DISCUSSION PAPER SERIES

DP12640

PERSUASION WITH CORRELATION NEGLECT: MEDIA POWER VIA CORRELATION OF NEWS CONTENT

Gilat Levy, Inés Moreno de Barreda and Ronny Razin

PUBLIC ECONOMICS

ISSN 0265-8003

PERSUASION WITH CORRELATION NEGLECT: MEDIA POWER VIA CORRELATION OF NEWS CONTENT

Gilat Levy, Inés Moreno de Barreda and Ronny Razin

Discussion Paper DP12640
Published 23 January 2018
Submitted 23 January 2018

PERSUASION WITH CORRELATION NEGLECT: MEDIA POWER VIA CORRELATION OF NEWS CONTENT

Abstract

We model the power of media owners to bias readers' opinions. In particular we consider readers that have "correlation neglect", i.e., fail to understand that content across news outlets might be correlated. We study how a media owner who controls several outlets can take advantage of the readers' neglect. Specifically, we show that the owner can manipulate readers' beliefs even when readers understand the informativeness of news outlet by outlet. The optimal strategy of the owner is to negatively correlate good news and positively correlate bad news. The owner's power is increasing in the number of outlets she owns but is constrained by the limited attention of readers. Importantly, our analysis suggests several new insights about welfare in media markets. First, measures of media bias have to take into account the correlation between news outlets. Second, media-market competition curbs the ability of owners to bias readers' beliefs. In particular, we show that readers always benefit from breaking media conglomerates, even when all the new media owners share the same bias. Finally, we highlight a potential cost of media diversity. When readers have correlation neglect, diversity in the interests of owners might lower the informativeness of news content.

JEL Classification: N/A

Keywords: N/A

Gilat Levy - g.levy1@lse.ac.uk
lse and CEPR

Inés Moreno de Barreda - ines.morenodebarreda@economics.ox.ac.uk
Oxford University

Ronny Razin - r.razin@lse.ac.uk
LSE and CEPR

Persuasion with Correlation Neglect: Media Power via Correlation of News Content

Gilat Levy, Ines Moreno de Barreda and Ronny Razin[1]

Abstract: We model the power of media owners to bias readers' opinions. In particular we consider readers that have "correlation neglect", i.e., fail to understand that content across news outlets might be correlated. We study how a media owner who controls several outlets can take advantage of the readers' neglect. Specifically, we show that the owner can manipulate readers' beliefs even when readers understand the informativeness of news outlet by outlet. The optimal strategy of the owner is to negatively correlate good news and positively correlate bad news. The owner's power is increasing in the number of outlets she owns but is constrained by the limited attention of readers. Importantly, our analysis suggests several new insights about welfare in media markets. First, measures of media bias have to take into account the correlation between news outlets. Second, media-market competition curbs the ability of owners to bias readers' beliefs. In particular, we show that readers always benefit from breaking media conglomerates, even when all the new media owners share the same bias. Finally, we highlight a potential cost of media diversity. When readers have correlation neglect, diversity in the interests of owners might lower the informativeness of news content.

[1]Levy: London school of Economics; Moreno de Barreda: University of Oxford; Razin: London school of Economics.
We thank Clement Minaudier for excellent research assistance. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No SEC-C413.

1 Introduction

A recent literature in economics studies echo chambers and their influence on the beliefs and attitudes in society.[2] One reason for such effects is that when confronted with several pieces of information, people do not fully understand that these pieces are repeated, or more generally correlated, which induces amplification of information and sometimes polarization. This arises with correlation neglect, a bias by which people treat pieces of information as if they were independent.[3] Ortoleva and Snowberg (2015) document how correlation neglect shapes political views. Eyster and Weizsaker (2011), Kallir and Sonsino (2009) and Enke and Zimmermann (2013) provide experimental evidence for correlation neglect.[4]

There is a good reason to think that consumers of news media are likely to suffer from correlation neglect to some extent. For one thing, news items are constantly copied and repackaged across outlets. Cagé et al. (2017) study copyright in news media, following pieces of news as they trickle through different outlets including social media. They document how pieces of news are often copied multiple times and across different outlets. In addition they find that only 32% of online content is original. Still, despite the prevalence of copying, media outlets hardly name the sources they copy.[5] Thus readers are exposed to repeated news, potentially without being aware of it.[6]

It is also clear that people read multiple sources of information. Individual-level survey data on 18 countries from Reuters Institute for the Study of Journalism shows that the average news consumer uses about five news sources per week.[7] Communication among individuals also implies that, indirectly, they are exposed to even more sources. Gentzkow et al. (2014) show that individuals' communication networks are segregated across work colleagues, friends, family and neighborhood associations, according to socioeconomic parameters and political preferences. This means that individuals communicate with those who read similar media outlets.[8] Such echo chambers imply that they will be exposed to repeated and correlated sources.

In this paper we analyze how media owners can design the news coverage across their outlets to influence readers by taking advantage of correlation neglect. Specifically, we analyze how strategic media owners, who own multiple news outlets, can use correlation of content strategically to manipulate readers who believe that news outlets are independent. The owner could be a media mogul that controls several media outlets, possibly across different platforms. Alternatively, this could be the owner or editor of a news aggregation website that decides which links and news to reproduce.[9]

Indeed, the intervention of owners in the editorial decisions of their news outlets has always been an important issue in the debate about the regulation of the media industry.[10] It is one of the reasons behind

[2]The debate is focused on explaining the increase in polarization in American politics. Glaeser and Sunstein (2009) and Sunstein (2017) argue that communication in segregated groups implies echo chamber effects and that increased use of technology, such as social media, are behind the increase in polarization of opinions. Gentzkow and Shapiro (2011) and Boxell et al. (2017) argue that social media is not the engine behind the increase in polarization but rather it is communication at work, with family and friends that are the main factors behind echo chamber effects.

[3]De Marzo et al. (2003), Golub and Jackson (2012) and Gagnon-Bartsch and Rabin (2015) study how correlation neglect affects the diffusion of information in social networks. Glaeser and Sunstein (2009) and Levy and Razin (2015a;b) explore the implications for group decision making in political applications.

[4]On the theory side, Ellis and Piccione (2017) provide an axiomatic characterization of individuals that cannot account for correlation (or complexity in their terminology). Levy and Razin (2017) discuss environments for which correlation neglect is more likely.

[5]They also show that reputation mechanisms allow to mitigate these issues, so that a 50% higher originality of content allows to have a 35% increase in sharing on facebook. Reputation motives appear to solve 40% of the copy right problems.

[6]News aggregation websites are another example of how media is copied and the sources of information are made harder to trace. These sites publish their own news as well as links to similar news in other sites, which therefore induce individuals to be exposed to repetition of news.

[7]See Kennedy and Prat (2017) for an analysis of these data. See also the "In Changing News Landscape, Even Television is Vulnerable: Trends in News Consumption: 1991-2012" a report by Pew Research, http://www.people-press.org/2012/09/27/in-changing-news-landscape-even-television-is-vulnerable/

[8]About two-thirds (63%) of Americans say family and friends are an important way they get news, whether online or offline; 10% see them as the most important. See http://www.journalism.org/2016/07/07/pathways-to-news/.

[9]More generally, one could think about correlating content across different pieces of news within the same outlet, such as different programs in a TV channel, although readers might be more aware of correlation in this case.

[10]Some insight into the nature of such interventions can be gleaned from the evidence given to the Leveson Inquiry in

a common call to have independent editorial boards. For example, in the UK, in June 2017, the culture secretary decided to refer 21st Century Fox's £11.7bn bid to seize full control of satellite broadcaster Sky to the Competition and Markets Authority, for a fuller, "phase two" investigation. The FT reports that behind this decision was the fact that "While Fox and News Corp are separate companies, the Murdoch Family Trust has material influence across both companies." To secure the deal 21st Century Fox has to take some measures that "[...] include setting up a separate editorial board with a majority of independent members to oversee Sky News and a commitment to maintain Sky-branded news for five years at current funding levels."[11]

We consider a basic framework with a binary state of the world. A news outlet provides a signal on the state of the world. A media owner, after observing the state, decides on the joint information structure over signals for the news outlets. The readers correctly interpret each outlet in isolation, i.e., they understand the accuracy of every individual outlet, but believe that the joint distribution over the content of all news outlets satisfies (conditional) independence. We analyze the cases of a monopolist media owner and of competition between several owners.

Our first result is that a media monopolist is able to change the expected posterior of the readers away from the prior by correlating the content across news outlets. This implies that even in cases where readers can detect the bias of each outlet, some influence still takes place through the correlation of news across outlets. This has important empirical implications. Media bias or slant is often measured outlet by outlet and therefore important effects such as echo chamber effects and the contribution of media coverage to polarization are under-estimated. Indeed, Bai et al (2015) show in an experiment that individuals understand the bias of government sources when presented in isolation but do not discount repeated information, even when they know the information comes from one biased government source. In this sense, the power that comes from owner concentration is more pronounced than the media bias outlet by outlet.

We show that the optimal correlation of news content for a biased owner depends on whether the true state of the world favors the owners's objective or not. For example, when the owner has linear utility over the posterior of the reader, she will in general negatively correlate news content when the state is in her favor and positively correlate unfavorable news when the state goes against her objective. The intuition for this result stems from both correlation neglect and Bayesian persuasion. Correlation neglect implies that the reader's posterior is monotone in the number of pieces of news pointing to a particular direction. The posterior, represented as a function of the number of outlets with favorable information, takes the familiar S-shape of the value of information function. The media owner prefers to strongly correlate bad news some of the times, in order to be able to provide mildly good news most of the times. We show that the optimal solution for the owner follows a similar concavication intuition as in Kamenica and Gentzkow (2011), so that the positive and negative correlation results above follow from the concavication of the different parts of the S-shaped value function.

We explore the implications of the above results to the informativeness of the media market. While different aspects of the media market might influence what the accuracy of each outlet might be, we analyze the optimal choice of accuracies by the media owner in the context of their ability to persuade. We show that media owners have an incentive to choose news outlets that are strictly informative but not too much. If news outlets are not informative, the owner cannot use correlation to create bias, as readers ignore what is written in each paper. If news outlets are too informative, the owner looses the flexibility to manipulate the posteriors of readers.

The power of the media owners to manipulate the receiver using correlation implies that owners have a strictly increasing incentive to add more news outlets to their arsenal. In addition, when owners can choose both the number of outlets, as well as the accuracy of each newspaper, they can fully manipulate

the UK. From http://www.independent.co.uk/news/uk/crime/the-world-according-to-rupert-murdoch-7679254.html: "Andrew Neil, who edited The Sunday Times between 1983 and 1994, recalled in Full Disclosure that although the proprietor did not expect to see his views repeated immediately in the next paper *"he had a quiet, remorseless, sometimes threatening way of laying down the parameters within which you were expected to operate... stray too far too often from his general outlook and you will be looking for a new job."* The former Times and Sunday Times editor Harold Evans said that Murdoch broke all of his promises of editorial independence after taking over titles."

[11] See https://www.ft.com/content/6dd32342-5ccd-11e7-b553-e2df1b0c3220?mhq5j=e3

the reader to choose their optimal action. Of course this result depends on the ability, as well as on the choice, of the reader to read many news outlets. More realistically, readers might have limited attention. We show that with limited attention the incentive to add more news outlets is curbed. In this case, introducing more news outlets creates more noisy reactions by readers and media owners would prefer to increase their control over reader responses by tailoring the number of outlets to the limited attention of the readers.

Competition in the media market has been on the agenda of many state regulators. We show that there are two ways in which competition in the media market affects readers' welfare. First, competition curtails the ability of media owners to use correlation. Breaking up media conglomerates into small entities, even if the owners have similar interests, limits the scope for correlating content. In the extreme, when each media outlet is owned by a different owner, information is aggregated correctly by the readers. As a result, we find that readers always benefit from breaking up media conglomerates. This result also supports the common call for independence of editorial boards.

The second effect of competition on readers' welfare is through diversity of ownership. Several recent reports have suggested the importance of diversity of opinion in the media market.[12] Our model highlights a potential cost of diversity. Because readers in our model do not process information correctly, competition between two opposing interests is not necessarily better than competition between aligned interests in the market. The desire of each interest to pull the reader in their direction, coupled with correlation neglect, induces a tug of war that leads to lower information aggregation. Thus, compared to the case of similar interests, a duopoly with opposing interests induces less bias in the readers' posterior but also a less information. As a result, readers sometimes prefer not to have diversity of opinion in ownership.[13]

2 Related Literature

There are several papers that have attempted to measure media bias.[14] Groseclose and Milyo (2005) use partisan references to document media bias in the US. Durante and Knight (2012) use airtime to show media bias in Italy, Puglisi (2011) finds bias in the NYtimes coverage using space devoted to partisan issues and Gentzkow and Shapiro (2010) use textual analysis to rate the bias of American newspapers. Prat (2017) surveys the different ways that media power has been modelled and measured and suggests a new measure of power which is based on the potential for owners to slant coverage. What all these papers have in common is that they measure or analyze the media bias of each news outlet separately. However, an implication of the type of media bias we analyze here is that bias might not necessarily be observable unless one looks at several co-owned news outlets.

