Economics Division
School of Social Sciences
University of Southampton
Southampton SO17 1BJ
UK

Discussion Papers in Economics and Econometrics

Mass Media: Constrained Information and Heterogenous Public

Maksymilian Kwiek

No. 0606

This paper is available on our website
http://www.socsci.soton.ac.uk/economics/Research/Discussion_Papers

Mass media: constrained information and heterogenous public

Maksymilian Kwiek*

May 9, 2006

Abstract

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

It is shown that the sender can optimally provide information that is somewhat useful to all agents, but not perfect to anybody in particular. Because all receivers observe only one coarse signal delivered by the same mass medium their bahaviour is pefectly correlated, positively or negatively, even if the underlying states of nature are independent. In addition, if the correlation between states of nature of any two players is sufficiently high, their behaviour is positively correlated. However, we may have a situation where all agents are symmetric, the correlation of states of nature is negative (positive), but the behaviour is positively (negatively) correlated.

The model can be used to explain the role of mass media in creating comovement among various industries during business cycle, or finacial contagion.

Keywords: Mass media, news, cheap talk, quantization, comovement, herding, contagion.

JEL Classification: L82, E32

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

1 Motivation

> "Journalists are writing us into a recession". *The Economist* Oct/04/2001.

My typical morning is to wake up, brush my teeth and turn on the TV to read headlines on the BBC teletext while I eat my breakfast. There are just 21 pages devoted to the news. On the other hand, one can assert that every morning the BBC receives hundreds of pieces of information from news agencies, own reporters, free-lancers, paparazzis, etc. The BBC faces a difficult question – which 21 pieces of information should be published on the teletext?

Obviously, in order to answer this question, the BBC needs to know what the audience wants to know. But the problem is that the audience has heterogenous preferences. For instance, I am interested in international politics, but I am not interested in controversies around hunting ban in England, or tropical storms in the Caribbean. However, I know for a fact that the hunting ban is very important for a great number of people, and that many others have either relatives in the Caribbean or plan to go there on holiday.

In this situation the BBC can generally do one of two things, either to provide information that is not perfect to anybody in particular, but somewhat informative to all – a strategy that could be called "unfocused" – or to concentrate on some issues but report them in great detail – this could be called a "focused" strategy. In short, the dilemma is whether 'to inform everybody about nothing' or 'to inform nobody about everything'. In the above BBC teletext example, the unfocused strategy would be to devote one page to one topic, covering 21 most important topics; the focused strategy would be to pick one topic, say hunting ban, and drag about it for all 21 pages.

The model presented in this paper can explain why actions may be highly correlated across agents even if agents' situations are very different – and a role that mass media (newspapers, television, popular internet sites) can play in facilitating this process. The aim of this paper is to explain this aspect of business cycle dynamics or financial contagion from a perspective that is usually ignored by economic analysis.

The model is rather simple. There is a single mass medium and a great number of agents (readers). The problem of an agent is to tailor her actions to the unknown optimal action, called the state of nature. Agents are heterogenous, in the sense that states of nature of different agents may be uncorrelated. Agents do not trade nor they engage in any other relationship with each other; the only common feature that they share is that they read the same newspaper. Mass medium is useful to an agent, because it knows the state of nature. However, the state of nature can be only *partially revealed* by mass medium. The most extreme assumption is made – that the mass medium can deliver only one of two messages to the agents.

Under the assumptions of the model below, the mass medium optimally chooses to inform all the agents in an imperfect way, rather than to focus on one agent and inform her perfectly. Since all agents have access to the same single signal, which is only partially informative, their actions are highly correlated, positively or negatively. This is true, even if the states of nature are independent among agents. In other words, the presence of such a mass medium amplifies the absolute value of correlations between actions of agents who otherwise have little in common.

The frictions in information transmission from the mass medium to the receiver is crucial in this analysis. Without them, the newspaper would reveal everything that it knows, allowing the readers to costlessly pick and choose whatever is relevant to them. The mass medium could then extract full surplus through the take-it-or-leave-it price. Only with such transmission constraints the medium's problem is nontrivial and the phenomenon of comovement can arise.

It is important to realize, however, that this constraint is not necessarily a physical constraint on the capacity of sender's communication channel like the one in the above BBC example. It is much easier to consider a constraint on the receivers' side. Suppose that messages consist of arbitrary many and arbitrary long codewords/messages consisting of binary digits. A sequence like this can reveal the true state, provided that the agent is patient enough to wait until the entire transmission is completed. However, assume that the agent has alternative uses for her time and reads only the first part of a message. This type of constraint would prompt a similar decision problem for the sender as discussed in this paper, since the value of the newspaper depends on the realization of the initial digits only. This interpretation seems to be particularly appealing in the case of mass media, where adding an additional page to a newspaper does not seem to be prohibitively expensive. It is rather the reader who has little time to read deeply and relies on headlines. Compare Graber (1988) for evidence of selective attention of TV audience.

