News Media and Delegated Information Choice

Kristoffer Nimark[1] Stefan Pitschner[2]

[1]Cornell University

[2]Uppsala University and Swedish House of Finance

April 25, 2018

News Media and Delegated Information Choice

The man who buys a newspaper does not know beforehand what will be in the news.

Jacob Marschak, 1960

News Media and Delegated Information Choice

Every day, a large number of events occur, each of them potentially relevant for the decisions of households and firms

- No individual firm or household has the resources to observe all of these events
- News media monitor the state of the world and report the subset of events that are deemed most newsworthy

The editorial aspect of news reporting is pervasive in reality, but has not been studied in the existing literature.

The plan

I. Present stylized facts about news coverage from a statistical topic model

- Empirically document specialization of newspapers and state dependence of reporting decisions

II. Propose formal framework to study the editorial function of news media

- Characterize news outlets by **news selection functions** and study how they affect agents beliefs and actions

III. Delegated information choice in a simple beauty contest model

- Study how strategic motives, distributions and news selections functions interact to determine actions in a simple coordination game

Measuring News Coverage

Measuring News Coverage using the LDA

Latent Dirichlet Allocation (LDA) can be used to extract topics from text data

- Originally appeared in Blei, Ng and Jordan (2003)

Statistical topic classification

- A topic is (endogenously) defined by a frequency distribution of words
- Documents probabilistically belong to every topic

Input from researcher:

- Text corpus partitioned into documents
- Number of topics

Advantages:

- Objective and easy to replicate
- Naturally measures importance of topics

The news data

The News Data

Use text from Dow Jones Factiva data base

- Contains historical content from news papers, wire services and online sources beginning in 1970.

Extract text snippets from front page articles

- Focus on events considered to be most newsworthy by individual papers

The sample is two 90-day periods that include major events

- September 11 terrorist attacks
- Lehman Brothers Bankruptcy

Topic classification applied across both periods

- Estimating single model on both periods allows for "timeless" news topics.

The number of topics is set to 10 in benchmark specification

Newspaper sources

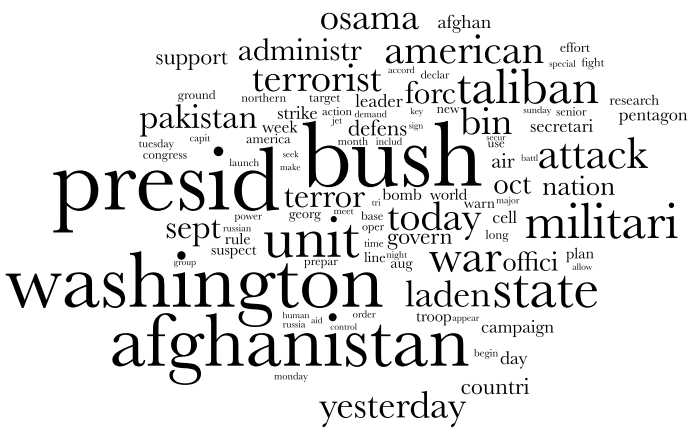
Newspaper Full Name	Short Name	Newspaper Full Name	Short Name
Atlanta Journal	AJ	The Las Vegas Review-Journal	LVR
Charleston Gazette	CG	The New York Times	NYT
Pittsburgh Post-Gazette	PPG	The Pantagraph	PG
Portland Press Herald	PPH	The Philadelphia Inquirer	PI
Sarasota Herald-Tribune	SHT	The Wall Street Journal	WSJ
St. Louis Post-Dispatch	SLP	The Washington Post	WP
Telegram & Gazette Worcester	TGW	USA Today	UT
The Boston Globe	BG	Winston-Salem Journal	WiSJ
The Evansville Courier	EC		

