Competition among mass media*

Maksymilian Kwiek†

February 13, 2010

Abstract

This paper investigates how mass media provide information to readers or viewers who have diverse interests. The problem of a mass medium comes from the fact that there is a constraint on how much information can be delivered.

It is shown that the mass medium optimally provides information that is somewhat useful to all agents, but not perfect to anybody in particular.

This benchmark model is then used to investigate competition among mass media with differentiated products. In the equilibrium of the example studied, mass media differentiate their news fully, as if they were monopolies on the subset of readers to which they tailor their news. However, prices are disciplined by competition.

Keywords: Mass media, product differentiation, news, cheap talk, quantization

JEL Classification: D83, L11, L82

*Draft. Parts of this paper have been circulated under the title "Mass media: constrained information and heterogeneous public".

†University of Southampton, M.Kwiek@soton.ac.uk

1 Introduction

There is a growing literature on why and how mass media may systematically distort their picture of reality. Mass media do report only some facts and omit others. The selection of topics, news, even selection of words is not random—and it is often taken for granted that this phenomenon is undesirable.

This study will provide a possible answer in what sense and how mass media reporting deforms the picture of reality. The model presented here will have multiple mass media taking very different positions from each other. However, this distortion may be efficient, given the preferences of the population of readers and various technological constraints.

The model studied in this paper focuses on the defining feature of a mass medium – that it provides information to masses. The two basic premises are that readers are impatient in accessing news and that they are heterogeneous in what type of news they want to learn. Impatience implies that a message of a mass medium must be short. Heterogeneity of the readers means that a short message cannot inform everyone perfectly. This creates a trade-off: should a mass medium focus on a few readers and provide a high quality of signal to them (and cease to be a mass medium) or present a shallow message, informative to all but only somewhat?

This trade-off is investigated in two contexts. The first one is the analysis of the optimal behavior of a monopolistic mass medium facing the above constraint on the information transmission. If the marginal cost of informing symmetrically heterogeneous readers is low enough, the mass medium optimally chooses to inform all the readers superficially, rather than to focus on one or a few of them. By extending the readership, the newspaper provides a signal that is less informative to the inframarginal readers, which creates two opposing effects. On one hand the newspaper sells to more readers, but on the other hand it extracts a lower price from an individual reader. This is a classical trade-off of the monopolist—and here the result is unambiguous—it is better to increase quantity of readers rather than the quality of a signal and price to an individual reader. In other words, the optimal information policy implies elastic demand. The surprising element is that the correlation of readers' situations does not have to be large enough; even if readers are completely statistically independent, the newspaper chooses to inform them all in an imperfect way.

The second context is a duopoly, where two mass media compete with each other in choosing information policies and prices, somewhat in spirit of Hotelling model of product differentiation. There is a multiplicity of equilibria, but in all of them newspapers behave as monopolies in choosing their informational policies for their readerships. The effect of competition is twofold. Firstly, prices are lower relative to monopoly. Secondly, the readership sizes cannot be too asymmetric. This is because otherwise, the more specialized newspaper would be able to capture some of the readers of the more popular and low quality newspaper and still extract a high enough price.

This study is able to illuminate a related and very rich discussion of whether and why mass media are biased. Mullainathan and Shleifer (2005) assume that individual readers derive utility from reading the news confirming their initial bias. Bias is a taste parameter and heterogeneity comes form the fact that different readers are biased differently. This transforms the model back into a Hotelling model of differentiated products. A monopolist would like to position its policy "in the middle" of the bias spectrum, while firms in duopoly would position themselves at the extremes. The authors conclude that competition, contrary to many opinions, may as well lead to greater bias in the news, measured as a distance between the real state of nature and the reported news.

Here the consumers are fully rational news readers. Each has her own guessing problem and purchases a newspaper expecting that it delivers a useful signal. Heterogeneity comes form the assumption that different readers are interested in different guessing problems, which may be somewhat correlated. For instance, one reader may be interested in the weather in Chicago, another in the weather in Boston, another in Cleveland and so on. Superficially, the results look similar to Mullainathan and Shleifer (2005): monopolistic newspaper delivers the weather report to all, even very different locations (Chicago, Boston, Cleveland and all other cities), while firms in duopoly tailor their news to their segments and deliver news that are specialized to their markets (maybe weather for Midwest and the East Coast). However, a duopolistic firm provides a signal that is objectively better to an individual reader than a single-newspaper monopolist does. This cannot be interpreted as biasing news, which is very different to the interpretation given by Mullainathan and Shleifer (2005).

One cannot say a priori that some readers receiving favourable news about Democratic party, and others receiving favorable news about Republican party is a worse outcome than the one in which all readers receive balanced news about both parties. This is for the same reason as there is no point of forcing football fans and motosport fans to read a bit about football and a bit about motosport. If people prefer to read slanted news then slanted news increase social welfare.

It is true that the worry about biased news may be justified in real world. But this is because of the externalities that news may have through actions of readers. Reading favorable news about Democratic party affects only private welfare, but if political action of the readers is affected by such news, then the newspapers slanting the message may create externality, and that is a concern. Since such an externality may be negative or positive, one cannot say if slanting the news is bad or good. Consider a newspaper providing slanted news about the environment or education. If that message leads the readers to pollute less or increase the general level of human capital, then the central planner should increase such bias.

In assessing social desirability of market's outcomes, this study assumes no externality and employs the usual efficiency criterion (and hence the results should be interpreted carefully when applying to situations with externalities, such as media bias in politics).

Thanks to the simplicity of the model, both positive results and welfare analysis are clear-cut. Since all readers buy at most one newspaper and since the entire market will be covered, monopolist's price is not affecting the total welfare. This means that the monopolist's problem of constructing revenue-maximizing information policy is the same as central planer's problem of constructing welfare-maximizing information policy. Consequently, monopolist's outcome is efficient. Two competing newspapers thus cannot provide better information than a monopolist which owns two newspapers. However, symmetric duopoly equilibrium will be able to match the efficient outcome. In addition to this, in any equilibrium duopoly results in a greater consumer surplus. On the other hand, increasing the number of newspapers from one to two is unambiguously good: a duopoly in the least efficient equilibrium is more efficient than a single-newspaper monopoly.

Remaining literature

Apart from already mentioned Mullainathan and Shleifer (2005), there is a number of recent articles analyzing media markets. For instance, Gentzkow and Shapiro (2006) study information bias caused by newspapers' concern about their reputation. Namely, newspapers may withhold information that is true but sounds implausible given readers' initial prior belief. Barron (2006) investigates media bias induced by journalists' career concern. These papers report other studies as well.

Quantization is the process of converting an input from a rich state space into a finite number of discrete values. The properties of various quantization methods are studied in information theory for purpose of coding, compressing or digitalization (see Gray and Neuhoff (1998), or Gersho and Gray (1991)). This literature is very closely related to the model in this paper, where the mass medium tries to represent a complex state of nature in a shorter form. Of a particular importance here is vector quantization, where state space is multi-dimensional; in this paper dimensions will represent different aspects of reality that are important to different agents. Information theory is mostly interested in asymptotic cases where transmission consists of "large" number of messages/codewords. This paper however opts for a (mostly) symmetric model with two messages and two states for each agent, in order to keep as many formulae explicit as possible.

Daw (1991) investigates a decision maker with bounded memory, who has to decide whether to remember a lot about the result of one experiment or rather to remember something about both. In contrast to this study, Daw's results are that the strategy focusing on one item is better. Apart from Daw there are very few papers in economics considering a constraint on information transmission. For instance, a number of studies emerged on persuasion, started by a few papers by Glazer and Rubinstein (for instance 2004). The key assumption in this literature is that there is a limit on how many pieces of evidence can be revealed – which ones are revealed is endogenous. This strand differs from the present study to the extent that here mass media are assumed to be unbiased. A similar feature is shared by a number of studies of "advice", where a complex state of the world is transformed in to a simple recommendation. For a recent study, see Gill and Sgroi (2008) and a literature overview there. Veldkamp (2006) analyzes financial frenzies induced by media and the importance of a trade-off between learning about one market or another. "The trade-off takes the form of a constraint on number of signals each reader can purchase. Such a constraint could be interpreted as limited space in newspapers or limited time to read each piece of information" (Veldkamp (2006) page 585).

Irmen and Thisse (1998) considers a general Hotelling model, not specific to mass media, where product characteristics are multidimensional. In their equilibrium the producers want to differentiate their products in only one dimension, but the product characteristics converge in all other dimensions. The model presented here is different and it brings a very different conclusion; the newspapers want to differentiate along all dimensions. Interestingly, Irmen and Thisse (1998) use the competition between Newsweek and Time as their real-life example.

2 Model: monopolistic newspaper

The set of all potential uniformed readers is $\mathcal{N} = \{1, ..., N\}$.[1] A reader (receiver, TV viewer, agent, consumer, internet surfer) of type $n = 1, ..., N$ cares only about dimension n of state $s \in S$. Thus a dimension n can be also called an issue or aspect of reality, that is relevant to reader n.