2 Literature

This model provides an additional way of explaining comovement of agents' actions – a feature strikingliy difficult to obtain in frictionless economic models populated by rational agents. Comovement is even one of the defining characteristics of the business cycle; yet, how exactly it arises remains a puzzle. See Christiano and Fitzgerald (1998) for an overview of comovement of industries' characteristics over the business cycle. Related and equally important is comovement across geographical regions, such as financial contagion, for instance. In this example, the crisis (and euphoria) spreads across countries that are otherwise unrelated by any fundamentals. For instance, Calvo and Mendoza (2000) identify contagion as one of the main phenomena that characterized disturbing financial crises of the 1990-ties (such as the Mexican crisis of 1994, Asian 1997, or Russian in 1998). They assert that costs of acquiring information may lead many investors to choose to stay uninformed; these investors may in turn incorrectly interpret actions of informed investors. To my knowledge, however, there are no studies that would place mass media in the center of this issue and explain their role as an element that may be responsible for high correlation of agents' actions.

This model is somewhat related to models of informational herding (Banerjee, 1992). Consider a sequence of agents who take actions one-by-one. A decision maker learns something useful about the world from two sources – her private signal and observed actions of her predecessors. Information that is conveyed through previous actions may be strong enough for an agent to choose the same action regardless of her own signal. Such an action does not provide any additional information to her successor, who – as she is in the same situation – will behave in the same way. Consequently, a disaggregate information, no matter how complete, may be lost. One reason why this occurs is because the action set is very coarse; a finer action space would allow agents to tailor their chosen actions closer to optimum, thus revealing their private signal to their successors.

In the model below we have one predecessor – the mass medium – and a large number of successors, all taking actions simultaneously. The coarse action set in sequential herding model corresponds to the constraint on how much information can be delivered by the mass medium in the model below. In the absence of this constraint, the mass medium would reveal true state of nature, agents would learn the aspects relevant to them and we would not observe any more correlation in actions than in states of nature.

Daw (1991) analyses the problem of a buyer looking for bargain prices, who understands that her memory is not perfect. The buyer has to decide what to remember if only one bit of information can be recorded for future reference. An interesting version of the model analyzed by Daw is where the decision maker faces two independent experiments, and has to decide whether to remember a lot about the result of one experiment or rather to remember something about both – again focused or unfocused strategy. Daw's results are that the focused strategy is better. His framework is somewhat different than presented below, but his conclusion is entirely opposite. In model below, the decision maker decides to use an unfocused strategy, and this result seems to be quite general.

Quantization is the process of converting an input from a rich state space into a finite number of discrete values (see Gray and Neuhoff (1998), or Gersho and Gray (1991)). The properties of various quantization methods are studied in information theory for purpose of coding, compressing or digitalization. This literature is very closely related to our application, where the mass medium tries to represent the state of nature in a shorter form. Of a particular importance here is vector quantization, where state space is multi-dimensional; in our case 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 or codewords. This paper is interested in exactly the opposite situation, where very little can be transmitted. This is supposed to represent a situation where agents are very distracted by other activities while reading or watching the news. In our case, the most extreme assumption possible is made, namely that only one of two messages can be delivered to the readers.

Another difference with the current quantization literature is that this literature is not concerned with the issue of a correlation between dimensions, but this question is exactly what we are asking. It should be obvious that as the number of codewords increases in a given environment, then the state of nature is represented better and better and correlation vanishes. This paper explains, however, that this correlation increases when the number of codewords goes down. There are some results from information literature hinting that output correlation is not zero in general even if the inputs are uncorrelated. In particular, Fejes Toth (1959) and Newman (1964) showed that even in the simple case of two independent uniform random variables, the performance achievable by a hexagonal partition of a unit square of the state space is strictly better than quantizer with a square partition.

This approach to economic analysis of mass media is new, to my knowledge. There is a number of recent articles analysing media markets (see Gentzkow and Shapiro (2006), Mullainathan and Andrei Shleifer (2005), Barron (2006)). But their focus is mostly media bias: how and why various media outlets choose sometimes very different informational policies. For instance, if the newsppaper faces a biased public then it can choose not to publish some bits of information (slanting). This paper assumes directly that the transmission channel is restricted, so that the medium has to choose what to publish, and The same aspects of this issue can be addressed in the model introduced in this paper.

To summarize the results: It is shown that in the simple model studied in this paper, the sender optimally provides unfocused information – somewhat useful to all agents, but not perfect to anybody in particular. Because all receivers observe only a binary signal delivered by the same mass medium, their behavior is perfectly correlated, positively or negatively, even if the underlying states of nature are independent. In addition, if the underlying correlation between states of nature of any two players is sufficiently high, their behavior is positively correlated. However, we may even have a situation where all agents are symmetric, the underlying pairwise correlation is negative (positive), but the behavior is positively (negatively) correlated.

3 Model

Suppose that 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 $s^m = (s_1^m, ..., s_N^m) \in S$ (we will ignore this index where possible). 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 s_n.

Assumption 1. Symmetry with respect to state: For any n, marginal distribution is uniform, that is

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

Uninformed agent (receiver, reader) of type $n = 1, ..., N$ cares only about dimension n of state $s \in S$. Agent'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 zero otherwise. Net utility in a quasilinear in expenditure.

