

Shaun Bevan

The UK Policy Agendas Project Media Dataset Research Note: *The Times (London)*

19-09-2011

Politics is a complex system of interactions and reactions from within and outside of government. One of the most important elements of the political process from outside of government is the media not just for what it represents, but for what it measures. Media attention captures both salience and events both of which factor heavily into the political decision-making process. However, how best to measure and conceptualize media attention, especially over time, is a difficult task as debates concerning the appropriateness of different sources, sampling techniques and content coding abound. This research note covers and explains the decisions made in relation to the UK Policy Agendas Project media database that consists of content coded front-page headlines from *The Times (London)* from every Wednesday edition over the period 1960 to 2008. It further discusses the broad patterns of attention in the resulting dataset including the number of policy versus non-policy stories, attention to different policy issues and the amount of domestic versus international attention in the paper. This introduction to the data is intended to allow its users to properly account for its unique nature in their own work.

Choosing a Source: Why *The Times (London)*?

The UK Policy Agendas Project focuses on the content and not the tone of political agendas whether they are government, public or media agendas. Therefore, when the decision was being made concerning which source to use for a historical database of media attention it was important to choose a source focused on content and not a particular political ideology or

lack thereof. In the UK two national papers are by their very nature focused on content, *The Times (London)* and *The Daily Telegraph* both of which are newspapers of record. Considered in the French to be *presse de reference*, these papers are consistently viewed as meeting a high standard of journalism nationally and internationally while maintaining the same traditions and coverage over time. While either paper matches our general criteria, the availability of a longer historical record for *The Times (London)* made it the natural choice for our purposes.¹ Furthermore, the appeal for *The Times (London)* to a wider assortment of political views than the overwhelmingly Conservative readership of *The Daily Telegraph* is an advantage in supporting its legitimacy as a single source even though the UK Policy Agendas Project is only focused on content (Ipsos MORI 2005).

A decision also had to be made in regards to how to measure media attention with two main alternatives. The first, a sample of media attention from the entire paper is appropriate and in many ways better measures events as both salient and non-salient issues are represented throughout the paper. The US Policy Agendas Project (www.policyagendas.org) used this sampling technique for its media dataset of the *New York Times Index*, but the broad nature of the data and the limited sample size are clear limitations and the resulting data performs poorly as a measure of media salience. The second alternative is a focus on front-page media attention. The front-page not only represents the media agenda and events, but the prioritization of issues by the media and the public. Because of this prioritization a front-page media agenda better matches the UK Policy Agendas Project's goals of capturing the attention in government and public agendas (Boydston 2008, pp. 30-32). The Project's media database therefore measures front-page media attention. However, while front-page media attention closely matches the goals

¹ While our database stops in 1960 due to practical concerns *The Times Digital Archive* contains searchable and machine readable stories from 1785 to 1985 allowing others in future work to fully extend a media dataset for over 200 years.

of the project it may not be appropriate for other studies concerned with overall media attention of both salient and non-salient events. Other researches should therefore take care to note if the data fits their conceptualization of media attention.

Finally, a sampling strategy also needed to be decided upon as the sheer number of stories on the front page over a 49 year period proved daunting. As nearly all of the UK Policy Agendas Project datasets focus on the yearly level the main requirement was that the sampling technique captured the overall media agenda throughout the year. However, prime minister's questions (PMQs) which make up one of the project's key datasets occur at a much shorter weekly timeframe. PMQs were asked weekly on every Wednesday when parliament is in session over the period the project studies (1997-2008). In order to control for the events leading up to PMQs the content of the front-page of the newspaper of that day is essential. Therefore, front page media attention for every Wednesday over the entire 1960 to 2008 period was gathered as it allowed for both a proper aggregation for the other datasets created by the project and a close match to the PMQs dataset. The headlines² of these stories were then blind-coded according to UK Policy Agendas Project major topic codes³ by two researchers and compared leading to 80% intercoder reliability with disagreements resolved by the project manager.

² The use of headlines instead of the complete text of newspaper stories was a practical decision based on available resources and the complexity of the data gathering process. However an in-depth pilot comparing the coding of complete stories vs. headlines alone showed that in more than 90% of cases the major topic codes assigned matched when coding the complete stories or headlines alone. The majority of disagreements between these two techniques occurred in non-policy topics such as arts, history & culture and entertainment (28) or Sports and Recreation (29) where more ambiguous headlines were often used.

