

# Media Power

Andrea Prat  
Columbia University\*

January 27, 2017

## Abstract

This paper defines the power of a media organization as its ability to induce voters to make electoral decisions they would not make if reporting were unbiased. It derives a robust upper bound to media power over a range of assumptions about the beliefs and attention patterns of voters. The paper then presents estimates of the power index for the US and shows how these can inform merger analysis and other policy debates.

## 1 Introduction

The media industry is different from other industries partly because it has an *indirect* effect on welfare through information externalities imposed on the policy process. Concentration may be damaging not only because it has a direct effect on prices and quantities but also because media owners may be able to manipulate democratic decision-making. The indirect effect looms large in the public debate (Leveson 2012).

Standard approaches to measuring the direct effects of market concentration are inappropriate for measuring indirect effects (Polo 2005, Noam 2009, Ofcom 2009). First, while direct competition mostly occurs between outlets on the same platform (e.g. radio stations), indirect effects encompass all platforms, because outlets in different platforms may be used by the same people and ultimately affect the same policy process. Second, an outlet's economic

---

\*This project developed from discussions with Mark Armstrong. I am greatly indebted to him. I am also thankful to Tim Bresnahan, Filipe Campante, Sylvain Chassang, Stefano DellaVigna, Ruben Enikopolov, Matt Gentzkow, Lisa George, Yosh Halberstam, Kei Kawai, Brian Knight, Leslie Marx, Eli Noam, Maria Petrova, Michele Polo, Riccardo Puglisi, Daniel Rappoport, Barbara Veronese, Simon Wilkie, Ali Yurukoglu, and seminar participants at Bocconi, CERGE-EI, the CIRPEE Conference in Political Economy, Columbia, The Federal Communications Commission, Harvard, Hong Kong University, the London School of Economics, the 2015 Media and Communications Conference (Chicago), NBER, the New York City Media Seminar, the 2013 North American Winter Meetings of the Econometric Society, the Princeton Political Economy Conference, and Rice University for helpful suggestions. Eric Hardy and Patrick Kennedy provided outstanding research assistance.

importance – its total share of audience or advertising revenue, say – need not correspond to its ability to affect policy.<sup>1</sup>

This paper suggests a theoretical approach to measure the indirect effects of media concentration. It takes as its unit of analysis not a particular media market but the mind of an individual voter. Once we determine the influence that a news source has on each voter, we can derive its overall influence by aggregating across voters, thus determining the vote share it controls.

Rather than trying to specify the influence of each source exactly, the paper takes a “robust bounds” approach (Chassang 2013, Madarasz and Prat 2016, Carroll 2014, and Chassang and Padro i Miquel 2013), characterizing the upper bound on the influence of a given media owner under a range of assumptions. This upper bound is called the owner’s “power”.

The starting point of the analysis is a media consumption matrix, which describes the news sources individual voters currently follow. Each voter can follow multiple sources belonging to different platforms. Voters are Bayesian and they use the information they receive from the media sources they follow to decide how to vote. Voters have subjective prior beliefs on the probability that media are captured. They also have a potentially bounded capacity to absorb information (*bandwidth*): they only observe or remember a certain number of news items from the various sources they follow. The relative quality of political candidates is stochastic and the media receive a large number of signals correlated with candidate quality.

As the goal is to characterize maximal influence, the analysis focuses on a media owner who is assumed to have a pure political motive: she wants a particular candidate to win this election, and she has no concerns for the short- and long-term commercial return or the journalistic reputation of the media companies she owns. In line with the worst-case spirit of the analysis, it is assumed that voters do not switch away from media sources that become biased.

The main technical result (Proposition 2) states that the worst-case scenario can be expressed as the solution of a polynomial equation. As one would expect, in the worst case scenario the media organization faces voters who believe they are not being manipulated.

The role of attention patterns is more subtle. Assuming all voters have the same bandwidth, the worst case is not necessarily the one in which bandwidth is minimal. However, the main substantive result of the paper (Proposition 3) shows that, if no outlet controls more than 50% of voter attention, maximal power is indeed achieved when voters have minimal bandwidth and the power of a media organization  $G$  is proportional to:

$$\frac{a_G}{1 - a_G},$$

where  $a_G$  is the share of the average voter’s news diet accounted for by outlets owned by  $G$ .

---

<sup>1</sup>A 2003 attempt by the US Federal Communications Commission to introduce a simple weighted cross-platform measure, the Media Diversity Index, was struck down by an appellate court in *Prometheus Radio Project v. FCC* because of “irrational assumptions and inconsistencies.” For the current regulatory situation in the United States, see Federal Communications Commission (2016).

Although in the baseline case voters differ only in terms of media consumption, the model can be extended to include an ideological dimension. The paper also reports or discusses a number of other extensions, including the effect of abstention.

The empirical part of the paper reports two sets of results: (1) The power of media owners in the US in 2012; and (2) A calibration exercise that uses existing estimates of media influence.

According to the index, the three most powerful organizations in 2012 were, in order of decreasing power, News Corporation (Fox News, *Wall Street Journal*), Comcast (NBC, MSNBC), and Time Warner (CNN, *Time Magazine*). The most powerful newspaper, the *New York Times*, was in tenth position behind the most powerful pure-internet source, Yahoo News, in sixth position. In practice this ranking based on power is similar (but not identical) to one based on reach, defined as the fraction of voters reached by an owner's outlets over a given time period. The ranking is quite different from one based on a concentration-based measure of media power proposed by Noam (2009).

The robustness of the index is probed along various directions: different criteria for inclusions of news sources (daily or weekly), different definitions of the index (worst case and minimal attention), different years (2010 and 2012, when all major media are included), and different assumptions about the distribution of voter ideology and the probability of abstention. While the absolute values of the power index vary with the specification chosen, the relative ranking of media organizations is quite stable.

The calibration exercise consists in utilizing available estimates of media power to restrict the set of models under consideration. The paper illustrates this approach with estimates from DellaVigna and Kaplan (2007) and Martin and Yurukoglu (2016) for broadcast media, and Gentzkow, Shapiro and Sinkinson (2011) and Chiang and Knight (2011) for the press. Two scenarios are considered, where combinations of the estimates in the cited papers represent upper bounds to media influence.

Finally, the power index is used to evaluate the effect of proposed or hypothesized mergers between a number of US media companies. The value of the index is extremely high for some of the television-based merged entities.

The present paper relates to a large and growing body of empirical research on media economics, and in particular on the influence of media on the democratic system (Anderson, Strömberg, and Waldfogel, 2016). Evidence about the effect of media bias on electoral outcomes is mixed.<sup>2</sup> Evidence on the motivations of media owners is mixed too. Durante and Knight (2006) document sudden and 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 theory side, owners may manipulate news for their own goals (Baron 2006, Besley

---

<sup>2</sup>DellaVigna and Kaplan (2007) and Martin and Yurukoglu (2016) find a significant effect of Fox News on US voting patterns, and Enikolopov, Petrova and Zhuravskaya (2011) find an even stronger effect of the entry of NTV into selected Russian regions. However, Gentzkow, Shapiro and Sinkinson (2011) rule out even moderate effects of entry and exit of partisan newspapers on party vote shares in the United States from 1869 to 2004. Moreover, there is also evidence that US newspaper readers show some sophistication in the way they handle media bias (Durante and Knight 2006, Chiang and Knight 2011).

and Prat 2006, Balan, De-Graba, and Wickelgren 2009, Duggan and Martinelli 2011, Anderson and McLaren 2012, Petrova 2012). Anderson and McLaren (2012) compare a media duopoly to a media monopoly, in the presence of politically motivated media owners, and analyze the effect of a merger. Brocas et al. (2010) characterize the effect of competition and ownership on diversity of viewpoint and informational efficiency. The main contribution of the present paper relative to past work is to provide a measure of the power of a given owner that is informative under a range of assumptions about the technology of voter persuasion.

Section 2 describes the model, reports the two main propositions, and discusses limitations and extensions of the theoretical analysis. Section 3 contains the empirical analysis, including various specifications of the power index computed with US data, the calibration exercise, and the merger simulation. Section 4 concludes.

## 2 Theory

The theory part of the paper is divided into three sections. The first section introduces the model and analyzes a benchmark case with no media bias (Proposition 1). The following section studies the more interesting situation with biased media and reports the two main results of the paper: Propositions 2 and 3. The last section contains a discussion of limitations and extensions.

### 2.1 Setup and Preliminaries

Let us begin by describing the model when all news sources are unbiased. There are two candidates,  $A$  and  $B$ . The relative quality of candidate  $B$  over candidate  $A$  is a random variable  $\sigma$ , distributed according to density function  $f$  with support  $[0, 1]$ . The function  $f$  is symmetric around  $\frac{1}{2}$  ( $f(\sigma) = f(1 - \sigma)$ ) and unimodal. There is a mass one of voters with homogeneous preferences. In expectation, the two candidates are equally attractive, but given  $\sigma$  voters prefer candidate  $B$  if and only if  $\sigma \geq \frac{1}{2}$ . Specifically, voters' payoff is  $\frac{1}{2}$  if they elect  $A$  and  $\sigma$  if they elect  $B$ . In Section 2.3 we discuss how the model can be extended to ideological voters.

However, voters do not observe the relative quality  $\sigma$  directly. They rely on the media for information. There is a set of news sources, which do not observe  $\sigma$  directly either but receive binary signals drawn from a binomial distribution with mean  $\sigma$ . Let  $\mathbb{M}$  denote the finite set of sources, with typical individual outlet denoted with  $m \in \mathbb{M}$ . Let  $x_m = (x_{m1}, \dots, x_{mN})$  denote a vector of  $N$  binary signals – *news items* – observed by source  $m$ , with  $\Pr(x_{mi} = 1) = \sigma$ . News items are, conditional on  $\sigma$ , independent within and across sources.

Voters may follow more than one news source. Let  $M \subset \mathbb{M}$  denote some subset of sources. Then voters are partitioned into *segments*, indexed by the subset  $M$  of sources they consume, and for each  $M \subset \mathbb{M}$  let  $q_M$  be the fraction of voters who consume (exactly) the subset  $M$ . Clearly,

$$\sum_{M \subset \mathbb{M}} q_M = 1 .$$

Segment	Share	Tv1	Tv2	Np1	Np2	Np3	Web1	Web2
1	20%	■						
2	10%	■		■				
3	10%			■				
4	10%		■					■
5	10%		■		■		■	■
6	10%		■		■		■	
7	10%		■			■	■	
8	10%					■	■	■
9	10%				■	■	■	■
Reach		30%	40%	20%	30%	30%	50%	40%
Attention		25%	14.1%	15%	8.3%	9.1%	15%	13.3%
$\Pi$		0.333	0.164	0.176	0.090	0.101	0.176	0.152

Table 1: Example of a media consumption matrix with seven media sources: two television networks (Tv1, Tv2), three newspapers (Np1, Np2, Np3), and two online sources (Web1, Web 2). Voters are divided into nine segments depending on the sources they follow. In each segment voters devote equal attention to each source indicated by a box. The *reach* of a source is the percentage of voters who follow that source. The *attention share* of a source is the weighted average of that source’s attention share in each segment. The variable  $\Pi$  is the power index of the correspondent source as defined in Proposition 3.

For simplicity, suppose that all voters see at least one outlet, so that  $q_\emptyset = 0$ . This makes no difference provided that voters who receive no messages vote randomly.

Table 1 contains a media consumption matrix. Voters belong to nine possible segments. There are seven news sources: two television channels, three newspapers, and two news websites. A solid square in a cell indicates that voters in the corresponding row follow the news source in the corresponding column. The table reports two possible measures of media penetration: the reach (the total share of voters who follow that source) and the attention share. The latter is defined as follows: for each segment, let  $m$ ’s attention share be zero if voters in that segment do not follow  $m$  and  $1/|M|$  if they do (where  $|M|$  is the number of outlets followed in that segment);  $m$ ’s aggregate attention share is the weighted average of  $m$ ’s attention share in each segment.

Unbiased media simply report all the  $N$  signals they receive. Thus media source  $m$  reports  $N$  binary numbers. A voter in group  $M$  is exposed to  $|M| \times N$  signals.

However, voters have potentially limited *bandwidth*. They observe or remember a limited number  $K \in \{1, \dots, N\}$  of news items, randomly selected among the set of  $|M| \times N$  items available. Bandwidth does not affect the electoral outcome in the unbiased case, but will be crucial once media can be biased.

After observing  $K$  binary news items, the voter computes their average  $s^i$ . Voter  $i$  prefers  $B$  if  $E[\sigma] \geq \frac{1}{2}$ . Under sincere voting, he casts his ballot for  $B$  if and only if  $s^i \geq \frac{1}{2}$ .