There is a monopolistic mass medium (sender, newspaper) who knows the true state $s \in S$. It is assumed that this newspaper can send only one of two possible messages, too few relative to the dimensionality N. This forms a constraint that lies in the heart of the model. In particular, the 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 optimally. Hence, let X be the set of all subsets of S. The focus of this note is to investigate the optimal action of the newspaper, $x \in X$.

Timing in the model is as follows.

- In the first stage, the newspaper publicly commits to a partition, x, and announces a take-it-or-leave-it price of the report p. If newspaper can price-discriminate, $p \in R_+^N$, otherwise the price has to be the same for each agent, $p \in R_+$.
- In the second stage, after agents observed the partition and the price, they decide whether to purchase the report or not.
- Then 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 $s \notin x$; agents who purchase the report, learn its content; and finally, each agent n takes action a_n, and payoffs are realized.

This timing structure 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 "line" and it is expected to follow this policy in the future.

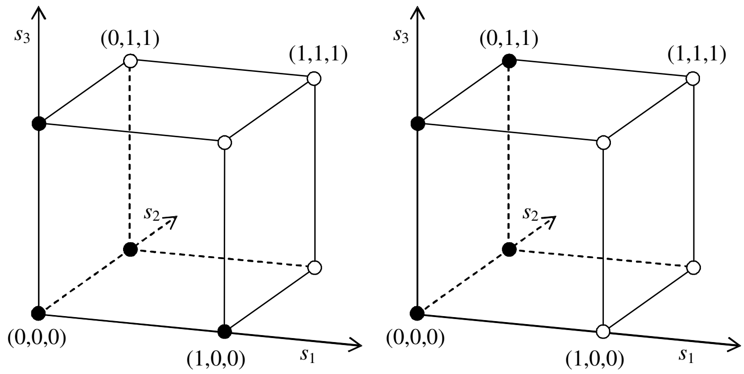
If $N = 1$, then the situation is trivial. Possible partitions are $X = \{\{\varnothing\}, \{0\}, \{1\}, \{0, 1\}\}$. The message space is rich enough to inform the agent about the state; take-it-or-leave-it offer can capture the entire surplus. Hence $x = \{0\}$ or $x = \{1\}$ are both optimal. If the agent buys the report then the loss is zero, since the agent can choose the action equal to a perfectly revealed state. Without the report, the expected loss is simply $\frac{1}{2}$. The difference between zero an half can be captured by the newspaper through a price $p = \frac{1}{2}$.

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

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

Later, we are going to call this partition diagonal w.r.t. $(0, 0, 0)$. Panel (b) shows a partition

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



(a) (b)

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

focused on agent $n = 1$. That is, newspaper tells only agent 1 his state, $s = x \Leftrightarrow s_1 = 0$. Nobody else learns anything about their states.

To follow this three-dimensional case, suppose that the distribution is uniform, $q = \frac{1}{8}$. We can easily find the expected loss and the value of the newspaper to each of three agents, and ultimately the revenue of a price-discriminating newspaper, for both cases shown on Figure (1).

Panel (a)

agent	$n = 1, 2, 3$
Loss cond. on x and a_n	$\frac{3}{4}a_n + \frac{1}{4}(1 - a_n)$
optimal action cond. on x	$a_n = 0$
Loss cond. on x	$\frac{1}{4}$
Loss cond. on y and a_n	$\frac{1}{4}a_n + \frac{3}{4}(1 - a_n)$
optimal action cond. on y	$a_n = 1$
Loss cond. on y	$\frac{1}{4}$
Uncond. loss, L	$\frac{1}{4}$
$p_n = \frac{1}{2} - L$	$\frac{1}{4}$

$$\text{Revenue } R = \frac{3}{4}$$

Panel (b)

agent	$n=1$	$n=2,3$
Loss cond. on x and a_n	a_n	$\frac{1}{2}a_n + \frac{1}{2}(1-a_n)$
optimal action cond. on x	$a_n = 0$	any
Loss cond. on x	0	$\frac{1}{2}$
Loss cond. on y and a_n	$1-a_n$	$\frac{1}{2}(1-a_n) + \frac{1}{2}a_n$
optimal action cond. on y	$a_n = 1$	any
Loss cond. on y	0	$\frac{1}{2}$
Uncond. loss, L	0	$\frac{1}{2}$
$p_n = \frac{1}{2} - L$	$\frac{1}{2}$	0

$$\text{Revenue } R = \frac{1}{2}$$

Even here we can see the preview of some of the results. Note that the unfocused partition in panel (a) is better for the newspaper than the focused partition (b). Moreover, in partition (a) all three agents take the same action as a consequence of reading the same newspaper. Hence, if this partition constitutes a policy of the newspaper, then the correlation between actions of agents is $+1$, even if their states of nature are completely independent.