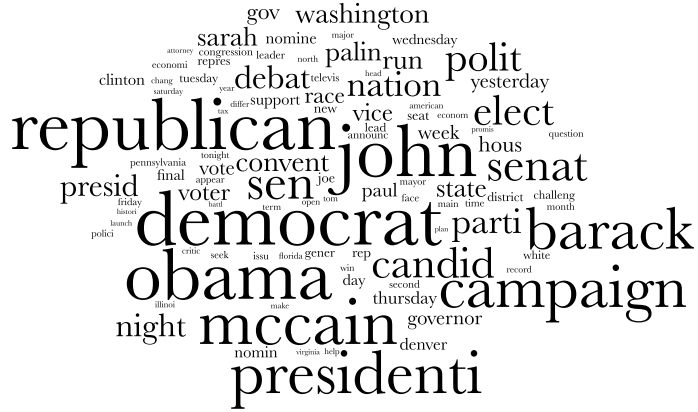
The estimated news topics

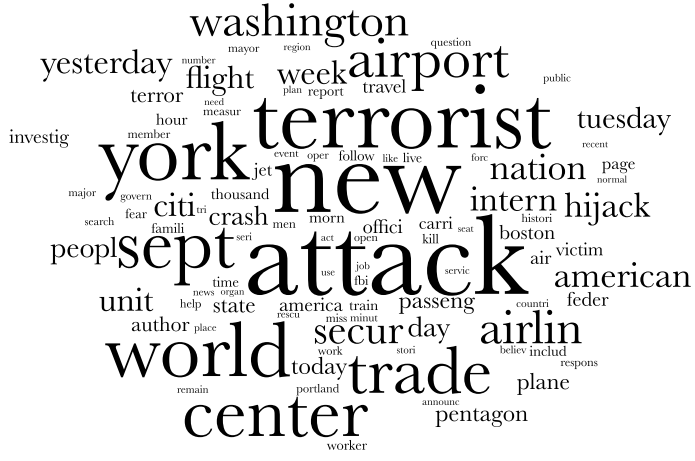
LDA topics classification

Topic	Words with the highest assigned probabilities (in descending order)
1	bush presid washington afghanistan unit state militari taliban war attack
2	democrat john republican obama mccain presidenti campaign barack sen senat
3	school year student counti high state univers review journal colleg
4	year old home ago time day just peopl like famili
5	financi washington billion market hous bush bank feder crisi govern
6	state million year plan new citi health compani say propos
7	mail daili staff charleston west counti said virginia st state
8	yesterday polic said offic anthrax court feder offici investig charg
9	attack new terrorist york world center sept trade airport airlin
10	citi new today palestinian aug georgia west day isra south