The literature has provided some mixed results about how bias in coverage actually affects the behavior of readers. Papers such as Della Vigna and Kaplan (2007) and Enikolopov et al. (2011) find significant effects of the entry of biased news outlets on voting behavior. On the other hand, Gentzkow et al. (2011) find no evidence of effects of entry and exit of partisan newspapers on party vote shares in the United States from 1869 to 2004. Evidence on the effects of ownership on bias is mixed too. Durante and Knight (2012) document significant changes in state television coverage in Italy when Silvio Berlusconi came to power. However, Gentzkow and Shapiro (2010) find that owner identity has no significant effect on newspaper slant in the US. On the other hand, Enikolopov et al. (2011) show how media is persuasive by comparing electoral outcomes of the 1999 parliamentary election in Russia across

[12] For a recent such report see the Leveson inquiry into the culture, practice and ethics of the press in the UK, http://webarchive.nationalarchives.gov.uk/20140122145147/http:/www.levesoninquiry.org.uk/.

[13] This result echoes some of the claims about the public debate before the Brexit referendum in the UK. There was a perception on both sides that the debate was not informative as one would expect from such an important public debate. The lack of informativeness of a balanced debate might also behind some of the results in Enikolopov et al. (2011), who show that people exposed to two-sided coverage in Russia were less likely to turn out as compared with those only exposed to pro-government propaganda in 1999.

[14] Among theoretical papers explaining such bias are Baron (2006), Besley and Prat (2006), Balan et al. (2009), Duggan and Martinelli (2011), Anderson and McLaren (2012) and Petrova (2012).

geographical areas that had more access to independent TV; they show that such exposure decreased aggregate vote for the incumbent party by 8.9%, increased overall voting for the opposition parties, and decreased turnout. These findings are consistent with our model of competition, where voters may be less biased but potentially less informed and hence less inclined to vote.

Our paper relates to Anderson and McLaren (2012) and Mullainathan and Shleifer (2005) who analyze competition in the media market. Anderson and McLaren (2012) study the comparison between a media duopoly to a media monopoly, in the presence of biased media owners, but rational readers.[15] Due to the rationality of readers, in their model, the duopoly market structure always aggregates all information. In contrast, we assume that readers have correlation neglect which implies that the duopoly market structure might induce additional noise as compared to the case of a monopoly. Mullainathan and Shleifer (2005) study the effect of competition on bias and information aggregation when readers have confirmation bias.[16] They show that whether competition aggregates information depends on whether readers have divergent beliefs. When readers share the same beliefs, competition does not aggregate information efficiently. In particular they conclude that reader heterogeneity is more important for accuracy in media coverage than competition. In Section 6.1 we introduce heterogeneous preferences among receivers and observe that although a small divergence of preferences improves the accuracy in media coverage, when the readers become too polarized, the media tailors a particular sector of the population and information decreases over all.[17]

Our optimal information structure shares some features with the optimal solution in Harbaugh et al. (2017) albeit through very different mechanisms. Harbaugh et al. (2017) analyze news' distortions when the receiver is rational but uncertain over the accuracy of the news and news are drawn independently. They show that when the news is mostly good, shoring up relatively bad news is most persuasive as it makes the good news appear more consistent and hence credible. But when the news is bad, exaggerating good news is better as bad news then appear less consistent.

Our paper also contributes to the literature on Bayesian persuasion. Our key contribution to this literature is to analyze persuasion when receivers have correlation neglect. Kamenica and Gentzkow (2011) identify the optimal signal that a sender can design to persuade a rational receiver. As we show, correlation neglect implies a new constraint on how the receiver updates their beliefs. In contrast with the rational receiver model, we show that an implication of this is that persuasion takes place even when the sender has a linear utility over the posteriors of the receiver. Meyer (2017) analyzes how the heterogeneity of receivers affects the optimal persuasion of a sender (see also Arieli and Babichenko (2016)). She shows that the sender benefits from receivers being homogenous when the signals are private. This is related to our model where we treat all information sources symmetrically as a simplifying assumption and we conjecture that this would hold in our model as well in which information is publicly provided.[18]

Bayesian persuasion has also been studied in the context of multiple competing senders. Gentzkow and Kamenica (2017) show that competition always increases the information of the receiver, as do Boleslavsky and Cotton (2014). As we discussed above, correlation neglect implies that competition sometimes introduces noise leading to less information.

3 The Model

In the basic model there is one media owner (sender) who observes the state of the world. The sender owns n (even, $n > 2$) news outlets. Each news outlet provides a signal about the state of the world. The sender can control, to some degree, the joint distribution of the signal realizations in all n outlets. The

[15] For a similar methodology see Brocas and Carrillo (2007).

[16] Balan et al. (2009) provided a similar comparison of market structures, albeit with a model in which the behaviour of readers is black boxed.

[17] Galperti and Trevino (2017) show how media outlets trade-off clarity and accuracy when competing for attention of readers, where more attention increases clarity.

[18] There are several papers which study persuasion with multiple audiences, such as Chan et al. (2016), Alonso and Câmara (2016), Wang (2015) and Bardhi and Guo (2016), where some restrict attention to private messages to different sets of audiences. In Section 6 we consider the case in which some readers are rational, and readers with different preferences. In our model the message of the sender is always public and the qualitative results of the paper are robust to these extensions.

reader (receiver) cares about the state of the world and tries to learn about it by reading the news outlets. The reader understands the informational content of each outlet in isolation, but treats all outlets as if they are independent.

We proceed with a simple informational model; we consider a binary state of the world and binary signals, one for each news outlet.[19] The model can be generalized, but this is the simplest environment to consider correlation neglect.

Specifically, consider an environment with two states of the world, $\omega \in \{0, 1\}$. We start with a uniform prior which is common knowledge. The sender observes the state and has at their disposal n binary signals or news outlets indexed by i. We denote by $p_i^\omega = Pr(s_i = 1|\omega)$ the accuracy (or marginal distribution) of signal s_i in state ω.

For simplicity, all the analysis in the main body of the paper focuses on identical accuracy across outlets ($p_i^\omega = p_j^\omega \equiv p^\omega$), symmetric marginals across states ($p^1 = 1 - p^0$), which we will denote simply by p (i.e., $p^1 = p$, $p^0 = 1 - p$) and equal prior. In Appendix B we extend the analysis to non-symmetric marginals (across the two states) and any prior and show that the results are qualitatively similar.

We assume that the receiver knows the marginal p. This assumption is motivated by the readers being familiar with the outlets they read over time, able to trace old news reports as well as previous states of the world and assess the accuracy of each news outlet as an information source. In that sense, the receiver is not fully naïve as she understands the process according to which individual signals are correlated with the state of the world. However, the receiver is also not fully sophisticated; we assume that she fails to recognize the interdependence among news outlets. Understanding correlation is indeed a much more complicated process. Thus, our assumption is that the receiver believes that the signals are (conditionally) independent from one another, and hence has correlation neglect.

Since all individual signals have the same marginals and the receiver treats the signals as independent, the number of signals with realization '1' is a sufficient statistic for the receiver's expectation about the state of the world.[20] We denote by $V(p, k, n)$ the expectation of the receiver when she observes k realizations '1' out of n signals with accuracy p. Given the receiver's correlation neglect we have:

$$V(p, k, n) = \frac{p^k(1-p)^{n-k}}{p^k(1-p)^{n-k} + (1-p)^k p^{n-k}}.$$

When this does not lead to confusion, i.e. when p and n are given and fixed, we will denote this function by V_k. The following is a helpful characterization of V_k,

Lemma 1. *The receiver's posterior, V_k, as a function of k, is symmetric around $k = n/2$, increasing, and concave (convex) above (below) $k = n/2$.*

All the proofs are found in the Appendix.

Intuitively, V_k represents the value of information. When a majority of signals has value 1 (0), then adding another 1 signal has decreasing (increasing) returns.

We assume that the sender cares about the posterior of the receiver but not about the state of the world. In particular, her utility, $U_S(V_k)$, is increasing in V_k and independent of the state.

As in the persuasion literature, the sender commits to an information structure (a joint distribution over signals) at any state ω. Such commitment arises in the media market when a media owner sets up the editorial systems, chooses editorial boards and hires journalists with specific ideologies or abilities.

For reasons outside our model, in many environments the accuracy of the information source may be determined through the competition in the market for news. Too low an accuracy may mean that readers will not consume this source. Too high an accuracy may mean that it is too costly for the owners.

[19]The restriction to binary signals is not qualitatively important in the context of correlation neglect. Even with more complicated structures a media owner would need $n > 2$ outlets to manipulate the reader.

[20]We are assuming full correlation neglect, that is, the receiver does not account for any level of correlation. Levy and Razin (2017) consider cases in which an agent consider bounded levels of correlation. Our results can be extended to such an environment as well; as long as the receiver considers bounded levels of correlation, the sender will be able to manipulate the receiver when designing information structures with full correlation.

We do not model these forces here, instead we consider the accuracy of the signals, p, as interior and exogenously given. In Sections 4.1 and 4.2 we analyze a more general formulation of the problem in which the sender designs a joint information structure, including the choice of parameters such as n and p, which she can commit to, and where the receiver knows n and p.

We restrict attention to anonymous distributions, i.e., distributions that allocate the same probability to any vector of signal realizations with the same sum $k \equiv \sum_{i=1}^{n} s_i$. This restriction is without loss of generality as anonymous distributions are optimal among all, including non-anonymous ones (see footnote 22). Therefore, we can denote by q_k^ω the probability assigned by the sender to some specific k signals having realization 1 and the others having realization 0, in state ω. All these events would have the same probability and thus we can summarize by $Q_k^\omega = \binom{n}{k} q_k^\omega$ the total probability in state ω of having exactly k signals with realization 1.

We can rewrite the accuracy constraint as follows:

$$\begin{aligned} p^\omega &= \Pr(s_i = 1|\omega) = \sum_{k=1}^{n} \Pr(\sum_j s_j = k|s_i = 1, \omega) \\ &= \sum_{k=1}^{n} \binom{n-1}{k-1} q_k^\omega = \sum_{k=1}^{n} \frac{k}{n} \binom{n}{k} q_k^\omega \\ &= \sum_{k=1}^{n} \frac{k}{n} Q_k^\omega \end{aligned}$$

The sender's problem then becomes:
For any ω, and given p^ω and n, the sender chooses $\{Q_k^\omega\}_{k=1}^{n}$ to maximize

$$\begin{aligned} &\textstyle\sum_{k=0}^{n} Q_k^\omega U_S(V_k), \\ &s.t. \\ &\textstyle\sum_{k=0}^{n} Q_k^\omega \frac{k}{n} = p^\omega \\ &\textstyle\sum_{k=0}^{n} Q_k^\omega = 1 \\ &Q_k^\omega \geq 0. \end{aligned}$$

The first constraint represents the marginal or accuracy constraint and the last two constraints correspond to the standard restrictions of probability measures.

4 The optimal monopoly solution

We now analyze the optimal solution for the monopolist when the accuracy of the signals is exogenously given. Assume first that the sender's utility is linear, $U_S(V_k) = V_k$. Figure 1 represents such utility as a function of k, the number of signals with realization 1.

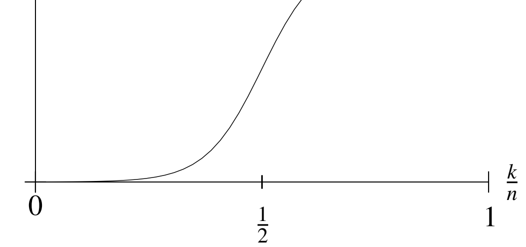
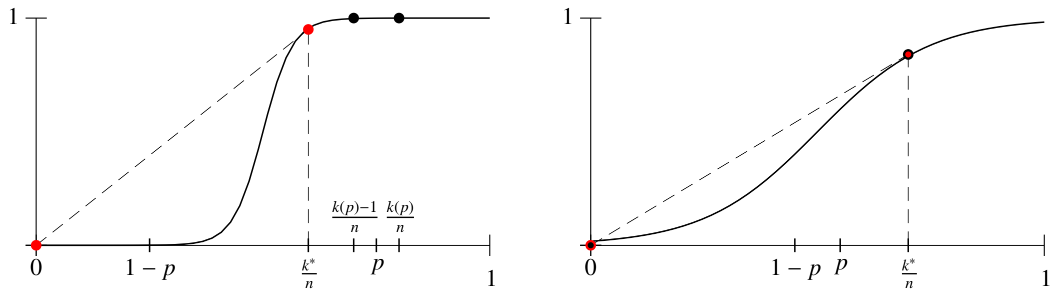


Figure 1: Linear utility

The optimal solution is achieved by finding the *concavication* of the V_k function and then, depending on the marginal, either concentrating all the weight around the marginal, or spreading it around the two closest points to p^{ω} that are at the intersection of the concavication and the function V itself.[21] This is illustrated in Figure 2 below.