There are two problems with the above simple comparison of partitions (a) and (b). Firstly, we still do not know if (a) is indeed optimal. Secondly, even if it is optimal, we can easily see that it is not uniquely optimal. Consider partition

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

It is almost the same as partition (a), except that it is rotated by 90° to the right. It can be easily shown that it generates the same revenue as (a), but the correlation between actions, although perfect in absolute terms, in reality is either $+1$ or -1.

4 The decision problem of an agent

To find an equilibrium in the model, we proceed backwards, starting with the problem of an agent purchasing the newspaper and taking an action in the last stage.

For each partition x, define $Q_n(x,0)$ to be the probability that true state is in x and in the same time s_n is 0. Define $Q_n(x,1)$, $Q_n(y,0)$ and $Q_n(y,1)$ in the same way:

$$Q_n(x,0) = \sum_{\{m: s^m \in x, s_n^m = 0\}} q(s^m) \quad Q_n(x,1) = \sum_{\{m: s^m \in x, s_n^m = 1\}} q(s^m)$$
$$Q_n(y,0) = \sum_{\{m: s^m \in y, s_n^m = 0\}} q(s^m) \quad Q_n(y,1) = \sum_{\{m: s^m \in y, s_n^m = 1\}} q(s^m)$$

Suppose that the agent n buys a newspaper and learns that the state is x, then the posterior probability that $s_n = 0$ is $\frac{Q_n(x,0)}{Q_n(x,0)+Q_n(x,1)}$. The expected loss conditional on x is

$$\frac{Q_n(x,0)}{Q_n(x,0)+Q_n(x,1)} a_n + \frac{Q_n(x,1)}{Q_n(x,0)+Q_n(x,1)} (1-a_n)$$

This is a linear objective function, so typically we are going to have a boundary solution. 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

$$\min\left\{\frac{Q_n(x,0)}{Q_n(x,0)+Q_n(x,1)}, \frac{Q_n(x,1)}{Q_n(x,0)+Q_n(x,1)}\right\}$$

Similarly, the optimal expected loss conditional on y is

$$\min\left\{\frac{Q_n(y,-1)}{Q_n(y,-1)+Q_n(y,1)}, \frac{Q_n(y,1)}{Q_n(y,-1)+Q_n(y,1)}\right\}$$

Since the probability of x is $Q_n(x,0)+Q_n(x,1)$ the unconditional optimal loss (before learning the content of a newspaper) is

$$\begin{aligned} L &= [Q_n(x,0)+Q_n(x,1)]\min\left\{\frac{Q_n(x,0)}{Q_n(x,0)+Q_n(x,1)}, \frac{Q_n(x,1)}{Q_n(x,0)+Q_n(x,1)}\right\} \\ &\quad + [Q_n(y,0)+Q_n(y,1)]\min\left\{\frac{Q_n(y,0)}{Q_n(y,0)+Q_n(y,1)}, \frac{Q_n(y,1)}{Q_n(y,0)+Q_n(y,1)}\right\} \end{aligned}$$

or

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

Assumption 1 (symmetry w.r.t. state) implies

$$\frac{1}{2} = Q_n(x,0)+Q_n(y,0) = Q_n(x,1)+Q_n(y,1)$$

Using these two equations to eliminate $Q_n(y,0)$ and $Q_n(y,1)$ from L, we obtain eventually

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

If the agent refrains from buying the newspaper, his expected loss is $\frac{1}{2}$. Hence the valuation that agent n attaches to the newspaper is

$$\begin{aligned} |Q_n(x,0) - Q_n(x,1)| &= \left| \sum_{m:s^m\in x, s_n^m=0} q(s^m) - \sum_{m:s^m\in x, s_n^m=1} q(s^m) \right| \\ &= \left| \sum_{m:s^m\in x} (1-s_n^m)\, q(s^m) - \sum_{m:s^m\in x} s_n^m q(s^m) \right| \\ &= \left| \sum_{m:s^m\in x} (1-2s_n^m)\, q(s^m) \right| \qquad (1) \end{aligned}$$

Notice that the above value, if divided by the probability of x, is the difference between $\frac{Q_n(x,0)}{Q_n(x)}$, the posterior probability that the state is zero conditional on x, and $\frac{Q_n(x,1)}{Q_n(x)}$, the posterior probability that the state is one conditional on x. The greater the wedge between these probabilities created by a newspaper, the greater its value. In other words, the value of a newspaper lies in its ability to surprise.

5 Price-discriminating newspaper.

If a newspaper can price-discriminate, then the price decision for a given partition is simple, just charge the exact surplus (1) generated by this newspaper. Corresponding total revenue is

$$R(x) = \sum_{n=1}^{N} \left| \sum_{m:s^m \in x} (1 - 2s_n^m)\, q(s^m) \right|$$

Consider a case of odd N. We are going to define a diagonal partition. Diagonal partition with respect to $s \in S$ contains all the points that differ from s in at most half of the dimensions.

Definition 1 *Partition x is diagonal with respect to $s \in S$ if*

$$x = \left\{ s' \in S : \sum_{n=1}^{N} |s_n - s'_n| \leq \frac{N-1}{2} \right\}$$

(Partition is diagonal if there exists s such that it is diagonal w.r.t. s).