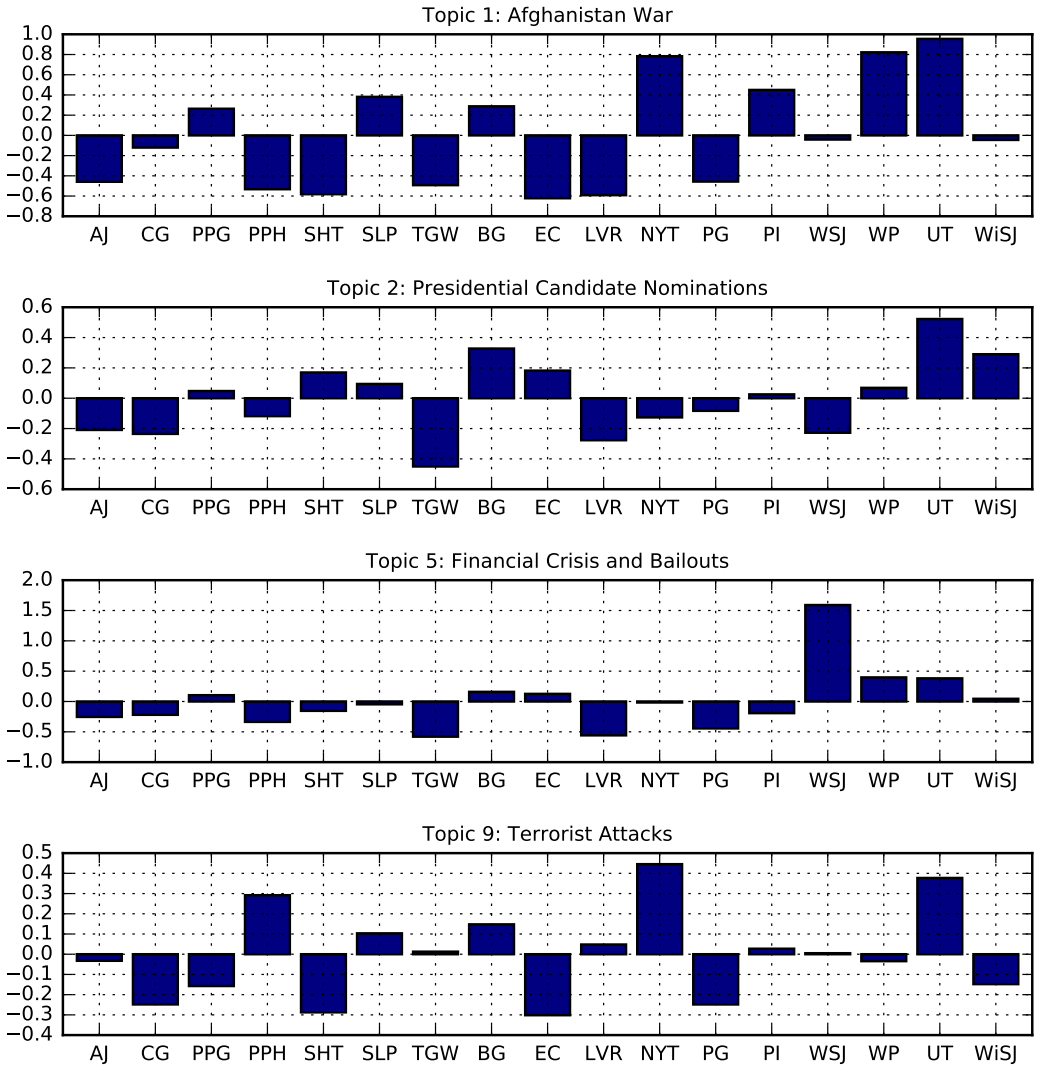
Topics 1,2,5 and 9 as word clouds







Specialization of newspapers



Two measures of news coverage over time