What is the probability that voter  $i$  in  $M$  votes for  $B$ ? Let  $s_m$  be the average of the  $N$

signals received by outlet  $m$ . With  $N \rightarrow \infty$ , we have  $s_m \rightarrow \sigma$  for all media  $m$ . In that case, the probability that voter  $i$  in  $M$  votes for  $B$  is equal to the probability that the sample mean of a binomial random variable with  $K$  realizations and mean  $\sigma$  is at least  $1/2$ . By the law of large numbers this probability is also the vote share within segment  $M$ . Thus the vote share in segment  $M$  is at least  $1/2$  if and only if  $\sigma$  is at least  $1/2$ . As this holds for every segment, we have verified that:

**Proposition 1** *With unbiased media, as  $N \rightarrow \infty$ ,  $B$  is elected if and only if  $\sigma \geq \frac{1}{2}$ .*

While the identity of the winning candidate in Proposition 1 is unaffected by assumptions on voter bandwidth, the margin of victory is affected by bandwidth – a fact that will play a crucial role in the next section. To illustrate this point, Figure 1 depicts the vote share of Candidate  $A$ ,  $p_A$ , as a function of candidate quality differential  $\sigma$  for four possible bandwidth values, from the smallest:  $K = 1$  to the limit as  $K \rightarrow \infty$ .

As one expects, the vote share (of  $A$ ) is decreasing in the quality (of  $B$ ). Independent of bandwidth,  $p_A$  is exactly  $1/2$  when  $\sigma = 1/2$ , as predicted in Proposition 1. However, bandwidth determines the slope of the vote share function. The probability that a voter chooses the wrong candidate, say  $A$  when  $\sigma > \frac{1}{2}$ , corresponds to the chance that he observes/recalls a higher number of signals favorable to  $A$  than to  $B$ . That probability decreases with  $K$  and in the limit it goes to zero. This explains why the vote share function becomes increasingly S-shaped as bandwidth increases.

## 2.2 The Media Power Index

Let us now entertain the possibility that an agent acquires control of a subset  $G$  of the set of active media  $\mathbb{M}$ . In the worst-scenario spirit, this agent – henceforth known as the “evil media owner” – has one goal only: she wishes to see candidate  $A$  elected, independent of the relative quality of the two candidates.

The notion of attention share introduced in the previous section extends to sets of media sources. The *attention share* of media group  $G$  in segment  $M$  is defined as

$$g_M = \frac{|M \cap G|}{|M|}.$$

and  $M$ 's overall attention share is:

$$a_G = \sum_M q_M g_M.$$

In the pessimistic view of the world that we must adopt to compute the upper bound of media influence, the evil owner faces no constraint to selective reporting. In particular, she can fail to report any or all the items that are favorable to  $B$ . Recall that each outlet receives an unboundedly large number of news items  $N$ . Hence, for any  $\sigma \in (0, 1)$ , the evil

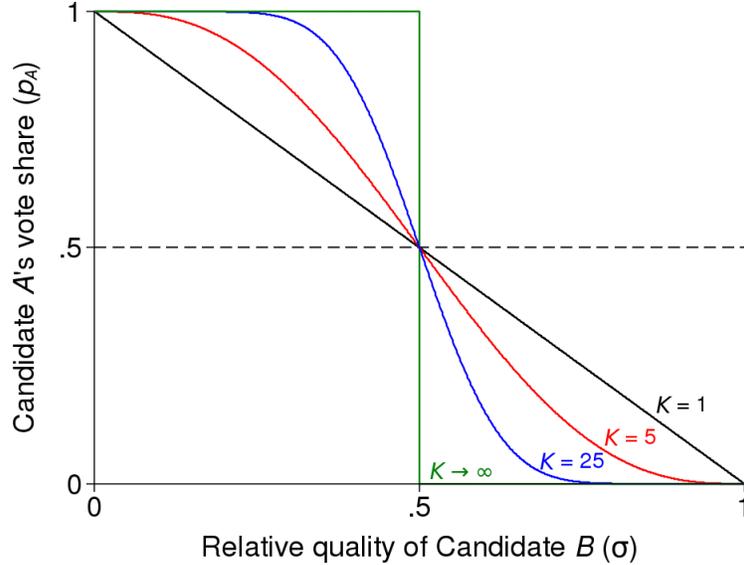


Figure 1: Candidate  $A$ 's vote share with unbiased reporting. Each of the four functions represents the relation between candidate quality and vote share for a different bandwidth value  $K$ . With unbiased reporting, all vote share functions intersect at  $(\frac{1}{2}, \frac{1}{2})$ , implying that, for any bandwidth level, Candidate  $A$ 's vote share is greater than 50% if and only if  $\sigma < \frac{1}{2}$ .

owner can find at least  $K$  items favorable to  $A$ . This means she can choose to report any share  $s \in [0, 1]$  of news items that are favorable to  $B$ . In one extreme case  $s = 1$  and all signals are reported, as in the unbiased case. In the other extreme  $s = 0$  and only signals that are favorable to  $A$  are reported.

A voter with bandwidth  $K$  observes/recalls  $K$  of the items that the biased media outlet reports. In the worst-case scenario, the voter does not see how many items the outlet actually reported.<sup>3</sup> If he did, he could try to deduce the presence of bias directly.

How do voters react to the possibility that the owner of  $G$  is evil? Let  $\beta \in (0, 1)$  be the prior probability that voters assign to the owner being evil. This parameter captures the voters' belief that  $G$  is under the effective control of a unitary owner and that such owner is biased in favor of candidate  $A$ . The parameter  $\beta$  should be viewed as a subjective belief rather than the objective probability that the owner is biased. Given prior probability  $\beta$ , voters are fully rational and they utilize Bayesian updating. In particular they update their belief  $\beta$  on the basis of the news stories they receive from their news sources.

We are now ready to analyze the model, with the objective of finding the upper bound to the electoral influence of media organization  $G$ . Recall that a voter with bandwidth  $K$  in group  $M$  receives/remembers a  $K$ -sized vector of signal realizations randomly drawn from the media outlets in group  $M$ . As before, the number of signals that come from a particular

<sup>3</sup>This assumption represents a deviation from the Bayesian persuasion set-up (Kamenica and Gentzkow (2011)).

outlet is random and the selection of signals within an outlet is random too. Now, however, the voter faces a more complex Bayesian updating process.

To analyze this, we begin by writing the probability that a voter  $i$  in group  $M$  observes a particular realization of the  $K$ -sized signal vector  $y^i$  he receives from media sources in  $M$ . The vector includes news items randomly drawn from outlets in  $M$ . This probability is computed according to the beliefs of the voter. Suppose the voter believes that the owner is evil with probability  $\beta$  and that an evil owner would use reporting strategy  $\hat{s}$ . Then, the probability of realization  $y^i = Y$  would be given by  $\Pr(y^i = Y|\sigma, \hat{s})$  and the voter would compute the expected value of candidate quality as follows:

$$E[\sigma|Y, \hat{s}] = \frac{\int_0^1 \Pr(y^i = Y|\sigma, \hat{s}) \sigma f(\sigma) d\sigma}{\int_0^1 \Pr(y^i = Y|\sigma, \hat{s}) f(\sigma) d\sigma}$$

This is standard Bayesian updating and it implicitly includes a revised probability  $\hat{\beta}$  that the owner of  $G$  is evil. The voter would vote for  $A$  if and only if  $E[\sigma|Y, \hat{s}] \leq \frac{1}{2}$ . Candidate  $A$  would be elected if at least half of the voters vote for her.

For every belief  $\hat{s}$ , this defines a maximization problem for the evil owner. How does she maximize the chance that  $A$  is elected given  $A$ 's true quality, given the media consumption matrix, and given voters' conjecture about her reporting strategy  $\hat{s}$ ? An equilibrium is a fixed point where voters' beliefs are consistent with the optimal reporting strategy. For most values of  $\beta$  this fixed point is difficult to compute. To see this, note that an evil owner will typically not want to choose news items that are "too biased" because it would increase suspicion from the voters and raise  $\hat{\beta}$  (see Dziuda 2011).

In what follows, we wish to identify the maximal effect, for any value of  $\beta$ , that an evil owner of any media group  $G$  can have on an electoral outcome. This question is equivalent to determining the highest  $\sigma$  such that  $A$  is elected. We call that maximal quality, which can depend on voter bandwidth  $K$ ,  $\bar{\sigma}(K)$ . The group  $G$  has no power if  $\bar{\sigma}(K) = \frac{1}{2}$ : the evil owner can only help to elect candidates that would be elected anyway under unbiased media. The group has maximal power when  $\bar{\sigma}(K) = 1$ :  $G$  always gets  $A$  elected.

In our setting, the lower bound to the quality of the worst-electable  $A$ -candidate corresponds to an upper bound to the welfare loss that the evil owner can inflict on the electorate. The utility voters get under unbiased media is simply  $\max(\frac{1}{2}, \sigma)$ . With an evil  $G$  owner, the utility can be as low as  $1 - \bar{\sigma}(K)$ . The maximal utility loss that voters can experience is when they receive  $1 - \bar{\sigma}(K)$  instead of  $\bar{\sigma}(K)$ . This leads to a natural definition for the index. The power index of group  $G$ , for a given bandwidth vector  $K$ , is

$$\Pi(K) = 2\bar{\sigma}(K) - 1$$

Thus,  $\Pi(K) \in [0, 1]$  with 0 denoting no power ( $G$  has no influence on elections) and 1 denoting absolute power ( $G$  controls all elections).

We proceed to characterize  $\Pi(K)$  for any  $G$  and any  $K$ :

**Proposition 2** *The upper bound to media power is reached when voter sophistication is lowest:  $\beta \rightarrow 0$ . For a given bandwidth  $K$ , the power of group  $G$  is  $\Pi(K) = 2\bar{\sigma}(K) - 1$ , where  $\bar{\sigma}(K)$  is the minimum between one and the smallest solution greater than  $1/2$  of :*

$$\sum_{M \subset \mathbb{M}} q_M p_A(g_M, K, \sigma) = \frac{1}{2},$$

where

$$p_A(g_M, K, \sigma) = \sum_{k=0}^{\lceil K/2 \rceil - 1} \binom{K}{k} ((1 - g_M)\sigma)^k (1 - (1 - g_M)\sigma)^{K-k}.$$

The proof of the Proposition, reported in the Appendix, proceeds in three steps. First, we identify a mathematical lower bound to  $E[\sigma|Y, \hat{s}]$  for any possible value of the belief  $\beta$ . The value of  $E[\sigma|Y, \hat{s}]$  can never be lower than the value achieved when all the biased outlets' news items are favorable to  $A$  and the voter believes that all media are unbiased. This bound is achieved when  $\beta \rightarrow 0$  and the evil owner reports only signals that are favorable to candidate  $A$ . Second, for each segment  $M$ , we show that this lower bound on the posterior of  $\sigma$  translates into an upper bound on the vote share of candidate  $A$  in that segment:  $p_A(g_M, K, \sigma)$ . Finally, we aggregate bounds on vote shares in individual segments to determine an upper bound on the quality  $\sigma$  that guarantees that at least half of the voters choose  $A$ .<sup>4</sup>

The Proposition is best illustrated through an example and a figure. Suppose we wish to compute the media power of Newspaper 2 in Table 1, for various values of  $K$ . Let us re-visit figure 1, which depicted  $A$ 's vote share in a segment with only unbiased media. In Proposition 2's notation, we would express the function depicted in Figure 1 as  $p_A(0, K, \sigma)$ .

Figure 2 draws the same type of plot for all nine segments under the assumption that the owner of Newspaper 2 is evil. Voters in segments 1-4, 7, and 8 do not read Newspaper 2 and they are not affected by its biased reporting: the plot in Figure 2(a) is exactly the same as in Figure 1.

Segments 5 and 9 are more interesting. Voters in those two segments use four media sources, including Newspaper 2. The vote share is given by  $p_A(\frac{1}{4}, K, \sigma)$ , which is depicted in Figure 2(b). Compared to (a),  $A$ 's vote share is still a decreasing function of  $\sigma$ , it is still s-shaped, and its steepness depends on bandwidth  $K$ . However, now the curves have all shifted to the right. Rather than intersecting the  $1/2$  horizontal line at  $\sigma = \frac{1}{2}$  as in the unbiased case, the intersection is now at  $\sigma = \frac{2}{3}$ . Media power is now visible: the evil owner can persuade a majority of voters in segment  $M$  to vote for  $A$  even when  $B$  is a superior candidate.

How are the vote shares in 2(b) obtained? By examining Proposition 2, we see that for each  $K$  the function  $p_A(g_M, K, \sigma)$  is the cumulative distribution function of a binomial distribution with  $K$  trials and parameter  $p = (1 - g_M)\sigma$ , evaluated at  $\lceil K/2 \rceil - 1$ . If  $K$

---

<sup>4</sup>The proposition keeps  $K$  constant and lets  $\beta \rightarrow 0$ . If bandwidth were allowed to increase while voters become more naive, voters might still be able to arrive at a truth with certainty even if some of the sources they receive are biased.

is odd,  $p_A(g_M, K, \sigma)$  is the probability that strictly more than half of the realizations are zeros, where the strict part is due to the assumption that ties are resolved in favor of  $B$ . The probability that an individual news item in segment  $M$  is equal to one is given by  $(1 - g_M)\sigma$ : all media sources in  $M$  always report zero, while all media sources outside  $M$  report 1 with probability  $\sigma$ . For instance, if we let  $K = 5$ ,  $A$ 's vote share in Segments 5 and 9 is given by the probability that at least 3 of the 5 news items that voters recall are zeros, with the probability that each individual item is one equal to  $(1 - g_M)\sigma = \frac{3}{4}\sigma$ . We therefore have

$$p_A\left(\frac{1}{4}, 5, \sigma\right) = \sum_{k=0}^2 \binom{5}{k} \left(\frac{3}{4}\sigma\right)^k \left(1 - \frac{3}{4}\sigma\right)^{5-k},$$

which is the  $A$ -vote share depicted for  $K = 5$  in Figure 2(b).