As an example, consider Figure (1), panel (a). It depicts a diagonal partition with respect to point $(0, 0, 0)$.

There are 2^N diagonal partitions in total – one for each point in S. Note, however, that a diagonal partition with respect to s is essentially the same as a diagonal partition with respect to $1 - s$, only with x and y exchanging their places, and gives the same surpluses to all agents. Taking this into account, there is 2^{N-1} nontrivial diagonal partitions.

The main result follows.

Proposition 1 *Let $q(s) > 0$ for all $s \in S$. Suppose that assumption 1 holds. If a newspaper can price-discriminate then the optimal partition is diagonal.*

Proof. Fix a partition x that is not diagonal. Let D^+ be a set of agents (dimensions) 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^- be the remaining set of agents. Define s^* be a 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, it is not diagonal with respect to s^*. Therefore, either there exists $s' \in x$ such that $\sum_{n=1}^N |s_n^* - s'_n| \geq \frac{N+1}{2}$, or there exists $s' \notin x$ such that

$$\sum_{n=1}^{N} |s_n^* - s'_n| \leq \frac{N-1}{2} \tag{2}$$

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

The revenue form including s' in the set x is

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

or simply

$$R(x \cup s') \geq R(x) + \left(\sum_{n \in D^+} (1 - 2s'_n) - \sum_{n \in D^-} (1 - 2s'_n) \right) q(s') \tag{3}$$

Furthermore, inequality (2) implies

$$\frac{N-1}{2} \geq \sum_{n \in D^+} s'_n + \sum_{n \in D^-} (1 - s'_n)$$

or

$$\begin{aligned} 1 &\leq -\sum_{n \in D^+} 2s'_n - \sum_{n \in D^-} (2 - 2s'_n) + N \\ 1 &\leq -\sum_{n \in D^+} 2s'_n - \sum_{n \in D^-} (2 - 2s'_n) + \sum_{n \in D^+} 1 + \sum_{n \in D^-} 1 \\ 1 &\leq \sum_{n \in D^+} (1 - 2s'_n) - \sum_{n \in D^-} (1 - 2s'_n) \end{aligned}$$

Therefore, the brackets in (3) is strictly positive. Since $q(\cdot) > 0$ as well, including s' in x will strictly increase the revenue, $R(x \cup s') > R(x)$. This proves that a nondiagonal x is not optimal. ■

This Proposition goes far in determining the optimal strategy of a newspaper. Still, there is 2^{N-1} candidates for optimal partition, each associated with different diagonal partition. Which diagonal partition is optimal will depend on distribution $q(\cdot)$. One special case is when the distribution is uniform. Additional results will be presented in the next section.

Corollary 1 *Let $q(s) = 2^{-N}$, for all $s \in S$. If the newspaper can price-discriminate or set uniform prices, then any diagonal partition is optimal.*

Meanwhile, we can state the main result regarding the comovement of agents' actions,

Corollary 2 *If the newspaper can price-discriminate, then the correlation between actions of any two agents is $+1$ or -1.*

6 Symmetric players

From now on, we are going to assume that the distribution is symmetric with respect to agents.

Assumption 2. Symmetry with respect to agents. For any two points $s, s' \in S$, such that $\sum_{n=1}^{N} s_n = \sum_{n=1}^{N} s'_n$, we have $q(s) = q(s')$.

Define $q_k = q(s)$ if $\sum_{n=1}^{N} s_n = k$, for $k = 0, ..., N$. By assumption 1 we have $q_k = q_{N-k}$. This vector q_k for $k = 0, ..., \frac{N-1}{2}$ completely determines the probability distribution. Since the total probability adds to one we have

$$\frac{1}{2} = q_0 + \sum_{k=1}^{\frac{N-1}{2}} \binom{N}{k} q_k \tag{4}$$

For any pair of agents, a correlation between their states is

$$\rho = 1 - 8 \sum_{k=1}^{\frac{N-1}{2}} q_k \binom{N-2}{k-1} \tag{5}$$

For example, if $q_k = 0$ for $k \geq 1$, then correlation is $+1$ and the distribution is uniform on two points, $q_0 = q_N = \frac{1}{2}$.

Now, let us focus on the revenue from a diagonal partition with respect to some point s^*. We want to see how Assumption 2 restricts the number of partitions that have to be checked.