The state space is $S = \{0,1\}^N$, where N is the number of potentially relevant aspects of reality, and each aspect of reality may be either low, represented by zero, or high, represented by one. Index $m = 1, ..., 2^N$ enumerates all vectors in S, so that m'th possible realization of the state is $s^m = (s_1^m, ..., s_N^m) \in S$ (This index will be ignored where possible). Vectors consisting of N ones or zeros will be denoted 1^N or 0^N respectively. The probability of s is given by $q(s)$. Let $Q_n(s_n) = \sum_{\{m: s_n^m = s_n\}} q(s^m)$ be the marginal distribution of nth dimension.

One property will be assumed about the distribution function for the entire presentation.

Assumption 1 *For any n, marginal distribution is uniform, that is*

$$\frac{1}{2} = Q_n(0) = Q_n(1)$$

This assumption merely says that in the absence of any additional information such as provided by the newspaper, a reader is uncertain about her state. If this marginal distribution is strongly skewed in one direction, then the newspaper is unlikely to change the view of the reader and hence has little value to her.

[1] Notation: the set $\mathcal{N}$ will have N elements, the set $\mathcal{N}_0$ will have N_0 elements etc.

Reader's action is $a_n \in \{0,1\}$, and her state-dependent and action-dependent loss function (negative utility) gives her one if she does not guess her state correctly and gives her zero otherwise. Net utility is equal to negative expected loss minus the payment.

The assumption that lies at the hart of this model is that readers are busy and will not read possibly long messages provided by the newspapers. Assume that even if the messages consist of long strings of binary digits, readers will only read the first digit. The problem is that the message space is too coarse relative to the dimensionality of state space and if the newspaper is read by multiple and heterogeneous readers, in general it will not be perfectly tailored to any one reader. Intuitively, the world as understood by the newspaper is too complicated to be pictured accurately in simple headlines, so the newspaper faces a nontrivial decision problem of how to communicate what it knows.

There are two elements of newspaper's strategy: the choice of informational policy offered to readers followed by the pricing decision. Because the readers will read only the first binary digit of the newspaper's communication, the informational policy boils down to a partition problem – newspaper is assumed to partition the state space into two elements, $x \subseteq S$ and $y = S \setminus x$, and report in which element of the partition the true state is located, in x or in y. The report is assumed to be sincere, and the true dilemma for the newspaper is how to partition S. Hence, the focus of this note is to investigate the optimal action of the newspaper, $x \in X$, where X is the set of all subsets of S.

The production and distribution technology is simple. At the time of writing the message, the monopolistic newspaper (sender, mass medium, TV station, website) will know the true state $s \in S$. This information is delivered to readers at a cost $c \geq 0$ per reader, interpreted as printing and distribution cost. There are no other costs.

Timing in the model is as follows.

1. The newspaper publicly commits to a partition, $x \in X$.

2. The newspaper announces a take-it-or-leave-it price of the report $p \in R_{+}^{N}$ (This assumes perfect price discrimination. Later, some results involving uniform pricing will be easy to develop)

3. After readers observed the partition and the price, they decide whether to purchase the report or not.

4. The uncertainty is resolved and payoffs are realized. In particular, the newspaper learns s and sincerely writes in the report that either $s \in x$ or $s \in y$; agents who purchase the report, learn its content; and finally, each reader n takes action a_n, and payoffs are realized.

This matches the timing of subscription. Agents subscribe to a newspaper knowing that the events to be reported did not even happen yet. They do so because the newspaper has certain informational policy and it is expected to follow this policy in the future.

2.1 Example

Figure (1) shows two of many possible partitions in the case of three dimensions, $N = 3$.[2] Panel (a) shows symmetric unfocused partition

$$x = \{(0,0,0), (0,0,1), (0,1,0), (1,0,0)\}$$

Later, this partition will be called diagonal w.r.t. $(0,0,0)$ or the *main diagonal* and newspaper will be said to *target* all three readers. Panel (b) shows a partition

$$x = \{(0,0,0), (0,1,0), (0,0,1), (0,1,1)\}$$

[2] If $N = 1$ then the message space is large enough to reveal the true state of nature exactly. The case of even number of dimensions (such as $N = 2$) will be mentioned later.

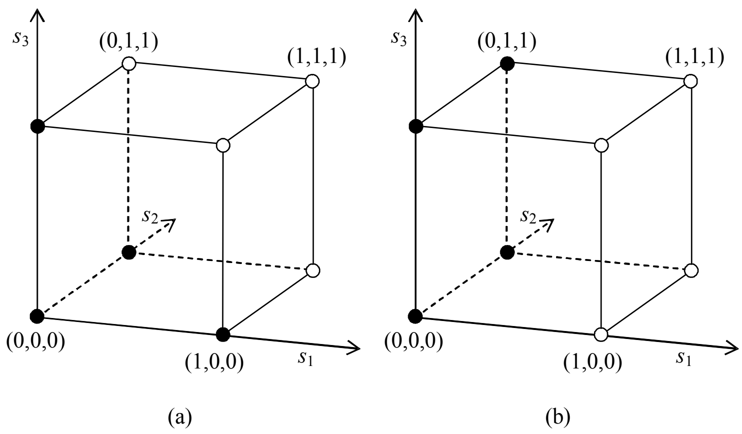

Figure 1: Case $N = 3$. Black dots represent set x. Panel (a) – unfocused strategy. Panel (b) – strategy focused on agent $n = 1$.

focused on reader $n = 1$. That is, newspaper tells only reader 1 his state, $s \in x \Leftrightarrow s_1 = 0$. Nobody else learns anything about their states beyond what is embedded in the correlation with the state of agent 1.

To continue this three-dimensional example, suppose that the distribution is uniform, $q(s) = \frac{1}{8}$ and that cost is zero, $c = 0$. One can easily find the expected loss and the value of the newspaper to each of three agents, and ultimately the gross value of this price-discriminating newspaper, for both cases shown on Figure (1).

Panel (a)	
reader	$n = 1, 2, 3$
Loss cond. on x	0.25
Loss cond. on y	0.25
Uncond. loss, L	0.25
$v_n = \frac{1}{2} - L$	0.25
Total gross value	$V = 0.75$

Panel (b)		
reader	$n = 1$	$n = 2, 3$
Loss cond. on x	0	0.5
Loss cond. on y	0	0.5
Uncond. loss, L	0	0.5
$v_n = \frac{1}{2} - L$	0.5	0
Total gross value	$V = 0.5$	

Note that the partition (a) is better for the newspaper than partition (b) as it creates higher total gorss value to the readers which can be extracted through take-it-or-leave-it prices.

2.2 The decision problem of a reader

To find an equilibrium in the general model, one proceeds backwards, starting with the problem of a reader purchasing the newspaper and taking an action in the last stage.

For each partition x, define $Q_n^x(y,0)$ to be the probability that $s \in y$ and in the same time $s_n = 0$.

$$Q_n^x(y,0) = \sum_{\{m : s^m \in y, s_n^m = 0\}} q(s^m)$$

Define $Q_n^x(y,1)$, $Q_n^x(x,0)$ and $Q_n^x(x,1)$ in the same way. For each choice of partition x, let $\bar{Q}_n^x(y) = Q_n^x(y,0) + Q_n^x(y,1)$ be the probability of set y.

Suppose that reader n buys a newspaper and learns that the state is x. Then the posterior probability that $s_n = 0$ is $\frac{Q_n^x(x,0)}{\bar{Q}_n^x(x)}$. The expected loss conditional on x is

$$\frac{Q_n^x(x,0)}{\bar{Q}_n^x(x)} a_n + \frac{Q_n^x(x,1)}{\bar{Q}_n^x(x)} (1 - a_n)$$

This is a linear objective function. If the first fraction is smaller that the second one, then the optimal choice is $a_n = 1$. Otherwise it is $a_n = 0$. In any case, the minimal expected loss conditional on x is just

$$\min \left\{ \frac{Q_n^x(x,0)}{\bar{Q}_n^x(x)}, \frac{Q_n^x(x,1)}{\bar{Q}_n^x(x)} \right\}$$

Similarly, one can define the optimal expected loss conditional on y.