(a) Marginal $p = 0.75$ **(b)** Marginal $p = 0.55$

Figure 2: Optimal joint distribution given the marginal constraint. The black (red) dots represent the positive probability mass points when the state is 1 (0).

In order to characterize the optimal solution, define k^* as the highest integer such that $V_k - V_0 \leq k(V_k - V_{k-1})$, and $k(p)$ as the integer such that $\frac{k-1}{n} \leq p < \frac{k}{n}$.

Suppose that the marginal probability is relatively high so that $pn \geq k^*$. In this case there are diminishing benefits of putting weight on some better realizations of the signals. This would imply loses due to the compensation with negative realizations in order to satisfy the marginal constraint. On the other hand, if the marginal probability is relatively low, $pn < k^*$, the benefits of providing good news some of the times more than compensates the diminishing loses of providing fully correlated negative news.[22] Overall, we expect good news to be *negatively correlated* (so that many signals are '1' but some are '0') in the good state, while bad news are often *positively correlated* (so that all are '0') in the bad state.

The discussion above is summarised in Proposition 1:

Proposition 1. *With linear utility, the optimal solution for the sender's problem has positive weight on at most two values of k and is given by:*

$$\omega = 1 \qquad \textit{If } p \geq \frac{k^*}{n}: \qquad\qquad \textit{If } p < \frac{k^*}{n}:$$

$$Q^1_{k(p)} = np + 1 - k(p) \qquad\qquad Q^1_{k^*} = \frac{np}{k^*}$$
$$Q^1_{k(p)-1} = k(p) - np \qquad\qquad Q^1_0 = 1 - \frac{np}{k^*}$$

$$\omega = 0 \qquad Q^0_{k^*} = \frac{n(1-p)}{k^*}$$
$$Q^0_0 = 1 - \frac{n(1-p)}{k^*}$$

[21] Intuitively, the results are related to those in Kamenica and Gentzkow (2011). Although in the current analysis the utility of the sender is linear in the receiver's posterior, correlation neglect implies that this posterior has the usual S-shape of the value of information function. In addition, the Bayesian plausibility constraint in Kamenica and Gentzkow (2011) is here replaced by the marginal constraints.

[22] From this discussion it is easy to argue that anonymous distributions, distributions that allocate the same weight across all realizations with the same number of 1s, are optimal even among non-anonymous ones. Note that the value function V only depends on p and not on the joint distribution chosen. Allocating different weights to different realizations with the same k, implies that in order to satisfy the marginal constraints (for non-anonymous distributions), some positive weight will need to be allocated to some realizations with different k, which can be shown not to be optimal given the convexity and concavity of the V function.

And the sender's indirect utility is then

$$W_S(p,n) = \begin{cases} \frac{1}{2}[V_{k(p)-1} + V_0 + (np+1-k(p))(V_{k(p)} - V_{k(p)-1}) + \frac{n(1-p)}{k^*}(V_{k^*} - V_0)] & if \quad p \geq \frac{k^*}{n} \\ \frac{1}{2}[2V_0 + \frac{n}{k^*}(V_{k^*} - V_0)] & if \quad p < \frac{k^*}{n} \end{cases}$$

where k^* and all the V_k depend on p.

Note that due to correlation neglect, the sender manages to change the ex-ante expectation of the receiver's posterior belief away from the prior. In other words, the ex ante expected action is higher than $\frac{1}{2}$. This can be seen as the actions are not symmetric across states. The sender can always use in state 0 the mirror image symmetric distribution to the one she uses in state 1, but she chooses not to, and hence can implement a higher expected action.[23]

One could easily generalize the analysis above to general sender utility functions $U_S(V_k)$ in which the solution is determined by the concavication of $U_S(V_k)$ as the following figure shows.

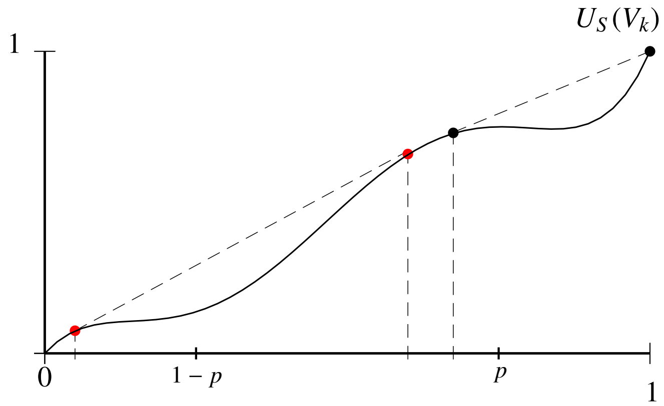


Figure 3: Optimal information structure under general utility: the black (red) dots represents the positive probability mass points when the state is 1 (0).

4.1 Strategic choice of the accuracy of news outlets

In our analysis above, we assumed that the accuracy of the signals was exogenously given. A wider context of competition for consumers would imply that media outlets are informative, although technological constraints and the cost of information may imply that they are not fully informative.

We abstract away from these reasons and consider the case in which the sender could endogenously choose the accuracy of their signals, while maintaining the assumption that the readers understand the accuracy of each outlet.

In particular, we assume that the sender can choose p optimally without any cost or competition considerations, and that as in the standard persuasion literature she can commit to this marginal.[24] We show that the optimal p chosen by the sender is interior. Specifically, we have:

Proposition 2. *The sender's indirect utility $W_S(p,n)$ is maximized at p^* satisfying $\frac{1}{2} < \frac{k^*}{n} \leq p^* < 1$.*

To manipulate the beliefs of readers with correlation neglect the sender has to choose informative but not fully informative marginals. Intuitively, both the fully informative and the least informative marginal will not allow the sender to manipulate the receiver. Moreover, the marginal will be such that indeed

[23]This differs from Kamenica and Gentzkow (2011) in which when the sender has linear preferences as here, and the receiver is rational, the sender does not manage to affect the expected action.

[24]It is possible of course that she chooses a different marginal for each source but this extension will not add much to the analysis. See also Meyer (2017) who shows how in similar environments homogeneity across outlets is best for the sender.

$pn \geq k^*$, and so we would observe the solution which involves only negative correlation in the good state.

Note that reasoning above is robust and does not depend on the fact that we assume symmetric marginals across the two states (so that $p^1 = 1 - p^0$). When we allow for different marginal constraints it is neither optimal to have $p^1 = p^0$, so that the signals are not informative, or $p^1 = 1$ and $p^0 = 0$, so that the signals are fully informative.

4.2 Optimal number of news outlets

We now consider the incentives of the owner to increase (or decrease) the number of news outlets. If an owner cannot correlate the information between outlets, adding more media outlets is not valuable: the expected utility of the owner given that the information structure satisfies independence is fixed at the prior. Having the ability to correlate across news outlets however introduces a benefit for the owner, as we now show:

Proposition 3. *The sender's indirect utility $W_S(p, n)$ increases in n, and moreover, when n is large enough, negative correlation of news increases in n.*

The result implies that senders can become more powerful when they control more media outlets. We show this below using a graphical illustration of the value of being able to correlate signals. When there is no ability to correlate, the marginal utility of increasing the number of media outlets is 0, as the expected utility of the sender from any n under independence is the same - at the prior. However, the ability to correlate creates positive marginal value:

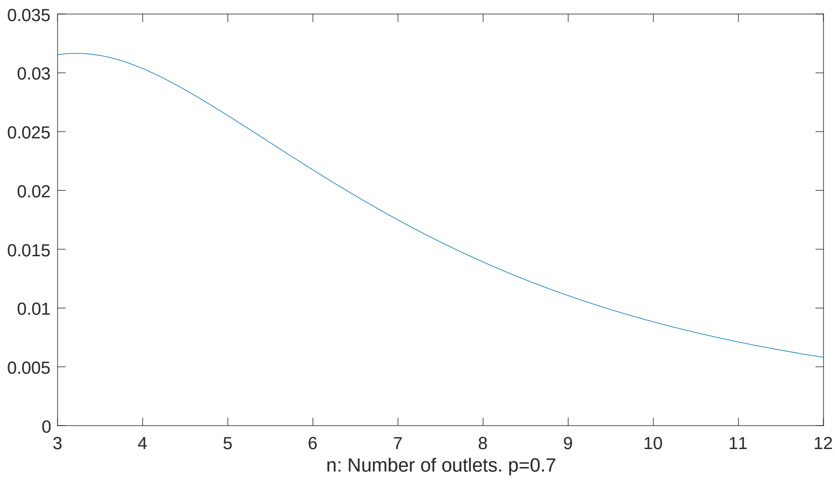


Figure 4: Sender's marginal benefit from controlling an extra outlet.

This is in line with the results in Prat (2017) who shows that when senders can choose to manipulate the signals as they wish (that is, receivers are fully naïve and take the sender's signals at face value), then their power is higher the more media outlets they control. In our case this arises even when the sender is not able to freely manipulate the signals as the receiver holds the sender to account via the marginal constraints.

How does the optimal solution change with the number of signals n? We show in the appendix that k^* decreases with n when n is large enough, implying that we expect to see negative correlation across information sources with a higher probability. Intuitively, when n is very large, a value of k^* slightly larger than $n/2$ is sufficient to induce the optimal belief from the point of view of the sender and is "cost effective" in the sense that more weight can be put on this in order to satisfy the marginal constraints.

Finally we note that when the owner can choose both the number and the accuracy of outlets, she can reach full persuasion. In particular, the sender can get close to an indirect utility of $W_S(p, n) = 1$, which implies that the receiver always believes that the state is '1' with probability one.

4.3 Limited attention

Surveys of news readers' habits consistently show that individuals consume many news sources but devote different amounts of times to different sources. In addition, it is difficult to know how much attention is actually paid to each source and how individuals internalize what is communicated to them. This implies that news outlet owners face uncertainty with respect to the attention that readers devote to their outlets.

We now model limited attention. Specifically, we assume that the receiver only pays attention to $l \leq n$ outlets. However, we assume that the sender does not know which l outlets of the n that she owns the receiver chooses to read. In other words, the sender cannot target specific l signals to the receiver.

Let $V_j^l = \frac{p^j(1-p)^{l-j}}{p^j(1-p)^{l-j}+(1-p)^j p^{l-j}}$ be the expectation of the receiver when a receiver pays attention to only l signals, out of which j have realization 1. We can therefore modify the sender's utility when she delivers n signals, of which k signals have realization 1, given the attention level l:

$$E(V_j^l \mid k, n) = \sum_{j=\max\{0,l-n+k\}}^{\min\{l,k\}} \frac{\binom{k}{j}\binom{n-k}{l-j}}{\binom{n}{l}} V_j^l$$

Lemma 2. *Fix $l < n$, $E(V_j^l \mid k, n)$ is S-shaped and $E(V_j^l \mid k, n) \leq (\geq) V_k^n$ for $k > \frac{n}{2}$ ($k < \frac{n}{2}$).*

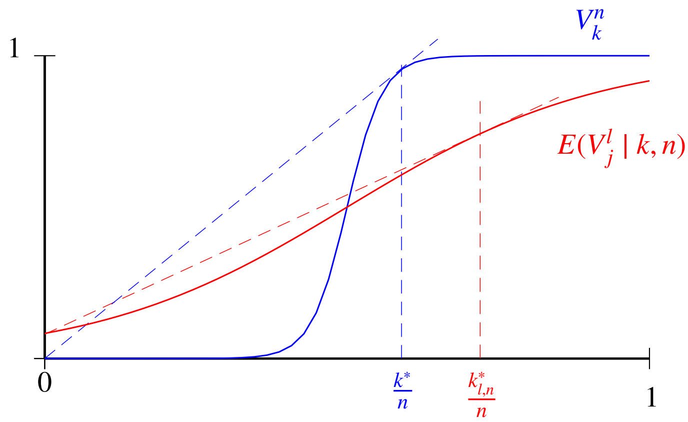


Figure 5: Limited Attention: the blue line corresponds to the sender's utility when she sends k signals with realisation 1 out of n signals, and the receiver pays attention to all of them. The red line corresponds to the sender's utility when she sends k signals with resliation 1, out of n signals, but the receiver has limited attention to l signals.

As a result, we can extend the solution to the basic problem for this case as well (see Figure 5). Intuitively, limited attention implies lower maximum beliefs because the receiver observes less signals, and moreover for any k, we are introducing uncertainty as the sender cannot target the receiver with specific l signals.

We now re-visit the question of optimal n. Note that if the sender knows for sure that she faces a receiver who can observe only $l < n$ signals, then sending more than l signals only introduces uncertainty. In other words, $V_k^l > E(V_j^l \mid k, n)$; V_k^l is the belief of the receiver when she observes k '1' realizations out of l signals (consider $k < l$), and it is therefore the highest belief that enters the expectation $E(V_j^l \mid k, n)$. Thus any optimal solution given n signals can be improved upon by reducing the number of signals to l. We then have:

Proposition 4. *When the receiver has limited attention and can only observe l outlets, the optimal number of signals is l.*

The result above highlights the importance of being able to target your audience. In the face of limited attention, when owners cannot target particular readers to particular newspapers, they face noise in the distribution of signals read by the readers, which reduces their utility.