Choose a point s^* that will be the reference point for a diagonal parition w.r.t. s^*. Let $\sum_{n=1}^{N} s_n^* = A$ be the number of ones in point s^*. Similarly, let $B = N - A$ be the number of zeros. With assumptions 1 and 2, the revenue will not change if we replace point s^* with another point which still has the same number A. In this sense, A fully determines the diagonal partition and revenue. This allows us to define x^A and R^A, respectively, to be the diagonal

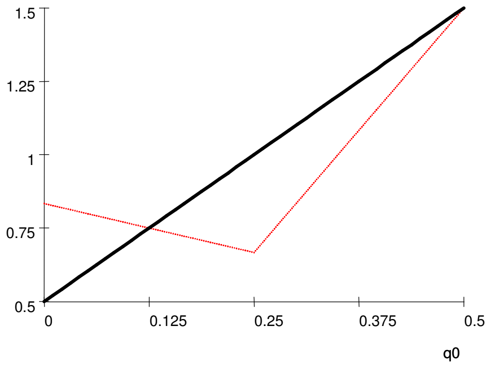


Figure 2: Revenues R^0 (bold) and R^1 (thin).

partition represented by a number A and the revenue generated by it, respectively. Moreover, replacing high states (ones) with low states (zeros) will not change the revenue, so that $R^A = R^{N-A}$. Hence, with Assumption 2, we only have to check revenues from diagonal partitions generated by

$$A \leq \frac{N-1}{2}. \tag{6}$$

There is only $\frac{N-1}{2} + 1$ of them, as opposed to 2^{N-1} of diagonal partitions identified in previous section.

Partition x^0 will be called the main diagonal partition.

The following example studies the problem of the newspaper in the three-dimensional case for any probability distribution satisfying Assumptions 1 and 2.

Example 1 *Suppose $N = 3$. Equation (4) gives $q_0 = \frac{1}{2} - 3q_1$, and by equation (5) the correlation is positive if $q_0 > \frac{1}{8}$. The revenue from the main diagonal partition is*

$$R^0 = 3(q_0 + q_1) = 2q_0 + \frac{1}{2}$$

If partition is x^1 then the agent's surplus is $p_i = (q_0 + q_1)$ for agent $i = 1, 2$; and it is $p_3 = |q_0 - 3q_1|$ for agent 3. The revenue is

$$R^1 = |q_0 - 3q_1| + 2(q_0 + q_1) = \left|2q_0 - \frac{1}{2}\right| + \frac{1}{3}(4q_0 + 1)$$

Conclusion: the unique optimal partition is x^0 if $q_0 \in \left(\frac{1}{8}, \frac{1}{2}\right)$ and x^1 if $q_0 < \frac{1}{8}$. If $q_0 = \frac{1}{8}$ or $\frac{1}{2}$ then both diagonal partitions are optimal. Bold line on the Figure (2) shows R^0 and the thin one shows R^1. The latter has a kink because of the absolute value.

Interestingly, the main diagonal partition is optimal if and only if the correlation is positive. This observation only partially generalizes to more dimensions.

Proposition 2 *Suppose that either the newspaper price-discriminates or sets uniform prices. Let the probability distributions satisfy Assumptions 1, 2 and* $q_k > 0$ *for all* $k = 0, ..., \frac{N-1}{2}$. *There is* $\rho^* < 1$ *such that for all distributions* $\{q_k\}_{k=0}^{(N-1)/2}$ *that have correlations* $\rho \in (\rho^*, 1)$ *the main diagonal partition is uniquely optimal.*

Proof is in the appendix.

In the extreme case of $\rho = 1$, the main diagonal partition is optimal, but so is any other diagonal partition. This is because perfect correlation implies that the distribution is uniform over two points $(0, ..., 0)$ and $(1, ..., 1)$, and any partition that distinguishes between these two points is optimal. The proposition shows that the main diagonal partition is uniquely optimal if correlation is less but close to 1.

The following is immediate,

Corollary 3 *Suppose that the conclusion of Proposition (2) holds. The correlation between actions of any two agents is* $+1$.

The above results hold for "high" enough correlation. However, for positive pairwise correlation the main diagonal partition may be sub-optimal, or for negative correlation it my be optimal, so that in this sense Example (1) does not generalize. The following two examples assert this.

Example 2 *Let* $N = 5$, $q_0 = 0.1$, $q_1 = 0$, *and* $q_2 = 0.04$. *This distribution gives us* $\rho = 0.04$. *The revenue from the main diagonal partition is* $R^0 = 0.9$, *while the revenue from the diagonal partition w.r.t.* $(0, 0, 0, 0, 1)$ *is* $R^1 = 1.02$. *If the newspaper uses this partition, then the correlation between actions of agents 1-4 is perfect positive, and between actions of agent 5 and any of the agents 1-4 is perfect negative.*

Example 3 *Let* $N = 5$, $q_0 = 0$, $q_1 = 0.046$, *and* $q_2 = 0.027$. *This distribution gives us* $\rho = -0.016$. *The main diagonal partition generates the revenue of* $R^0 = 0.96$. *This is higher than the revenues from other diagonal partitions,* $R^1 = 0.9$ *and* $R^2 = 0.954$.

These examples show that one may have a completely symmetric environment, where the newspaper is optimally revealing information in an asymmetric way. Moreover, for given two agents the correlation between their states may be positive (negative), but voluntarily purchased newspaper that is somewhat informative about these states may make their actions negatively (positively) correlated.