1. **Fraction of total news** devoted to topic k on day t

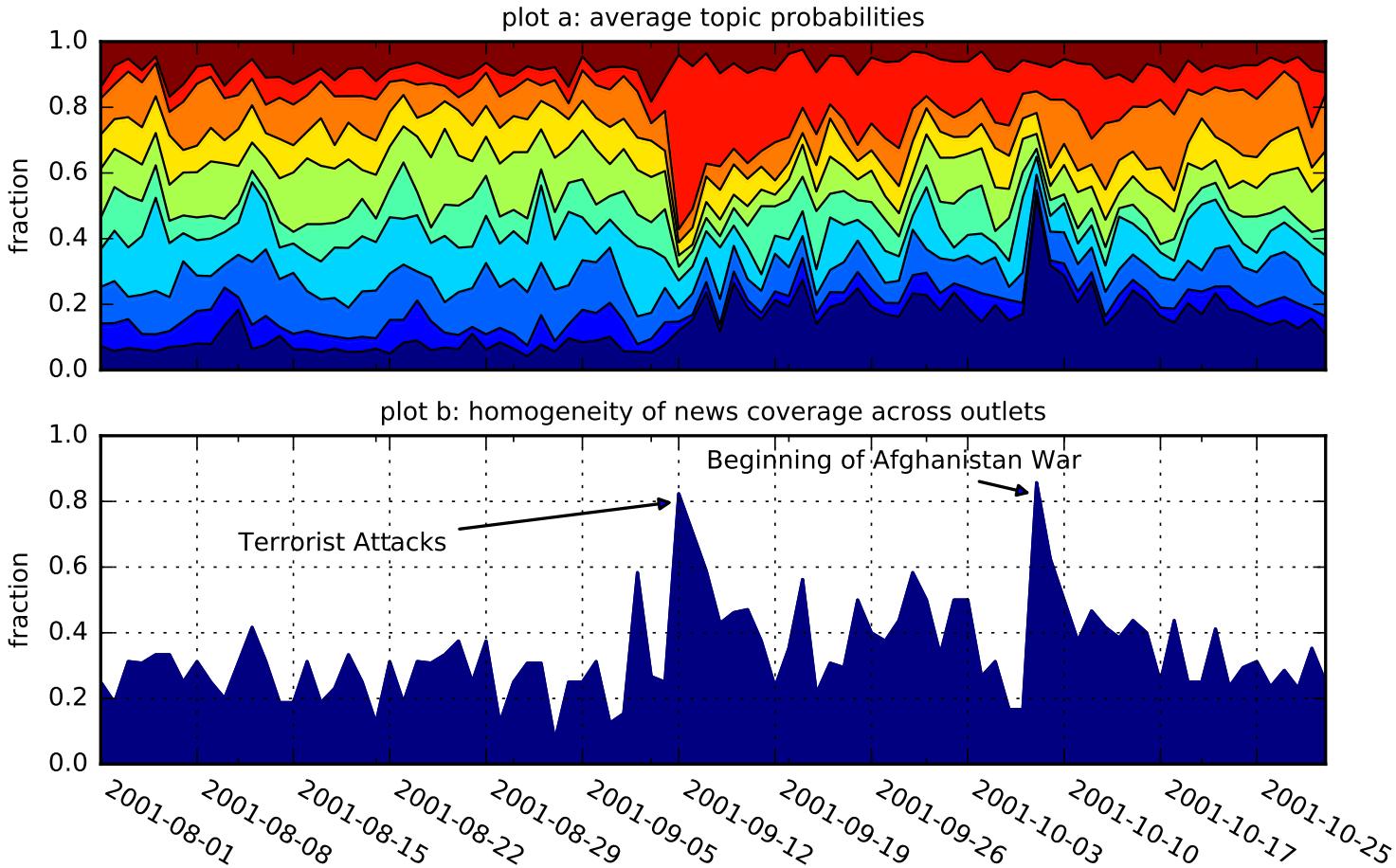
$$F_{t,k} \equiv \frac{\sum_d \theta_{t,d,k}}{D_t}$$