Finally, in segment 6 Newspaper 2 is one of three outlets that voters follow. This makes it more influential, as illustrated by the rightward shift of the the vote share function when we move from Figure 2(b) to Figure 2(c).

Putting everything together, we see that  $A$ 's overall vote share is given by the weighted sum of her vote shares in the ten segments. Figure 2(d) depicts this expression as a function of  $\sigma$ . The intersection between the function and the  $1/2$  horizontal line identifies the  $\bar{\sigma}$  at which the two candidates get 50% of the votes each. For any lower value of  $\sigma$ , candidate  $A$  is elected. The power index of Newspaper 2 is given by  $\Pi = 2\bar{\sigma} - 1$ .

From Proposition 2, we see that the value of  $\bar{\sigma}(K)$  can be computed as the solution of a polynomial equation of degree  $K$ . For Newspaper 2, this yields:  $\bar{\sigma}(25) = 0.5398$ ,  $\bar{\sigma}(5) = 0.5445$ , and  $\bar{\sigma}(1) = 0.5454$ . We also see that  $\lim_{K \rightarrow \infty} \bar{\sigma}(K) = 0.5$ . Therefore, the values of the power index for Newspaper 2 are

$$\Pi(1) = 0.091, \Pi(5) = 0.089, \Pi(25) = 0.079, \Pi(\infty) = 0.$$

The media power index has the interpretation introduced above. For instance, the 8.9% power index for  $K = 5$  indicates that the maximal welfare loss that voters can experience – expressed as the percentage of the difference between a perfect candidate ( $\sigma = 1$ ) and the worst possible candidate ( $\sigma = 0$ ) – as a result of Newspaper 2's biased reporting is 8.9%.

The media power index in Proposition 2 depends on voter bandwidth  $K$ . This parameter may be difficult to quantify. It may also be unstable across time with voters paying more or less attention to domestic politics. We may therefore want to find an upper bound to the upper bound, namely a  $K$ -independent upper limit to the power of a media group.

The example of Newspaper 2 discussed above might lead one to conjecture that power is decreasing in  $K$  and it is always maximized when  $K = 1$ . Unfortunately, one can construct examples where the value of the power index is greatest for  $K \rightarrow \infty$  and examples where the index is maximized for some finite  $K > 1$ .

However, there is a simple condition that guarantees that the power index is maximized when  $K = 1$ :

**Proposition 3** *If the share of voters who follow at least one of the sources owned by  $G$  is not greater than 50%, then  $\bar{\sigma}(K) \leq \bar{\sigma}(1)$  for all  $K$ , and an upper bound to the power of group  $M$  is given by*

$$\Pi(1) = \min\left(1, \frac{a_G}{1 - a_G}\right).$$

The broad intuition behind Proposition 3 comes from the fact that a decrease in bandwidth  $K$  corresponds to an increase in the probability that individual voters make random mistakes. This creates noise in the ballot choices of voters who follow unbiased sources. Noise is beneficial to the evil owner if she controls less than 50% of the attention share.

The result is immediate to see in the special case where voters are neatly split into a segment that follows only sources belonging to  $G$  and another segment with only unbiased sources. In the first segment all votes go to  $A$ . In the second segment, for any  $\sigma > \frac{1}{2}$ , the share of voters who choose  $A$  is decreasing in  $K$ , because a greater bandwidth reduces the probability that a voter observes a majority of items in favor of the ‘wrong’ candidate. Hence,  $A$ ’s overall vote share is decreasing in  $K$ . If the reach of  $G$  is smaller than 50%, the decreasing vote share translates into a lower power index  $K$  (if the reach is greater than 50%, then in this simple case  $A$  is always elected).

For the general case, the proof requires a number of steps because unfortunately even under the 50% condition it is not true that  $A$ ’s vote share given  $\sigma$  is monotonic in  $K$ . This can be seen in Figure 2(d), where for certain values of  $K$ , the vote share is smaller for  $K = 25$  than for  $K \rightarrow \infty$ .

The proof instead proceeds by showing that, if the 50% condition is satisfied, the following condition (which says that if  $K = 1$  and  $\sigma = \bar{\sigma}(1)$   $A$ ’s vote share is one half),

$$\sum_{M \subset \mathbb{M}} q_{MPA}(g_M, 1, \bar{\sigma}(1)) = \frac{1}{2},$$

implies that, for any bandwidth, if  $\sigma = \bar{\sigma}(1)$ ,  $A$ ’s vote share cannot be more than one half:

$$\sum_{M \subset \mathbb{M}} q_{MPA}(g_M, K, \bar{\sigma}(1)) \leq \frac{1}{2} \text{ for all } K. \quad (1)$$

As discussed above, the left-hand side of (1) is the value of the cumulative distribution function of a binomial distribution. The proof relies on properties relating to the s-shape of the binomial distribution. Intuitively, a smaller bandwidth creates uncertainty as to how individual voters would vote, which enhances the ability of a media group to manipulate – on average – voting probabilities. As  $K$  increases, this causes the vote share to fall below one half.

## 2.3 Discussion

The set-up we have just analyzed contains a number of assumptions made in a worst-case spirit: the owner is purely partisan; all sources outside  $G$  are unbiased; the owner faces no

restrictions in her ability to select news; voters do not observe the number  $N$  of news items; and voters do not react to the risk of bias by changing the set of sources they follow. This implies that the media consumption matrix can be taken as stable, a crucial advantage in the empirical analysis.

We look at the influence of one individual owner over electoral outcomes. One could imagine that this owner is acting in explicit or tacit concert with other like-minded owners (*the coalition case*) or is acting against other owners who are trying to influence electoral outcomes in the opposite direction (*the opposition case*). Both cases could be analyzed within our approach. In the opposition case, the damage the owner produces is smaller than in our analysis. In fact, the addition of, say, a right-wing biased source in a world of left-leaning sources may improve welfare (this is simply because the other biased source re-empowers voters that are not affected by it). In line with our bounds approach, we therefore disregard the opposition case: it is just one possible reason why the upper bound is not reached. The coalition case is instead more relevant. The analysis can easily be extended to encompass sets of like-minded media owners: we simply define the set of biased media sources as those owned by owners in the set.<sup>5</sup> However, of course, defining sets of independent but like-minded media owners is a difficult and subjective task. Instead, ownership is an objective criterion that is already used by regulatory agencies. So, the present paper focuses exclusively on the latter.

Four additional assumptions have been made for analytical convenience: (a) All voters have the same bandwidth  $K$ ; (b) Voters have homogeneous preferences; (c) Voters devote equal attention to all the sources they follow; (d) There is no abstention. These assumptions are not essential to the methodology. All of them can be removed and power indices can be computed through a similar approach. In the rest of this section, we discuss how we can extend the model to get rid of (a), (b), and (c). The possibility of abstention is explored in the empirical analysis.

## Heterogeneous Bandwidth

One could conceive of even more pessimistic worst-case scenarios. One could allow the bandwidth to be different in different segments: namely  $K_M$  in segment  $M$  could be different from  $K_{M'}$  in segment  $M'$ . We could then determine the maximal power of media organization  $G$  under any possible vector of bandwidths  $\mathbf{K}$ . Intuitively, this will be achieved when the bandwidth is high in segments where  $G$  is dominant and low in segments where  $G$  is weak or absent.

This case is explored in Online Appendix A. Proposition 5 provides a theoretical characterization of the worst-case media power index for any possible vector of bandwidths  $\mathbf{K}$ . One can also conceive of cases that are intermediate between a fixed  $K$  and the case studied in the appendix. For instance, attention could be scaling smoothly with the number of sources

---

<sup>5</sup>For instance, Herman and Chomsky (1988) argued that most US news sources have a built-in bias in favor of a free-market capitalistic ideology. Supporters of that view can identify the set of sources that belong to the “propaganda system” and use the present model to compute the propaganda system’s media power index.

a voter follows, or varying randomly across segments.

One of the robustness checks in the empirical analysis (the next section) will utilize Proposition 5 to compute index values for the United States. In theory, using this measure could lead to different relative rankings of media organizations. In practice the ranking of all top US media organizations is unchanged.

### **Ideological Voters**

The analysis can be readily extended to a situation where voters have ex-ante opinion differences. The ideology of a voter is modeled as an array of signals that the voter received in the past. Ideological signals are binary and perform exactly the same mathematical function as signals generated by unbiased media. We can endow a particular voter with any number of *A*-leaning ideological signals and any number of *B*-leaning ideological signals. For any assumption we make about voter ideology, we can provide a suitable re-statement of Proposition 5.

While the analysis can be extended to any assumption about the ideological distribution of voters, a particularly tractable case is when voters can be divided in three groups: *A*-extremists, *B*-extremists, and moderates. The two extremist groups have the same size. *A*-extremists already have so many ideological signals in favor of *A* that, given their bandwidth, they would vote for *A* no matter what signals they receive from *B*. *B*-extremists are defined analogously. Moderates have no ideological signals. In that case, Proposition 5 applies, as stated, with the proviso that the set of voters is restricted to moderates, as those are the ones who will decide the election. The next section will also report power indices computed under the assumption that US voters can be divided into Democrats, Republicans, and independent voters.

### **Unequal Attention Allocation**

The current definition of media consumption is binary: either a voter follows a news source or she does not. One can instead allow agents to allocate more or less attention to different sources: each voter would be described by an attention vector, with weights summing up to one. The media power index can be easily adapted and the analysis would be similar. If one had a dataset with this information covering all major media platforms, one could re-compute the media power indices under the news specification.<sup>6</sup>

## **3 Media Power in the United States**

This section applies the media power index to the US media industry. We compute index values for most major media organization in two ways, both suggested by the theory. We

---

<sup>6</sup>Currently, there exists information for specific platforms (e.g., television usage) but, to the best of my knowledge, there is no publicly available dataset comprising multiple platforms: e.g., the time individuals devote to newspapers, television news, online news, etc.

first compute theoretical upper bounds, under an array of possible specifications. We then calibrate the model on the basis of existing estimates of media influence and compute power indices in that scenario. The section concludes by estimating the effect of hypothetical mergers between news organizations on media power.

### 3.1 Data

To implement the methodology introduced in the previous section, we need a media consumption matrix, namely information about what news sources individual voters follow similar to the example in Table 1. For the US, this data is provided by the Media Consumption Survey, which has been conducted every other year since 1994 by the Pew Research Center. In 2012, this telephone-based survey covered approximately 1500 US residents and included 103 questions.<sup>7</sup>

Questions refer to news consumption only, not entertainment. When covering a generalist media outlet like CBS, NBC, or ABC, the interviewer asks explicitly about news programs, mentioning them by name, for example *ABC World News* with Diane Sawyer. The respondent is asked to specify whether he or she follows that particular source “regularly, sometimes, hardly ever or never.”

The survey covers television, radio, printed media, websites, and social media. The respondent can indicate the three news sites he or she visits most often. “Google” indicates “Google News,” etc. We also have information on whether a particular news source is accessed through its traditional platform or through its website (e.g. *New York Times* and [www.nytimes.com](http://www.nytimes.com) or Fox News and [www.foxnews.com](http://www.foxnews.com)). We combine this information under the same source. Every respondent indicates whether they read a newspaper. In addition they are asked whether they read any of these four dailies: *New York Times*, *Wall Street Journal*, *Washington Post*, and *USA Today*.

A “media conglomerate” is defined as a corporate entity that owns, directly or indirectly, a controlling stake in the companies that own the individual media sources on the basis of the situation in 2012. We identify three conglomerates: News Corporation (Fox News and *Wall Street Journal*), Comcast (NBC and MSNBC), Time Warner (CNN and *Time Magazine*). Unfortunately, the fact that newspaper titles outside the major four are not identified individually means that we are unable to compute the power of press conglomerates such as Gannett and Tribune Publishing.

One subtle question is to what extent certain new media sources should be considered original sources or neutral aggregators of content provided by other sources (George 2008, George and Hogendorn 2013). Google News and MSN News are pure aggregators while Yahoo News produces original content.<sup>8</sup> For the purpose of the present exercise, we consider

---

<sup>7</sup>The Media Consumption Survey includes about 3000 respondents but only half of them (randomly selected) get the complete set of media consumption question.

The Media Consumption Survey has been discontinued. There are no 2014 or 2016 editions.

<sup>8</sup>Still, in our model a news aggregator could influence voter information by modifying the underlying algorithm to favor certain types of news. If this ability is unfettered, the aggregator has the same ability to affect reporting as any other news source.

them as independent news sources.

A key decision is to define what constitutes media consumption. Some sources are used more frequently than others. As we do not have quantitative information on usage (see suggestions for further research in the Conclusion), we rely on the qualitative frequency data collected by Pew. For the baseline results, we include sources that are used daily or almost daily. This imposes two requirements: the source must be updated on a continuous or daily basis, like a website, a daily news program on radio or television, or a daily newspaper, and the respondent must report that he or she follows the source “regularly”.