It seems that the above results are robust in the sense that the message of Corollary (3) extends to the situation with more than two messages. In particular, assume a continuous state and action space, as well as loss function equal to the square distance between action and state. Suppose that there is a mass medium which has $M > 2$ messages that can be delivered to 2 heterogenous agents each interested in different dimension, with a pairwise correlation coefficient between states equal to $\rho < 1$. Then the conjecture is that with ρ high enough but still less than one, the correlation of actions among agents who read the uniquely optimal newspaper would be $+1$ exactly.

7 Concluding remarks

This paper makes two contributions. Firstly, it models mass medium in a novel way. Mass medium is like a coding device that faces a nontrivial task – how to pack as much information as possible into a constrained channel leading to the brains of the readers. This modelling approach opens a host of new possibilities for the analysis of the mass media industry in the spirit of the industrial economics. What if there is an oligopoly of mass media outlets, competing in prices and informational policies? To what extent there is a product (information) differentiation? How does the outcome on the mass media market depend on the "landscape" of agents preferences such as their hetero- and homogeneity? Why some types of information are provided to the receivers via mass market in a form of simplified headlines, while other types of information are provided via highly specialized channels which are used by only very few receivers. Such a model of this industry would encompass agent's time allocation problem involoving decision whether to read, media decision to enter and exit and competition among existing media in terms of price and information policy, possibly an information gathering technology.

Second contribution is an application of this model of mass medium to describe some properties of the aggregate behaviuor of its clients. It should not be claimed that this model gives the whole story of comovement, but it provides a striking new way of analysis.

From the normative perspective, business cycle may be considered inefficient, but one aspect of business cycle – comovement – is not inefficient per se in the context of the above model. It can be easily seen that even if constrained mass media amplify comovemnt, they are beneficial to the society. Facilitating communication would eliminate comovement and that would be good. But eliminating mass media would also eliminate comovement, and that would be bad.

It is not entirely clear how one would test this model. But one can try to exploit some historical and contemporary changes in publication technology. Monopolistic mass medium reached its maturity in XIX and XX centuries, with its mass produced newspapers or broadcasting. We could be witnessing a shift in this technology towards more on-demand and customized media, such as internet or satelite TV, and more competition among firms. This could be interpreted as a natural experiment that could be used to test the above theory of media-induced comovement. The implication of these technological changes is that the level of comovement should go down as readers reley less on one-size-fits-all media.

8 Appendix: Proof of Prop (2)

Step 1. Let us focus on a price-discriminating newspaper. The revenue is

$$R^A = \sum_{n=1}^{N} \left| \sum_{m:s^m \in x^A} (1 - 2s_n^m)\, q(s^m) \right| \tag{7}$$

Consider the expression inside the absolute value. Note that equation (6) implies that $(0, ..., 0)$ is always – and $(1, ..., 1)$ is never – an element of any diagonal x under consideration. Hence, we have

$$\begin{aligned} \sum_{m:s^m \in x^A} (1 - 2s_n^m)\, q(s^m) &= q_0 + \sum_{m:s^m \in x^A \setminus (0,...,0)} (1 - 2s_n^m)\, q(s^m) \\ &> q_0 - \sum_{m:s^m \in S \setminus ((0,...,0) \cup (1,...,1))} q(s^m) \\ &= q_0 - (1 - 2q_0) \\ &= 3q_0 - 1 \end{aligned}$$

Hence for all $q_0 \geq \frac{1}{3}$ the expression inside the absolute value is positive. There is a threshold ρ^* such that for all $\rho > \rho^*$ we have $q_0 \geq \frac{1}{3}$. In this case, no matter what diagonal partition x satisfying equation (6) is used, the revenue can be written without the absolute values.

Define $x_k^A = \left\{ s \in x^A, \sum_{n=1}^{N} s_n = k \right\}$. Using this new notation and assuming $\rho > \rho^*$ the revenue is

$$\begin{aligned} R^A &= \sum_{n=1}^{N} \sum_{k=0}^{N} \sum_{m:s^m \in x_k^A} (1 - 2s_n^m)\, q_k \\ &= \sum_{k=0}^{N} q_k \sum_{m:s^m \in x_k^A} (N - 2k) \\ &= \sum_{k=0}^{N} q_k \,(N - 2k) \left(\# x_k^A\right) \end{aligned}$$

Only $\# x_k^A$ needs to be explained – the number of elements in set x_k^A.

Step 2. Consider a diagonal partition w.r.t. $s^* = (0, ..., 0, 1, ..., 1)$, represented by $A \geq 1$. A point 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, $l_B = \sum_{n=1}^{B} s_n$, and let l_A be the number of dimensions where point s has zero, and s^* has one, $l_A = \sum_{n=B+1}^{N} (1 - s_n)$. The total number of differences between s and s^* is the usual $l_B + l_A = \sum_{n=1}^{N} |s_n - s_n^*|$.

The total number of ones in any such point s is

$$k = A - l_A + l_B \tag{8}$$

There is exactly $\binom{B}{l_B}\binom{A}{l_A}$ points that are represented by a given pair (l_B, l_A). Or, if we eliminate l_B using equation (8), there is $\binom{B}{k+l_A-A}\binom{A}{l_A}$ points for a given k and l_A. In order to get the total number of points in x_k^A we need to sum these numbers over all possible l_A for a given k.