5 Competition and Welfare

In this section we consider the welfare of readers and how it is affected by regulation of media owners and by competition.

Our key observation is that both enhanced competition among media owners, and regulation reform for the independence of editorial boards, reduce the ability of owners to correlate. If the n media outlets are owned by n different owners, correlation becomes more difficult. Similarly, if regulators make sure that media outlets are not under the influence of owners, and their editorial boards are independent, again, the ability to correlate is reduced.

Discussion among regulators focuses on lowering concentration of ownership not only in order to lower market power but also in order to enhance plurality or "diversity of opinions".

Below we analyze our model in the case of competition, when there are two media owners. We do so for two cases: when owners have the same interests and when they have diverse or opposing interests. We then use this to ask whether it is always better to break up ownership, and if so, whether diverse interests induce higher or lower welfare to the reader compared with the case of the same interests.[25]

To focus on the effects of correlation neglect, our model of competition is simplified in the sense that we abstract away from questions of how readers choose what to read, how much attention they pay, how competition affects profits of media owners, etc. Our analysis can be generalized to include these other elements. Here we assume as above that readers are exposed to all information provided and that the only motive of media owners is to influence opinion. In Section 6.1 we provide an extension in which the media owner can target specific readers, where these differ in their preferences.

For welfare calculations, we need to specify the utility of the receiver. Assume for simplicity that the receiver has to choose an action $a \in \{0, 1\}$. The receiver gains a utility of 1 if she chooses the action a that corresponds with the state of the world, $a = \omega$, and 0 otherwise. With this utility, the receiver will choose 1 iff $V_k \geqslant \frac{1}{2}$. Similarly, we set the sender's utility to be binary.[26] Let then $U_S(V_k) = \begin{cases} 0 \text{ if } V_k < \frac{1}{2} \\ 1 \text{ if } V_k \geqslant \frac{1}{2} \end{cases}$. These preferences are represented in Figure 6.

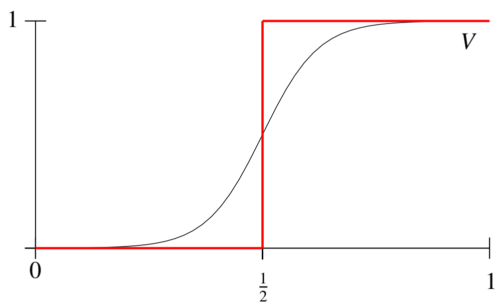


Figure 6: Binary utility.

For simplicity of notation we will assume that np is itself an integer. Moreover, when the receiver is indifferent between the two actions ($k = \frac{n}{2}$), we assume that the receiver takes the action which is in line with the state of the world. This assumption is for technical reasons.[27]

[25] Note that an intervention in the market may affect the marginal constraint p and the number of outlets n: We proceed now with a fixed p and n but later we show that perfect competition is best for all n.

[26] If the sender cares about the receiver's action, as sometimes is assumed, this will follow from the receiver's preferences.

[27] This is not an important assumption in the case of a monopoly or for the perfect competition case. The assumption guarantees existence of equilibrium in the analysis of a duopoly with divergent interests.

Applying the same technique as in Section 4, the optimal solution for a monopolist with n signals who prefers the highest action, will be:[28]

$$Q^1_{np} = 1 \qquad Q^0_{\frac{n}{2}+1} = \frac{(1-p)n}{\frac{n}{2}+1}$$
$$Q^0_0 = 1 - \frac{(1-p)n}{\frac{n}{2}+1}$$

The payoff to the receiver when she faces a monopolist, $W_R^{Mon}(p,n)$, is easily described as the probability that the right decision is taken. This is given by:

$$W_R^{Mon}(p,n) = 1 - \frac{1}{2}\left(\frac{(1-p)n}{\frac{n}{2}+1}\right) = \frac{np+2}{n+2}$$

Finally, note that $\frac{\partial}{\partial n} W_R^{Mon}(p,n) = -\frac{2(1-p)}{(n+2)^2} < 0$. Thus the welfare of the receiver decreases with n.

5.1 Competition

We continue to assume now that there are n media outlets, but that the regulator can break up the monopoly by ensuring that more owners split these n outlets. The key idea behind a competitive market in our framework is that each media owner needs to satisfy the marginal constraint for each of their outlets; fewer outlets will imply then a lower ability to bias beliefs through correlation.

Formally, a pure strategy for a player M_i consists of an information structure over the m_i outlets under the player's control $\{Q^\omega_{k_i}\}_{k_i=0}^{m_i}$, such that the marginal constraint $\sum_{k_i=0}^{m_i} \frac{k_i}{m_i} Q^\omega_{k_i} = p^\omega$ is satisfied. For concreteness, we analyze the simplest case of a duopoly, with each media owner having $\frac{n}{2}$ news outlets. Our results will hold in the more general case (see Section 5.2.3).

In equilibrium, each owner designs a feasible information structure which is a best response to that of the other owner. Since the owners design their content based on the realization of the state of the world, we can solve for the equilibria of the game separately for each state.

5.1.1 Similar interests

Assume first that both owners, M_1 and M_1', have the same preferences as the monopolist in Section 3, preferring action $a = 1$. Note that there can be many equilibria in this case. As the owners have the same preferences, we can choose the Pareto dominant equilibrium. We then have:

Proposition 5. *When the two owners have the same interests, the Pareto dominant equilibrium satisfies:*

$$M_1: \qquad Q^1_{\frac{n}{2}p} = 1 \qquad Q^0_{\frac{n}{2}(1-p)} = 1$$

$$M_1': \qquad Q^1_{\frac{n}{2}p} = 1 \qquad Q^0_{\frac{n}{2}p+1} = \frac{n(1-p)}{np+2}$$
$$Q^0_0 = 1 - \frac{n(1-p)}{np+2}$$

This equilibrium results in an expected utility for the media owners of $W_S^{D-Sim}(p,n) = \frac{1}{2}\frac{n+2}{np+2}$ and an expected utility for the reader of $W_R^{D-Sim}(p,n) = 1 - \frac{1}{2}\frac{n(1-p)}{np+2}$.

Note that as in the case of a monopolist, in state $\omega = 1$ the owners always convince the reader to take the correct action. However, in state $\omega = 0$ the receiver chooses action 1 with a lower probability than in the case of a monopolist. As a result, the utility of the duopolists is lower than the one of a monopolist. This is purely a consequence of the inability of the two owners to fully coordinate their actions.

[28]Note that given the binary preferences of the sender, this optimal structure is not unique. In fact any information structure that in state 1 guarantees action $a = 1$ (i.e. $Q^1_j = 0$ if $j < \frac{n}{2}$) while satisfying the marginal constraint will also be optimal. The information structure in state 0 is nevertheless unique.

5.1.2 Opposing interests

We now consider the case of a duopoly in which the owners are biased in opposite direction. Player M_1 as before is interested in maximising the probability of the receiver taking $a = 1$, while player M_0 has the opposite preferences.

Proposition 6. *The unique equilibrium of the duopoly game consists of both players setting the same information structure described below:*

$$Q^1_k = \frac{(1-p)}{n+2} \quad \textit{for} \quad 0 \le k \le \frac{n}{2} - 1 \qquad\qquad Q^0_0 = 1 - \frac{n}{2}\frac{(1-p)}{n+2}$$

$$Q^1_{\frac{n}{2}} = 1 - \frac{n}{2}\frac{(1-p)}{n+2} \qquad\qquad Q^0_k = \frac{(1-p)}{n+2} \quad \textit{for} \quad 1 \le k \le \frac{n}{2}$$

The expected utility of each media owner is $W_S^{D-Opp}(p,n) = \frac{1}{2}$ and the expected utility for the reader is $W_R^{D-Opp}(p,n) = 1 - 2(1-p)^2\frac{n}{n+2}$.

Note that both owners use the same strategy. In state $\omega = 1$ ($\omega = 0$) they uniformly mix over all k's besides $k = \frac{n}{2}$ ($k = 0$), on which they place a higher weight so as to satisfy the marginal constraint. This strategy, used by both, implies a high degree of noise in equilibrium.

The key idea here is that competition between the owners leads to a zero sum game, each owner's payoff being $\frac{1}{2}$ minus the other owner's payoff. Given the correlation neglect, the reader chooses action $a = 1$ whenever more than half of the signals are positive. As a result, one owner tries to increase k while the other tries to reduce it. In equilibrium we must end up with a distribution over realzations that contain large levels of noise. This is because if one owner chooses an interior number of signals 1 with high probability, the other owner can always choose an strategy that counteracts this and pulls the reader's decision in her direction. This cat and mouse between the owners eliminates the overall bias in the information structure, however this comes at the cost of a very uninformative decisions by the readers.

5.2 Regulation and Welfare

We now use the results in Proposition 5 and 6 above to derive three welfare implications. First, we show that breaking up the market is always beneficial, both in the case of similar preferences and in the case of opposing preferences. Second, we show that sometimes diversity of preferences is actually worse for the reader than similar preferences. Finally, we show that full independence of editorial boards is optimal even if the number of news outlets in the market changes as a response to regulation.

5.2.1 Breaking up monopolies

Competition, disregarding the interests of the owners, implies a lower ability to correlate. As one would expect, the marginal constraints are more binding as they apply to smaller sets of signals.

For the case of similar interests, this results in the overall lower ability of the senders to manipulate the receiver in state 0, and hence results in higher utility for the readers, $W_R^{Mon}(p,n) < W_R^{D-Sim}(p,n)$.

For the case of opposing interests competition has two effects: it leads to less bias but also to more noise in news coverage. By comparing the utility of the reader in both scenarios, it is easy to see that $W_R^{Mon}(p,n) = \frac{np+2}{n+2} < W_R^{D-Opp}(p,n) = 1 - 2(1-p)^2\frac{n}{n+2}$, and therefore the reduction in the bias has a stronger effect than the increase in the level of noise. To see why note that the duopoly behaves equally well (or bad) in both states of the world, leading to the wrong action whenever the two owners' uniform randomization over news end up on the bad end. This happens with probability $2(1-p)^2\frac{n}{n+2}$. On the other hand a monopolist leads always to the right action when the state is in his favor, but manages to fool the reader with probability $2(1-p)\frac{n}{n+2}$ on the unfavorable state. For any $p > \frac{1}{2}$, the mistakes on the duopoly case are less probable than the manipulation in the case of a monopolist.

These results are summarised in the following proposition.

Proposition 7. *Breaking up a monopolist to a duopoly increases the welfare of the reader, both when the duopolists have the same interest and when they have opposing interests.*

Note that while we consider breaking up the monopoly to two owners, the results can be extended to other forms of competition.

5.2.2 The cost of diversity

We now consider the welfare difference between a duopoly with the same interests and one with opposing interests. We have shown that if owners have different biases, competition under correlation neglect eliminates the bias but introduces a high degree of noise. Moreover, the lower the accuracy of the news, the higher the weight on the uniform part of the distribution and hence the noisier the solution becomes. We find that when the accuracy of the news is relatively low, the noisier information transmission means that the reader prefers to face a duopoly with similar interests rather than a duopoly with opposing interest:

Proposition 8. *For any $n \geq 4$ there exists a $\frac{1}{2} < p(n) < 1$, such that if $p < p(n)$ the reader prefers a duopoly with the same interest to a duopoly with opposing interests.*

This result shows that regulatory policy has to be more nuanced when calling for diversity of opinions, and that previous results derived in the literature showing the benefits of competition do not necessarily hold when readers have correlation neglect.

5.2.3 Robust regulation: independent editorial boards

When considering regulation, one often has to forecast how this will change incentives in the market. For example, breaking up a monopolist may result in more or less news outlets overall, as each owner now may adjust the number of outlets to maximize their utility. Thus, for a robust regulation, agencies need to consider how market structure will potentially change.

We conclude this section by showing that when the market is perfectly competitive, then disregarding the number of outlets and the type of ownership structure, the reader gains the highest utility. This is achieved with an extreme market structure in which ownership is completely diffused. Alternatively, this could also be achieved when there is one owner but regulators are able to enforce full independence of editorial boards.

Suppose that there are n news outlets, each completely independent. For each news outlet the only way to fulfil the marginal constraint is to deliver a signal according to the Bernoulli distribution with parameter p. Thus, given conditional independence, the probability that the right decision is taken is:

$$\sum_{k=n/2}^{n} \binom{n}{k} p^k (1-p)^{n-k}.$$

It is obvious that the welfare of the receiver increases with n. However, in the case of a monopolist, or in the case of duopolists with different interests we had analyzed above, the receiver's welfare decreases with n. It is then easy to show that full independence is better for any number of outlets. In other words, it is a regulatory policy that is robust to changes in the market structure:

Proposition 9. *Full independence with at least three news outlets dominates a monopoly or a duopoly with or without diversity of opinions and any number of news outlets.*

Thus, if strong regulation can be enforced, this is preferred no matter how the market ownership will end up.