## 3.2 Upper Bounds

Table 3 reports media power measures. The main specification, found in column (4), is computed according to Proposition 3.

The highest value of the index corresponds to News Corporation (Fox News and the *Wall Street Journal*). The four top slots are taken by conglomerates that are mainly or exclusively television-based. New media occupy four positions: Yahoo News is sixth, MSN is eighth, Google News is 12th, and AOL is 15th.

The press appears to be less powerful than television. The most powerful newspaper, the *New York Times*, is only ninth. *USA Today* is 14th. The *Washington Post* and the *Wall Street Journal* (considered in isolation) are below *USA Today*. All the other US dailies (if they had a common owner) would have a power index of 0.214, lower than News Corporation on its own.

In terms of absolute magnitude, the 0.221 of News Corporation means that it could control a vote share of 11 percentage points. While this theoretical upper bound is large, it is of the same magnitude order as the most recently available empirical estimate of the effect of Fox News on the Republican vote share: approximately 6% in 2008 (Martin and Yurukoglu 2016). Comcast’s power is approximately two thirds of News Corporation, and it comes in roughly equal proportions from NBC (the most powerful of the Three Big Networks: ABC, CBS, NBC) and MSNBC (the third-most powerful news channel after Fox News and CNN).

The key pattern we observe – the dominance of television and in particular of News Corporation – is due to the combination of two factors. First, television reaches many more voters than other platforms. As we see in Column (1), 31% of respondents (who use at least one news source regularly) follow Fox News as opposed to 8% who read the *New York Times* (including website access). Second, television viewers follow fewer other sources, on average. This becomes apparent in column (3), which reports the percentage of users of a particular news source who say they use at least five other sources regularly. This share varies greatly, from Viacom’s 48% (they are Comedy Central viewers) to Yahoo News’ 16%. The *New York Times* has the second-highest value, almost twice as large as that of Fox News. The presence of other sources determines how a media organization’s reach translates into attention shares (reported in column (2), which in turn converts into media power. So, to sum up, News Corporation has a large value of  $\Pi$  because it has a large reach and because it commands a large attention share among its users. Other sources, like the *New York Times*, do worse on both counts.

The media power index computed here can be compared to standard market-share based measures of concentration. Column (5) reports the Company Power Index (CPI) as defined in Noam (2009), namely the average value of the squared market share in the various media industries (including internet media) weighted by industry revenue. As the squared market share is also the building block of the Herfindahl-Hirschmann Index (HHI), CPI is closely related to standard competition policy measures. CPI and media powers are related: the correlation index between (4) and (5) is 0.40. However, there are also large discrepancies. The most striking is that our most powerful media organization, News Corporation, is now ranked sixth. Comcast is first, thanks to its dominance of the cable industry. Google is second, thanks to its dominance of the internet market.

The media power index can be used to explore the effect of voter turnout and voter ideology on relative media influence. It can also be tailored to features of specific elections, such as state-level races and the presence of the electoral college in US presidential elections.

### 3.2.1 Voter Turnout

The model assumed no abstention – or, equivalently, that the turnout rate is independent of media consumption. In practice, however, different news sources may have followers that are more or less likely to vote. If we knew the turnout probability for each voter, the attention share of each media source would simply be weighted by the turnout probability of followers of that source. One (rough) proxy for the probability of voting is voter registration – one of the questions asked in the Pew Survey. This may provide some information on the direction of the change in media power once we allow for differential turnout rates.

Column (6) presents the new results. Users of traditional platforms are more likely to be registered voters. The power of television tends to be higher among registered voters than in the overall population, with the notable exception of CNN. New media tend to have lower index values. Newspapers are subject to two effects. The print readership is higher among registered voters and the digital readership is lower. The overall effect on the four newspapers for which we have individual data is ambiguous and close to zero.

### 3.2.2 Ideology

One important robustness check concerns ideology. As we discussed at the end of Section 4, the model can be extended to an environment with ideological voters. The simplest way of doing this is to assume that there are partisan voters and moderates. At the end of the Pew survey voters are asked to identify as Democrats, Republicans, or Independents. For the purpose of this robustness check, we assume that partisan voters always vote for their party, while moderates decide on the basis of the information they receive. Thus, our model applies as stated after we restrict attention to Independents.<sup>9</sup>

---

<sup>9</sup>This exercise assumes that the share of Democrats (who vote) is exactly equal to the share of Republicans (who vote), and therefore their weights cancel out in the election. One could add a correction for this difference.

Column (7) reports power index values computed for independent voters only. The figures are quite similar to those in column (4). They tend to be slightly lower, which is a consequence of the fact that independent voters have a slight tendency to follow a larger number of news sources than partisan voters. Change in the power of an individual source depends on whether the source is more or less likely to be followed by independent voters. Mainstream television appears to be underrepresented among independent voters (for instance, MSNBC’s reach is 15.9% in the general population and 12.3% among independents). This also applies to some niche sources like the *Rush Limbaugh Show*. The sources that are over-represented among Independents tend to be public-service media (NPR, PBS, BBC) and financial sources (*Wall Street Journal*, Bloomberg, Reuters). As a result of this, Comcast and Time Warner lose a little power while, thanks to *WSJ*, News Corporation gains some ground.

### 3.2.3 Electoral College

All the previous analysis focused on vote shares. However, US presidential elections are not decided on the basis of the the popular vote but rather by the electoral college. As is well known, in practice the winner is decided by a small number of “swing states,” whose voters are more or less evenly split between Democrats and Republicans. This feature can be accommodated in the present set-up.

Column (8) reports media power indices computed for a set of swing states: Colorado, Florida, Iowa, Nevada, New Hampshire, North Carolina, Ohio, Virginia, and Wisconsin.<sup>10</sup> The most striking difference with the baseline case is the increased influence of Comcast, mostly due to NBC, not MSNBC. The power index of the *New York Times* decreases by two thirds.

## 3.3 Calibration

In the previous section, we computed the power of media organizations on the basis of purely theoretical upper bounds. An alternative, which we explore in this section, is to tighten those theoretical upper bounds with values obtained from empirical estimates. This approach corresponds to assuming – in a sense that we will formalize below – that the influence of the media on the political process can never be greater than it was in the episodes that gave rise to the empirical estimates.

Recall that in the model  $\beta$  represents the sophistication level of an individual voter, namely the prior probability that he knows that a media owner is biased. To simplify the analysis, we consider a slightly different definition of sophistication. We assume that a share  $b$  of voters are completely sophisticated (and therefore  $\beta = 1$ ) and a share  $1 - b$  is completely

---

<sup>10</sup>Swing states are chosen on the basis of the vote share difference between Obama and Romney in the 2012 presidential election. One could use two reference points, either 0 or the overall margin by which Obama won (4%). We use both and add a 4% margin on each side. Therefore we include all states that Romney won with a margin of 4% or less and all states that Obama won with a margin of 8% or less.

naive ( $\beta = 0$ ). The vote share of candidate  $A$  in segment  $M$  is thus

$$p_A = (1 - b) ((1 - g_M)(1 - \sigma) + g_M) + b(1 - \sigma),$$

and the power index is

$$\Pi_G = \frac{(1 - b) a_G}{1 - (1 - b) a_G}, \quad (2)$$

which is the same expression as in Proposition 3 except that the attention share of media group  $a_G$  is now scaled down by the share of naive voters  $1 - b$ .

The calibrated upper bounds approach requires an empirical estimate of the power of a particular media organization  $G$ , namely an estimate of the share of votes that  $G$  is able to swing:  $\tilde{\Pi}_G$ . That figure, together with the attention share  $a_G$ , can be replaced in expression (2). This yields a calibrated value of  $b$ , which denotes the share of sophisticated voters that is consistent with the empirical estimate of media influence:

$$\hat{b} = \frac{a_G (\tilde{\Pi}_G + 1) - \tilde{\Pi}_G}{a_G (\tilde{\Pi}_G + 1)}. \quad (3)$$

As one would expect, the share of sophisticated voters  $\hat{b}$  needed to rationalize a certain observed influence  $\tilde{\Pi}_G$  is decreasing in  $\tilde{\Pi}_G$  (because only naive voters can be influenced) and in the attention share  $a_G$  (because news sources with larger audiences can swing more votes even if they have a more sophisticated audience).

Once an estimate of  $\hat{b}$  is obtained, it can in turn be used to compute upper bounds to media power for other media organizations that we believe have similar or higher levels of voter sophistication.

The literature on the effect of US media bias indicates that television and the press have different effects.

For television, DellaVigna and Kaplan (2007) estimated the influence of Fox News on the 2000 US presidential election. Between 1996 and 2000, Fox News was introduced in cable broadcasting in localities that comprise 35% of the US population. A difference-in-differences approach, controlling for locality fixed effects and voting trends, indicates that in areas where Fox News was introduced the Republican vote share increased by 0.4-0.7 percentage points, depending on the specification. Taking the upper bound of the interval and recalling that the vote swing is twice the increase in the vote share, we obtain  $\tilde{\Pi}_G = 0.014$ . A second estimate of the Fox News effects comes from Martin and Yurukoglu (2016). Using a different approach based on channel position, the authors estimate that removing Fox News from cable television during the 2008 election cycle would have reduced the Republican vote share by 6.34 percentage points, corresponding to a power index of  $\tilde{\Pi}_G = 0.127$ .

By contrast, the effect of the press appears to be zero. Gentzkow, Shapiro, and Sinkinson (2011) finds a precisely estimated null effect of entry and exit of partisan newspapers on vote shares. Chiang and Knight (2011) decompose the probabilistic effect of a newspaper's endorsement of political candidates into an informational effect, which depends on the credibility of the newspaper, and a pure "persuasion" effect. They cannot reject the hypothesis

that the persuasion effect is zero and all of the effect of endorsements comes from information provision.<sup>11</sup> In both cases, this corresponds to  $\tilde{\Pi}_G = 0$  and, by (3), to  $\hat{b} = 1$ . Namely, all newspaper readers are sophisticated.

Suppose we wish to adopt the approach of Gentzkow, Shapiro, and Sinkinson (2011) and Chiang and Knight (2011) for newspapers (and other text-based media such as websites) and DellaVigna and Kaplan (2007) or Martin and Yurukoglu (2016) for television (and other non-text-based media such as radio). The question arises of how to treat individuals who get their news from both classes of media. We assume that being exposed to written news makes them fully sophisticated even when they watch television.

We begin with the values of  $\tilde{\Pi}_G = 0.014$  under DellaVigna and Kaplan (2007), found above. We also need an estimate of the attention share of Fox News in 2000, split into viewers that are also newspaper readers and pure viewers. The reach of Fox News in 2000 according to Pew was 16.7% in 2000 and 27.6% in 2012.<sup>12</sup> If we assume that the attention share grew at the same rate in the same period because the average number of alternative sources used by Fox News viewers did not change, the 2012 attention share of 16.3% results in a 2000 attention share of 9.9%. In 2012, we observe that about 42% of the attention share of Fox News came from people who also read newspapers and must therefore be dropped. If we apply the same correction to 2000, we obtain an effective attention share  $a_G = 0.057$ .

Once we plug the values of  $\tilde{\Pi}_G$  and  $a_G$  in (3), we obtain an estimate of the sophistication parameter:  $\hat{b} = 0.76$ . Namely, the influence observed in DellaVigna and Kaplan is rationalized in the present model if 24% of (non-reader) Fox News viewers are naive. Under the assumption that the share of naive users for other visual sources is the same, we can plug  $\hat{b} = 0.76$  back into (2) to obtain the power index of any other media organization. Of course, the attention share must now be restricted to followers that are not readers (the shares are reported in Column (3) of Table 4). We thus obtain the media power index values in Column (4) of Table 4. The values are approximately one magnitude order lower than the theoretical maximum reported in Column (1) of the same table.

It is important to understand that the calibrated power index in (2) is a true upper bound to media influence only if one is willing to make three assumptions that guarantee that the event that gave rise to the estimate is indeed a worst-case scenario. First, no set of media users in 2012 are more naive than Fox News viewers were in 2000. Second, Fox News in 2000 chose its reporting strategy to maximize its influence on viewers. Third, the voters have indeed minimal bandwidth  $K = 1$ .<sup>13</sup> If one of these three assumptions is violated, then

<sup>11</sup>The upper bound of Chiang and Knight’s (2011) 95% confidence interval is 0.0132, meaning that someone who owned all US media could increase the vote share of her preferred candidate by 1.32% through endorsements (assuming away any informational effect). This translates into a power index of 0.026 for a media monopolist ( $G = \mathbb{M}$ ). The share of sophisticated voters consistent with this value is  $b = 0.975$ . This means that no more than 2.5% of voters can be naive.

<sup>12</sup>Pew surveyed usage of Fox News in 2000 but it missed a large number of other sources such as dailies. Therefore, the attention share must be estimated indirectly.

<sup>13</sup>This requires that in the electoral event that generated the observed  $\tilde{\Pi}_G$ , media organization  $G$  was pushing as hard as possible to get its favorite candidate elected. This in turn requires the assumption that the candidate favored by  $G$  had a lower valence than the other one, otherwise he could have won without support.

it is possible that some news sources in 2012 have greater potential influence than the values in Column (2).