Since the probability of x is $\bar{Q}_n^x(x)$, the unconditional optimal loss (before learning the content of a newspaper) is

$$L_n = \bar{Q}_n^x(x) \min\left\{\frac{Q_n^x(x,0)}{\bar{Q}_n^x(x)}, \frac{Q_n^x(x,1)}{\bar{Q}_n^x(x)}\right\} + \bar{Q}_n^x(y) \min\left\{\frac{Q_n^x(y,0)}{\bar{Q}_n^x(y)}, \frac{Q_n^x(y,1)}{\bar{Q}_n^x(y)}\right\}$$

or

$$L_n = \min\{Q_n^x(x,0), Q_n^x(x,1)\} + \min\{Q_n^x(y,0), Q_n^x(y,1)\}$$

Assumption 1 implies $\frac{1}{2} = Q_n^x(x,0) + Q_n^x(y,0)$ and $\frac{1}{2} = Q_n^x(x,1) + Q_n^x(y,1)$. Using these two equations to eliminate $Q_n^x(y,0)$ and $Q_n^x(y,1)$ from L_n, one obtains eventually

$$\begin{aligned} L_n &= \min\{Q_n^x(x,0), Q_n^x(x,1)\} + \frac{1}{2} - \max\{Q_n^x(x,0), Q_n^x(x,1)\} \\ &= \frac{1}{2} - |Q_n^x(x,0) - Q_n^x(x,1)| \end{aligned}$$

If the reader refrains from buying the newspaper, his expected loss is $\frac{1}{2}$. Hence the gross value that reader n attaches to the newspaper characterized by x is

$$v_n = |Q_n^x(x,0) - Q_n^x(x,1)| = \left|\sum_{m:s^m \in x} (1 - 2s_n^m)\, q(s^m)\right| \tag{1}$$

2.3 Price-discriminating monopolistic mass medium.

If a newspaper can price-discriminate, then the price decision for a given partition is simple. Just charge the exact gross value (1) generated by this newspaper. Corresponding revenue is equal to the total gross value $V(x) = \sum_{n=1}^{N} v_n$.

Diagonal partition with respect to $s \in S$ contains all the points that differ from s in less than half of the dimensions.

Definition 1 *Partition x is* diagonal with respect to $s \in S$ on set $\mathcal{N}_0 \subseteq \mathcal{N}$

if

$$\sum_{n\in\mathcal{N}_0} |s_n - s'_n| < \frac{N_0}{2} \Rightarrow s' \in x \text{ and}$$
$$\sum_{n\in\mathcal{N}_0} |s_n - s'_n| > \frac{N_0}{2} \Rightarrow s' \in y$$

Partition is diagonal on set $\mathcal{N}_0$ *if there exists* s *such that it is diagonal w.r.t.* s *on set* $\mathcal{N}_0$. *Partition is* main diagonal on set $\mathcal{N}_0$ *if it is diagonal with respect to* 0^{N_0}. *If a partition is main diagonal on set of readers* $\mathcal{N}_0$ *then the newspaper* targets set $\mathcal{N}_0$.

As an example, consider panel (a) of Figure (1), which depicts the main diagonal partition on $\mathcal{N}$. In this case, half of the dimensions is $\frac{3}{2}$, so partition cell x should contain all points that differ from the reference point $(0,0,0)$ in at most 1 position. Indeed, x contains point $(0,0,0)$, which does not differ from the reference point, and all three points that differ from it in one position, namely points $(0,0,1)$, $(0,1,0)$ and $(1,0,0)$. This newspaper targets all three readers.

If N_0 is even then for a given reference point there may be many diagonal partitions. In this case, the definition of a diagonal partition with respect to s does not specify in what cell of the partition the points s' such that $\sum_{n\in\mathcal{N}_0} |s_n - s'_n| = \frac{N_0}{2}$ should be located in. It may be in x or in y. If N_0 is odd, then there are no such points, and so there is only one diagonal partition for each reference point.

In total there are 2^{N_0} points of reference for diagonal partitions. Note, however, that a family of diagonal partitions with respect to s is essentially the same as a family of diagonal partitions with respect to $1^{N_0} - s$, only with x and y exchanging their places – and gives the same surpluses to all agents. Taking this into account, there is 2^{N_0-1} nontrivial reference points for diagonal partitions.

The main result of this section follows. All proofs are in the Appendix.

Proposition 1 *Suppose that Assumption 1 holds. Let $q(s) > 0$ for all $s \in S$. If the newspaper sells to a nonempty subset $\mathcal{N}_0$ of readers, then*

1. *Every optimal partition is diagonal on set $\mathcal{N}_0$,*
2. *If N_0 is even, then there is a reference point s^* so that optimal x is diagonal w.r.t. s^* and all diagonal partitions w.r.t. s^* are optimal.*

This Proposition deals with the case of sunk costs of cN_0, so optimality here means maximization of the revenue, or total gross surplus from $\mathcal{N}_0$ readers. The first part states the necessary condition for optimality of a partition – that the partition must be diagonal. The second part clarifies the case of even number of readers; the states that are "on the border" of an optimal x can be included in any cell of the partition and that will not change the gross value and hence the profit.

This goes far in determining the optimal strategy of a newspaper, but still, there is 2^{N_0-1} candidates for optimal partition (if N_0 is odd), each associated with a different reference point. Which diagonal partition is optimal will depend on distribution $q(\cdot)$. Next section adds extra assumption that will provide uniqueness.

2.4 Symmetric players and useful parametrization

It will be convenient to further restrict the family of distributions and consider the following distribution parametrized by a single coefficient $\rho \in [0,1]$,

Assumption 2 *The distribution is*

$$q(s) = \begin{cases} q_0 = \rho\frac{1}{2} + (1-\rho)\frac{1}{2^N} & \text{if } s = 0^N \text{ or } s = 1^N \\ q_1 = (1-\rho)\frac{1}{2^N} & \text{otherwise} \end{cases}$$

Parameter ρ is the correlation coefficient for any two dimensions. If $\rho = 0$ then the distribution is uniform and readers' states are independent; if $\rho = 1$

then the probability is uniform on the two extreme states, 0^N and 1^N, and readers' states are perfectly correlated.

It is easy to find the marginal distribution on the subset of all readers and this will be useful later. Let $\mathcal{N}_0 \subset \mathcal{N}$ denote this set of readers. Define the marginal distribution on $\mathcal{N}_0$, denoted $q^{\mathcal{N}_0}(s_1, ..., s_{N_0}) = \sum_{s_n : n \notin \mathcal{N}_0} q(s_1, ..., s_N)$. Obviously,

$$q^{\mathcal{N}_0}(s) = \begin{cases} q_0^{\mathcal{N}_0} = \rho \frac{1}{2} + (1-\rho)\frac{1}{2^{N_0}} & \text{if } s = 0^{N_0} \text{ or } s = 1^{N_0}, \\ q_1^{\mathcal{N}_0} = (1-\rho)\frac{1}{2^{N_0}} & \text{otherwise} \end{cases}$$

Under this more specific assumption on distribution, the unique optimal partition is the one which is diagonal with respect to 0^{N_0}, or in other words, targets a set $\mathcal{N}_0$. Note that the partition is unique for a given set of readers $\mathcal{N}_0$. But the readers are all identical so a different set of readers having the same number of elements would also lead to the same total value and therefore the same profit. This type of multiplicity is ignored here and hereafter.

Firstly, define a function

$$h(N) = \begin{cases} \binom{N-1}{\frac{N-1}{2}} & \text{if } N \text{ is odd} \\ \binom{N-1}{\frac{N}{2}-1} & \text{if } N \text{ is even} \end{cases}$$

The following result provides the explicit formulae for gross value created by the optimizing newspaper.

Proposition 2 *Suppose that Assumption 2 holds. Let $0 < \rho < 1$. If the newspaper sells to a nonempty subset of readers $\mathcal{N}_0$, then a partition is optimal if and only if the newspaper targets set $\mathcal{N}_0$. The average individual gross value of reader n is*

$$v_n(N_0) = \begin{cases} \hat{v}_n(N_0) = \rho \frac{1}{2} + (1-\rho)\frac{1}{2^{N_0}} h(N_0) & \textit{if } n \textit{ is targeted} \\ \rho \frac{1}{2} & \textit{if } n \textit{ is not targeted} \end{cases}$$

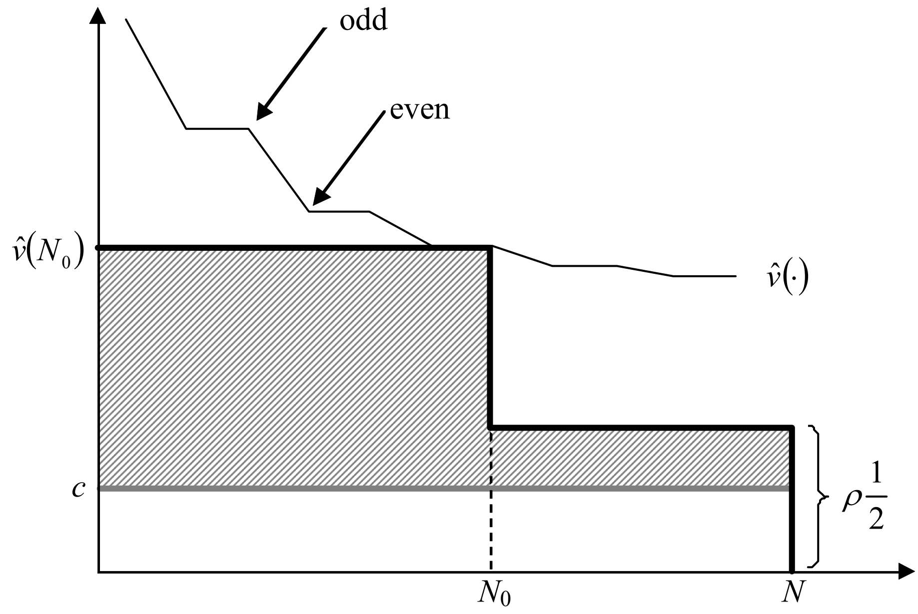

Figure 2: Function $\hat{v}_n(\cdot)$ and the profit of the newspaper.