2. **Homogeneity** of news coverage

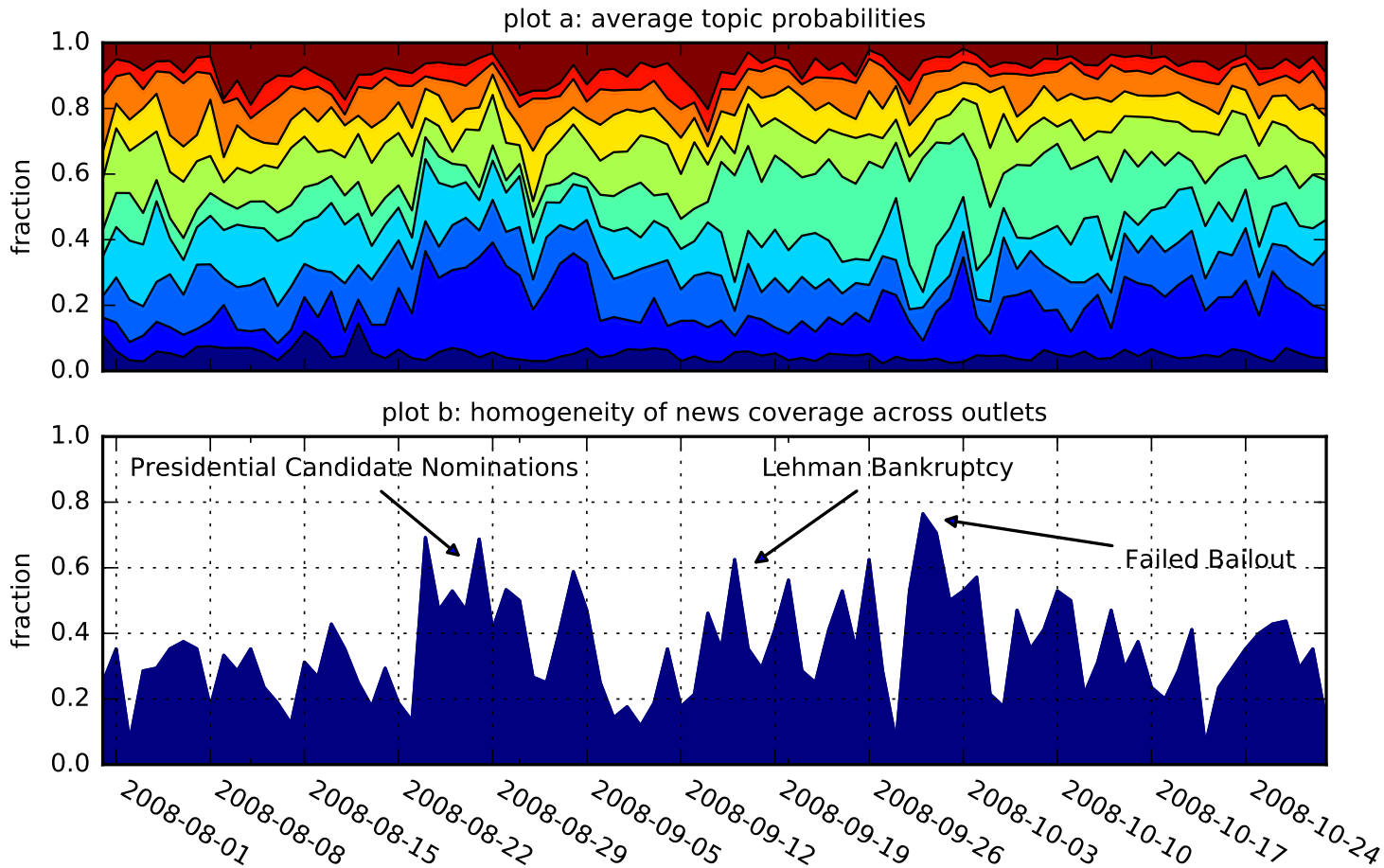
$$H_t \equiv \frac{\sum_m \mathcal{I}\left(\arg\max_k F_{t,m,k} = \arg\max_k F_{t,k}\right)}{M}$$

Newspapers are indexed by $m \in 1, 2, ..., M$.

Editorial decisions around 9/11



Editorial decisions around Lehman bankruptcy



3 stylized facts about news coverage

Newspapers provide specialized content

- Different papers tend to cover different topics

State dependence of news coverage

- Topic weights within a newspaper vary over time

Major events increase homogeneity of news coverage

- Events like the September 11 terrorist attacks and the Lehman bankruptcy made news coverage more homogenous across newspapers

What are the implications for agents beliefs and actions?

Delegated Information Choice

Formalizing the editorial role of news media

Model editorial decision making by introducing **news selection functions**.

Show that

1. Delegated state-dependent information choice can always reduce posterior entropy
2. News selection functions affect beliefs beyond the immediate content of what is reported
3. News selection functions determine degree of common knowledge about events

News selection functions

State of the world is $\omega \in \Omega_1 \times \Omega_2 \times ... \times \Omega_n \equiv \Omega$

Definition: *A news selection function* $\mathcal{S} : \Omega \to s \in \{0,1\}^n$ *is a mapping from possible states of the world* Ω *to an n-dimensional indicator vector s. The m-dimensional vector of reported outcomes is defined as* $\omega^s \equiv \{\omega_i : s_i = 1\}$.

Non-trivial selections are those such that $0 < m < n$.

What is the use of delegated information choice?