We follow the same approach for Martin and Yurukoglu (2016). The estimate of the attention share of Fox News in 2008 is 0.154. Applying the same 42% pure-viewer correction, we obtain  $a_G = 0.089$ . Such attention share is not sufficient to rationalize the  $\tilde{\Pi}_G = 0.127$  estimated by Martin and Yurukoglu (2016). The closest we can get is by assuming that all pure viewers are perfectly naive:  $\hat{b} = 0$ . This yields a power index for Fox News in 2008 of  $\Pi_G = 0.094$ . If we transpose these assumptions to 2012 we obtain the power indices in Column (5) of Table 4. The calibrated power index of broadcast-based media organizations is now much higher than in Column (4). For Fox News, it is almost half the value of the theoretical upper bound.

To sum up, the main message of this calibration exercise is that using existing empirical estimates to tighten the theoretical upper bounds does not overturn the relative power ranking of US media organization. This is mainly because the platform for which estimated coefficients are lower – the press – has already limited power. The magnitude of the change in the index depends on which estimate of the Fox News effect one uses.

### 3.4 Merger Analysis

This section applies the media power index to two hypothetical sets of mergers: one based on the numerical example we have used throughout the paper, and one involving proposed or hypothesized mergers between existing US media companies.

The first example revisits Table 1. Suppose that newspaper Np3 is for sale and both television network Tv2 and newspaper Np1 have expressed an interest in acquiring it. Which of the two buyers poses a larger risk?

Under current US guidelines (Federal Communications Commission 2016), the merger between Tv2 and Newspaper 3 is prohibited because it involves cross-ownership of a daily newspaper and a broadcast station. Setting aside this consideration, a standard analysis based on market shares would also yield the conclusion that Newspaper 1 is a less dangerous prospect because the reach of an Np1-Np3 pair would be 50% while the reach of Tv2-Np3 would be 60%. (see Table 2, Panel A). Media power indices paint a different picture. The total attention share of Tv2-Np3 is lower than that of Np1-Np3. Therefore, the potential risk is higher if Np3 is bought by the other newspaper rather than the TV station.

The second example focuses on four US media conglomerates with a large television presence. It has been suggested that CBS, News Corporation, Time Warner, and Viacom are particularly likely to be involved in pairwise consolidation.<sup>14</sup> Panel B of Table 2 reports the power indices, computed according to Proposition 3, for every possible pairwise merger.<sup>15</sup>

According to the table, the hypothetical merged entities have very different index values. The most powerful new conglomerate would be formed by News Corporation and Time

---

<sup>14</sup>See for instance: <http://blogs.reuters.com/breakingviews/2015/12/31/viacom-time-warner-among-2016-merger-wannabes/>

<sup>15</sup>As the Pew data stops in 2012, we use that year for the analysis.

A	merging sources			merged entities	
	Tv2	Np1	Np3	Tv2+Np3	Np1+Np3
Reach	40%	20%	30%	60%	50%
Attention	14.1%	15%	9.1%	23.2%	24.1%
II	0.165	0.176	0.100	0.304	0.318

B	merging sources				merged entities					
	CBS	NewsCorp (NC)	Time Warner (TW)	Viacom	CBS NC	CBS TW	CBS Viacom	NC TW	NC Viacom	TW Viacom
Reach	8.7%	31.1%	20.9%	6.3%	36.8%	26.4%	14.5%	43.8%	35.5%	24.8%
Attention	3.5%	18.1%	7.9%	2.8%	21.6%	11.4%	6.2%	26.0%	20.9%	10.7%
II	0.036	0.221	0.086	0.028	0.275	0.128	0.066	0.352	0.264	0.120

Table 2: Effect of hypothetical mergers:

*Panel A.* Effect of the acquisition of Newspaper 3 (Np3) by either Television Network (Tv2) or Newspaper 1 (Np1), based on the the media consumption matrix in Table 1.

*Panel B.* Hypothetical pairwise mergers between CBS, NewsCorp, Time Warner, and Viacom, based on 2012 Pew data, as reported in Table 3

*Definitions.* Reach: percentage of voters following that source or set of sources. Attention: Attention share of that source or set of sources. II: Power index of that source or set of sources (as defined in Proposition 3)

Warner, and its power index would be five times as large as the power index of the least powerful merged entity, the one involving a re-merger of CBS and Viacom. As News Corporation and Time Warner are ranked respectively first and third in term of media power, it is not surprising that their joint index value would be extremely high. The resulting entity would be more than twice as powerful as the runner-up, Comcast. While in the hypothetical example in Panel A, the ranking of merged entities by reach differs from the raking of merged entities by media power, in the real-world example of Panel B, the two indices rank merged entities in the same way.

Online Appendix II examines the effect of a hypothetical merger between Gannett and Tribune Publishing. Both companies own a number of newspaper titles.

## 4 Conclusion

This paper has developed a media power notion based on two principles: the most natural unit of analysis is the mind of voters, and the power index is computed on the basis of the worst case over a set of possible assumptions about the beliefs and attention patterns of voters. The resulting media power index can be calculated for the United States on the basis of existing media consumption data.

The media power indices highlight two lessons. First, the potential influence of some

large US media conglomerates on the political process may be quite large. The size of these upper bounds supports the criticism by Polo (2005) and others (see introduction) of the standard approach to measuring media concentration. The problem is that, while most US media industries appear relatively competitive according to standard market-based definitions, this does not translate into individual-level diversity in news consumption: a large share of the electorate get their political information from a small number of news sources, typically television networks. The proposed index, which documents this form of media concentration, highlights the need for media regulators to complement standard market-centered concentration measures with a voter-centered approach.

Second, the idea that new media are changing the traditional power structure finds limited support in our media indices. One traditional platform stands out in terms of media power. The four most powerful media organizations in the US are mainly television providers. The most powerful pure internet source (Yahoo) is sixth, and the most powerful pure press source (the *New York Times*) is tenth. This finding is highly robust to different specifications and survives even if we focus only on individuals under 40 years of age. As confirmed by the merger analysis in the previous section, media regulators should be extremely wary of mergers involving large television organizations.

## References

- [1] Simon P. Anderson and John McLaren, Media Mergers and Media Bias with Rational Consumers, *Journal of the European Economic Association*, 2010.
- [2] Simon P. Anderson, David Strömberg, and Joel Waldfogel (eds.). *Handbook of Media Economics*, North Holland, 2016.
- [3] David Balan, Patrick DeGraba, and Abraham L. Wickelgren. Ideological Persuasion in the Media. Working paper, 2009.
- [4] David Baron. Persistent Media Bias. *Journal of Public Economics* 90: 1–36. 2006.
- [5] Timothy Besley and Andrea Prat. Handcuffs for the Grabbing Hand? The Role of the Media in Political Accountability. *American Economic Review*, 96(3): 720-736, June 2006.
- [6] Isabelle Brocas, Juan D. Carrillo, and Simon Wilkie. A Theoretical Analysis of the Impact of Local Market Structure on the Range of Viewpoints Supplied, *Federal Communications Commission Review of Media Ownership Rules*, Media Study No 9, 2010.
- [7] Gabriel Carroll. Robustness and Linear Contracts. *American Economic Review*, 105(2): 536-63, 2015.
- [8] Sylvain Chassang. Calibrated Incentive Contracts. *Econometrica* 81(5): 1935–1971.

- [9] Sylvain Chassang and Gerard Padro i Miquel. “Corruption, Intimidation and Whistleblowing: a Theory of Inference from Unverifiable Reports.” NBER Working Paper 20315, 2014..
- [10] Fang Chiang and Brian Knight. Media Bias and Influence: Evidence from Newspaper Endorsements, July 2011, *Review of Economic Studies*, 78(3), 795-820.
- [11] John Duggan and Cesar Martinelli. A Spatial Theory of Media Slant and Voter Choice, *Review of Economic Studies*, vol. 78 (2011) 640-666.
- [12] Ruben Durante and Brian Knight. Partisan Control, Media Bias, and Viewer Responses: Evidence from Berlusconi’s Italy. *Journal of the European Economic Association*. June 2012.
- [13] Wioletta Dziuda. Strategic Argumentation. *Journal of Economic Theory* 146(4): 1362–1397, July 2011.
- [14] Ruben Enikolopov, Maria Petrova, and Ekaterina Zhuravskaya. Media and Political Persuasion: Evidence from Russia. *American Economic Review*, December 2011, 111(7): 3253-85.
- [15] Federal Communications Commission. Review of the Broadcast Ownership Rules. August 25, 2016. Available on: <https://www.fcc.gov/document/2014-quadrennial-review-commissions-broadcast-ownership-rules> (accessed on October 14, 2016).
- [16] Matthew Gentzkow and Jesse Shapiro, Media Bias and Reputation. *Journal of Political Economy* 114(2): 280-316. April, 2006.
- [17] Matthew Gentzkow and Jesse Shapiro, ”What Drives Media Slant? Evidence from U.S. Daily Newspapers”, *Econometrica* 78, 2010, 35-71.
- [18] Matthew Gentzkow, Jesse M. Shapiro, and Michael Sinkinson. The Effect of Newspaper Entry and Exit on Electoral Politics, *American Economic Review* 101 (7). December 2011.
- [19] Matthew Gentzkow and Emir Kamenica, Competition in Persuasion. Working paper. University of Chicago. September 2012.
- [20] Lisa M. George. The Internet and the Market for Daily Newspapers. *BE Journal of Economic Analysis & Policy* 8:(1) 1935-1982. July 2008.
- [21] Lisa M. George and Christiaan Hogendorn. Local News Online: Aggregators, Geo-Targeting and the Market for Local News. Working paper, 2013.
- [22] Edward S. Herman and Noam Chomsky. *Manufacturing Consent: The Political Economy of the Mass Media*. Pantheon Books, 1988.

- [23] Emir Kamenica and Matthew Gentzkow. Bayesian Persuasion. *American Economic Review*. 101 (6). October 2011.
- [24] Leveson Inquiry. An Inquiry into the Culture, Practices, and Ethics of the Press: Executive Summary, London. Available from, [nationalarchives.gov.uk](http://nationalarchives.gov.uk), 20t:appendix\_power\_indices12.
- [25] Kristof Madarasz and Andrea Prat. Screening with an Approximate Type Space. *Review of Economic Studies*, forthcoming.
- [26] Gregory Martin and Ali Yurukoglu. Bias in Cable News: Real Effects and Polarization. Working Paper, Stanford University, December 2016.
- [27] Eli M. Noam. *Media Ownership and Concentration in America*. Oxford University Press. 2009.
- [28] Ofcom (2009). Second Public Service Broadcasting Review: Putting Viewers First. Available on [www.ofcom.org.uk](http://www.ofcom.org.uk), January 2009.
- [29] Maria Petrova. Mass Media and Special Interest Groups”, *Journal of Economic Behavior and Organization*, September 2012, 84(1), pp. 17-38
- [30] Michele Polo. Regulation for Pluralism in the Media Markets, in P.Seabright J. von Hagen (eds.) *Regulation of Media Markets*, Cambridge University Press. 2005;
- [31] Riccardo Puglisi. Being the New York Times: the Political Behavior of a Newspaper.STICERD; London School of Economics, Political Economy and Public Policy (PEPP) Working Paper n. 20. 2006.