Figure (2) shows the distribution of values created by a newspaper targeting a set of $\mathcal{N}_0$ readers. An important observation is that readers who are not even targeted by a newspaper may still find it somewhat useful. The reason is that their state is correlated with the state of a targeted reader. As the readers become more statistically independent, $\rho \to 0$, the value created by the newspaper to readers who are not targeted vanishes. On the other hand, as the readers are more and more statistically correlated, $\rho \to 1$, the value of readers who are not targeted converges to the value of targeted readers.

Some of the properties of these values can be easily computed.

Lemma 1 *Suppose Assumption 2 holds.*

1. *Suppose N_0 is even, then adding one more reader leaves the individual*

average value of a targeted reader unchanged

$$\hat{v}_n(N_0) = \hat{v}_n(N_0 + 1)$$

2. *Suppose N_0 is odd, then adding one more reader increases the total value of all targeted readers by $\rho\frac{1}{2}$*

$$\hat{v}_n(N_0)N_0 + \rho\frac{1}{2} = \hat{v}_n(N_0 + 1)(N_0 + 1)$$

3. *Suppose N_0 is odd, then adding one more reader leaves the total value of all readers unchanged*

$$\hat{v}_n(N_0)N_0 + \rho\frac{1}{2}(N - N_0) = \hat{v}_n(N_0 + 1)(N_0 + 1) + \rho\frac{1}{2}(N - (N_0 + 1))$$

This lemma shows that the case of two readers is not that interesting because newspaper's revenue is exactly the same as with one reader, if $\rho = 0$. The simplest nontrivial case involves three readers, as in the above illustrative example. One can also note that the individual gross value $\hat{v}_n(N_0)$ to a targeted reader is decreasing in odd N_0 down to $\frac{1}{2}\rho$, but that the total gross value $\hat{v}_n(N_0)N_0$ to odd N_0 is increasing.

The analysis so far focused on optimal partition for a given readership N_0. The last step is to investigate how big the readership should be. The newspaper seeks to find the best readership N_0 to maximize profit. It turns out that the behavior of this profit as a function of N_0 depends on the relationship between the correlation coefficient and average cost. If $c < \frac{\rho}{2}$ then the profit is increasing; if $\frac{\rho}{2} < c$ then the profit is decreasing for all N_0 large enough, but may be increasing for small N_0. The following $N^*(c)$ will be the candidate for optimal readership in this case

$$N^*(c) = \arg\max_{N_0 \in \{0,1,3,5,\ldots\}} (\hat{v}_n(N_0) - c)N_0$$

(Also define N^o to be N if N is odd and $N-1$ if N is even). The following proposition clarifies all cases.

Proposition 3 *Suppose assumption 2 holds.*

1. *If $c < \frac{\rho}{2}$, then profit*
$$\left(\hat{v}_n\left(N_0\right) - c\right) N_0 + \left(\frac{\rho}{2} - c\right)\left(N - N_0\right)$$
is strictly increasing. The action is optimal if and only if the newspaper targets all N (even or odd) readers.

2. *If $\frac{\rho}{2} < c$, then $N^*(c)$ exists and is generically unique. Profit $\left(\hat{v}_n\left(N_0\right) - c\right) N_0$ as a function of odd N_0 is strictly increasing for $N_0 < N^*(c)$ and strictly decreasing for $N_0 > N^*(c)$. The action is optimal if and only if the newspaper targets odd* $\min\{N^o, N^*(c)\}$ *readers.*

3. *If $c = \frac{\rho}{2}$, then profit $\left(\hat{v}_n\left(N_0\right) - c\right) N_0$ is strictly increasing between even N_0 and odd $N_0 + 1$, but is constant between odd N_0 and even $N_0 + 1$. If N is odd, then the action is optimal if and only if the newspaper targets N readers. If N is even, then targeting N readers and $N-1$ readers is optimal.*

3 Duopoly

In the first stage, newspapers publicly and simultaneously commit to their partitions. In the second stage, after observing the realization of the profile of partitions, (x_1, x_2), each announces a discriminatory price of the report, p_{jn}. This section obeys Assumption 2 throughout and focuses on a case of $\frac{N}{2}$ being an odd integer. Follow the convention that if $c \leq \frac{\rho}{2}$ (so that $N^*(c)$ does not exist), then $N^*(c) = \infty$.

The demand side works similarly as before. In particular, each reader can read only one message that can take two values, so that a reader will choose one newspaper or none at all, but will never buy two. Let $v_{jn} = \left|Q_n^{x_j}(x_j, 0) - Q_n^{x_j}(x_j, 1)\right|$ be the value that newspaper x_j creates to reader n.

Lemma 2 *For any value profile and price profile, the following reader's behavior maximizes her payoff*

1. *If $v_{jn} - p_{jn} < 0$ then n will not buy j.*
2. *Otherwise, n buys j if $v_{jn} - p_{jn} > v_{in} - p_{in}$.*
3. *If that is equal, then n buys j if $v_{jn} > v_{in}$.*
4. *If that is equal too, then choose newspaper randomly.*

From now on, the reader is assumed to behave according to this lemma. This formally defines a two-stage extensive form game with two newspapers as players. The next section analyses the Subgame Perfect Nash Equilibria of this game.

3.1 Equilibria

The equilibrium behavior in the second stage is described by the following lemma

Lemma 3 *In any equilibrium, for any gross value profile $\{v_{1n}; v_{2n}\}_{n=1,...,N}$ newspaper $i = 1, 2$ announces prices*

$$p_{in} = c + \max\{v_{in} - \max\{v_{jn}, c\}, 0\}$$

for all readers $n = 1, ..., N$, and $j \neq i$.

Intuitively, a form of Bertrand competition for each separate reader ensues. In any equilibrium, the surplus of a seller generating the lower value must be "competed away". As an illustration assume that $c = 0$ and reader

n has values such that $v_{1n} > v_{2n}$. Then newspaper 1 would announce price $p_{1n} = v_{1n} - v_{2n}$ and newspaper 2 price $p_{2n} = 0$. The net surplus of this reader is equal to v_{2n} across newspapers; so by point (3) of Lemma 2, the reader buys newspaper 1. One might say that the price (or per reader profit) of newspaper 1 is equal to the "gross value" of the newspaper to this reader, v_{1n}, like in the monopoly case, minus the "price concession" that must be granted in order to convince him to choose this newspaper, v_{2n}. If $v_{1n} = v_{2n}$ then prices of both newspapers are equal to the cost of delivery zero.

The next lemma describes behavior in the first stage. It observes that the partition that is the best response has a familiar form from the previous section,

Lemma 4 *Suppose that newspaper i best responds to a $\{v_{jn}\}_{n=1,\dots,N}$ by choosing a partition, such that there exist readers with values $v_{in} \geq \max\{v_{jn}, c\}$. Denote this set of readers by $\mathcal{N}_i$. Then newspaper i targets $\mathcal{N}_i$.*

The intuition behind this key result is simple. Since profit can be viewed as the difference between gross value and the price concession – the former depending on own partition and the latter on the partition of the opponent – maximizing profit must involve maximization of the gross value. The choice of the partition of the opponent is irrelevant.

So far the conclusion is that in any equilibrium there can be readers targeted by newspaper i only, by j only, by neither or by both. Let $f\left(N_j^e\right) = N^*\left(\hat{v}_{jn}\left(N_j^e\right)\right)$. The following proposition asserts that no reader can be targeted by both newspapers and furthermore clarifies the magnitudes of equilibrium readerships of the newspapers. It is the main result of this section.

Proposition 4 *Equilibrium in the first stage.*

1. *Let $N \leq 2N^*(c)$. Partition profile (x_1, x_2) is an equilibrium in the first stage if and only if newspaper $i = 1, 2$ targets set $\mathcal{N}_i^e$ readers such that*

   (a) $\mathcal{N}_1^e \cup \mathcal{N}_2^e = \mathcal{N}$

(b) $\mathcal{N}_1^e \cap \mathcal{N}_2^e = \varnothing$

(c) $f\left(N_j^e\right) \leq N_i^e \leq N^*(c)$, *where* $i \neq j$.

2. *Let* $2N^*(c) < N$. *Partition profile* (x_1, x_2) *is an equilibrium in the first stage if and only if newspaper* $i = 1, 2$ *targets set* $\mathcal{N}_i^e$ *readers such that*

   (a) $\mathcal{N}_1^e \cap \mathcal{N}_2^e = \varnothing$

   (b) $N_i^e = N^*(c)$.