5.2.4 Behavioural responses

In many media markets regulation is difficult and slow. Ensuring that editorial boards are independent is hard to enforce.[29] Alternatively, one behavioural response is to adjust the cutoff prescribed in the binary

[29] See the recent discussion on News Corp bid to overtake Sky news. https://www.ft.com/content/6dd32342-5ccd-11e7-b553-e2df1b0c3220?mhq5j=e3.

model. For example, readers who observe that overtime they are more likely to take action 1 rather than 0, despite their equal prior, may correct for this by changing their behaviour and taking action 1 only when at least $k > \frac{n}{2}$ signals have realization 1. This will reduce the instances in which action 1 is taken and will make them less manipulable by a monopolist who prefers such an action. Of course the monopolist will design information structures that will be optimal vis a vis the new behaviour of the reader; still, it can be shown that such behavioural response will increase the utility of the reader even when the monopolist's adjustment is taken into account.

6 Extensions

Throughout the paper we have assumed that the sender faced a unique receiver that had correlation neglect. An alternative interpretation of the model is that the sender faces a population of receivers with similar preferences, all of them having correlation neglect. In this section we check the robustness of the analysis when we relax two of the assumptions of the model: that all the receivers have homogeneous preferences, and that they all have correlation neglect. Throughout this section we come back to the case of a monopolist that wants action $a = 1$ to be implemented.

6.1 Heterogeneous preferences

Consider the case in which the receivers differ in their preferences. In particular we consider the case in which a section of the population is keener on action $a = 1$ and hence needs less positive evidence in order to choose such an action. The other section of the population prefers action $a = 0$ and hence demands more convincing evidence in order to choose the high action.[30]

To make the analysis simple and symmetric, suppose that half of the population of receivers is *opposed* to the sender and requires a proportion $\alpha > \frac{1}{2}$ of signals '1' to choose action $a = 1$ whereas the other half is *partisan* and requires a proportion $1 - \alpha < \frac{1}{2}$ of signals '1' to choose action $a = 1$. We assume for convenience that $n\alpha$ is an integer.

The sender then faces a step function as described in Figure 7. Depending on the divergence of preferences, that is, on how big α is, the monopolist would target one section or another of the population.

If α is small ($\alpha < \min\{\frac{2}{3}, p\}$), so the preferences of the readers are not far apart, then the monopolist will try to convince all the receivers by choosing an information structure that achieves action 1 for sure in state 1 and convinces the whole population of receivers with positive probability in state 0 (see Figure 7a).

If α is big ($\alpha > \frac{2}{3}$) but no higher than p, the monopolist targets the receivers that are easier to convince, those who are partisans. The optimal information structure achieves action 1 for sure in state 1 and convinces the partisans receivers in state 0 with positive probability whereas those opposed to her never choose action 1 in state 0 (See Figure 7b).

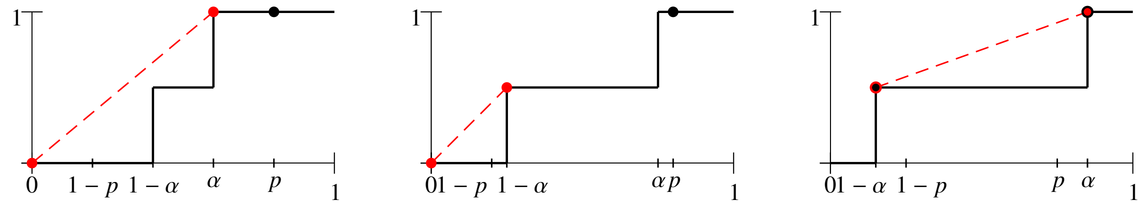
Finally, if $\alpha > p$ (whether $\alpha > \frac{2}{3}$ or not) then the partisan section of the receiver are so biased towards action 1 that the sender can target the receivers opposed to her views. In the optimal information structure, the partisans always choose action 1 and the receivers opposed to her choose action 1 with positive probability in both states, but never with probability one (See Figure 7c).

In particular we can compute how the divergence of preferences among receivers affects the ultimate probability of choosing the correct action.

$$\Pr(a = \omega) = \begin{cases} \frac{1}{2} + \frac{1}{2}\frac{\alpha-(1-p)}{\alpha} & if \quad \alpha < \min\{\frac{2}{3}, p\} \\ \frac{1}{2} + \frac{1}{4}\frac{1-p-2\alpha}{1-\alpha} & if \quad \frac{2}{3} < \alpha < p \\ \frac{1}{4} + \frac{1}{2}\frac{p-(1-\alpha)}{2\alpha-1} & if \quad \alpha > p \end{cases}$$

As long as $\alpha < \min\{\frac{2}{3}, p\}$ the diversity of preferences improves the receivers action choice because it disciplines the media owner who wants to persuade the whole population and hence needs to report

[30] An alternative explanation is that the population has heterogenous priors beliefs about the state of the world.



(a) $\alpha < \min\{\frac{2}{3}, p\}$ **(b)** $\alpha > \frac{2}{3}, \alpha < p$ **(c)** $\alpha > p$

Figure 7: Heterogeneous receivers

negative news more often when $\omega = 0$. However, when $\alpha > \min\{\frac{2}{3}, p\}$ the probability of choosing the correct action decreases with α. When $\frac{2}{3} < \alpha < p$ the media owner tailors the persuasion towards partisans. The more extreme partisans are, the easier it is to persuade them to take the incorrect action in state $\omega = 0$. When $\alpha > p$, the media concentrates its persuasion on the receivers with opposite views, however, the more sceptic the receiver is the more difficult it is to convince those readers even in the state in which they agree.

Note that when the divergence of preferences is not too big (Figure 7a) the solution is close to the one described in Section 5 and our analysis is robust to this sort of heterogeneity.

6.2 Rational receivers

Consider now the case in which the sender faces a mixed audience that includes both naïve individuals as analyzed above, as well as rational individuals that understand the correlation in the information structure generating the signals.

A rational receiver that perfectly understands the information structure will update beliefs in a way that the expected posterior always equals the prior. Therefore, when the sender's utility is linear on the receivers' beliefs as in Section 3 and the sender faces a rational receiver, any information structure leads to exactly the same expected payoff. In particular the optimal solution proposed in Proposition 1 remains optimal when some of the receivers are rational. In that sense, the naïve solution can be seen as a refinement on the set of possible solutions for rational receivers.

When the sender's preferences are not linear on the receiver's beliefs, the signal structure does affect the sender's payoff even when facing rational receivers. To see that, consider the binary model described in Section 5. The sender can convince a rational receiver to take action 1 by producing k signals '1' as long as the information structure in state $\omega = 1$ allocates more weight to k signals '1' than the information structure in state $\omega = 0$, i.e. as long as $Q^1_k > Q^0_k$. In particular, if the information structures in the two states never produce the same number of signals '1', then the receiver perfectly learns the state.

The optimal information structure for a sender that faces a rational receiver is given as follows:

Lemma 3. *When facing a rational receiver, the optimal strategy of the sender is to use the following information structure:*

$$Q^1_{\frac{n}{2}+1} = \frac{n(1-p)}{\frac{n}{2}-1} \qquad Q^0_{\frac{n}{2}+1} = \frac{(1-p)n}{\frac{n}{2}+1}$$
$$Q^1_n = 1 - \frac{n(1-p)}{\frac{n}{2}-1} \qquad Q^0_0 = 1 - \frac{(1-p)n}{\frac{n}{2}+1}$$

And the sender's indirect utility is $\frac{1}{2} + \frac{1}{2}\frac{n(1-p)}{\frac{n}{2}+1}$.

Note that as stated in Footnote 28 this information structure is also optimal when the receiver is naïve, therefore a sender that uses this information structure when facing naïve receivers is immune to the presence of some rational receivers.

Corollary 1. *The optimal information structure is robust to the introduction of some rational receivers in the audience.*

7 Conclusion

We have presented a model in which media owners can take advantage of the readers' cognitive bias of correlation neglect. We have shown that in this environment, breaking up media conglomerates as well as insuring independence of editorial boards is always beneficial, but that diverity of opinions among media owners is not necessarily so. The next step in the theoretical analysis will be to combine our model with more traditional models of readership specifically those that allow readers to choose their information sources on the basis of e.g., political bias.

In terms of the empirical literature, our paper suggests that a good way to proceed is to examine the combined bias of several news outlets (which are related through ownership). While the literature has focused so far on measuring and analyzing the bias of each news outlet separately, our analysis shows that such bias may arise only when considering many such outlets together.

A Appendix

A.1 Proof of Lemma 1

The concavity (convexity) of V_k for $k > \frac{1}{2}$ ($k < \frac{1}{2}$) follows from Lemma 7 in Appendix B when $\mu = \frac{1}{2}$ and $q = 1 - p$. The symmetry around $k = \frac{1}{2}$ follows from $V(p, \frac{n}{2}, n) = \frac{1}{2}$ and $V(p, k, n) + V(p, n-k, n) = 1$.

A.2 Proof of Proposition 1

Preliminary step: Suppose that we have $Q_x, Q_y, Q_z > 0$ for $x < y < z$. Suppose we increase by ε the weight on y, and hence reduce the weight on x and z together by $-\varepsilon$. To satisfy the marginal constraint, we need

$$-x\beta + y\varepsilon - z(\varepsilon - \beta) = 0 \quad \text{which implies} \quad \beta = \frac{\varepsilon(z-y)}{z-x}$$

Note that we cannot have $\beta = 0$ or $\beta = \varepsilon$. In terms of utility then, this is worth:

$$\Delta_{x,y,z} = -V_x(\frac{z-y}{z-x}) + V_y - V_z(1 - \frac{z-y}{z-x})$$

If V is concave on this region $\Delta_{x,y,z}$ is positive, if convex $\Delta_{x,y,z}$ is negative.

Lemma 4. $\Delta_{0,k-1,k} > (<)0$ *is sufficient to characterize the optimal solution.*

Note that $\Delta_{0,k-1,k} = -V_0(\frac{1}{k}) + V_{k-1} - V_k(1 - \frac{1}{k}) = V_{k-1} - V_k + \frac{1}{k}(V_k - V_0)$. Therefore, $\Delta_{0,k-1,k} > 0$ iff $V_k - V_0 > k(V_k - V_{k-1})$.

Proof of Proposition 1 and Lemma 4: First note that by the above, we will not have three $Q's$ or more which are strictly positive. If so, we can shift weight either to a middle one from two extreme values or to two extreme values from a middle one and we get in one of these ways an improvement in utility. We cannot have one value only which is positive because of the constraints, so there are two and only two $Q's$ which are positive.

Second assume that the two $Q's$ are Q_k, Q_l for $k - 1 > l \geq \frac{n}{2}$. Assume now we shift weight to $k - 1$, we get $\Delta_{l,k-1,k} > 0$ by concavity.

By the constraints, $Q_k, Q_{k-1} > 0$ is feasible when $\frac{k-1}{n} \leq p \leq \frac{k}{n}$.

Next assume that $\frac{n}{2} > l > 0$, and $k > l$. Here $\Delta_{0,l,k} < 0$. Moreover, $\Delta_{0,n/2-1,n/1} < 0$. Finally note that as $\Delta_{0,k,k+l} > 0$ for $k > k^*$, then whenever $k, 0$ is feasible for $k > k^*$ then it is preferred to $k + l, 0$ by concavity. Thus such a solution will involve the first k that satisfies $\Delta_{0,k-1,k} = 0$.

So any optimal solution involves either $k, 0$ or $k, k-1$, for $k > \frac{n}{2}$ so as stated in Lemma 1, $\Delta_{0,k-1,k} > (<)0$ will be sufficient to characterize the optimal solution. Note further more that $k, 0$ is possible whenever $p \leq \frac{k}{n}$.

For any $\frac{k-1}{n} \leq p \leq \frac{k}{n}$, we can choose either $k, 0$ or $k, k-1$: but we know that whenever $pn \leq \frac{n}{2} + 1$, for sure we choose $\frac{n}{2} + 1, 0$ rather than $n/2, 0$ or $n/2, n/2 + 1$, and whenever $p > \frac{n-1}{n}$, we choose $n, n-1$ as $\Delta_{0,n-1,n} > 0$.

Intuitively/graphically, the function becomes less concave when p goes down. Thus once we move to the convex solution of $k^*(p), 0$, we remain at this form of solution.

Suppose at some $p \leq \frac{k}{n}$, for the first time, we have chosen to move to $k, 0$. Let us denote this by k^*. and so k^* is the highest integer at which $(V_k - V_{k-1})(k-1) + V_0 - V_{k-1} = 0$ and p falls in the bounds so that $k(p)$ and $k^*(p)$ identify.

A.3 A smooth presentation

In what follows it will sometimes be useful to work with the smooth version of the value of information V:

$$V(p,x,n) = \frac{p^{nx}(1-p)^{n(1-x)}}{p^{nx}(1-p)^{n(1-x)} + p^{n(1-x)}(1-p)^{nx}}$$

In this representation x correspond to k/n, i.e. the share of signals that are equal to 1. We allow x to be a continuous variable. This would be a good approximation in the case in which we have a large number of signals n. But more generally this will help us simplify the exposition.

