Election Unspun, Unspun
Martin Moore and Gordon Ramsay

From Monday 5\textsuperscript{th} January 2015, the day UK election campaigning really started, up to Sunday 26\textsuperscript{th} April, less than a fortnight before polling day, there was not a single national news article about Harrogate & Knaresborough constituency. Not one article from over 450,000 articles published across the national press, or the BBC, Channel 4 or Sky online. This despite it being a marginal constituency (2010 Conservative majority: 1,039), and despite it being situated in a local government area containing no daily local paper. It was one of 54 constituencies that were not referred to at all from January through April. Exactly half of all UK constituencies – 325 – were covered in six or fewer articles across that 16-week period.

We know this because we analysed over 450,000 articles using a digital content analysis tool we developed to monitor political and policy news, called Steno. Like the seventeenth century geologist from whom it takes its name, Steno digs through layer upon layer of online news articles from across the national press and broadcasters. With the data it gathers and analysis it enables, we have been able to produce, for each of the 18 weeks leading up to the 2015 UK election, a week-by-week analysis of news coverage of the election – called Election Unspun (at electionunpun.net).

Two years ago we could not have done this. Like many other academics in the UK and around the world, if we wanted to do statistical news content analysis, we had to do it the old-fashioned way – head-deep in text, video or audio. In our case the turning-point was an analysis of news coverage of the Leveson Inquiry and its aftermath. This involved reading over 2,000 articles from across the nationals and ‘tagging’ each article by hand according to a series of criteria including: publication, publisher, headline, byline, day/date published, sources referenced, issues covered, and tone.\footnote{http://mediastandardstrust.org/mst-news/media-standards-trust-leveson-coverage-analysis-launched/}

This was not only exhausting, it took a very long time. It also meant that the analysis was necessarily constrained. Criteria had to be set out in advance. It was difficult to play around with new hypotheses. It was hard to see patterns over time or within publications. But this was generally standard procedure for quantitative content analysis.

We were convinced that there had to be a better way. In a digital world, when most print content is also published online (along with a whole lot more), it seemed peculiarly twentieth century to be hand counting. Journalisted.com, a site we developed in 2007, had showed us it was possible to identify articles published on news website and link them to a journalist’s profile to produce a sort of souped up LinkedIn for journalists. Churnalism.com, an online service we

developed in 2011, showed what we could do through textual comparison of millions of news articles and press releases. Putting together what we learned from developing these we built Steno.

Steno is able to collect articles from almost any news site (in any language that uses an alphabetical system) and identify the basic information about each article – its author, publisher, date/time published, as well as – crucially – the full text of the article, shorn of captions, adverts, sidebar links, subheadings, etc. We can then use it to filter the articles – for example according to which section of the site they were published - /sport, /fashion, /politics. And we can exclude the content we are not interested in analyzing (such as sponsored microsites).

But Steno’s most important asset is its capacity to tag. Using Steno we are able to tag every article based on its content. We can tag, for example, every article that refers to Ed Miliband, or to Nigel Farage, or to Leanne Wood. We can, using tags, distinguish between articles that refer to David Miliband, but not to Ed Miliband. We can look for every article that mentions Leanne Wood, Plaid Cymru and the NHS. And, we can tag every article that refers to any of the UK’s 650 Parliamentary constituencies. Since tagging can be based on any string of text (including in headlines or URLs, as well as the main body of text), there is no limit on what can be selected, tagged and analysed.

The UK election campaign 2015 is Steno’s baptism of fire. This is the first time the tool has been used to do fast, weekly analysis of ongoing news coverage and then publish that analysis every Thursday, confident that the numbers are accurate.

We ensure high levels of accuracy by building in a function to flag up articles about which there may be a degree of uncertainty (i.e. an article containing “Labour” and “NHS” may be about childbirth or the Royal Baby, rather than the party’s health policies). We devoted a good few hours each week ensuring that any mis-tagged articles were corrected.

Knowing that numbers rarely speak for themselves, we decided to work with Soapbox, a data visualization company specializing in politics, to help communicate our findings and spread them online and via twitter.

It has not always been straightforward. We tried to measure the personalization of the campaign by studying the adjectives used to describe party leaders. After wading through a mixture that included nasty, clueless, arrogant, gormless, passionless, hapless, etc, we quickly realized there were more than we could identify, and that any comprehensive tonal analysis would also have to include nouns and adverbs. We abandoned that one for now, although tonal analysis is the logical next step for Steno.

At the same time we have stumbled upon unexpected but important findings. We discovered, for example, that based on polling numbers, UKIP were being under-represented in the press while the LibDems were being over-represented. While UKIP were polling at around 13-15%, only 8% of Party political coverage was
about UKIP. By contrast, though the LibDems were polling at below 8%, they were receiving almost 12% of Party political coverage.

Other findings have been less surprising but still significant. On partisanship, for example, the Conservatives benefited from almost three times as many positive front page leads as Labour from 5th January to 3rd May (80 vs 30). These included ‘Return of the Good Life’ (Daily Telegraph, 15th April) and ‘11 Days to Save Britain’ (Daily Express, 26th April). Over the same period over fifty front page leads attacked Labour or Ed Miliband, as against twenty-one that attacked the Conservatives. These included ‘Labour’s £1,000 Tax on Families’ (The Times, 24th April), ‘Red Ed’s Zero Hours Hypocrisy’ (The Daily Mail, 2nd April), and ‘Cam’s Bribes Will Cost Every Family £1,439 A Year’ (The Mirror, 16th April).