³ Several major topic codes such as weather (26), fire (27) and human interest stories (34) were added to the codebook to cover all the non-policy related stories covered by the media. These additions follow the strategy employed by the US Policy Agendas Project's *New York Times Index* dataset.

The result of these decisions and the content coding process was a database of 21,854 front page headlines from every Wednesday edition of *The Times (London)* over the 1960 to 2008 period.⁴

Understanding the Dataset: Changes in *The Times (London)* Over Time

From 1960 to 2008 two key changes in the format of *The Times (London)* occurred. These include a change to a standard broadsheet format on the 3rd of May 1966 and a change to a tabloid format in November 2003. Capturing front-page attention during the broadsheet era was completely straight forward with the front page conceptualized as all headlines printed on the first page of the paper ranging from 8 to 11 headlines on average. Prior to the change to a broadsheet format *The Times (London)* took its role as a paper of record quite seriously and presented information on births, deaths, marriages, sports scores, play and concert reviews, and even TV schedules on the first several pages. As a result, the section headed as “News” started anywhere between 5 and 17 pages into the paper. A review of the stories that headed the “News” section clearly demonstrated that these stories matched what we would normally think of as front-page news with the biggest and most important headlines of the day being listed at the start of this section. Therefore in the pre-broadsheet era the front page was conceptualized as all “News” stories on the first full page of “News” in the paper each day.⁵

The change to a tabloid format for the paper in November 2003 presented its own unique issues as no clear delineation between front-page news and the remainder of the news existed as with the broadsheets or the pre-broadsheet era. Often, the front page of the tabloid format of the

⁴ A break in this dataset occurred from December 1978 to November 1979 and this break is discussed in detail in the next section.

⁵ “News” stories included before the first full page of “News” in the paper are also included in the dataset leading to a slight oversampling in this period occurring largely at random based on the amount of other front page attention.

paper consisted of a single, but in general no more than three headlines on any particular day. A brief “Inside” overview also printed on the first page highlights different stories throughout the paper and not just front-page news as it is normally conceptualized. The second page of the paper in the tabloid years contains two or more editorial pieces covering general concerns of the day, human interest stories and other matters and does not represent front-page news in the manner we normally think of it. The third and fourth pages contain many additional “News” headlines and can best be considered an extended front page. In order to gather front page media attention all stories from the first, third and fourth pages were gathered leading to a noticeable oversample in these years. However, the exclusion of the fourth page led to a systematic under-sampling when compared to previous years which presents an even larger issue for statistical analyses. Nevertheless studies focused on the count of stories should be aware of this change in the baseline in the tabloid years and address it as necessary for any statistical or qualitative analyses.

A Break in Media Coverage: *The Times* Goes on Strike

Unfortunately the use of *The Times (London)* does not allow for a continuous account of media attention for the entire 1960 to 2008 period. *The Times* suffered from the economic downturn and an era of labour disputes like so many of the businesses and professions in the late 70s and management was forced to shut down the paper from 1st of December 1978 to the 12th of November 1979.⁶ For a nearly one full year no stories were published and this is reflected in the dataset. This fact prevents a detailed weekly, monthly or quarterly analysis of the data when including this period. An aggregation to a yearly measure only masks the issue as the inclusion

⁶ The shutdown of *The Times* was not unique with the paper itself reporting on the striking and shutdowns of numerous other papers up until its own shutdown.

of just six front pages in 1979 leads to dramatically lower levels of attention and variance than in complete years. The use of percentages of attention instead of raw counts is a partial fix, but 1979 remains a problem year in the data and researchers should take care when making conclusions that use this portion of the data. However, unlike many data issues in the social sciences this data is not actually missing and represents true zeros in the number of stories in *The Times (London)* over this period.

With that said, missing data techniques for the strike period may be appropriate depending on how the researcher conceptualizes the data. For example, if it is conceptualized as a measure of salience or events it can be treated as missing, but if it is conceptualized as a measure of media attention this period should be treated as real zeros as no media attention from *The Times* actually existed. However, techniques for imputing missing data points will no doubt undercount the amount of attention to labour issues and to the media during this time period as both the number of strikes and the concern over a striking media will not be captured. With this downward bias in attention to these issues in mind the use of an imputed series for this period can be properly used for quantitative and qualitative analyses.