Consider an agent that is constrained in the number of news stories or events he can observe. He can either:

1. Decide ex ante which variables/events are most important to him and look them up himself
2. Delegate what to get information about to an information provider that can condition on the state before deciding what to report

By choosing (2) an agent can avoid spending resources on learning about events that are ex post not useful

- Clearly, state dependent news selection has to be delegated: One cannot both condition the choice on the entire state and not pay the cost of observing the entire state

What is useful to know depends on the specific setting, but it is still possible to show that delegated information choice generally optimal

Ex ante information choice

Definition: *An ex ante information choice function* $\overline{\mathcal{S}}$ *is defined by an n-dimensional indicator vector* $\overline{s} \in \{0,1\}^n$ *with the* $\overline{m}$*-dimensional random vector* $\omega^{\overline{s}}$ *of observed outcomes is defined as* $\omega^{\overline{s}} \equiv \{\omega_i : \overline{s}_i = 1\}$. *The indicator vector* $\overline{s}$ *is independent of the state* ω.

Delegated information choice can always reduce posterior entropy

Proposition: For any given ex ante information choice function $\overline{\mathcal{S}}$ such that $\overline{m} < n$ there exists a news selection function $\mathcal{S}^*$ with $m^* = \overline{m}$ that achieves a lower posterior entropy.

Proof:

1. Fix an ex ante information choice $\overline{s}$.
2. Define the candidate news selection function $\mathcal{S}^*$ so that $s^*(\omega) = \overline{s}$ in every state except ω'.
3. In state ω', for an i and j such that $\overline{s}_i = 1$ and $\overline{s}_j = 0$, set $s_i^*(\omega') = 0$ and $s_j^*(\omega') = 1$.
4. Since $s_j^* = 1$ only in the state ω', $\mathcal{S}^*$ thus reveals the vector $\omega^{\overline{s}}$ if $\omega \neq \omega'$ and the entire state vector if $\omega = \omega'$.
5. Conditioning on additional information reduces entropy

$$H\left(\omega \mid \omega^{\overline{s}}, s_j^*\right) < H\left(\omega \mid \omega^{\overline{s}}\right).$$

since ω is not independent of s_j^* given $\omega^{\overline{s}}$.

News selection is by itself informative

A news report is always informative about the immediate content of the report

- However, if news selection is state-dependent and understood by agents, the selection is by itself informative.

The news selection function $\mathcal{S}$ associates a pair $\{\omega^s, s\}$ with each state of the world.

- Given prior beliefs $p(\omega)$ and agent who observes ω^s and knows the function $\mathcal{S}$ has posterior beliefs $p(\omega \mid \omega^s, s)$

But in general $p(\omega \mid \omega^s) \neq p(\omega \mid \omega^s, s)$

Beliefs and news selection functions

Proposition: Posterior beliefs about the unreported stories $\omega^{\not s}$ coincide with $p\left(\omega^{\not s} \mid \omega^{s}\right)$ only if the probability of reporting about $\omega^{\not s}$ is conditionally independent of ω^{s}. That is

$$p\left(\omega^{\not s} \mid \omega^{s}, s\right) = p\left(\omega^{\not s} \mid \omega^{s}\right)$$

only if

$$p\left(s \mid \omega^{s}\right) = p\left(s \mid \omega^{s}, \omega^{\not s}\right).$$

Proof: Follows immediately from Bayes' rule

$$p\left(\omega^{\not s} \mid \omega^{s}, s\right) = \frac{p\left(s \mid \omega^{s}, \omega^{\not s}\right)}{p\left(s \mid \omega^{s}\right)} p\left(\omega^{\not s} \mid \omega^{s}\right).$$

Implications of state dependent news selection

Corollary 1: Reporting of the least newsworthy outcomes leads to the largest reductions in posterior uncertainty, reporting of the most newsworthy outcomes leads to the smallest reduction in uncertainty.
Example: Very few events could have crowded out the news coverage of the 9/11 terrorist attacks

Corollary 2: Two individuals who learn about the same outcomes from different information providers may draw different inference about non-reported events.
Example: If neither Fox News nor MSNBC report new revelations of Trump-Russia collusion, only the audience of MSNBC can draw the conclusion that there are no new revelations to report.

Public information and common knowledge

In the literature, **public information** means information that is **common knowledge**

- E.g. Morris and Shin (AER 2002), Angeletos and Pavan (Econometrica 2007), Angeletos, Hellwig and Pavan (Econometrica 2007), Amador and Weill (JPE 2010), Cespa and Vives (REStud 2012), Hellwig and Veldkamp (REStud 2009).