The range of l_A needs to be sorted out. There is a number of restrictions involving l_B, l_A, B, A and k. Conditions $C1 - C4$ are obvious, $C5$ is simply a definition of a diagonal partition x with respect to s^*,

$$\begin{array}{lc} C1 & l_A \leq A \\ C2 & l_B \leq B \\ C3 & l_A \geq 0 \\ C4 & l_B \geq 0 \\ C5 & l_B + l_A \leq \frac{B+A-1}{2} \end{array}$$

Firstly, note that $C3$ and $C5$ imply that $l_B \leq \frac{B+A-1}{2}$. By equation (6) we know that $B \geq \frac{B+A-1}{2}$, hence we may disregard $C2$. Secondly, $C4$ and equation (8) imply that $l_A \geq A - k$. Thirdly, $C5$ and (8) mean that

$$l_A \leq \frac{1}{2}\left(A - k + \frac{B+A-1}{2}\right)$$

All these imply that the range of l_A is

$$\max\{0, A-k\}, ..., \min\left\{A, \frac{1}{2}\left(A - k + \frac{B+A-1}{2}\right)\right\}$$

In other words, the number of elements in x_k^A is

$$\#x_k^A = \sum_{l_A=\max\{0,A-k\}}^{\min\left\{A,\frac{1}{2}\left(A-k+\frac{B+A-1}{2}\right)\right\}} \binom{B}{k+l_A-A}\binom{A}{l_A} \tag{9}$$

Step 3. We are going to show that the following difference is strictly positive

$$R^0 - R^A = \sum_{k=0}^{N} q_k (N-2k)\left(\#x_k^0 - \#x_k^A\right)$$

for any $A \geq 1$.

Recall that x^0 is the main diagonal partition, having $A = 0$. Equation (9) boils down to

$$\#x_k^0 = \begin{cases} \binom{B+A}{k} & k \leq \frac{B+A-1}{2} \\ 0 & k \geq \frac{B+A+1}{2} \end{cases}$$

Now let $A \geq 1$. Case 1. Suppose that $k \leq \frac{B+A-1}{2} - A$. Then the summation in (9) goes up to A, and we have

$$\#x_k^0 - \#x_k^A = \binom{B+A}{k} - \sum_{l_A=\max\{0,A-k\}}^{A} \binom{B}{k+l_A-A}\binom{A}{l_A}$$

It can be shown that in both cases, $A \geq k$ and otherwise, a standard formula in combinatorics[1] can be used to conclude that this expression is zero, $\#x_k^0 - \#x_k^A = 0$.

Case 2. If $k > \frac{B+A-1}{2} - A$ and $k \leq \frac{B+A-1}{2}$, then the summation does not go all the way up to A. Hence, by the same formula it must be that

$$\#x_k^0 - \#x_k^A = \binom{B+A}{k} - \sum_{l=\max\{0,A-k\}}^{\frac{1}{2}A-k+\frac{B+A-1}{2}} \binom{B}{k+l-A}\binom{A}{l} > 0$$

Step 4. The difference in revenues for a price-discriminating newspaper is strictly positive, if $A \geq 1$:

$$\begin{aligned} R^0 - R^A \quad &= \sum_{k=0}^{\frac{N-1}{2}-A} q_k\,(N-2k)\underbrace{\left(\#x_k^0 - \#x_k^A\right)}_{\text{zero}} + \\ &+ \sum_{k=\frac{N-1}{2}-A+1}^{\frac{N-1}{2}} q_k \underbrace{(N-2k)}_{\text{positive}}\underbrace{\left(\#x_k^0 - \#x_k^A\right)}_{\text{positive}} + \\ &+ \sum_{k=\frac{N+1}{2}}^{N} q_k \underbrace{(N-2k)}_{\text{negative}}\underbrace{\left(-\#x_k^A\right)}_{\text{nonpositive}} \end{aligned}$$

Step 5. Since the price-discriminating newspaper chooses uniform prices, the newspaper which cannot price-discriminate would also behave in the same way.

9 Literature

Banerjee, A., 1992, A simple model of herd behavior, *Quarterly Journal of Economics*, Vol. 107, pp. 797-817.

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

Calvo, Guillermo A. and Enrique G. Mendoza, 2000, Capital-Markets Crises and Economic Collapse in Emerging Markets: An Informational Frictions

[1] $\binom{B+A}{k} = \sum_{r=0}^{k} \binom{B}{r}\binom{A}{k-r}$.

Approach, *The American Economic Review*, Vol. 90, No. 2, Papers and Proceedings of the One Hundred Twelfth Annual Meeting of the American Economic Association, pp. 59-64.

Christiano, Lawrence J. and Terry J. Fitzgerald, 1998, The business cycle: It's still a puzzle, *Economic Perspectives*, Vol.22, pp. 56-83.

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.

Graber, Doris (1988): *Processing the News: How People Tame the Information Tide*, (2nd edn.) New York: Longman

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.

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.