Patterns of Media Attention

The media dataset gathered for the UK Policy Agendas Project as discussed above exhibits several interesting and informative patterns of attention in relation to policy content and domestic vs. international stories. The remainder of this research note focuses briefly on a few of these patterns of attention providing an overview of the types of stories contained in the dataset.

The Times (London) like all modern media outlets contains a mix of both domestic and international stories. While the UK Policy Agendas Project has several international focused

major topic codes⁷ an additional level of domestic versus international headline coding proved necessary.⁸ A large number of stories fitting domestic major topic codes, such as the performance of the US stock market (banking, finance and domestic commerce [15]) or the prices set by OPEC (energy [8]), were clearly focused overseas, but had a domestic impact on the UK. Similarly, many international stories focused on the place of the UK in world affairs, such as involvement in the Iraq War (defence [16]), while many others focus on purely international events such as human rights abuses in mainland China (international affairs and foreign aid [19]). To account for this a dummy variable indicating if the UK, through its people or government, was directly involved in a story (coded 0) or not (coded 1) was created with the default assumption being that a story involved the UK (coded 0). A comparison of the total number of stories with a domestic vs. an international focus according to this dummy variable is presented in Figure 1.

[insert Figure 1 about here]

In Figure 1 it is clear that the main focus of *The Times (London)* is on issues the UK is directly involved in both at home and abroad. However, with the change to a broadsheet format with a traditional front-page in 1966 an increase in internationally focused stories occurred. This change was gradual and not immediate, from a low baseline in the pre-1966 period to a new and fairly steady high of roughly 100 stories in 1968. This represents an actual change in the content of the paper and not an artifact of the sampling strategy for the pre-1966 period discussed in detail above. Furthermore, starting in the mid-1980s the number of purely internationally focused

⁷ Namely defence (16) foreign trade (18) and international affairs and foreign aid (19).

⁸ See Table A1 in the appendix for a complete list of major topic codes used in the media dataset

stories begins to decline and even declines during the over-sampled tabloid period. There could be many reasons for this change from a change in the focus of the paper, to less direct newspaper headlines. However, the decline of the British Empire hit a major turning point in the late 70s with the independence of the majority of the remaining British colonies leading to a greater focus on international and world affairs rather than colonial affairs in agendas such as the Speech from the Throne and Acts of Parliament. The change in the media agenda reflects these changes to Britain and British politics then and is a likely explanation.

The media does not just focus on government and on policy. Many major headlines and front page stories instead focus on events such as fires, extreme weather and of course entertainment and celebrities and *The Times* is no different than any other media outlet in that regard. The share of policy related content in *The Times* is however quite important for the UK Policy Agendas Project and for others interested in policy. Figure 2 presents the number of policy and non-policy stories contained in *The Times* according to UK Policy Agendas Project major topic codes. The majority of policy stories have nothing to do with any particular piece of legislation, but instead focus on an issue area, such as healthcare (3) or social welfare (13), that is related to the policies that government attends to. Non-policy stories on the other hand cover sporting results, celebrity marriages and divorces and other events that parliament does not address while in session.

[insert Figure 2 about here]

In Figure 2 it is clear that the majority of the content in *The Times (London)* concerns policy on some level although the amount of attention to policy issues was on the decline until

the start of the tabloid era with occasional spikes in the amount of policy content such as in the mid-80s with high levels of attention to defence (16) and international affairs and foreign aid (19). While the spike in the tabloid era is in part due to sampling the increase is actually far too high to be due to the change in paper format alone and suggests that despite the name of the format change meaning *The Times (London)* has become more policy oriented in recent years. This is further demonstrated by the lack of a similar increase in the number of non-policy stories with the change in format. Of course this may be due to a restructuring of the paper itself, but nevertheless the majority of front-page attention in the tabloid era remains focused on policy issues.