Define $x^*(p,n)$ as the solution to:

$$\frac{\partial V}{\partial x}(p,x,n) = \frac{V(p,x,n) - V(p,0,n)}{x}. \tag{1}$$

Now the optimal solution (using the smoothness) is:

If $p \geq x^*(p,n)$:

$$\begin{cases} Q^1_p = 1 \\ Q^1_j = 0, \quad j \neq p \end{cases} \qquad \begin{cases} Q^0_{x^*} = \frac{(1-p)}{x^*} \\ Q^0_0 = 1 - \frac{(1-p)}{x^*} \\ Q^0_j = 0, \quad j \notin \{nx^*, 0\} \end{cases}$$

If $p < x^*(p,n)$:

$$\begin{cases} Q^1_{x^*} = \frac{p}{x^*} \\ Q^1_0 = 1 - \frac{p}{x^*} \\ Q^1_j = 0, \quad j \notin \{x^*, 0\} \end{cases} \qquad \begin{cases} Q^0_{x^*} = \frac{(1-p)}{x^*} \\ Q^0_0 = 1 - \frac{(1-p)}{x^*} \\ Q^0_j = 0, \quad j \notin \{x^*, 0\} \end{cases}$$

And the sender's expected utility is

$$W_S(p,n) = \begin{cases} \frac{1}{2}[V_p + V_0 + \frac{(1-p)}{x^*}(V_{x^*} - V_0)] & if \quad p > x^* \\ \frac{1}{2}[2V_0 + \frac{1}{x^*}(V_{x^*} - V_0)] & if \quad p < x^* \end{cases}$$

A.4 Proof of Proposition 2

We start with a preliminary result:

Lemma 5. *For any $0 \leq \gamma \leq 1$, and $x, y > 0$, the function $f(\gamma, x, y) = x^{n\gamma}y^{n(1-\gamma)}$ is supermodular in (x, y).*

Proof.

$$\frac{\partial^2 f(\gamma,x,y)}{\partial x \partial y} = n^2\gamma(1-\gamma)x^{n\gamma-1}y^{n(1-\gamma)-1} > 0$$

□

Remark 1. *For any $p > \frac{1}{2}$ and any $\gamma \in [0,1]$,*

$$p^n + (1-p)^n > p^{n\gamma}(1-p)^{n(1-\gamma)} + (1-p)^{n\gamma}p^{n(1-\gamma)}$$

Proof of Proposition 2: Recall that the sender's utility is:

$$W_S(p,n) = \begin{cases} \frac{1}{2}[V_p + V_0 + \frac{(1-p)}{x^*}(V_{x^*} - V_0)] & if \quad p > x^* \\ \frac{1}{2}[2V_0 + \frac{1}{x^*}(V_{x^*} - V_0)] & if \quad p < x^* \end{cases}$$

Taking a derivative of this w.r.t. p, and using the Envelope Theorem so that the indirect effect through the change in the optimal solution is zero, we have:

$$\frac{\partial W_S(p,n)}{\partial p} = \begin{cases} \frac{1}{2}[(\frac{\partial V_x}{\partial p})_{x=p} + \frac{1-p}{x^*}\frac{\partial V_{x^*}}{\partial p} + (1 - \frac{1-p}{x^*})\frac{\partial V_0}{\partial p} - \frac{1}{x^*}(V_{x^*} - V_0)] & if \quad p > x^* \\ \frac{1}{2}[\frac{1}{x^*}\frac{\partial V_{x^*}}{\partial p} + (2 - \frac{1}{x^*})\frac{\partial V_0}{\partial p}] & if \quad p < x^* \end{cases}$$

We first show that for $p < x^*$, $\frac{\partial W_S(p,n)}{\partial p} > 0$.
In order to see that note that

$$\frac{\partial V_x}{\partial p} = \frac{p^{n-1}(1-p)^{n-1}n(2x-1)}{(p^{xn}(1-p)^{n(1-x)} + p^{n(1-x)}(1-p)^{xn})^2} \tag{2}$$

$$\frac{\partial V_0}{\partial p} = -\frac{p^{n-1}(1-p)^{n-1}n}{((1-p)^n + p^n)^2} \tag{3}$$

Thus,

$$\begin{aligned}
sign\left\{\frac{\partial W_S(p,n)}{\partial p}\right\} &= sign\left\{\frac{1}{x^*}\frac{\partial V_{x^*}}{\partial p} + (2 - \frac{1}{x^*})\frac{\partial V_0}{\partial p}\right\} \\
&= sign\left\{\frac{1}{x}p^{n-1}(1-p)^{n-1}n(2x-1)(\frac{1}{(p^{xn}(1-p)^{n(1-x)}+p^{n(1-x)}(1-p)^{xn})^2} - \frac{1}{((1-p)^n+p^n)^2})\right\}_{x=x^*} \\
&= sign\left\{((1-p)^n + p^n)^2 - (p^{xn}(1-p)^{n(1-x)} + p^{n(1-x)}(1-p)^{xn})^2\right\}_{x=x^*} > 0
\end{aligned}$$

where the third equality follows because $x^* > \frac{1}{2}$ and the inequality follows by Remark 1.

Note also that when $p \to 1$, $\frac{\partial V_x}{\partial p}\Big|_{x=p}$, $\frac{\partial V_0}{\partial p}$ and $1-p$ converge to 0 while $-\frac{1}{x^*}(V_{x^*} - V_0)$ is negative and bounded away from 0. In that case $\frac{\partial W_S(p,n)}{\partial p}\Big|_{p\to 1} < 0$ and so the solution is bounded away from 1. We therefore conclude that the solution is in $[x^*(p,n), 1)$.

A.5 Proof of Proposition 3

To see this, using the smooth representation, it is easy to show that

$$\frac{\partial V}{\partial n}(p,x,n) = \frac{p^n(1-p)^n(ln(p) - ln(1-p))(2x-1)}{[p^{nx}(1-p)^{n(1-x)} + p^{n(1-x)}(1-p)^{nx}]^2}, \tag{4}$$

so for $p > \frac{1}{2}$, V is increasing in n if $x > \frac{1}{2}$ and decreasing otherwise. In particular for any $x > \frac{1}{2}$, $\frac{V(p,x,n)-V(p,0,n)}{x}$ is increasing in n.
Note that

$$x^* = \arg\max_x \frac{V(p,x,n) - V(p,0,n)}{x}$$

and

$$\frac{\partial V}{\partial x}(p,x^*,n) = \max_x \frac{V(p,x,n) - V(p,0,n)}{x}$$

hence, $\frac{\partial V}{\partial x}(p,x^*,n)$ is also increasing in n. Given (1), we can write the ex-ante utility of the sender as:

$$W_S(p,n) = \frac{1}{2}[V_p^n + V_0^n + (1-p)\frac{\partial V}{\partial x}(p,x^*,n)]$$

We have shown that $\frac{\partial V}{\partial x}(p,x^*,n)$ is increasing in n, so it is enough to show that $V_p^n + V_0^n$ is also increasing in n. But,

$$\frac{\partial V_p^n + V_0^n}{\partial n} = p^n(1-p)^n(ln(p)-ln(1-p))\left[\frac{(2p-1)}{[p^{np}(1-p)^{n(1-p)} + p^{n(1-p)}(1-p)^{np}]^2} - \frac{1}{(p^n + (1-p)^n)^2}\right] > 0,$$

where the sign follows by Remark 1.

Therefore we have established that the sender wants as many signals as possible.

We can also compute an upper bound for the sender's welfare $W_S(p,n)$ by noting that when n goes to infinity, $\frac{\partial V}{\partial x}(p,x^*,n)$ converges to 2 (as x^* converges to $\frac{1}{2}$), and therefore,

$$\lim_{n\to\infty} W_S(p,n) = \lim_{n\to\infty} \frac{1}{2}[V_p^n + V_0^n + (1-p)\frac{\partial V}{\partial x}(p,x^*,n)] = \frac{3}{2} - p.$$

Lemma 6. $\partial x^*/dn < 0$ *for a large enough n*

Proof. We now show that x^* decreases in n.

Recall that x^* solves equation (1). Define $F(x,n) = V_x - V_0 - x\frac{\partial V_x}{\partial x}$. Then $F(x^*,n) = 0$ Using total differentiation, we have that:

$$\frac{dx^*}{dn} = -\frac{\frac{dF}{dn}|_{x=x^*}}{\frac{dF}{dx}|_{x=x^*}}$$

Note that

$$\frac{dF}{dx}|_{x=x^*} = \frac{\partial V_x}{\partial x}|_{x=x^*} - \frac{\partial V_x}{\partial x}|_{x=x^*} - x^*\frac{\partial^2 V_x}{\partial x^2}|_{x=x^*} = -x^*\frac{\partial^2 V_x}{\partial x^2}|_{x=x^*}$$

By concavity, $\frac{\partial^2 V_x}{\partial x^2}|_{x=x^*} < 0$ and hence $\frac{dF}{dx}|_{x=x^*} > 0$. Using equation (4) we have:

$$\begin{aligned}\frac{dF}{dn}|_{x=x^*} &= \frac{\partial V_x}{\partial n}|_{x=x^*} - \frac{\partial V_x}{\partial n}|_{x=0} - x^*\frac{\partial^2 V_x}{\partial x \partial n}|_{x=x^*} \\ &= p^n(1-p)^n \ln(\frac{p}{1-p})[\\ &\quad \frac{(2x^*-1)}{[p^{nx^*}(1-p)^{n(1-x^*)} + p^{n(1-x^*)}(1-p)^{nx^*}]^2} + \frac{1}{[p^n + (1-p)^n]^2} \\ &\quad -x^*\frac{\partial}{\partial x}\frac{(2x-1)}{[p^{nx}(1-p)^{n(1-x)} + p^{n(1-x)}(1-p)]^2}]\end{aligned}$$

Where the last partial derivative equals:

$$\frac{\partial}{\partial x}\frac{(2x-1)}{[p^{nx}(1-p)^{n(1-x)} + p^{n(1-x)}(1-p)]^2} = \frac{2}{[p^{nx}(1-p)^{n(1-x)} + p^{n(1-x)}(1-p)]^2} - \frac{2(2x-1)n\ln(\frac{p}{1-p})(p^{nx}(1-p)^{n(1-x)} - p^{n(1-x)}(1-p))}{[p^{nx}(1-p)^{n(1-x)} + p^{n(1-x)}(1-p)]^3}$$

So the sign of $\frac{dF}{dn}|_{x=x^*}$ is the same as the sign of:

$$\begin{aligned}&\frac{(2x^*-1)}{[p^{nx^*}(1-p)^{n(1-x^*)} + p^{n(1-x^*)}(1-p)^{nx^*}]^2} + \frac{1}{[p^n + (1-p)^n]^2} \\ &-\frac{2x^*}{[p^{nx^*}(1-p)^{n(1-x^*)} + p^{n(1-x^*)}(1-p)^{nx^*}]^2} + \frac{2x^*(2x^*-1)n\ln(\frac{p}{1-p})(p^{nx}(1-p)^{n(1-x)} - p^{n(1-x)}(1-p))}{[p^{nx^*}(1-p)^{n(1-x^*)} + p^{n(1-x^*)}(1-p)^{nx^*}]^3} \\ = &\frac{1}{[p^{nx^*}(1-p)^{n(1-x^*)} + p^{n(1-x^*)}(1-p)^{nx^*}]^2}\left[2x^*\ln\frac{p}{1-p}^{(2x^*-1)n}\frac{\frac{p}{1-p}^{(2x^*-1)n} - 1}{\frac{p}{1-p}^{(2x^*-1)n} + 1} - 1\right] + \frac{1}{[p^n + (1-p)^n]^2}\end{aligned}$$

Now for sufficiently large n, x^* approaches $1/2$, and hence it has to be that $\frac{p}{1-p}^{(2x^*-1)n} \to \infty$ (as $V_{x^*} = \frac{\frac{p}{1-p}^{(2x^*-1)n}}{\frac{p}{1-p}^{(2x^*-1)n}+1} \to 1$) so $\frac{dF}{dn}|_{x=x^*} > 0$ and x^* is decreasing in n. □

A.6 Proof of Proposition 4

As stated in the main text, the proof of Proposition 4 follows trivially from Lemma 2 which is proven below.

Proof of Lemma 2:

Consider $k > \frac{n}{2}$ (the case $k < \frac{n}{2}$ is analogous). Assume that we have $l < n$ signals. To save on notation we denote $E(V_j^l \mid k, n)$ by $V_k^{l,n}$, i.e., $V_k^{l,n}$ is the expected belief of the receiver when the sender sends k signals 1 out of n, and the receiver only pays attention to l signals. $V_k^{l,n}$ is given by:

$$V_k^{l,n} = \sum_{j=\max\{0,l-n+k\}}^{\min\{l,k\}} \frac{\binom{k}{j}\binom{n-k}{l-j}}{\binom{n}{l}} \frac{p^j(1-p)^{l-j}}{p^j(1-p)^{l-j} + (1-p)^j p^{l-j}}$$

There are three possible cases to consider.