Condition $f\left(N_j^e\right) \leq N_i^e$ deserves some explanation. Consider a profile of readerships that partitions the total set of readers and in which N_i is small and N_j is large. Newspaper j creates a low value for the readers it targets. Newspaper i may want to deviate by targeting some of these readers in addition to its own readers. After such a deviation, a reader who is targeted by both newspapers is in a really good situation, because the newspapers engage in a more aggressive price war in this market. Such reader will ultimately buy rather newspaper i, whose price is equal to

$$p_{in} = c + \hat{v}_{in}(N_i) - \hat{v}_{jn}\left(N_j^e\right)$$

This means however, that the price-cost margin of newspaper i from this reader is equal to $\hat{v}_{in}(N_i) - \hat{v}_{jn}\left(N_j^e\right)$. The profit of a newspaper deviating in such a way behaves exactly like a profit of a monopolist with unit 'cost' equal to $\hat{v}_{jn}\left(N_j^e\right)$. The number $N^*\left(\hat{v}_{jn}\left(N_j^e\right)\right)$ exists and the further away N_i is from it, the lower the profit after such a deviation. So, if initial readership N_i^e is already greater or equal to $N^*\left(\hat{v}_{jn}\left(N_j^e\right)\right)$, then targeting even more new readers that are already targeted by j will lead to lower profit, and therefore cannot constitute a profitable deviation. On the other hand, if readership N_i^e is lower than $N^*\left(\hat{v}_{jn}\left(N_j^e\right)\right)$, then targeting new readers will increase the profit and hence (N_1^e, N_2^e) is inconsistent with equilibrium.

It is straightforward to conclude that function f is an increasing step function with odd values. As it turns out, however, f can be approximated by a linear function f_a, if the argument N_j^e is large. In particular, note that $f\left(N_j^e\right) \leq N_i^e$ is equivalent to the condition that marginal gross value from increasing the readership of newspaper i from odd N_i^e to $N_i^e + 2$ does not exceed the marginal 'cost' of increasing this readership, equal to $\hat{v}_{jn}\left(N_j^e\right)$, or in other words

$$\frac{\left(N_i^e+2\right)\hat{v}_{in}\left(N_i^e+2\right) - N_i^e \hat{v}_{in}\left(N_i^e\right)}{2} \leq \hat{v}_{jn}\left(N_j^e\right)$$

Plugging in the formulae for values, one obtains

$$\frac{1}{2}\frac{1}{2^{N_i^e+1}}\frac{\left(N_i^e+1\right)!}{\left(\frac{N_i^e+1}{2}!\right)^2} \leq \frac{1}{2^{N_j^e-1}}\frac{\left(N_j^e-1\right)!}{\left(\frac{N_j^e-1}{2}!\right)^2} \tag{2}$$

Factorial can be approximated using Sterling's formula[3] leading to

$$N_i^e \geq f_a\left(N_j^e\right) = \frac{1}{4}N_j^e - \frac{5}{4}$$

Note that expression (2) implies the following result:

Corollary 1 *Function f is independent of parameters ρ and c.*

There is a convenient way to illustrate the equilibria graphically in the space where readerships are measured on two axes, as in Figure 3. A point in this space depicts readerships of newspapers j and i. Given the total number of potential readers N and provided that $N \leq 2N^*(c)$ newspapers can achieve readerships N_j^e and N_i^e only if $N_j^e + N_i^e = N$. In other words, equilibrium readerships can lie only on a a line with slope -1, such as line AB on Figure 3. The further the line lies from the origin, the higher N it represents.

In addition to that, readerships cannot be too asymmetric. Newspaper i does not have any incentives to increase her readership N_i^e as long as it

[3] $\frac{N!}{\left(\frac{N}{2}!\right)^2} \approx \sqrt{\frac{2}{\pi}}\frac{1}{\sqrt{N}}2^N$ for large N.

is above $f\left(N_j^e\right)$; that is, above point A on that line. Otherwise newspaper i would like to target some of the readership of newspaper j. On the other hand, newspaper i does not want to decrease her readership if it stays below $N^*(c)$; that is, it is below point B. Since the situation is symmetric across players, newspaper j is subjected to the same incentives. All pairs of odd readerships on a line between points A and C are equilibrium readerships.

If the economy has a little more potential readers, as represented by a line DE, then the same incentives are at work, but the condition $N_i^e \leq N^*(c)$ begins to operate before the condition $f\left(N_j^e\right) \leq N_i^e$. As the result all profiles of odd readerships on the line between E and D represent equilibria.

If the economy has more than $2N^*(c)$ readers, then in all equilibria the readership is precisely $N^*(c)$, graphically represented by point F.

3.2 Some comparative statics

The effect of change of cost c is fairly straightforward. The lower c, the cheaper the delivery of a newspaper is. In case of $N \leq 2N^*(c)$, as c falls, $N^*(c)$ goes up and the set of equilibrium readership profiles expands. If c goes below $\frac{1}{2}\rho$ then $N^*(c)$ becomes infinity. On the other hand, if $2N^*(c) < N$ then the unique equilibrium readership profile increases in line with $N^*(c)$. On the other hand, Corollary 1 makes it clear that condition $f\left(N_j^e\right) \leq N_i^e$ is not affected by c.

The change of ρ is more interesting. One effect is on $N^*(c)$ – the higher ρ, the higher $N^*(c)$.

Secondly, there may be an effect on function f. Since ρ describes the 'closeness' between any two readers or markets, one may expect that its change affects the competition between the newspapers. The larger this correlation coefficient is, the easier it should be for one newspaper to capture the readers targeted by the other newspaper. However, Corollary 1 proclaims that function f is independent of ρ. Indeed increasing ρ affects positively the incentives of newspaper i to capture the readers targeted by the opponent

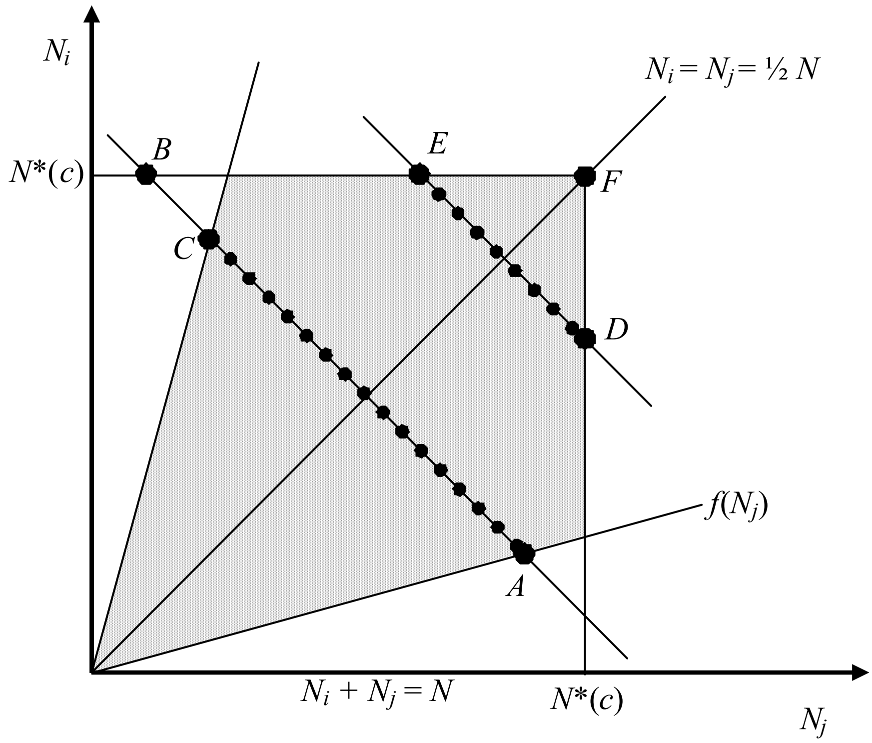

Figure 3: Duopoly. Readership profiles in equilibrium.

j in the sense that individual gross value form captured readers increases. However, the individual gross value created by newspaper j also increases, so that the price concession that newspaper i must grant after the deviation increases too. The net effect of these two opposing forces is zero.

This observation means that there is a discontinuity around point $\rho = 1$. With perfect correlation, a newspaper is as good to a targeted reader as to the one that is not targeted. Consequently, prices are equal and readers may randomize between newspapers with various probabilities. Any readership profile satisfying $N_1^e + N_2^e = N$ can be supported by an equilibrium, if $c < \frac{1}{2}$. But as soon as ρ falls to below one, condition $f\left(N_j^e\right) \leq N_i^e$ begins to have a bite, knocking out all readership profiles that are too asymmetric.

On the other hand, the competition does become fiercer as the correlation between readers increases, in the sense that for any equilibrium readership profile prices go down to c as ρ increases to 1.

4 Discussion and extensions

This section discusses a few easy extensions that revolve around classical issues in the industrial organization literature such as uniform prices, entry, collusion and efficiency. All these exercises can be performed with relatively few modifications, and proofs are skipped. More challenging modeling choices are discussed in the last part of this section.

4.1 Uniform prices

The analysis above assumes that newspapers can tailor prices to particular readers or markets. This is justified if arbitrage is not possible, for instance if newspapers are sold in different countries, or if a mass medium is not a newspaper but for instance a coded TV channel. In many cases however, one would expect that all readers can purchase a newspaper for the cheapest price around.