Common knowledge is a much stronger assumption than the everyday meaning of **publicly available**

- Not all information that is publicly available is observed by everybody and not all information that is observed by everybody is known to be observed by everybody...and so on.

The degree to which knowledge about an event is common among agents is important in strategic settings

- Bank runs, currency attacks, political regime change, price setting and production decisions in macroeconomic models etc

Common knowledge and news selection functions

In the data we saw that some events were widely reported

- 9/11, Afghan War, Lehman Bros bankruptcy etc

These events were arguably not only mutual knowledge but common knowledge among large groups of people after they happened.

But **what** made these events common knowledge?

The sets of more and less interesting states

Definition: (Set of more newsworthy states)
$\mathcal{M}\left(\omega_j^{s*}\right) \equiv \{\omega : s_j = 0\}$ is the set of states that contain m outcomes more newsworthy than ω_j^{s*}.

Definition: (Set of less or equally newsworthy states)
$\mathcal{L}\left(\omega_j^{s*}\right) \equiv \{\omega : s_j = 1\}$ is the set of all states that do not contain m outcomes more newsworthy than ω_j^{s*}.

The set $\mathcal{L}\left(\omega_j^{s*}\right)$ is the complement to $\mathcal{M}\left(\omega_j^{s*}\right)$ in Ω and contains the set of all states that are consistent with ω_j^{s*} being included in the reported vector ω^{s*}.

Delegated Information Choice and Common Knowledge

Definition: (Common knowledge) An outcome ω_j^* is common knowledge if every agent $a \in \mathcal{A}$ assign probability one to $\omega_j = \omega_j^*$ in all states that any agent believe is possible.

Lemma: The outcome ω_j^* is common knowledge if and only if for all states $\omega \in \bigcup_{a \in \mathcal{A}} \mathcal{L}^a \left(\omega_j^{s*} \right)$ we have that $\omega_j = \omega_j^*$.

Sufficient conditions for approximate common knowledge

Proposition

(i) If $p(\omega_i \mid \omega_j) = p(\omega_i)$ for all $i \neq j$, and

(ii) if for every i and any ω_i^* such that

$$\lim_{\omega_i \to \omega_i^*} \mathcal{M}^a(\omega_i) = \varnothing \forall i, a$$

we have that

$$\lim_{\omega_i \to \omega_i^*} p(\omega_i) = 0,$$

then the value of ω_i approaches approximate common knowledge in the limit $\omega_i \to \omega_i^*$.

A beauty contest model with news selection

A beauty contest model with delegated information choice

The model is an abstract coordination game in the spirit of Morris and Shin (2002).

- Agents want to take an action that is close to a fundamental variable and the action of other agents

Two essential differences relative to original model:

1. Agents have heterogeneous interests
2. Agents delegate the information choice to specialized benevolent information providers that monitor state but report only subset of outcomes

The model incorporates these features in an as simple set-up as possible

A beauty contest model with delegated information choice

1. News selection functions are endogenous
2. Degree of common knowledge is endogenous and depends on both preferences and distributions
3. Agents respond stronger to events that believe are closer to common knowledge
4. With continuous distributions and quadratic loss, extreme events become approximate common knowledge
5. Delegated information choice introduces correlation in actions of the same sign as the strategic motive in utility functions

Conclusions

Documented stylized facts about news coverage

- Different newspapers provide specialized content and tend to cover different topics to different degrees
- Major events increase homogeneity of news coverage.

Formalized the editorial service provided by news media

- Provide formal justification for delegated state dependent information choice
- News selection function influence what agents infer from a given reported event
- News selection functions determine degree of common knowledge

Different perspective from ex ante information choice literature

- e.g. Sims (2003), Mackowiak and Wiederholt (2009, 2010), Alvarez, Lippi and Paciello (2011), Matejka (forthcoming), Matejka and McKay (2015), Stevens (2014),Grossman and Stiglitz (1980), Veldkamp (2006a,2006b), Van Nieuwerburgh and Veldkamp (2009, 2010)