Case 1, $l < n - k$:
In this case j can go from 0 to l. An upper bound on the utility is, by concavity, the expectation at the average number of j signals. The average is easily computed as $\sum_{j=0}^{l} \frac{\binom{k}{j}\binom{n-k}{l-j}}{\binom{n}{l}} j = kl/n$. Then in this case we can compute the belief of getting kl/n 1 signals out of l signals. This belief is lower than V_k if:

$$\begin{aligned} V^l_{kl/n} &= \frac{p^{kl/n}(1-p)^{l-kl/n}}{p^{kl/n}(1-p)^{l-kl/n} + p^{l-kl/n}(1-p)^{kl/n}} < \frac{p^k(1-p)^{n-k}}{p^k(1-p)^{n-k} + p^{n-k}(1-p)^k} = V^n_k \Leftrightarrow \\ kl/n + n - k &< k + l - kl/n \Leftrightarrow 2k > n, \end{aligned}$$

which holds.

Case 2, $n - k < l < k$:
We proceed by induction on l. Step 1: $V^n_k \geq V^{n-1,n}_k = (1 - \frac{k}{n})V^{n-1}_k + \frac{k}{n}V^{n-1}_{k-1}$

Note that $(1 - \frac{k}{n})V^{n-1}_k + \frac{k}{n}V^{n-1}_{k-1} \leq V^{n-1}_{(1-\frac{k}{n})k+\frac{k}{n}(k-1)}$ (by concavity)$= V^{n-1}_{\frac{k}{n}(n-1)} \leq V^n_k$

Step 2: Assume by induction, $V^n_k \geq V^{l,n}_k$. We now show $V^n_k \geq V^{l-1,n}_k$, by showing that $V^{l,n}_k \geq V^{l-1,n}_k$.

For short, let $q^l_j \equiv \frac{\binom{k}{j}\binom{n-k}{l-j}}{\binom{n}{l}}$.

But we know $V^{l,n}_k = \sum_{j=l-(n-k)}^{k} q^l_j V^l_j \geq \sum_{j=l-(n-k)}^{k} q^l_j((1 - \frac{j}{l})V^{l-1}_j + \frac{j}{l}V^{l-1}_{j-1})$ from step 1 applied to l instead. We will show that $\sum_{j=l-(n-k)}^{k} q^l_j((1 - \frac{j}{l})V^{l-1}_j + \frac{j}{l}V^{l-1}_{j-1}) = V^{l-1,n}_k = \sum_{j=l-1-(n-k)}^{k} q^{l-1}_j V^{l-1}_j$. In $\sum_{j=l-(n-k)}^{k} q^l_j((1-\frac{j}{l})V^{l-1}_j + \frac{j}{l}V^{l-1}_{j-1})$ we have in $V^{l,n}_k$ a distribution over V^{l-1}_j ranging over $\{l-1-(n-k), .., k\}$ and in $V^{l-1,n}_k$ a distribution V^{l-1}_j also over $\{l-1-(n-k), .., k\}$, but now we compare the weights.

In $\sum_{j=l-(n-k)}^{k} q^l_j((1 - \frac{j}{l})V^{l-1}_j + \frac{j}{l}V^{l-1}_{j-1})$ the weight over V^{l-1}_j is $q^l_j(1 - \frac{j}{l}) + q^l_{j+1}\frac{j+1}{l}$. On the other hand, the weight in $V^{l-1,n}_k$ is q^{l-1}_j (for interior ones, excluding $l - 1 - (n - l)$, and k).

Let us compare these weights now.

$$q^l_j(1-\frac{j}{l})+q^l_{j+1}\frac{j}{l} = \frac{\binom{k}{j}\binom{n-k}{l-j}}{\binom{n}{l}}(1-\frac{j}{l})+\frac{\binom{k}{j+1}\binom{n-k}{l-j-1}}{\binom{n}{l}}\frac{j+1}{l} = \frac{\binom{k}{j}\binom{n-k}{l-1-j}}{\binom{n}{l-1}} \frac{\frac{k!}{j!(k-j)!}\frac{n-k!}{l-j!(n-k-l+j)!}}{\frac{n!}{l!(n-l)!}}(1-\frac{j}{l})+\frac{\frac{k!}{j+1!(k-j-1)!}\frac{n-k!}{l-j-1!(n-k-l+j+1)!}}{\frac{n!}{l!(n-l)!}}\frac{j+1}{l}$$

$$= \frac{\frac{k!}{j!(k-j)!}\frac{n-k!}{l-j-1!(n-k-l+j+1)!}}{\frac{n!}{l-1!(n-l+1)!}} = q^{l-1}_j, \text{ as required.}$$

For $l - 1 - (n - k)$, in $V^{l,n}_k$ the weight is $q^l_{l-(n-k)}\frac{l-(n-k)}{l} = q^{l-1}_{l-1-(n-k)}$, and for k, in $V^{l,n}_k$ the weight is $q^l_k(1 - \frac{k}{l}) = q^{l-1}_k$

Case 3, $n - k < k < l$:
The same previous argument can be replicated if the maximum is instead $l > k$.

Now note that due to the symmetry, $V^{l,n}_{n/2} = V^{n,n}_{n/2}$. Thus as $V^{l,n}_k$ reaches a lower maximum and is lower at any point, it will be less concave than V^n_k.

To see that this yields a solution with more positive correlation, note that for a large enough n, the solution will be converging to $n/2$ and thus for all relatively small l, a higher k^* will arise. Also, we can use the comparative statics from the previous proof.

A.7 Proof of Proposition 5

We first check that the strategies described in the proposition are feasible and are indeed an equilibrium.

In state $\omega = 1$, the two media owners choose $Q^1_{\frac{np}{2}} = 1$ which is indeed feasible given the marginal p and that each owns $\frac{n}{2}$ outlets. This strategy leads to the optimal solution and hence cannot be improved upon.

In state $\omega = 0$, one of the owners choose $Q^0_{\frac{n(1-p)}{2}} = 1$ which is clearly feasible. In order to convince the reader with some probability the other owner has to put weight on $k \geq \frac{np+2}{2}$ (Note that the reader chooses action 1 only if $k > \frac{n}{2}$). Given the convexity of the reader's belief over that region, the optimal way to do so is to put weigh on $k = \frac{np+2}{2}$ and $k = 0$, which is what the solution for M_1' depicts. It is clear that the M_1's strategy is optimal given the strategy of M_1'.

Finally, the proposed equilibrium is the one that reaches higher probability of action 1 in state $\omega = 0$ and hence Pareto dominates all others.

A.8 Proof of Proposition 6

We provide the proof for $\omega = 1$. The case $\omega = 0$ is symmetric. We first show that the strategies proposed constitute an equilibrium. We start by showing that the strategies are feasible, that is, they satisfy the probability constraints and the marginal constraint. It is easy to see that if $p > \frac{1}{2}$, $0 < Q_k < 1$ for all $0 \leq k \leq l$ and that $\sum_{k=0}^{l} Q_k = 1$, so the strategies are indeed an informations structure. Denote by $Q = \frac{2(1-p)}{l+1}$ and note that $Q_l = 1 - lQ$. The marginal constraint becomes:

$$\sum_{k=0}^{l-1} Q\frac{k}{l} + Q_l = p$$

Replacing Q and Q_l by their values,

$$\begin{aligned}
\frac{Q}{l}\sum_{k=0}^{l-1} k + 1 - lQ &= \frac{Q(l-1)}{2} + 1 - lQ \\
&= 1 - \frac{Q(l+1)}{2} = p
\end{aligned}$$

Finally, to see that this is an equilibrium note that the expected utility of M_1 when she sets k_1 positive signals, and M_0 plays according with the strategy is:

$$\begin{aligned}
V_{k_1} &= \Pr(k_0 + k_1 \geq l) \\
&= \sum_{s=l-k_1}^{l} Q_s \\
&= 1 - \frac{2(l-k_1)(1-p)}{l+1}
\end{aligned}$$

which is linear in k_1. To see that there are no profitable deviations from this strategy, consider another information structure for M_1, $\{Q'_{k_1} = Q_{k_1} + \Delta_{k_1}\}_{k_1=0}^{l}$. For the deviation to be feasible, it needs to satisfy the following two constraints:

$$\sum \Delta_{k_1} = 0, \quad \sum \frac{k}{l}\Delta_{k_1} = 0.$$

But then the change in the expected utility for M_1 given the deviation is:

$$\begin{aligned}
\sum \Delta_{k_1} V_{k_1} &= \sum \Delta_{k_1} - \sum \Delta_{k_1}\frac{2(l-k_1)(1-p)}{l+1} \\
&= \sum \Delta_{k_1} - \frac{2l(1-p)}{l+1}\sum \Delta_{k_1} + \frac{2l(1-p)}{l+1}\sum \frac{k_1}{l}\Delta_{k_1} \\
&= 0
\end{aligned}$$

So there is no profitable deviation for M_1. An analogous computation shows that the expected utility of M_0 as a function of k_0 given what M_1 does and given the tie breaking rule, is also linear in k_0.

Finally, we show that this equilibrium is unique.

Claim 1: (i) If for M_1 , $Q_k = 0$ for some $k < l$ then for M_0, $Q_{l-(k+1)} = 0$. (ii) If for M_0, $Q_k = 0$ for some $k < l$ then for M_1, $Q_{l-k} = 0$.

Proof. Note that (ii) will follow from (i) as they are mirror images. So we prove (i). Assume without loss that the state is one. Suppose that for some $k < l$, for M_1, $Q_k^{M_1} = 0$. Note that by the tie breaking rule, whenever the sum of signals with relaization '1' across both M_1 and M_0 are l or higher the sender's utility is 1. This implies that M_0 get the same level of utility from having $l - k$ signals with realization '1' to $l - (k + 1)$. As the expected utility of M_0 is decreasing in the number of signals with realization '1', this implies that $l - (k + 1)$ can never be part of an optimal solution for M_0; no concavication of his expected utility will include $l - (k + 1)$ as a solution. Therefore, $Q_{l-(k+1)}^{M_0} = 0$. □

An implication of Claim 1 is that all equilibria are of the form whereby there exists a k^*, $l \geqslant k^* > pl$ such that the support of the strategy of M_1 is $\{0, 1, ..., k\}$ and that of M_0 is $\{l - k, l - k + 1, ..., l\}$.

Claim 2: Fix and equilibrium with k^*, for all $k < k^*$, $Q_k^{M_1} = \alpha$, $Q_{l-k}^{M_0} = \beta$.

Proof. We provide a Graphical proof. We know that the solution for M_1 is to concavify his expected utility $V_k^{M_1} = \Pr(\tilde{k}_0 \geq l - k) = \sum_{s=l-k}^{l} Q_s^{M_0}$. the only way an optimal solution of this problem results in a support of $\{0, 1, ..., k\}$ is if $V_k^{M_1}$ is linear on $\{0, 1, ..., k\}$ which is what we need to prove. □

Consider M_0 and assume that $k^* < l$. Since $V_{l-k*}^{M_0} = -Q_{k^*}^{M_1} < 0$ and $V_k^{M_0} = -\sum_{s=l-k}^{k^*} Q_s^{M_1}$ for $k > l - k^*$ is linear, a convexication for M_0 will imply an optimal solution with strictly positive weights on $l - k^* - 1$ and l only. This contradicts our premise about the equilibrium being characterized by k^*.

We conclude that the equilibrium has to have $k^* = l$ and by Claim 2 must take the form given in the proposition.

A.9 Proof of Propositions 8

The proof follows from comparing the reader's expected utility in the two duopolies:
$W_R^{D-Sim} \geq W_R^{D-Opp}$ if and only if

$$1 - \frac{1}{2}\frac{n(1-p)}{np+1} \geq 1 - 2(1-p)^2\frac{n}{n+2}$$

Which for $p > \frac{1}{2}$ occurs whenever $p \leq p(n)$ where $p(n) = \frac{n-2+\sqrt{2(2+n)}}{2n}$. Note further that $\frac{\partial p(n)}{\partial n} = \frac{2\sqrt{2(2+n)}-4-n}{2n^2\sqrt{2(2+n)}} < 0$, for any $n \geq 0$ so as n becomes larger it is more likely that plurality of opinion benefits the reader.

A.10 Proof of Proposition 9

As the welfare of the receiver increases with the number of outlets n under perfect competition and decreases with n (under a duopoly), it is enough to focus on the minimal number of outlets to show the result. Suppose that there are 4 outlets (2 per each in a duopoly, 4 for a monopolist, and 4 in perfect competition) which is the minimal even number of media outlets for which our model is interesting. In this case for all p,[31]

$$\sum_{k=2}^{4}\binom{4}{k}p^k(1-p)^{4-k} \geq 1 - \frac{4}{3}(1-p)^2 \geq \frac{2p+1}{3}$$

Therefore, perfect competition is always preferred by the receiver.