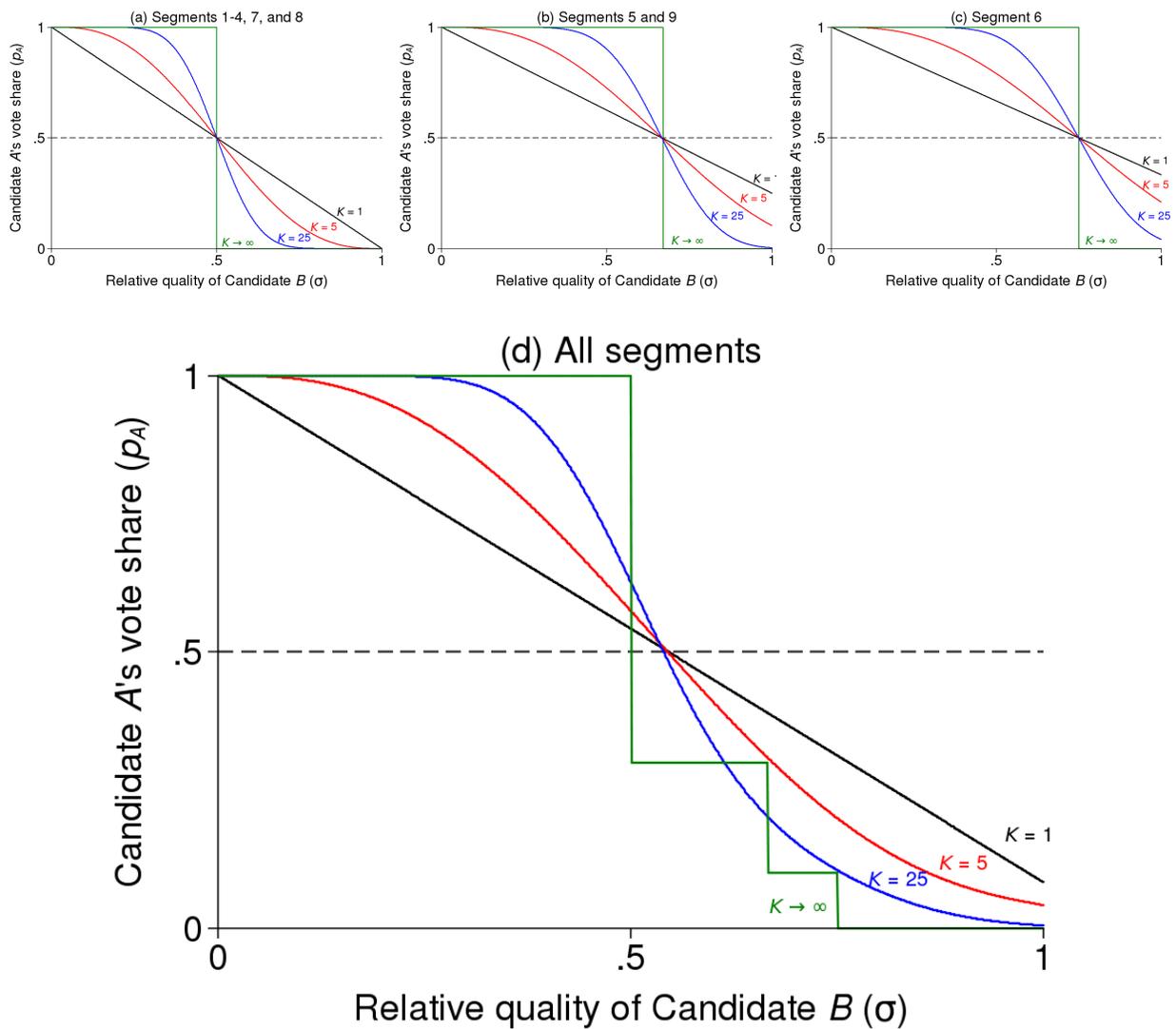


Figure 2: Candidate  $A$ 's vote share under the assumption that Newspaper 2 in the example of Table 1 is biased. In every plot, each of the four vote functions represents the relation between candidate quality and vote share for a different bandwidth value  $K$ . Plot (a) displays the vote functions for the four segments where Newspaper 2 is not read and therefore reporting is unbiased: the plot is identical to the one in Figure 1. Plot (b) depicts the two segments where Newspaper 2 is one of four sources used by voters. Plot (c) includes the segment where Newspaper 2 is one of three sources. Plot (d) displays the overall vote share function, obtained as a weighted average of the vote functions in the nine segments: the intersection of the overall vote function with the 50% horizontal line identifies the maximal value of  $\sigma$ ,  $\bar{\sigma}(K)$  for which the evil owner can get  $A$  elected.

Table 3: Media Power Indices

Media Organization	(1) Reach	(2) Attention Share	(3) 5+ Sources	(4) $\Pi$	(5) CPI	(6) Registered Voters	(7) Independent Voters	(8) Swing Voters
News Corporation (Fox, WSJ)	0.311	0.181	21.8	0.221	19	0.225	0.215	0.228
Comcast (NBC, MSNBC)	0.268	0.133	41.8	0.153	166	0.161	0.127	0.189
Time Warner (CNN, Time)	0.209	0.079	32.9	0.086	24	0.081	0.082	0.082
ABC	0.143	0.062	25.2	0.066	4	0.066	0.048	0.069
NPR	0.141	0.056	22.6	0.059	26	0.061	0.076	0.061
Yahoo	0.126	0.054	16.4	0.057	n.a	0.047	0.066	0.042
CBS	0.087	0.035	25.0	0.036	1	0.035	0.034	0.029
MSN	0.094	0.032	32.1	0.033	25	0.031	0.035	0.031
New York Times	0.079	0.030	41.2	0.031	5	0.032	0.028	0.010
Viacom	0.063	0.028	48.6	0.028	0	0.025	0.042	0.024
PBS	0.079	0.024	31.3	0.025	8	0.025	0.024	0.025
Rush Limbaugh	0.063	0.021	25.0	0.022	n.a.	0.023	0.014	0.024
Google	0.055	0.021	22.5	0.022	88	0.018	0.027	0.017
USA Today	0.047	0.019	28.6	0.020	5	0.020	0.017	0.020
AOL	0.031	0.013	17.8	0.013	.7	0.014	0.019	0.006

Source: Pew Media Consumption Survey, 2012.

1. Share of media users who follow that source regularly.
2. Attention share of that media source (Average individual attention share, where the attention share of that source for an individual who follows  $N$  sources is  $1/N$  if he follows that source and zero otherwise).
3. Percentage of users of that source who follow at least 5 sources.
4. Media power index  $\Pi$  as in Proposition 3, computed on the basis of daily usage for all the U.S.
5. Company Power Index (CPI): Average Herfindahl squared market share across platforms weighted by media platform share, computed in 2011 or later year (Noam 2009, p. 365): range: 0 (perfect competitor) to 10,000 (monopolist).
6. As in (4) but restricted to the sample of respondents who report being registered voters.
7. As in (4) but restricted to respondents who self-identify as independent.
8. As in (4) but restricted to residents of swing states: CO, FL, IA, NE, NH, NC, OH, VA, and WI.

Table 4: Calibration

<b>Media Organization</b>	(1) Baseline II	(2) Total Attention	(3) Non-readers Attention	(4) DVK	(5) MY
News Corporation (Fox, WSJ)	0.221	0.181	0.519	0.018	0.104
Comcast (NBC, MSNBC)	0.153	0.133	0.414	0.010	0.058
Time Warner (CNN, Time)	0.086	0.079	0.455	0.007	0.037
ABC	0.066	0.062	0.552	0.006	0.035
NPR	0.059	0.056	0.502	0.005	0.029
Yahoo	0.057	0.054	0.686	0.007	0.038
CBS	0.036	0.035	0.551	0.004	0.019
MSN	0.033	0.032	0.603	0.004	0.019
New York Times	0.031	0.030	0.000	0.000	0.000
Viacom	0.028	0.028	0.543	0.003	0.015
PBS	0.025	0.024	0.417	0.002	0.010
Google	0.022	0.021	0.616	0.002	0.013
Rush Limbaugh	0.022	0.021	0.561	0.002	0.012
USA Today	0.020	0.019	0.000	0.000	0.000
AOL	0.013	0.013	0.556	0.001	0.007
Share of sophisticated readers	0%			100%	100%
Share of sophisticated non-readers	0%			76%	0%

Source: Pew Media Consumption Survey, 2012.

1. Media power index II as in Proposition 3 (same as Column 4 in Table 3)
2. Attention share of media organizations (same as Column 2 in Table 3)
3. Attention share of media organization excluding readers
4. Calibrated media power index based on estimates by Della Vigna and Kaplan (2007) for broadcast media and Gentzkow, Shapiro, and Sinkinson (2011) and Chiang and Knight (2011) for newspaper readers.
5. Calibrated media power index on the basis of Martin and Yurukoglu (2016) for broadcast media and Gentzkow, Shapiro, and Sinkinson (2011) and Chiang and Knight (2011) for newspaper readers.

# Appendix A: Proofs

## 4.1 Proposition 2

We begin by writing the probability that a voter in group  $M$  observes a particular realization of the  $K$ -sized signal vector  $y^i$  he receives from media outlets in  $M$ . The vector includes news items randomly drawn from outlets in  $M$ . Let  $y_k^i$  denote the  $k$ th realization of the vector and let  $m(k)$  denote the media outlet it is drawn from.

This probability is computed according to the beliefs of the voter. Suppose the voter believes that the owner is evil with probability  $\beta$  and that an evil owner would use reporting strategy  $\hat{s}$ . Then, the probability of realization  $y^i = Y$  would be given by:

$$\begin{aligned} & \Pr(y^i = Y | \sigma, \hat{s}) \\ &= \sigma^{N_1(M/G)} (1 - \sigma)^{N_0(M/G)} \left( (1 - \beta) \sigma^{N_1(G)} (1 - \sigma)^{N_0(G)} + \beta (\hat{s}\sigma)^{N_1(G)} (1 - \hat{s}\sigma)^{N_0(G)} \right) \end{aligned}$$

where  $N_y(M/G)$  is the number of signals with value  $y$  coming from unbiased outlets, while  $N_y(G)$  is the same variable for potentially biased outlets (it must be that  $N_0(M/G) + N_1(M/G) + N_0(G) + N_1(G) = K$ .)

The voter computes the expected value of candidate quality as follows:

$$E[\sigma | Y, \hat{s}] = \frac{\int_0^1 \Pr(y^i = Y | \sigma, \hat{s}) \sigma f(\sigma) d\sigma}{\int_0^1 \Pr(y^i = Y | \sigma, \hat{s}) f(\sigma) d\sigma}$$

and votes for  $A$  if and only if  $E[\sigma | Y, \hat{s}] \leq \frac{1}{2}$ .

We now compute a lower bound to posterior  $E[\sigma | Y, \hat{s}]$ .

**Lemma 4** *For any vector of signals  $Y$ , let  $N_1(M/G)$  be the number of positive signals from unbiased media, let  $N_0(M/G)$  be the number of negative signals from unbiased media, and let  $K_G = N_0(G) + N_1(G)$  be the number of signals from biased media. The voter posterior  $E[\sigma | Y, \hat{s}]$  is bounded below by*

$$\frac{\int_0^1 \sigma^{N_1(M/G)} (1 - \sigma)^{N_0(M/G) + K_G} \sigma f(\sigma) d\sigma}{\int_0^1 \sigma^{N_1(M/G)} (1 - \sigma)^{N_0(M/G) + K_G} f(\sigma) d\sigma} \quad (4)$$

**Proof.** Define

$$\begin{aligned} g(\sigma) &= \sigma^{N_1(M/G) + N_1(G)} (1 - \sigma)^{N_0(M/G) + N_0(G)} ; \\ h(\sigma, \hat{s}) &= \sigma^{N_1(M/G)} (1 - \sigma)^{N_0(M/G)} (\hat{s}\sigma)^{N_1(G)} (1 - \hat{s}\sigma)^{N_0(G)} . \end{aligned}$$

Therefore

$$E[\sigma | Y, \hat{s}] = \frac{\int_0^1 ((1 - \beta) g(\sigma) + \beta h(\sigma, \hat{s})) \sigma d\sigma}{\int_0^1 ((1 - \beta) g(\sigma) + \beta h(\sigma, \hat{s})) d\sigma} .$$

We have

$$\text{sign} \left( \frac{d}{d\beta} E[\sigma|Y, \hat{s}] \right) = \text{sign} \left( \frac{\int_0^1 h(\sigma, \hat{s}) \sigma d\sigma}{\int_0^1 h(\sigma, \hat{s}) d\sigma} - \frac{\int_0^1 g(\sigma) \sigma d\sigma}{\int_0^1 g(\sigma) d\sigma} \right) \quad (5)$$

Now note that  $\frac{h(\cdot, \hat{s})}{\int_0^1 h(\sigma, \hat{s}) d\sigma}$  is a probability density function and it satisfies a reverse monotone likelihood ratio property (MLRP) in  $\hat{s}$  an  $\sigma$ . Recall that MLRP implies first-order stochastic (FOS) dominance. Therefore, our (reverse) MLRP implies that, if  $\hat{s}'' > \hat{s}'$ , then

$$\frac{h(\cdot, \hat{s}')}{\int_0^1 h(\tilde{\sigma}, \hat{s}') d\tilde{\sigma}} \text{ FOS-dominates } \frac{h(\cdot, \hat{s}'')}{\int_0^1 h(\tilde{\sigma}, \hat{s}'') d\tilde{\sigma}}.$$

If we set  $\hat{s}'' = 1$  and  $\hat{s}' = \hat{s}$ , we have that for any  $\hat{s} \in [0, 1]$ :

$$\frac{h(\cdot, \hat{s})}{\int_0^1 h(\tilde{\sigma}, \hat{s}) d\tilde{\sigma}} \text{ FOS-dominates } \frac{g(\cdot)}{\int_0^1 g(\tilde{\sigma}) d\tilde{\sigma}}.$$

Given that FOS dominance implies expected value ranking, we have:

$$\frac{\int_0^1 h(\sigma, \hat{s}) \sigma d\sigma}{\int_0^1 h(\sigma, \hat{s}) d\sigma} \geq \frac{\int_0^1 g(\sigma) \sigma d\sigma}{\int_0^1 g(\sigma) d\sigma}.$$

Combined with (5), this proves that

$$\frac{d}{d\beta} E[\sigma|Y, \hat{s}] \geq 0.$$

This in turn implies that, for any given  $N_1(M/G)$ ,  $N_1(G)$ ,  $N_0(M/G)$ , and  $N_0(G)$ , the value of  $E[\sigma|Y, \hat{s}]$  is minimized when  $\beta = 0$ . In that case we have

$$E[\sigma|Y, \hat{s}] = \frac{\int_0^1 \sigma^{N_1(M/G)+N_1(G)} (1-\sigma)^{N_0(M/G)+N_0(G)} \sigma f(\sigma) d\sigma}{\int_0^1 \sigma^{N_1(M/G)+N_1(G)} (1-\sigma)^{N_0(M/G)+N_0(G)} f(\sigma) d\sigma},$$

and it is immediate to see that for any  $N_0(G) + N_1(G) = K_G$ , the value of  $E[\sigma|Y, \hat{s}]$  is minimized when  $N_1(G) = 0$ . This proves the statement. ■

The lemma states that, given  $N_0(M/G)$  and  $N_1(M/G)$ , the value of  $E[\sigma|Y, \sigma]$  can never be lower than the value achieved when all the biased outlets' news items are favorable to  $A$  and the voter believes that all media are unbiased.