Assume that the model is the same as above except that prices must be the same for all readers. The result for the monopoly is immediate:

Proposition 5 *Suppose that there is one monopolistic newspaper. Choice of partition and price is optimal in the model with discriminatory prices if and only if it is optimal in the model with uniform prices.*

Since the optimal discriminatory price is uniform this result is straightforward. It strongly relies on the Assumption 2 of symmetry between readers.

However, the situation in the model of duopoly is slightly more complicated.

Proposition 6 *Suppose that there are two newspapers. If a strategy profile forms an equilibrium in the model with discriminatory prices then it forms an equilibrium in the model with uniform prices.*

This is straightforward. If no newspaper has incentives to deviate if it can choose any discriminatory price, then it cannot improve if it is restricted to use only uniform prices. On the other hand, the converse is not true – one may have an equilibrium in a model with uniform prices that is not an equilibrium in a version with discriminatory pricing. The reason is that an attempt by a newspaper to invade the readership of the other newspaper is much more difficult with uniform prices. A deviating newspaper has to offer a discount. With uniform prices, the discount has to be granted to all readers, even to high-value core readers. This effect was absent in discriminatory pricing model. As the result, the set of equilibrium readerships with uniform prices would contain the shaded set of equilibrium readerships on Figure 3 although this set would still be a proper subset of the square $[0, N^*(c)] \times [0, N^*(c)]$.

4.2 Collusive behavior and efficient newspapers

Consider an optimal collusive action of two newspapers, which not only involves collusive pricing, but also collusive choice of partitions. Alternatively, one can think of a monopolist owning two newspapers.

Proposition 7 *Every optimal collusive strategy profile is to divide all agents into two groups of size* $\frac{N}{2}$ *and then create a newspaper targeting each group. The price of the newspaper targeting a reader is the same as a monopolistic full surplus extraction. The price of the newspaper not targeting a reader is set at a level higher than the newspaper's value to this reader.*

It is not surprising that collusive choice of prices is to keep them as high as possible, to extract the entire surplus created to the readers. It is interesting, however, that collusive partition is also one of many equilibrium partitions in duopoly – the symmetric one. In other words, if an optimal cartel of newspapers is split by a regulator, then the average quality of newspapers will not improve (and may deteriorate, on average, if the duopoly ends up in asymmetric equilibrium), although the price may go down.

The behavior of the monopolistic newspaper in discriminatory or uniform pricing cases is socially efficient as it maximizes the value created by the newspaper in the information transmission. Moreover, the corollary following from the proposition is that among many equilibria in duopoly case, the symmetric ones are efficient as well.

4.3 Other modifications and future work

This study has a capacity to serve as a starting point to a few extensions that will address a number of natural and important questions. The first class of questions relate to the analysis of a model with readers' actions creating an externality. This type of model would be able to assess the merits of media regulation, for instance in the context of potentially important influence of mass media on political or social actions (see the discussion in the Introduction above). The second interesting issue relates to media that are not only interested in revenue maximization from paying readers, but also have an agenda or additional interest in a particular action of readers. This type of model would be able to say something on commercial advertising or political influence of media owners.

5 Appendix

5.1 Proof of Proposition 1

Part 1. Fix a partition x, not diagonal on set $\mathcal{N}_0$. Let $D^+ \subseteq \mathcal{N}_0$ be a set of agents (dimensions) purchasing the newspaper for whom the probability difference $Q_n(x, 0) - Q_n(x, 1)$ is nonnegative:

$$D^+ = \left\{ n : \sum_{m:s^m \in x} (1 - 2s_n^m)\, q(s^m) \geq 0 \right\}$$

Let $D^- \subseteq \mathcal{N}_0$ be the set of remaining purchasing agents. Define $s^* \in \{0,1\}^{N_0}$ to be the point such that

$$s_n^* = \begin{cases} 0 & \text{if } n \in D^+ \\ 1 & \text{if } n \in D^- \end{cases}$$

As x is not diagonal on set $\mathcal{N}_0$, it is not diagonal with respect to s^*. Therefore, either there exists $s' \in x$ such that $\sum_{n=1}^{N_0} |s_n^* - s_n'| > \frac{N_0}{2}$, or there exists $s' \in y$ such that

$$\sum_{n \in \mathcal{N}_0} |s_n^* - s_n'| < \frac{N_0}{2} \tag{3}$$

Without loss of generality, assume that the second case applies.

The revenue form including s' in the set x is

$$\begin{aligned}
R\left(x \cup s'\right) &= \sum_{n \in D^{+}} \left| \sum_{m: s^{m} \in x \cup s'} \left(1-2 s_{n}^{m}\right) q\left(s^{m}\right) \right| + \sum_{n \in D^{-}} \left| - \sum_{m: s^{m} \in x \cup s'} \left(1-2 s_{n}^{m}\right) q\left(s^{m}\right) \right| + \sum_{n \notin \mathcal{N}_{0}} 0 \\
&\geq \sum_{n \in D^{+}} \left(\sum_{m: s^{m} \in x \cup s'} \left(1-2 s_{n}^{m}\right) q\left(s^{m}\right) \right) + \sum_{n \in D^{-}} \left(- \sum_{m: s^{m} \in x \cup s'} \left(1-2 s_{n}^{m}\right) q\left(s^{m}\right) \right) \\
&= \sum_{n \in D^{+}} \sum_{m: s^{m} \in x} \left(1-2 s_{n}^{m}\right) q\left(s^{m}\right) + \sum_{n \in D^{-}} \left(- \sum_{m: s^{m} \in x} \left(1-2 s_{n}^{m}\right) q\left(s^{m}\right) \right) + \\
&\quad + \sum_{n \in D^{+}} \left(1-2 s_{n}'\right) q\left(s'\right) - \sum_{n \in D^{-}} \left(1-2 s_{n}'\right) q\left(s'\right)
\end{aligned}$$

or simply

$$R\left(x \cup s'\right) \geq R(x) + \left(\sum_{n \in D^{+}} \left(1-2 s_{n}'\right) - \sum_{n \in D^{-}} \left(1-2 s_{n}'\right) \right) q\left(s'\right) \tag{4}$$

Furthermore, inequality (3) implies

$$\begin{aligned}
\frac{N_{0}}{2} &> \sum_{n \in D^{+}} s_{n}' + \sum_{n \in D^{-}} \left(1-s_{n}'\right) \\
0 &< - \sum_{n \in D^{+}} 2 s_{n}' - \sum_{n \in D^{-}} \left(2-2 s_{n}'\right) + N_{0} \\
0 &< - \sum_{n \in D^{+}} 2 s_{n}' - \sum_{n \in D^{-}} \left(2-2 s_{n}'\right) + \sum_{n \in D^{+}} 1 + \sum_{n \in D^{-}} 1 \\
0 &< \sum_{n \in D^{+}} \left(1-2 s_{n}'\right) - \sum_{n \in D^{-}} \left(1-2 s_{n}'\right)
\end{aligned}$$

Therefore, the brackets in (4) is strictly positive. Since $q(\cdot) > 0$ as well, including s' in x will strictly increase the revenue, $R\left(x \cup s'\right) > R(x)$. This proves that a x that is non-diagonal on set $\mathcal{N}_{0}$ is not optimal.

Part 2. Given any diagonal partition x with respect to s, define

$$\check{y} = \left\{ s' \in y : \sum_{n \in \mathcal{N}_0} |s_n - s'_n| = \frac{N_0}{2} \right\}$$

Let x be an optimal partition, and hence diagonal with respect to some reference point s. Denote this partition $x(s)$ and let the revenue from this partition be $R(x(s))$. Define D^+, D^- and point s^* like above.

Note that $x(s)$ must be diagonal w.r.t. s^* too (by the same argument as in Part 1 above.) Hence we can call this partition $x(s^*)$ and note that trivially $R(x(s^*)) = R(x(s))$.

Finally note that any diagonal partition w.r.t s^* is optimal. That is, consider (w.l.o.g.) $s' \in \check{y}$ and consider partition $x(s^*) \cup s'$ which is also a diagonal partition w.r.t. s^*. Then by the same derivation as in Part 1 leading to equation 4 we have $R(x(s^*) \cup s') \geq R(x(s^*))$. Since $x(s^*)$ is optimal, these revenues are equal.

5.2 Proof of Proposition 2

Suppose that N_0 is odd. Given N_0, the cost of cN_0 is sunk. Therefore, newspaper's optimal partition must maximize the total gross value from a set $\mathcal{N}_0$.

The necessary condition for the total gross value from these readers to be maximal is that a partition is diagonal with respect to $\mathcal{N}_0$. Consider a diagonal partition w.r.t. $s^* = (0, ..., 0, 1, ..., 1)$, represented by $A = \sum_{n \in \mathcal{N}_0} s^*_n \geq 1$ a number of ones in s^*. Let $B = N_0 - A$ be the number of zeros in s^*. Such partition will be denoted by x^A, its elements have N_0 dimensions and the relevant probability is the marginal $q^{\mathcal{N}_0}$. There are two types of readers. Readers $n = 1, ..., B$ (so that $s^*_n = 0$) are type one and readers $n = B + 1, ..., N_0$ (so that $s^*_n = 1$) are type two.