[31]The result also holds for $n = 3$, that is, when we consider uneven number of signals for the monopolist, and any even number of signals l for the duopoly. In such case the receiver prefers a monopoly over a duopoly iff

$$\frac{np+2}{n+2} - 1 + 2(1-p)^2\frac{l}{l+1} > 0,$$

which holds iff

$$l > \frac{n}{4(1-p) - n(2p-1)},$$

and this holds for a low p, low n, and high enough l.

A.11 Proof of Lemma 3

In order to induce action 1 in state 0 we need the two information structure to produce the same number of signals '1' with some probability. We therefore look at the following information structure:

$$Q^1_{rn} = x \qquad Q^0_{rn} = y$$

$$Q^1_{qn} = 1 - x \quad Q^0_{sn} = 1 - y$$

with $q \neq s$.
Given this information structure, the receiver chooses action 1 when observing qn and action 0 when observing sn. For the receiver to choose 1 upon observing rn we need,

$$\frac{x}{x+y} \geq \frac{1}{2}$$

The marginal constraint for $\omega = 1$ implies:

$$rx + q(1-x) = p$$

The marginal constraint for $\omega = 0$ implies:

$$ry + s(1-y) = 1 - p$$

The utility of the sender from this information structure is: $\frac{1}{2} + \frac{1}{2}y$ and so she is trying to maximize y.
Claim 1: Either $x = y$ or $s = 0$.
Proof: Assume that $x > y$ and that $s > 0$ then we can increase y and decrease s so as to keep constraints satisfied.

Claim 2: Either $q = 0$ or $s = 0$.
Proof: Assume both are strictly positive, increase both x and y so as not to affect $\frac{x}{x+y}$ and lower q and s to maintain constraints.

Suppose $q = 0$. In particular the marginal constraint implies that $x = \frac{p}{r}$ and $r \geq p > 1 - p > s$. Moreover, since $q \neq s$ this implies that $s > 0$ and from above that $y = x = \frac{p}{r}$. But then the marginal constraint in state 0 cannot be satisfied, since

$$r\frac{p}{r} + s(1 - \frac{p}{r}) = p + s(1 - \frac{p}{r}) > p > 1 - p.$$

Therefore $q > 0$ and $s = 0$. This implies that the problem becomes:

$$\max_y y \quad s.t \quad \frac{x}{x+y} \geq \frac{1}{2}, \quad rx + q(1-x) = p, \quad ry = 1 - p.$$

Which is solved for $r = \frac{1}{2} + \frac{1}{n}$, $q = 1$, $x = \frac{n(1-p)}{\frac{n}{2}-1}$ and $y = \frac{n(1-p)}{\frac{n}{2}+1}$.

B Asymmetric marginals, and arbitrary priors.

In this Section we consider general priors and asymmetric marginals and show that qualitatively all the previous results hold. Denote by μ the probability of the state $\omega = 1$, and denote by p and q the marginal probability of a signal 1 given the realization of the state:

$$\Pr(s = 1|\omega = 1) = p \geq q = Prob(s = 1|\omega = 0).$$

The receiver's response is then given by :

$$V(k, p, q)) = \frac{\mu p^k (1-p)^{n-k}}{\mu p^k (1-p)^{n-k} + (1-\mu) q^k (1-q)^{n-k}}$$

When it does not lead to confusion we will denote the receiver's response by $V(k)$ omitting the reference to p, and q.

Lemma 7. *$V(k)$ is convex in k for $k < \bar{k}$ and concave for $k > \bar{k}$*

Proof. Taking the first and second derivatives with respect to k:

$$\frac{\partial V(k)}{\partial k} = \frac{\mu(1-\mu)p^k(1-p)^{n-k}q^k(1-q)^{n-k}(ln(\frac{p}{1-p})-ln(\frac{q}{1-q}))}{(\mu p^k(1-p)^{n-k}+(1-\mu)q^k(1-q)^{n-k})^2}$$

$$\frac{\partial^2 V(k)}{\partial k^2} = \frac{\mu(1-\mu)p^k(1-p)^{n-k}q^k(1-q)^{n-k}(ln(\frac{p}{1-p})-ln(\frac{q}{1-q}))^2}{(\mu p^k(1-p)^{n-k}+(1-\mu)q^k(1-q)^{n-k})^3}((1-\mu)q^k(1-q)^{n-k} - \mu p^k(1-p)^{n-k})$$

So $\frac{\partial^2 V(k)}{\partial k^2} < 0$ if and only if $\mu p^k(1-p)^{n-k} > (1-\mu)q^k(1-q)^{n-k}$, if and only if $V(k) > \frac{1}{2}$. Since V is increasing in k there is a threshold $\bar{k}$ such that V is convex for $k < \bar{k}$ and concave for $k > \bar{k}$ □

Working with the smooth representation of V (See Appendix A.3), we define x^* as before, i.e., the x that solves:

$$\frac{\partial V}{\partial x}(x^*) = \frac{V(x^*) - V(0)}{x^*} \tag{5}$$

Lemma 8. *The optimal information structure given the marginals p and q is given by:*

If $p > q \geq x^*$:

$$\begin{cases} Q^1_p = 1 & \\ Q^1_j = 0 & j \neq p \end{cases} \qquad \begin{cases} Q^0_q = 1 & \\ Q^0_j = 0 & j \notin \{nx^*, 0\} \end{cases}$$

If $p \geq x^* > q$:

$$\begin{cases} Q^1_p = 1 & \\ Q^1_j = 0 & j \neq p \end{cases} \qquad \begin{cases} Q^0_{x^*} = \frac{q}{x^*} & \\ Q^0_0 = 1 - \frac{q}{x^*} & \\ Q^0_j = 0 & j \notin \{x^*, 0\} \end{cases}$$

If $x^* > p > q$:

$$\begin{cases} Q^1_{x^*} = \frac{p}{x^*} & \\ Q^1_0 = 1 - \frac{p}{x^*} & \\ Q^1_j = 0 & j \notin \{x^*, 0\} \end{cases} \qquad \begin{cases} Q^0_{x^*} = \frac{q}{x^*} & \\ Q^0_0 = 1 - \frac{q}{x^*} & \\ Q^0_j = 0 & j \notin \{x^*, 0\} \end{cases}$$

And the sender's expected utility is

$$W_S(p,q) = \begin{cases} \mu V_p + (1-\mu)V_q & if \quad p > q > x^* \\ \mu V_p + (1-\mu)[V_0 + \frac{q}{x^*}(V_{x^*} - V_0)] & if \quad p > x^* > q \\ V_0 + [\mu\frac{p}{x^*} + (1-\mu)\frac{q}{x^*}](V_{x^*} - V_0) & if \quad x^* > p > q \end{cases}$$

Proof. The proof replicates the proof of Proposition 1 in which we replace $\frac{n}{2}$ by the concavity threshold $\bar{k}$ that solves $\mu(\frac{p}{q})^{\bar{k}} = (1-\mu)(\frac{1-q}{1-p})^{n-\bar{k}}$. □

References

Alonso, R. and Câmara, O. (2016). Persuading Voters, *American Economic Review* (forthcoming).

Anderson, S. P. and McLaren, J. (2012). Media Mergers and Media Bias with Rational Consumers, *Journal of the European Economic Association* **10**(4): 831–859.

Arieli, I. and Babichenko, Y. (2016). Private Bayesian Persuasion, *Mimeo* .

Balan, D., DeGrava, P. and Wickelgren, A. L. (2009). Ideological Persuasion in the Media, *Mimeo* .

Bardhi, A. and Guo, Y. (2016). Modes of Persuasion Toward Unanimous Consent, *Mimeo* .

Baron, D. P. (2006). Persistent Media Bias, *Journal of Public Economics* **90**: 1–36.

Besley, T. and Prat, A. (2006). Handcuffs for the Grabbing Hand? The Role of the Media in Political Accountability, *American Economic Review* **96**(3): 720–736.

Boleslavsky, R. and Cotton, C. (2014). Limited Capacity in Project Selection: Competition through Evidence Production, *Economic Theory* pp. 1–37.

Boxell, L., Gentzkow, M. and Shapiro, J. M. (2017). Is the internet causing political polarization? Evidence from demographics, *Mimeo* .

Brocas, I. and Carrillo, J. D. (2007). Influence through Ignorance, *The RAND Journal of Economics* **38**(4): 931–947.

Cagé, J., Hervé, N. and Viaud, M.-L. (2017). The Production of Information in an Online World: Is Copy Right?, *CEPR Dicussion Paper 12066* .

Chan, J., Gupta, S., Li, F. and Wang, Y. (2016). Pivotal Persuasion, *Mimeo* .

De Marzo, P., Vayanos, D. and Zwiebel, J. (2003). Persuasion Bias, Social Influence and Unidimensional Opinions, *The Quaterly Journal of Economics* **118**(3): 909–968.

Della Vigna, S. and Kaplan, E. (2007). The Fox News Effect: Media Bias and Voting, *The Quaterly Journal of Economics* **122**(3): 1187–1234.

Duggan, J. and Martinelli, C. (2011). A Spatial Theory of Media Slant and Voter Choice, *Review of Economic Studies* **78**: 640–666.

Durante, R. and Knight, B. (2012). Partisan Control, Media Bias, and Viewer Responses: Evidence from Berlusconi's Italy, *Journal of the European Economic Association* **10**: 451–481.

Ellis, A. and Piccione, M. (2017). Correlation Misperception in Choice, *American Economic Review* **107**(4): 1264–92.

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

Enke, B. and Zimmermann, F. (2013). Correlation Neglect in Belief Formation, (Mimeo).

Eyster, E. and Weizsaker, G. (2011). Correlation Neglect in Financial Decision-Making, (Mimeo).

Gagnon-Bartsch, T. and Rabin, M. (2015). Naive Social Learning, Mislearning, and Unlearning, (Mimeo).

Galperti, S. and Trevino, I. (2017). Coordination Motives and Competition for Attention in Information Markets, (Mimeo).

Gentzkow, M. and Kamenica, E. (2017). Competition in Persuasion, *Review of Economic Studies* **84**(1): 300–322.

Gentzkow, M. and Shapiro, J. (2010). What Drives Media Slant? Evidence from U.S. Daily Newspapers, *Econometrica* **78**: 35–71.

Gentzkow, M. and Shapiro, J. (2011). Ideological Segregation Online and Offline, *The Quarterly Journal of Economics* **126**: 1799–1839.

Gentzkow, M., Shapiro, J. and Sinkinson, M. (2011). The Effect of Newspaper Entry and Exit on Electoral Politics, *American Economic Review* **7**(101).

Gentzkow, M., Shapiro, J. and Sinkinson, M. (2014). Competition and Ideological Diversity: Historical Evidence from US Newspapers, *American Economic Review* **104**(10).

Glaeser, E. L. and Sunstein, C. R. (2009). Extremism and Solcial Learning, *Journal of Legal Analysis* **1**(1).

Golub, B. and Jackson, M. O. (2012). How Homophily Affects the Speed of Learning and Best-Response Dynamics, *The Quarterly Journal of Economics* **127**(3): 35–71.

Groseclose, T. and Milyo, J. (2005). A Measure of Media Bias, *Mimeo* .

Harbaugh, R., Maxwell, J. and Shue, K. (2017). Consistent Good News and Inconsistent Bad News, *Mimeo* .

Kallir, I. and Sonsino, D. (2009). The Neglect of Correlation in Allocation Decisions, *Southern Economic Journal* **75**(4): 1045–1066.

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

Kennedy, P. and Prat, A. (2017). Where Do People Get Their News?, *Mimeo* .

Levy, G. and Razin, R. (2015a). Correlation Neglect, Voting Behaviour and Information Aggregation, *American Economic Review* .

Levy, G. and Razin, R. (2015b). Does Polarization of Opinions Lead to Polarization of Platforms? The Case of Correlation Neglect, *Quarterly Journal of Political Science* **10**(3): 321–355.

Levy, G. and Razin, R. (2017). Combining Forecasts: Why Decision Makers Neglect Correlation, *Working Paper, LSE* .

Meyer, M. (2017). Information Design: Insights from Orderings of Dependence and Heterogeneity, *Mimeo* .

Mullainathan, S. and Shleifer, A. (2005). The Market for News, *American Economic Review* **95**: 1031–1053.

Ortoleva, P. and Snowberg, E. (2015). Overconfidence in Political Behavior, *American Economic Review* **105**(2): 504–35.

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

Prat, A. (2017). Media Power, *Journal of Political Economy* (Forthcoming).

Puglisi, R. (2011). Being The New York Times: the Political Behaviour of a Newspaper, *The B.E. Journal of Economic Analysis & Policy* **11**(1): 1–34.

Sunstein, C. R. (2017). *#Republic: Divided Democracy in the Age of Social Media*, Princeton University Press.

Wang, Y. (2015). Bayesian Persuasion with Multiple Receivers, (Mimeo).