using symmetry, we've demonstrated that this set of news sources to be better for $t_m - \frac{\pi}{2n} + \theta$. But this implies that the news sources can be shifted to the left to the get a better learning strategy. $\square$

Lemma A9 implies that when firms locate equidistantly, and choose the same τ, with n news sources, each firm serves $\frac{\kappa}{n}$ of the market. Each firm is competing over threshold types with neighbors that are κ to the right an left. Hence, existence of symmetric equilibria and the comparative results can be shown following the same strategy for the $\kappa = 1$ case taking the adjustment with respect to the threshold types into account. $\square$