The share of the paper devoted to particular policy issues over time is another important pattern. As a paper of record *The Times* is viewed as reporting the news as it happens, but not all events are newsworthy. Agricultural (4) news for instance is unlikely to receive high levels of media attention except during a crisis like the outbreaks of foot-and-mouth disease. Other issues such as wars and other defence (16) related issues will likely always be high on the agenda. A different source for media attention would of course have a different pattern of attention though, such as *The Financial Times* which would focus more on business issues and economic affairs. The similarity between the content of different national newspapers is staggering though as anyone who has seen a run-down of headlines on a comedy panel show can attest to. Figure 3 presents the pattern of attention in *The Times (London)* by policy oriented major topic codes over time. While such an investigation does not confirm or deny *The Times* as a suitable source for media attention it does allow for an assessment of the issues and how they are covered by the paper. A fact that is important to grasp in any study that uses the source.

[insert Figure 3 about here]

Figure 3 shows the patterns of attention over time to the nineteen policy oriented major topic codes. Some events such as the labour (5) disputes of the 1970s are quite evident from Figure 3, others such as foot-and-mouth disease, an agricultural (4) issue, are less pronounced and clear. From the figure it is also clear that certain issues, namely defence (16), international affairs and foreign aid (19), government operations (2), take up a large share of the media agenda. So too do law, crime and family issues (12) as well as transportation (10) that includes reports of road accidents, criminal driving offenses and road safety in general. Other issue areas are not as well represented, but tend to come on the media agenda in line with major events, such as with concerns over environmental (7) issues in the mid-2000s. Importantly the oversampling in the tabloid era appears to have only resulted in increased attention in two issue areas, namely health (3) and law, crime and family issues (12) suggesting that controlling for the oversampling should be done on an issue by issue basis.

Conclusion

The UK Policy Agendas Project media dataset focuses on *The Times (London)* front-page headlines from 1960 to 2008 and as such does an excellent job of capturing the content of the salient media agenda. A break due to strikes in the late 70s and the directed construction towards the salient agenda does however mean that the dataset is not suitable for all research. Those interested in using the data for their own work should take care to understand the construction and patterns of attention of the dataset to determine if it suits their purposes. However, in relation

to the prioritized agendas that are the focus of the UK Policy Agendas Project the dataset serves its purpose well with a robust and common measure of salience and salient events over time.

References

- Boydston, Amber E. 2008. "How Policy Issues Become Front-Page News." Doctoral Dissertation, The Pennsylvania State University.
- Ipsos MORI. 9 March 2005. "Voting intention by newspaper readership". Retrieved 16 September 2011 from: <http://www.ipsos-mori.com/researchpublications/researcharchive/755/Voting-Intention-by-Newspaper-Readership.aspx>.