Let us translate the lower bound on the posterior into an upper bound on the vote share that candidate  $A$  can receive. The lower bound in (4) is greater or equal to  $\frac{1}{2}$  if and only if

$$N_1(M/G) \geq N_0(M/G) + K_G$$

In other words, the voter selects candidate  $B$  if and only if the number of signals in favor of Candidate  $A$  is weakly larger than the number of signals in favor of  $B$ , including signals

from both unbiased and potentially biased outlets. The “weakly” part comes from the fact that  $\beta > 0$ . If the two candidates are supported by exactly the same number of signals, the voter would be exactly indifferent if  $\beta = 0$ . But for any strictly positive  $\beta$ , he must prefer  $B$ .

The probability that the voter selects  $B$  is thus equal to:

$$\Pr(N_1(M/G) \geq N_0(M/G) + K_G) = \Pr\left(\frac{N_1(M/G)}{N_1(M/G) + N_0(M/G) + K_G} \geq \frac{1}{2}\right)$$

The probability that an individual signal takes value 1 is  $(1 - g_M)\sigma + g_M \cdot 0$ . The probability that a particular voter selects  $A$  is given by the cumulative distribution of a binomial with parameter  $(1 - g_M)\sigma$ , with  $K$  possible realizations, evaluated at the highest integer that is strictly smaller than  $K/2$ . For  $K = 1$  it is 0, for  $K = 2$  it is 0, for  $K = 3$  it is 1, etc). Let  $\lceil K/2 \rceil$  denote the ceiling of  $K/2$ , namely the smallest integral that is at least as large as  $K/2$ . Then:

$$p_A(g_M, K, \sigma) = \sum_{k=0}^{\lceil K/2 \rceil - 1} \binom{K}{k} ((1 - g_M)\sigma)^k (1 - (1 - g_M)\sigma)^{K-k} \quad (6)$$

By the law of large numbers,  $p_A(g_M, K, \sigma)$  is the share of  $A$  votes in segment  $M$ . If  $p_A(g_M, K, \sigma)$  is an upper bound to the vote share that  $A$  can achieve under any voter belief, this means that in equilibrium  $A$ 's vote share in  $M$  can be higher than  $p_A(g_M, K_M, \sigma)$ . Furthermore, we can easily see that this bound is tight by finding one particular set of beliefs that achieves the bound. To see this just assume that the evil owner uses a strategy of reporting only zeros. When  $\beta \rightarrow 0$ , it is easy to verify that the vote share in  $M$  does indeed tend to  $p_A(g_M, K_M, \sigma)$  for any  $K_M$ .

Given any bandwidth  $K$ , the power of group  $G$  corresponds to the highest value of  $\bar{\sigma}(K)$  such that the  $A$ -vote share is at least  $1/2$ , namely the solution to

$$\sum_{M \subset \mathbb{M}} q_M p_A(g_M, K, \bar{\sigma}(K)) = \frac{1}{2}.$$

By replacing the expression for  $p_A$  obtained in (6), we complete the proof of the proposition.

## 4.2 Proposition 3

First, note that the summation over  $k$  in  $p_A$  in Proposition 2 goes from 0 to  $\lceil K/2 \rceil - 1$ . If  $K$  is an odd number,  $\lceil K/2 \rceil = \lceil (K + 1)/2 \rceil$ . For an odd  $K$ , this implies that  $p_A(g_M, K, \sigma) \geq p_A(g_M, K + 1, \sigma)$ . As we wish to maximize  $\sum_{M \subset \mathbb{M}} q_M p_A$ , the rest of the proof restricts attention to odd values of  $K$ . With this restriction,  $A$ 's vote share becomes

$$S(K, \sigma) = \sum_{M \subset \mathbb{M}} q_M \sum_{k=0}^{(K-1)/2} \binom{K}{k} ((1 - g_M)\sigma)^k (1 - (1 - g_M)\sigma)^{K-k}$$

Define:

$$p_M = (1 - g_M) \bar{\sigma}(1)$$

For clarity, replace  $\mathbb{M}$  with  $\{0, \dots, n\}$  and  $M$  with  $i$ . Also, order the  $p$ 's by decreasing value (if two segments have the same  $p$ , merge them):

$$p_0 > p_1 > \dots > p_n.$$

Each  $p_i$  is associated with a weight  $q_i$ . The assumption of the proposition is that  $q_0 \geq \frac{1}{2}$ . Also, by the definitions of  $S(K)$  and  $\bar{\sigma}(1)$ :

$$S(1, \bar{\sigma}(1)) = \sum_{i=0}^n q_i (1 - p_i) = \frac{1}{2}.$$

Given any odd  $K$ , define

$$b(p_i) = \sum_{k=0}^{(K-1)/2} \binom{K}{k} p_i^k (1 - p_i)^{K-k}$$

The function  $b(p_i)$  corresponds to the probability that the realization of the random variable  $k$  is lower than the midpoint  $K/2$ . We have

$$S(K, \bar{\sigma}(1)) = \sum_{i=0}^n q_i b(p_i).$$

To prove that  $\bar{\sigma}(K) \leq \bar{\sigma}(1)$ , it is sufficient to show that

$$\sum_{i=0}^n q_i b(p_i) \leq \frac{1}{2}. \tag{7}$$

It is easy to check that  $b(\cdot)$  is continuous, twice differentiable, and decreasing, with  $b(0) = 1$  and  $b(1) = 0$ . It also satisfies the following properties:

- (i) Antisymmetry: for any  $p \in [0, 1]$ ,  $b(p) = 1 - b(1 - p)$  (implying that  $b(\frac{1}{2}) = \frac{1}{2}$ );
- (ii) S-shape: If  $p < \frac{1}{2}$ ,  $b''(p) < 0$  and if  $p > \frac{1}{2}$ ,  $b''(p) > 0$ .

Property (i) is obtained by replacing  $k$  with  $K - k$  and  $p$  with  $1 - p$ . Property (ii) is proven by noting that:

$$b'(p_i) = \sum_{k=0}^{(K-1)/2} \binom{K}{k} k p_i^{k-1} (1 - p_i)^{K-k} - (K - k) p_i^k (1 - p_i)^{K-k-1}$$

and

$$\begin{aligned}
& b''(p_i) \\
&= \sum_{k=0}^{(K-1)/2} \binom{K}{k} k(k-1) p_i^{k-2} (1-p_i)^{K-k} - 2k(K-k) p_i^{k-1} (1-p_i)^{K-k-1} \\
&+ (K-k)(K-k-1) p_i^k (1-p_i)^{K-k-2} \\
&= \frac{1}{4} (K^2 - 1) p_i^{\frac{1}{2}K - \frac{3}{2}} (1-p_i)^{\frac{1}{2}K - \frac{3}{2}} (2p_i - 1),
\end{aligned}$$

which is positive if and only if  $p_i > \frac{1}{2}$ , thus proving property (ii).

Next, partition the set  $\{0, \dots, n\}$  into subsets  $I_A = \{i : p_i < \frac{1}{2}\}$  and  $I_B = \{i : p_i \geq \frac{1}{2}\}$ . Because of property (ii),  $b(\cdot)$  is concave for all  $p_i$  with  $i \in I_A$ . Therefore,

$$\sum_{i \in I_A} q_i b(p_i) \leq q_A b(p_A), \tag{8}$$

where  $q_A$  and  $p_A$  are defined such that

$$p_A q_A = \sum_{i \in I_A} p_i q_i \text{ and } q_A = \sum_{i \in I_A} q_i.$$

Let  $p_B = 1 - p_A$ . As  $q_0 \geq \frac{1}{2}$  and  $\sum_{i=0}^n p_i q_i = \frac{1}{2}$ , it must be that  $p_0 \leq p_B$ . Again, because of property (ii),  $b(\cdot)$  is convex for all  $p_i$  with  $i \in I_B$ . Therefore,

$$\sum_{i \in I_B} q_i b(p_i) \leq q_C b\left(\frac{1}{2}\right) + q_B b(p_B) = \frac{1}{2} q_C + q_B b(p_B), \tag{9}$$

where  $q_C$  and  $q_B$  are defined such that

$$\frac{1}{2} q_C + q_B p_B = \sum_{i \in I_B} q_i p_i \text{ and } q_C + q_B = \sum_{i \in I_B} q_i.$$

Note that by construction

$$q_A p_A + \frac{1}{2} q_C + q_B (1 - p_A) = \frac{1}{2} \text{ and } q_A + q_C + q_B = 1,$$

which together imply  $q_A = q_B$ .

Putting together (8) and (9), we have

$$\begin{aligned}
\sum_{i=0}^n q_i b(p_i) &\leq q_A b(p_A) + \frac{1}{2} q_C + q_B b(p_B) \\
&= q_A b(p_A) + \frac{1}{2} q_C + q_A (1 - b(p_A)) \\
&= \frac{1}{2},
\end{aligned}$$

which corresponds to (7).

## Appendix B: Additional Specifications of the Media Power Index

This section reports media power index values for five possible extensions of the baseline specification. The results are reported in Table 5. For comparison, the first column of the table lists the value of the baseline media power index.

### Previous Years

Although we have data from 2000 to 2012, only the 2010 and 2012 editions of the survey cover all major news platforms: daily and weekly press, radio, television, and websites. Previous years are less complete and quite sparse in the earliest editions. In 2000, the survey covered only a limited set of sources: news-only TV channels, radio stations, and magazines. In 2002, the survey was extended to the three major networks. In 2004, daily newspapers were added (as an aggregate source). In 2008, the survey added websites. In 2010 it began asking about the four major newspapers.

Column (2) in Table 5 reports the results for 2010. The differences with 2012 are relatively small. Interestingly, the power of all four daily newspapers for which we have individual data is increasing, and the increase is all due to internet access. The *Huffington Post* went from 0.009 to 0.025. News aggregators appear to lose ground, possibly because internet users are more likely to access original sources directly.

### Usage Frequency

For the baseline results, consumption was defined in terms of daily, or almost daily, usage. One could instead define consumption less stringently, for instance in terms of weekly usage. We therefore include all daily sources that are followed “regularly” or “sometimes” and all weekly sources that are followed “regularly”.

Column (3) reports the power index  $\Pi$  for weekly usage. There are two distinct effects. First, new weekly sources are considered, like *Time* (which in 2012 belonged to Time Warner) or *The Economist*.<sup>16</sup> However, weekly magazines have low readership. Second, the power of daily sources may increase or decrease, depending on whether the number of casual followers is relatively larger than the number of regular followers.

Moving to weekly usage has a limited effect. The relative strength of the three conglomerates is slightly larger, and, as one would expect, the power of daily newspapers is slightly lower. The power index of a hypothetical owner of all US newspapers but the major four would go down from 0.262 to 0.159.

---

<sup>16</sup>The questions on weekly magazines are of the form “*Time*, *Newsweek*, or similar?” Individual shares cannot be disentangled. We assign all readers in the category to *Time*. Therefore, the media power of Time Warner is overestimated.

## Extreme Worst Case

The media power index obtained in Proposition 3 is the worst case for any possible constant bandwidth  $K$  but things could be even worse if  $K$  were allowed to differ across voters. This possibility was mentioned at the end of Section 3 and it is explored in depth in Online Appendix I.

Column (4) reports the values of this “extreme” worst case power index. The values are much higher than in column (4) especially so for smaller organizations (using the extreme index has no effect for voters that only used one source). However, the relative ranking is almost unchanged, with the exception of some sources that were already close, like CBS and MSN.

## New Generations

Media consumption habits depend on age. Older generations, accustomed to print media, radio, and television, are less likely to consult new media for political information. This may provide some guidance on the direction of the evolution of media power in future years, as the weight of older generations is reduced. One basic way of approaching this question is to ask what media power patterns are if we restrict attention to voters who are younger than 40. This corresponds to assuming that all older voters are no longer around and that the media consumption patterns of under 40s remain stable.

A number of patterns – reported in Column (5) – emerge. The three traditional networks – CBS, NBC, and ABC – as well as Fox News and MSNBC are much less powerful among younger generations. The only two broadcast media who do better are CNN and NPR. As one would expect, new media fares much better, with Yahoo News in third position and MSN, Google, Facebook, and the *Huffington Post* gaining ground. Dailies overall are much less popular in younger generations. However, perhaps surprisingly, the two leading national dailies – the *New York Times*, and the *Wall Street Journal* – do well. The *Times* becomes the 7th most powerful source. This is due to their digital readership. Other high-brow news sources, like Bloomberg, Associated Press, and Reuters, also gain ground.