The individual value from a partition represented by A is given by equation 1, and can be written as

$$v_n(A) = \left| q_0^{N_0} + q_1^{N_0} \sum_{m: s^m \in x^A \setminus 0^{N_0}} (1 - 2s_n^m) \right|$$

The value of $\sum$ on the far right counts all states in $x^A \setminus (0, ..., 0)$ that have zero at nth dimension and subtracts the number of states in $x^A \setminus (0, ..., 0)$ that have one there. The strategy of the proof is to find these two numbers for both type of readers, and then compute the values $v_n(A)$ for both type of readers.

Consider type one reader first; without loss of generality let it be reader $n = B$. Any point $s \in S$ differs from s^* in a number of dimensions. Let l'_B be the number of dimensions where point s has one, and s^* has zero except for reader $n = B$, $l'_B = \sum_{n=1}^{B-1} s_n$. Let l_A be the number of dimensions where point s has zero, and s^* has one, $l_A = \sum_{n=B+1}^{N_0} (1 - s_n)$.

Case one: $s_n = 0$. The total number of differences between s and s^* is the total number of differences on dimensions $1, ..., B-1$ (l'_B of them), on dimension $n = B$ (no differences here) and on dimensions $B+1, ..., N_m$ (l_A of them); the total number of differences is therefore $l'_B + l_A = \sum_{n=1}^{N_0} |s_n - s_n^*|$. If this point s is to be in x^A then it must be that $l'_B + l_A \leq \frac{N_0 - 1}{2}$.

There is exactly $\binom{B-1}{l'_B}\binom{A}{l_A}$ points that are represented by a given pair (l'_B, l_A). Therefore, the total number of points that are in set x^A and have $s_n = 0$ is

$$\sum_{l_A = 0, ..., A} \; \sum_{l'_B = 0, ..., \frac{N_0 - 1}{2} - l_A} \binom{B-1}{l'_B}\binom{A}{l_A}$$

The total number of points that are in set $x^A \setminus 0^{N_0}$ and have $s_n = 0$ is

$$\sum_{l_A = 0, ..., A} \; \sum_{l'_B = 0, ..., \frac{N_0 - 1}{2} - l_A} \binom{B-1}{l'_B}\binom{A}{l_A} - 1$$

Case two: $s_n = 1$. Everything is the same except that there is one more difference between s and s^*, namely at dimension $n = B$. Such a point s will be in x^A if $l'_B + l_A + 1 \leq \frac{N_0-1}{2}$. Therefore, the total number of points that are in x^A (and in $x^A \setminus 0^{N_0}$) and have $s_n = 1$ is

$$\sum_{l_A=0,\ldots,A} \sum_{l'_B=0,\ldots,\frac{N_0-1}{2}-l_A-1} \binom{B-1}{l'_B}\binom{A}{l_A}$$

Hence, the difference between these two cases is

$$\begin{aligned} \sum_{m:s^m \in x^A \setminus 0^{N_0}} (1-2s_n^m) &= \sum_{l_A=0,\ldots,A} \binom{B-1}{\frac{N_0-1}{2}-l_A}\binom{A}{l_A} - 1 \\ &= \binom{N_0-1}{\frac{N_0-1}{2}} - 1 \end{aligned}$$

and the gross value of this type of reader is

$$\begin{aligned} v_n(A) &= \left(q_0^{N_0} - q_1^{N_0}\right) + q_1^{N_0}\binom{N_0-1}{\frac{N_0-1}{2}} \\ &= \rho\frac{1}{2} + (1-\rho)\frac{1}{2^{N_0}}\binom{N_0-1}{\frac{N_0-1}{2}} \end{aligned} \tag{5}$$

Now consider a reader of type two, without loss of generality let this be reader $n = N_0$. Any point s is characterized by: l_B, which is the number of differences between s and s^* on dimensions $1, \ldots, B$; by l'_A, which is the number of differences on dimensions $B+1, \ldots, N_0 - 1$: and by the value of the last dimension s_n.

Case one: $s_n = 0$; then the number of differences between s and s^* is $l_B + l'_A + 1$. This number must be no more than $\frac{N_0-1}{2}$ if a point s is to be in x^A; or $l'_A \leq \frac{N_0-1}{2} - l_B - 1$. There is exactly $\binom{B-1}{l_B}\binom{A-1}{l'_A}$ points that are represented by a given pair (l_B, l'_A). Therefore, the total number of points that are in $x^A \setminus 0^{N_0}$ and have $s_n = 0$ is

$$\sum_{l_B=0,\dots,B} \sum_{l'_A=0,\dots,\frac{N_0-1}{2}-l_B-1} \binom{B}{l_B}\binom{A-1}{l'_A} - 1$$

Case two: $s_n = 1$; the number of differences between s and s^* is $l_B + l'_A$. Therefore, the total number of points that are in $x^A \setminus 0^{N_0}$ and have $s_n = 1$ is

$$\sum_{l_B=0,\dots,B} \sum_{l'_A=0,\dots,\frac{N_0-1}{2}-l_B} \binom{B}{l_B}\binom{A-1}{l'_A}$$

Subtracting these two numbers from each other give

$$\begin{aligned} \sum_{m:s^m \in x^A \setminus 0^{N_0}} (1-2s_n^m) &= -\sum_{l_B=0,\dots,B} \binom{B}{l_B}\binom{A-1}{\frac{N_0-1}{2}-l_B} - 1 \\ &= -\binom{N_0-1}{\frac{N_0-1}{2}} - 1 \end{aligned}$$

The resulting value of this type of reader is

$$\begin{aligned} v_n(A) &= \left| \left(q_0^{N_0} - q_1^{N_0}\right) - q_1^{N_0}\binom{N_0-1}{\frac{N_0-1}{2}} \right| \\ &= \left| \rho\frac{1}{2} - (1-\rho)\frac{1}{2^{N_0}}\binom{N_0-1}{\frac{N_0-1}{2}} \right| \end{aligned}$$

It can be concluded immediately that this value is strictly less than of type one reader in equation 5. In other words, the total gross value would be maximized if the number of type two readers was zero, $A = 0$.

Now consider case that N_0 is even. By the second part of Proposition 1 there is diagonal partition w.r.t. point s^* so that point s belongs to x if it differs from s^* in at most $\frac{N_0}{2} - 1$ dimensions. The rest of the proof follows exactly the same steps, but with $\frac{N_0-1}{2}$ replaced by $\frac{N_0}{2} - 1$.

5.3 Proof of Lemma 2

By direct comparison of available options: buy newspaper 1, buy newspaper 2, not buy.

5.4 Proof of Lemma 3

Let

$$\bar{p}_{in} = c + \max\{v_{in} - \max\{v_{jn}, c\}, 0\}$$

First step is to confirm that $p_{jn} = \bar{p}_{jn}$ for $j = 1, 2$ and all readers n is an equilibrium. Suppose that player j is using this pricing. Changing newspaper's i price to a reader has only consequences for profit from this reader only, while the profit from other readers remains unchanged.

Consider a reader n for which $v_{in} > \max\{v_{jn}, c\}$. By sticking to $\bar{p}_{in}$, newspaper i gets a net profit of $v_{in} - \max\{v_{jn}, c\} > 0$. By deviating to a higher price for this reader, this newspaper looses this reader to newspaper j. By deviating to a lower price for this reader, this newspaper gets a lower revenue.

Now consider a case $v_{in} \leq \max\{v_{jn}, c\}$. Then by sticking to $\bar{p}_{in} = c$, newspaper i gets payoff of zero. By deviating to a higher price, this newspaper will not get to sell to the reader, so the profit stays at zero. Deviating to a lower price is obviously not profitable.

The second step is to show that there are no other equilibria. The logic follows the same lines as above and the proof is omitted.

5.5 Proof of Lemma 4

Suppose that newspaper i best responds to a $\{v_{j1}, ..., v_{jN}\}$ by choosing a partition generating $\{v_{i1}, ..., v_{iN}\}$ so that

$$v_{in} - \max\{v_{jn}, c\} \geq 0 \text{ if and only if } n \in N_i$$

and suppose that it does not target set N_i.

The revenue of newspaper i is equal to

$$\sum_{n\in N_i} (v_{in} - \max\{v_{jn}, c\}) \tag{6}$$

Note that the deviation to a partition that targets N_i strictly increases the total gross value of readers in N_i, that is $\sum_{n\in N_i} \hat{v}_{in}(N_i) > \sum_{n\in N_i} v_{in}$. As a consequence, the revenue from the original partition in (6) is strictly lower than the left-hand side of

$$\sum_{n\in N_i} (\hat{v}_{in}(N_i) - \max\{v_{jn}, c\}) \leq \sum_{n\in N_i} \max\{\hat{v}_{in}(N_i) - \max\{v_{jn}, c\}, 0\}$$

The right-hand side of the above equation is a revenue from targeting N_i readers (in case of N_i being even, this is the revenue coming from a symmetric partition targeting N_i, so that $\hat{v}_{in}(N_i)$ is not only the average value of each targeted reader, but the actual value, equal across readers)

This proves that the original partition was not optimal.