Figure 1: Quarterly *The Times* (London) Domestic and International Headlines

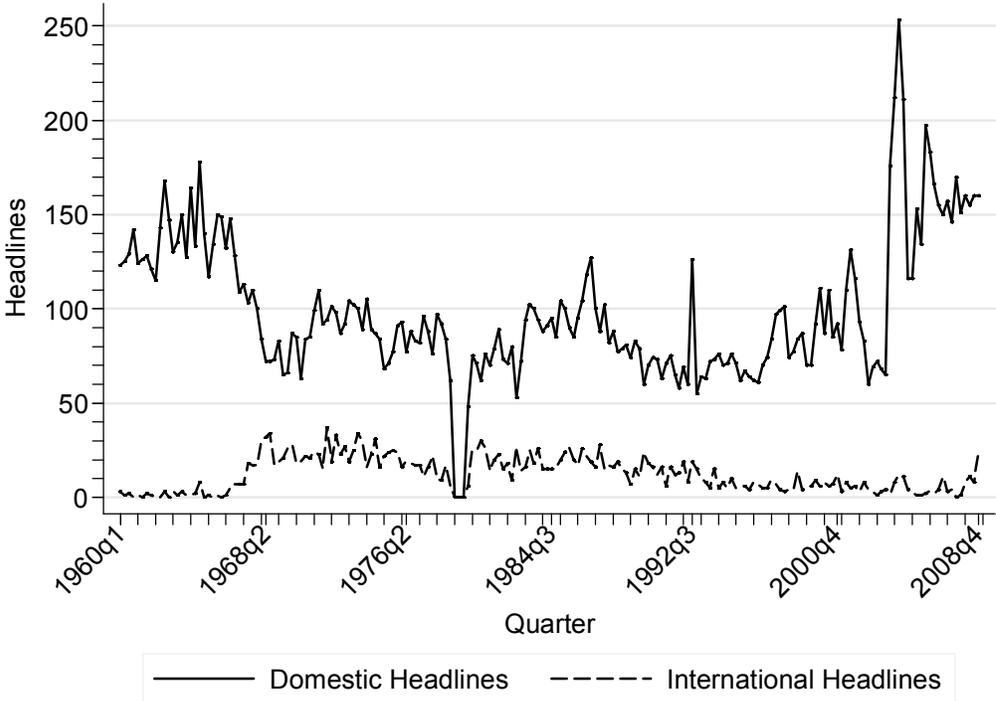


Figure 2: Quarterly *The Times* (London) Policy and Non-Policy Headlines

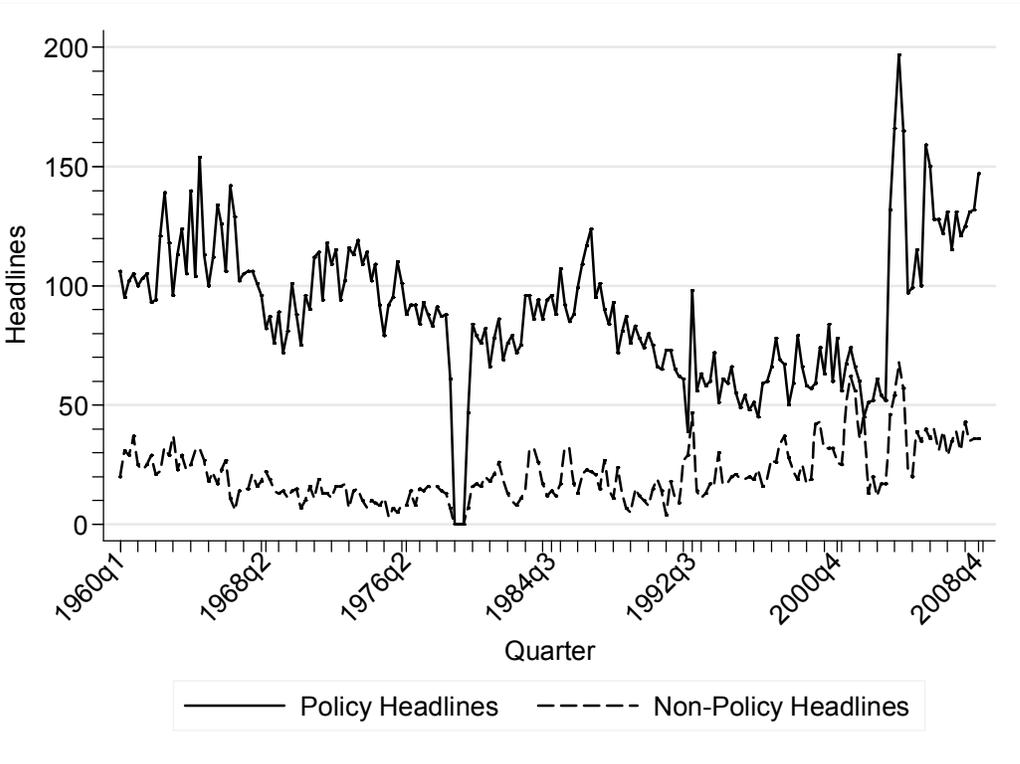
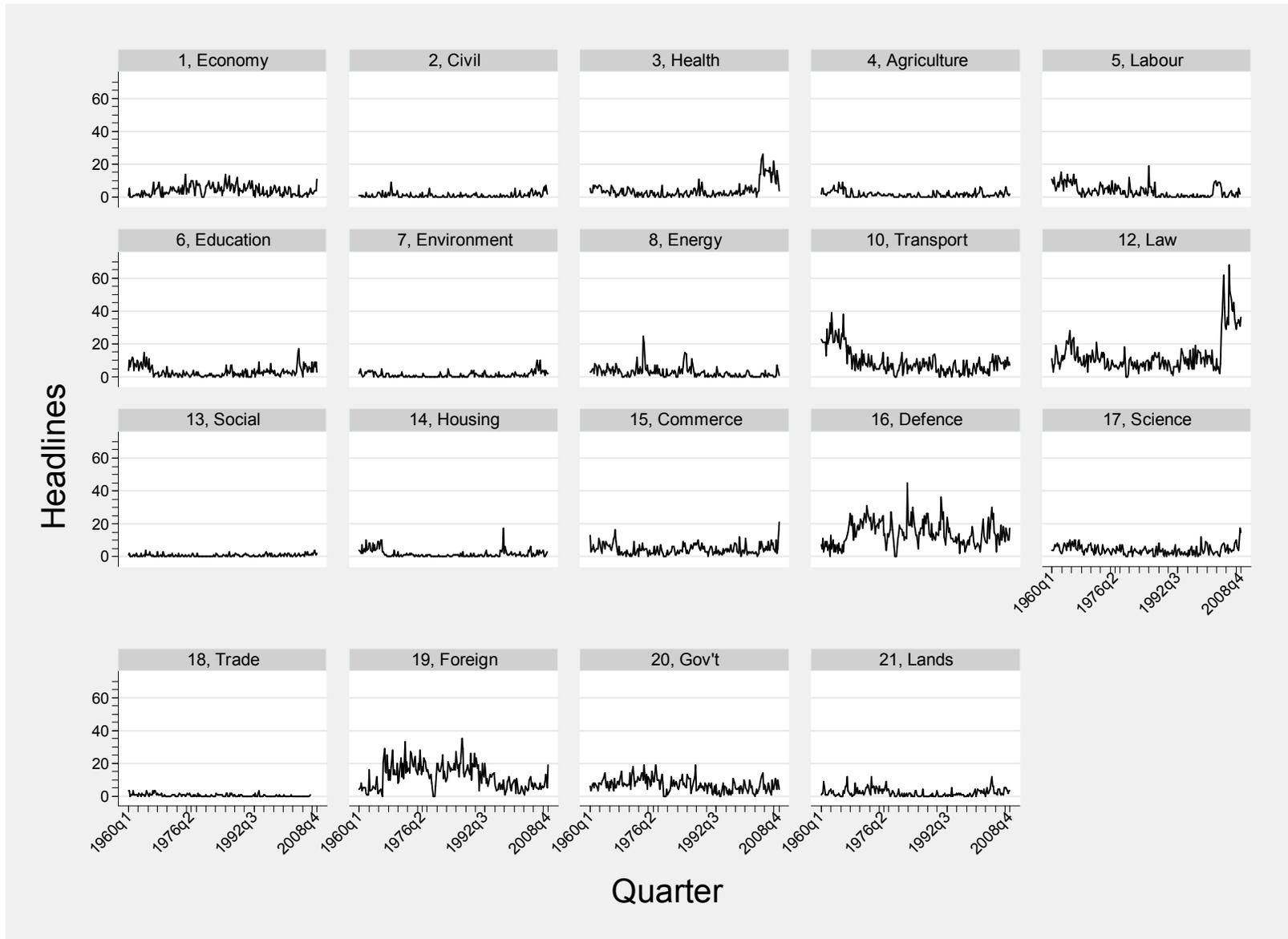