This said, even the dramatic assumption that all US voters over 40 are disenfranchised (the age of the median American voter is approximately 45) is not enough to revolutionize the US media power landscape. Television conglomerates still occupy three of the four top spots. While younger generations have different media consumption habits from the old, this difference alone does not lead to a radical subversion of power rankings.

## Media Power in State Elections

All the power indices above are computed for the whole of the United States. They measure the influence of media organizations on federal elections. One may wonder what media sources exert influence over elections for officials elected in a particular state. This question can be answered in a similar way by restricting attention to the set of Media Consumption Interview subjects who reside in the state of interest. This point can be illustrated by looking

at New York State.<sup>17</sup> Restrict attention to the fraction of respondents who are New York residents (approximately 7%) and perform the same analysis of column (1) on this restricted sample.

The results, reported in column (6), highlights two sets of effects. First, we see that local New York sources have greater index values. The *New York Times* more than doubles its power and is now the fourth most powerful media organization, behind the three conglomerates. Second, in line with the state's political leaning, we see a shift away from right-leaning media sources. News Corporation is now ranked third behind CNN and Comcast. *Rush Limbaugh* loses 80% of its power.<sup>18</sup>

---

<sup>17</sup>The Media Consumption Survey asks detailed questions about local newspapers but does not identify them by name, except for the *New York Times*, the *Washington Post*, *USA Today*, and the *Wall Street Journal*. Therefore, the only state where we can actually identify two important daily sources is New York.

<sup>18</sup>The state-level power indices must be interpreted with caution. As the media sources in the Pew sample are chosen because they are national in scope, they may provide only limited information on state-level politics. This will underestimate the power of local sources and overestimate the power of national sources.

Table 5: Appendix: Media Power Indices

<b>Media Organization</b>	(1) Baseline II	(2) 2010	(3) Weekly Usage	(4) Extreme Worst Case	(5) Younger than 40	(6) New York State
News Corporation (Fox, WSJ)	0.221	0.228	0.244	0.419	0.174	0.120
Comcast (NBC, MSNBC)	0.153	0.144	0.196	0.333	0.103	0.159
Time Warner (CNN, Time)	0.086	0.096	0.107	0.240	0.107	0.134
ABC	0.066	0.063	0.073	0.170	0.025	0.049
NPR	0.059	0.049	0.058	0.170	0.084	0.028
Yahoo	0.057	0.067	0.021	0.162	0.108	0.061
CBS	0.036	0.035	0.045	0.111	0.017	0.055
MSN	0.033	0.043	0.014	0.122	0.056	0.031
New York Times	0.031	0.026	0.032	0.101	0.052	0.067
Viacom	0.028	0.034	0.069	0.081	0.077	0.032
PBS	0.025	0.017	0.043	0.102	0.007	0.023
Rush Limbaugh	0.022	0.016	0.031	0.082	0.022	0.004
Google	0.022	0.023	0.008	0.071	0.047	0.036
USA Today	0.020	0.015	0.042	0.060	0.024	0.031
AOL	0.013	0.020	0.004	0.039	0.012	0.031

Source: Pew Media Consumption Survey. All data for the year 2012 except where noted.

1. Media power index II as in Proposition 3, computed on the basis of daily usage for all the U.S.
2. As in (1) but computed for 2010.
3. As in (1) but computed on the basis of weekly usage.
4. As in (1) but using the extreme worst-case media power index computed under the assumption that bandwidth can differ across segments.
5. As in (1) but restricted to respondents under 40 years of age.
6. As in (1) but restricted to residents of New York State.

# Online Appendix

## Online Appendix I: Power with Unknown Bandwidth

Proposition 3 characterizes the worst-case power index under the assumption that the bandwidth  $K$  is the same in all segments. We now turn to the worst-case scenario under the assumption that the vector of bandwidth  $K$  is potentially different for different segments. To compute the power of media group  $G$ , we must ask what vector of bandwidths maximizes the value of  $\sigma$  such that Candidate  $A$  is still elected.

We begin by re-stating Proposition 2 in a way that allows for different segments to have different bandwidth. Let  $K_M$  be the bandwidth of voters in segment  $M$ .<sup>19</sup> Let  $\mathbf{K}$  be the vector of all bandwidths.

Proposition 2 says that the power of group  $G$  is  $\Pi(\mathbf{K}) = 2\bar{\sigma}(\mathbf{K}) - 1$ , where  $\bar{\sigma}(\mathbf{K})$  is the minimum between one and the smallest solution greater than  $1/2$  of :

$$\sum_{M \subset \mathbb{M}} q_M p_A(g_M, K_M, \sigma) = \frac{1}{2},$$

where

$$p_A(g_M, K_M, \sigma) = \sum_{k=0}^{\lceil K/2 \rceil - 1} \binom{K_M}{k} ((1 - g_M)\sigma)^k (1 - (1 - g_M)\sigma)^{K_M - k}. \quad (10)$$

The key observation – which is shown formally in the proof of Proposition 5 – is that, for every value of  $g_M$  and  $\sigma$ , the maximal value of  $A$ 's vote share  $p_A(g_M, K_M, \bar{\sigma})$  is achieved when either  $K_M = 1$  or  $K_M \rightarrow \infty$ . This means that the upper envelope of  $A$ 's vote share over  $K_M$  is:

$$\max(p_A(g_M, 1, \sigma), p_A(g_M, \infty, \sigma))$$

This property becomes apparent in Figure 2. For every value of  $\sigma$ , the largest value of  $A$ 's vote share corresponds to either  $K_M = 1$  or  $K_M \rightarrow \infty$ . While this property simplifies the analysis, it is useful to keep in mind that it holds for a particular candidate quality  $\sigma$  and particular media attention share  $g_M$ .

Let  $\bar{\sigma}$  be the highest quality of candidate  $B$  for which  $M$  can still get candidate  $A$  elected. This is the same definition as in the previous section, except that now the maximal value is computed over all possible  $K$ -vectors. Similarly, the worst-case power index is defined as  $\bar{\Pi} = 1 - \bar{\sigma}$ . We are now ready to state the second main result of the paper:

**Proposition 5** *The power of group  $G$  is given by  $\bar{\Pi} = 1 - \bar{\sigma}$  where  $\bar{\sigma}$  is the minimum between one and the largest solution of*

$$\sum_{M \subset \mathbb{M}} q_M \max(p_A(g_M, 1, \sigma), p_A(g_M, \infty, \sigma)) = \frac{1}{2}$$

---

<sup>19</sup>If a segment contains voters with different bandwidths, it can in turn be split into multiple segments.

with

$$p_A(M, 1, \sigma) = 1 - (1 - g_M) \sigma$$

and

$$p_A(M, \infty, \sigma) \equiv \begin{cases} 0 & \text{if } (1 - g_M) \sigma \geq \frac{1}{2} \\ 1 & \text{if } (1 - g_M) \sigma < \frac{1}{2} \end{cases}$$

**Proof.** We wish to show that, for any given  $g_M$ ,  $K_M$ , and  $\sigma$ , the value of  $p_A(g_M, K_M, \sigma)$  is maximized either when  $K_M = 1$  or when  $K_M \rightarrow \infty$ .

Recall from the proof of Proposition 3 that, for any odd  $K$ ,  $p_A(g_M, K, \sigma) \geq p_A(g_M, K + 1, \sigma)$ . Therefore, if we wish to maximize  $A$ 's vote share we can focus on odd values of  $K$ .

Suppose we start with an odd positive integer  $K_M$  and we increase it by two to  $K_M + 2$ . As  $K_M$  is odd, we can write  $\lceil K_M/2 \rceil - 1 = (K_M - 1)/2$ : as we focus on odd values of  $K_M$ , we have

$$p_A(g_M, K_M, \sigma) = \sum_{k=0}^{(K_M-1)/2} \binom{K_M}{k} ((1 - g_M) \sigma)^k (1 - (1 - g_M) \sigma)^{K_M-k}$$

if we add two more signals, the probability that the majority of signals are in favor of  $B$  becomes

$$p_A(g_M, K_M + 2, \sigma) = \sum_{k=0}^{(K_M+1)/2} \binom{K_M + 2}{k} ((1 - g_M) \sigma)^k (1 - (1 - g_M) \sigma)^{K_M+2-k}$$

Let  $x$  be the number of signals favorable to  $B$  that the voter observed with  $K_M$ . If either  $x \leq (K_M - 3)/2$  or  $x \geq (K_M + 3)/2$ , then the two new signals cannot change the voter's decision.

If  $x = (K_M - 1)/2$ , the voter changes his decision (votes for  $B$  instead of  $A$ ) if he gets two signals in favor of  $B$ , which happens with probability  $((1 - g_M) \sigma)^2$ . If  $x = (K_M + 1)/2$ , the voter changes his decision (votes for  $A$  instead of  $B$ ) if he gets two signals in favor of  $A$ , which happens with probability  $(1 - (1 - g_M) \sigma)^2$ . With the two additional signals, the probability that  $A$  is elected decreases if and only if

$$\begin{aligned} & \Pr(x = (K_M - 1)/2 | K_M \text{ signals}) ((1 - g_M) \sigma)^2 \\ & - \Pr(x = (K_M + 1)/2 | K_M \text{ signals}) (1 - (1 - g_M) \sigma)^2 \\ & > 0 \end{aligned}$$

The inequality above corresponds to

$$\begin{aligned} & \binom{K_M}{K_M/2 - 1} ((1 - g_M) \sigma)^{K_M/2-1} (1 - (1 - g_M) \sigma)^{K_M/2+1} ((1 - g_M) \sigma)^2 \\ & > \binom{K_M}{K_M/2 + 1} ((1 - g_M) \sigma)^{K_M/2+1} (1 - (1 - g_M) \sigma)^{K_M/2-1} (1 - (1 - g_M) \sigma)^2. \end{aligned}$$

Noting that

$$\binom{K_M}{(K_M - 1)/2} = \binom{K_M}{(K_M + 1)/2},$$

and performing some simplifications, we re-write the inequality as

$$((1 - g_M) \sigma)^{K_M/2+1} (1 - (1 - g_M) \sigma)^{K_M/2+1} (2(1 - g_M) \sigma - 1) > 0.$$

The two additional signals increase the the probability that  $B$  is elected if and only

$$(1 - g_M) \sigma > 1/2.$$

As this condition does not depend on  $K_M$ , for every  $M$ , every  $g_M$ , and every  $\sigma$ , the maximal  $K_M$  must be either the lowest or the highest possible value. ■

One can also compute the worst-case power index for each other news source, which is reported in the table below along with the minimal-bandwidth and maximal-bandwidth indices:

Index	Tv1	Tv2	Np1	Np2	Np3	Web1	Web2
$\bar{\Pi}(1)$	0.333	0.164	0.176	0.090	0.101	0.176	0.152
$\bar{\Pi}(\infty)$	0	0	0	0	0	0	0
$\bar{\Pi}$	0.429	0.481	0.250	0.333	0.333	0.500	0.154

Given the discontinuous nature of the upper envelope function used to compute  $\bar{\Pi}$ , we were unable to find an analytical method for finding the value of the power index. However, as the function is monotonic and piecewise linear, a numerical method provides a fast and accurate approximation of the value of the index.<sup>20</sup>

## Online Appendix II: Hypothetical Newspaper Merger

The third example relates to the press. In April 2016, Gannett made a bid to acquire Tribune Publishing. Tribune Publishing owns the *Chicago Tribune*, the *Los Angeles Times*, and several other local titles. Gannett owns *USA Today*, the 3rd largest US newspaper by circulation and the 14th most powerful US source according to Table 3. It also owns dozens of other local US newspapers, the largest of which is the *Arizona Republic*.<sup>21</sup>

As discussed above, the Pew survey only allows respondents to name four US dailies. *USA Today* is one of them, but the other newspapers involved in the potential merger are not. To provide a rough estimate of the power of the merged company, we can assume that the ratio between circulation and attention is similar across all those titles. We know the value of

<sup>20</sup>The Stata approximation algorithm is available from the author upon request.

<sup>21</sup>The circulation data comes from Alliance for Audited Media, and it refers to weekday circulation in 2015. We exclude titles with a circulation of less than 50,000: this leads to 18 titles for Gannett and 14 titles for Tribune Publishing.

that ratio for the four newspapers in the Pew survey. Every million copies correspond to  $x$  percentage points of attention share, where  $x$  is 1.35 for the *New York Times*, 0.75 for *USA Today*, 1.35 for the *Wall Street Journal*, and 1.23 for the *Washington Post*. We therefore assume that the ratio is in the 0.75-1.35 range for all newspapers owned by Gannett and Tribune (except for *USA Today* for which we use the actual ratio).

The results of this exercise are summarized in Table 6. The low scenario is based on an attention-circulation ratio of 0.75, while the high scenario is based on a ratio of 1.35. In the high scenario, the merged entity would have a power index of almost 0.1, which is quite high but still smaller than all the television merged entities examined above, with the exception of CBS-Viacom.

		II	
	Circulation	Low	High
Gannett	4,274,663	0.033	0.044
Tribune Publishing	3,337,116	0.026	0.047
Merged entity	7,611,779	0.061	0.098

Table 6: Hypothetical merger between two US newspaper companies.