5.6 Proof of Proposition 4

The proof is conducted in a series of steps.

Step 1. In any equilibrium $N_i^e \leq N^*(c)$.

Proof. Suppose that $N^*(c) < N_i^e$. Then, by lowering the set of targeted readers to $N^*(c)$, newspaper i will strictly increase the total gross value created to its readers, hence the revenue and hence the profit. ■

Step 2. In any equilibrium no reader is targeted by both newspapers.

Proof. Consider an equilibrium in which newspapers target sets $\mathcal{N}_i$ and $\mathcal{N}_j$ such that $N_i \leq N^*(c)$ and $N_j \leq N^*(c)$ respectively and contrary to the lemma suppose that the set $\mathcal{N}_i \cap \mathcal{N}_j$ contains at least one reader. Without loss of generality assume that $c < \hat{v}_{in}(N_i) \leq \hat{v}_{jn}(N_j)$, or $N_j \leq N_i$. The implication of this is that the revenue of newspaper i from all readers in set

$\mathcal{N}_i \cap \mathcal{N}_j$ is zero, because it sells only to $N - N_j$ readers who are not targeted by j (if $\hat{v}_{in}(N_i) < \hat{v}_{jn}(N_j)$) or the price is equal to cost (if $\hat{v}_{in}(N_i) = \hat{v}_{jn}(N_j)$).

Note that there are no readers who are not targeted at all. If there were such readers n', then newspaper i could cease targeting reader n and start targeting n'. The only result of this change would be that newspaper i obtains an additional revenue of $\hat{v}_{in}(N_i) - \frac{1}{2}\rho$, which is strictly positive. From now on, assume that there are no readers who are not targeted at all.

Case 1. Consider equilibria in which set $\mathcal{N}_i \cap \mathcal{N}_j$ contains two readers or more. Newspaper i can cease targeting two readers belonging to $\mathcal{N}_i \cap \mathcal{N}_j$, strictly increase the value of $\hat{v}_{in}(\cdot)$ and hence the price to all readers that it sells to in set $\mathcal{N} \setminus \mathcal{N}_j$.

Case 2. Consider equilibria in which set $\mathcal{N}_i \cap \mathcal{N}_j$ contains precisely one reader.

Suppose that N_i is even. Then newspaper i can cease targeting this one reader belonging to $\mathcal{N}_i \cap \mathcal{N}_j$, strictly increase the value of $\hat{v}_{in}(\cdot)$ and hence the price to all readers that it sells to in set $\mathcal{N} \setminus \mathcal{N}_j$.

If N_i is odd then $N_j < N_i$ must be even (This is because $\frac{N}{2}$ is an integer by assumption, so $N + 1 = N_i + N_j$ is an odd number). Newspaper j can increase the readership by one reader already targeted by newspaper i, so that it targets odd readers $N_j + 1$. This will not affect the value of $\hat{v}_{jn}(\cdot)$ and will bring the additional revenue from this reader of $\hat{v}_{jn}(N_j) - \hat{v}_{in}(N_i) > 0$. ■

These results imply that in any equilibrium there are two sets of readers – set $\mathcal{N}_1^e$ is targeted by newspaper 1, while $\mathcal{N}_2^e$ is targeted by newspaper 2, such that $N_1^e + N_2^e \leq N$.

Step 3. If $N_1^e + N_2^e < N$ then $N_i^e = N^*(c)$.

Proof. Direct implication of Proposition 3 ■

Step 4. Consider any equilibrium readership profile (N_1^e, N_2^e) and suppose that $N_1^e + N_2^e = N$. Then N_i^e is odd.

Proof. Suppose not: let there be an equilibrium in which N_i^e is even. Since

$\frac{N}{2}$ is assumed to be an integer, $N = N_i^e + N_j^e$ is even, so is N_j^e. One of the newspapers must have strictly higher readership; without loss of generality assume $\hat{v}_{in}(N_i^e) > \hat{v}_{jn}\left(N_j^e\right)$. Then newspaper i would have incentives to target an additional reader already targeted by j. The value of function $\hat{v}_{in}(\cdot)$ does not change, so the additional profit is $\hat{v}_{in}(N_i^e) - \hat{v}_{jn}\left(N_j^e\right) > 0$. ∎

Step 5. Consider any readership profile (N_1^e, N_2^e) and suppose that $N_1^e + N_2^e = N$. If $N_i^e \leq N^*(c)$ then newspaper i has no incentives to lower its readership to any alternative $N_i < N_i^e$.

Proof. Direct implication of Proposition 3 ∎

Step 6. Consider any readership profile (N_1^e, N_2^e) and suppose that $N_1^e + N_2^e = N$. If

$$N^*\left(\hat{v}_{jn}\left(N_j^e\right)\right) \leq N_i^e \tag{7}$$

then newspaper i has no incentives to increase its readership to any alternative $N_i > N_i^e$. If $N_i^e < N^*\left(\hat{v}_{jn}\left(N_j^e\right)\right)$ then newspaper i has strictly positive incentives to increase the readership from N_i^e, and hence this N_i^e cannot be a part of an equilibrium.

Proof. Consider any readership profile (N_1^e, N_2^e) such that $N_1^e + N_2^e = N$. Newspaper i may want to change the behavior by targeting some extra readers from the opponent's set $\mathcal{N}_j^e$ in addition to those already targeted in set $\mathcal{N}_i^e$. Let $\mathcal{N}_i \supset \mathcal{N}_i^e$ be the set of readers targeted by deviating newspaper i. This deviation has chances of generating additional profit only if $\hat{v}_{in}(N_i) > \hat{v}_{jn}\left(N_j^e\right)$.

After a deviation, reader $n \in \mathcal{N}_i \cap \mathcal{N}_j^e$ targeted by both newspapers will buy rather from newspaper i, whose price, by lemma 3, is equal to

$$p_{in} = c + \hat{v}_{in}(N_i) - \hat{v}_{jn}\left(N_j^e\right)$$

This means however, that the price-cost margin of newspaper i from this reader is equal to $\hat{v}_{in}(N_i) - \hat{v}_{jn}\left(N_j^e\right)$. The profit of a newspaper deviating in such a way behaves exactly like a profit of a monopolist with unit 'cost' equal to $\hat{v}_{jn}\left(N_j^e\right)$. By Proposition 3, a finite and odd number $N^*\left(\hat{v}_{jn}\left(N_j^e\right)\right)$ exists and the further away N_i is from it, the lower the profit after the deviation. So, if initial readership N_i^e is already greater or equal than $N^*\left(\hat{v}_{jn}\left(N_j^e\right)\right)$, then targeting even more new readers that are already targeted by j will lead to lower profit, and therefore cannot constitute a profitable deviation.

On the other hand, if readership N_i^e is lower than $N^*\left(\hat{v}_{jn}\left(N_j^e\right)\right)$, then targeting new readers will increase the profit and hence (N_1^e, N_2^e) is inconsistent with equilibrium. ■

6 Literature

Baron, David P., 2006, Persistent media bias, *Journal of Public Economics*, Volume 90, Issues 1-2, Pages 1-36.

Daw, James, 1991, Search Decisions with Limited Memory, *The Review of Economic Studies*, Vol. 58, No. 1, 1-14.

Fejes Toth, L. 1959, Sur la representation d'une population infinie par un nombre fini d'elements, *Acta Math. Acad. Sci. Hung.*, vol. 10, pp. 76-81.

Gentzkow, Matthew, and Jesse M. Shapiro, 2006, Media Bias and Reputation, *Journal of Political Economy*, Volume 114, Number 2.

Gill David, and Daniel Sgroi, 2008, The Optimal Choice of Pre-launch Reviewer: How Best to Transmit Information using Tests and Conditional Pricing, manuscript, Southampton University.

Glazer, Jacob and Ariel Rubinstein, On Optimal Rules of Persuasion, *Econometrica*, Vol. 72, No. 6, (Nov., 2004), pp. 1715-1736.

Gray, Robert N. and David L. Neuhoff, 1998, Quantization, *IEEE Transactions on Information Theory*, Vol. 44, No. 6., pp 1-63.

Gersho, Allen and Robert M. Gray 1991, *Vector Quantization and Signal Compression*, Springer.

Irmen, Andreas and Jacques-François Thisse, 1998, Competition in Multi-characteristics Spaces: Hotelling Was Almost Right, *Journal of Economic Theory* 78, 76-102.

Mullainathan, Sendhil, and Andrei Shleifer, 2005, The Market for News, *American Economic Review* Vol. 95, No. 4.

Newman, D. J., 1964, The hexagon theorem, *Bell Laboratories Technical Memorandum*. Published in the special issue on quantization of the *IEEE Trans. Inform. Theory*, vol. 28, pp. 137-139, Mar. 1982.

Veldkamp, Laura L., 2006, Media Frenzies in Markets for Financial Information, *American Economic Review*, Vol. 96, No. 3.