Figure 3: Quarterly *The Times* (London) Headlines by Policy Area



Appendix

Table A1: UK Policy Agendas Project Media Dataset Major Topic Codes

Topic	Abbreviation	Name
1	Economy	Macroeconomics
2	Civil	Civil Rights, Minority Issues (*Immigration) and Civil Liberties
3	Health	Health
4	Agriculture	Agriculture
5	Labor	Labor and Employment
6	Education	Education
7	Environment	Environment
8	Energy	Energy
10	Transport	Transportation
12	Law	Law, Crime and Family Issues
13	Social	Social Welfare
14	Housing	Community Development, Planning and Housing Issues
15	Commerce	Banking, Finance, and Domestic Commerce
16	Defence	Defence
17	Science	Space, Science, Technology and Communications
18	Trade	Foreign Trade
19	Foreign	International Affairs and Foreign Aid
20	Gov't	Government Operations
21	Lands	Public Lands, Water Management, Colonial and Territorial Issues
24‡	Regional	Regional and Local Government Administration
26‡	Weather	Weather and Natural Disasters
27‡	Fire	Fires, accidents and other manmade disasters
28‡	Arts	Arts, History & Culture and Entertainment
29‡	Sport	Sports and Recreation
30‡	Death	Deaths, Death Notices and Obituaries
31‡	Church	Churches and Religion
33‡	Parties	Political Parties
34‡	Interest	Human Interest
99‡	Other	Other, Miscellaneous and Uncodable

‡ Indicates a Non-Policy Major Topic