If the newspaper front pages have been opinionated this election, the leader columns have been even more so. There were 335 anti-Labour leader articles between 5th January and 3rd May, as compared to 207 anti-Conservative. Over the same period there were 217 pro-Conservative leader articles as compared to 87 pro-Labour. The SNP hold the record on one-sided leaders however. Of 72 references in leader columns, 71 were negative. It should be noted that this does not include the Scottish press, including the Scottish Sun.

To set this in historical context, we compared partisanship in the 2015 UK election with partisanship in the 1992 UK election (using the microfiche available at the British Library Newsroom). 1992’s press coverage is often held up as being a benchmark of partisan campaign coverage. Leading the charge then was the Sun. Many will still remember its front page on election day: ‘If Kinnock wins today will the last person to leave Britain please turn out the lights’.

Yet it turns out that, based on leader columns, the 2015 Sun has been even more partisan than the Sun in 1992. In 1992 79% of all leader columns published in the Sun in the official campaign period up to the Sunday before polling day contained an anti-Labour comment. Leaders included ‘Danger of Labour’s Open Door’ (4th March, 1992), ‘Perils of Trusting Socialism’ (9th March, 1992) and ‘Labour’s Lesson in Lying’ (27th March, 1992). In 2015 this rose to 95%, with titles such as ‘Clueless Ed’ (31st March, 2015), ‘Lefty Stitch-Up’ (20th April, 2015), and Shameful Mili (25th April, 2015).
Partisanship was not restricted to the right-leaning sections of the press, however. The most partisan paper (by the measure of party support and criticism in leader columns) was the *Mirror*, which published the most leaders critical of any one party - 109 anti-Conservative leaders - and, jointly with the *Telegraph*, the largest number of leaders supporting a given party, 55 pro-Labour leaders.

Throughout the 2015 UK election campaign we have tried to remain conscious that most people are not nearly as anorak-ish about politics as we are. Newspapers and broadcasters know this too and so continue to devote considerable space to sport, celebrity, fashion and other news. To place the campaign in perspective, we contrasted the coverage of significant campaign events with celebrity or entertainment news coverage over the same period. We called this ‘Campaign vs Kardashian’, in deference to the publicity juggernaut of the Kardashian family. Almost every week, news references to Kim Kardashian outweighed references to the biggest campaign events. Often by some distance (see Figure 2).

---

*Figure 1: Sun partisanship in leader columns, 39-day run-up to 1992 and 2015 elections. Sample ends 4 days before respective polling days.*

Partisanship in Sun Leader columns, 1992 and 2015 Election Campaigns

- **1992**
  - Anti-Labour: 79%
  - Anti-Lab Leader: 44%
  - Pro-Conservative: 33%

- **2015**
  - Anti-Labour: 95%
  - Anti-Lab Leader: 77%
  - Pro-Conservative: 46%
Each Election Unspun analysis – of the issues covered, the party leaders mentioned, the partisanship of the leader columns - stands on its own, but at the same time sits within a broader academic context. As well as monitoring the dynamics of campaign coverage we are looking at the large question of who has been setting the issue agenda for the UK 2015 election.

Over forty years ago Maxwell McCombs and Donald Shaw staged their famous analysis of Chapel Hill voters’ policy priorities in the 1968 US Presidential election, which concluded that the news media plays a critical role in setting the voters’ agenda. Although they may not be successful in telling voters what to think, McCombs and Shaw found that news media had much greater success telling voters what to think about.

A lot has changed since 1968, not least the massive expansion of news sources and the emergence of social media. This makes the question, ‘who sets the agenda?’ even more complex and difficult to answer. To what extent is the agenda set by the political parties themselves, with their carefully constructed election grids and increasingly choreographed campaign events? Is the agenda still led by the mainstream news media, online or offline? Has social media – particularly Twitter – taken over? Or are all three following the direction set by public opinion itself?

Digital content analysis can help to answer this question, especially when correlated with analysis of public opinion, of party political election programmes, and social media. This is why, throughout the campaign, we have been working with Ipsos-MORI to ensure that our analysis maps to the opinion polling they are doing in the lead up to the election. At the same time we have been capturing set piece party events and statements, and – in the six weeks prior to polling day – analyzing candidates and political influencers on twitter.

It is too early to make conclusions based on the agenda-setting analysis, but certain findings are already clear. Immigration has been covered far less during
this campaign than might have been expected given its relative importance to the public. In February to March 2015 47% of people said that immigration would be very important when deciding which Party to vote for.\textsuperscript{2} Of the issues the public referred to, immigration was second only to health/the NHS, with the economy third. Yet, in terms of news coverage, immigration has come seventh after the economy, health, education, defence, crime & justice, and Europe.

On the eve of the election, Wednesday 6\textsuperscript{th} May, the \textit{Guardian}, the \textit{Financial Times} and the \textit{Sun’s} in-front-of-the-paywall election site, #sunnation all published stories based on analysis done using Steno. Fast turn-around digital content analysis is still nascent, but Election Unspun has shown us that it is both possible and powerful. There is no reason why, having been used for the UK election, it could not be used in other elections both within the UK and internationally. Equally, it could be extended to provide more in-depth analyses of coverage of policy areas like the EU or foreign policy. Our next hurdle is figuring out how to make a non-commercial, grant funded research tool sustainable and widely available to researchers, academics and journalists to drive understanding of how politics, policies and society are represented in digital media. Suggestions on a postcard welcome.

\textsuperscript{2} Source: Ipsos MORI; nationally representative sample of 4,574 British adults 16+ (Feb-March 2015), from ‘Immigration: How Views Change’, April 2015