

## 5

# International hyperlinks in online news media

Josh Cowls and Jonathan Bright

### Introduction: international news coverage on- and off-line

The disjuncture between ‘the world outside and the picture in our heads’ remains as relevant today as it was when Walter Lippman described it in 1922. At that time, a dramatically more efficient means of mass communication – radio – was making the spreading of news much faster and easier compared with what had come before; news of the outbreak of the First World War spread in hours and days, rather than weeks. Yet, Lippman argued, this quantitative shift in speed of transmission was not matched by a qualitative shift in the nature of mass communication. At the outbreak of war there was still an interval in which the picture of Europe held in people’s heads did not correspond to the new reality of Europe at war (Lippmann, 1922).

Lippman’s notion has been refined in the decades since, in parallel with the emergence of still more efficient communication platforms. As research has continued, however, attention has also started to shift from the *chronological* interval between the occurrence of an event and the transmission of news of it to the *contextual* interval between the myriad potentially newsworthy events occurring every day, and the far smaller number which are actually reported to the public at large. Media networks play a powerful ‘gatekeeping’ role by deciding which stories are given valuable airtime and column inches (Shoemaker, 1991); in turn, these decisions are shown to have a substantial effect on the formation of public opinion (McCombs and

Shaw, 1972, 1993) in a variety of areas, including their perceptions of international affairs.

The most dramatic contemporary shift in patterns of news consumption has been the emergence of online media as important locations for where people receive their news. Traditional media organizations have, by and large, persisted in the internet era: major print, radio and television outlets from the pre-internet age have been largely successful in transferring their success online. And this transfer has brought readers with them: a recent survey found that 68% of those who read newspapers do so online at least some of the time (OXIS, 2013: 45).

One important yet understudied consequence of this transfer is the embedding of hyperlinks in news content, which have the potential to direct readers to other sources of which they were not necessarily aware. Whilst most news organizations naturally want readers to remain on their websites, they also frequently link out to other websites, often to provide factual support or background to a story, 'placing news events in a thematic frame' (Coddington, 2012). The choices these news organizations make about which sites to link out to are, potentially, of huge significance, as the volume of web readership they receive means that they can boost the profile of another website considerably if they choose to link to it. For example, the BBC News website, the focus of this chapter, was viewed by just under 100 million people a month in 2014<sup>1</sup>: a link outwards from the BBC website therefore has considerable value in terms of directing traffic. Such choices are particularly important in terms of the debate about Lippman's notion of the world outside, as they can lead individuals straight to websites about the countries referred to in a given article. In other words, they provide the potential to enhance what Norris and Inglehart (2009) called 'cosmopolitan communications'.

This chapter seeks to explain patterns in news website outlinking practices, placing a particular focus on the country of origin of the website being linked to. It is structured in the following way. In the next section, we develop elements of a theory of outlinks, which necessarily relates to international news coverage patterns in general. We then describe the data used in our study: a large collection of news articles collected from a web archive of the BBC News website. We then test two analytical models which attempt to explain outlinking practices. Our results are discussed in the context of an internet which helps to globalize patterns of information consumption and news reading.

## Theorizing international outlinking patterns

As Coddington has highlighted, the major reason for news websites to link outwards to other websites, especially in different countries, is as a way of providing context or support to a story they are publishing themselves. Hence any theory of international outlinking also has to take into account explanations for patterns in international coverage.

Many studies have set out to investigate international news coverage, seeking to explain what factors drive coverage of different stories in different countries around the world. Chang et al. (1987) draw the useful distinction between ‘context-oriented’ factors – that is, factors relating to the national and international context within which particular events occur – and ‘event-oriented’ factors, which relate to the nature of the news events themselves. We will review each of these factors in turn here.

Within the remit of *contextual* factors, previous studies have investigated the effect of an array of variables – economic, demographic, geographic and political – on the distribution of coverage. Multiple studies show that economic factors are especially significant in predicting coverage. A large-scale analysis of news coverage on multiple media platforms in 38 countries (Wu, 2000) found that the volume of trade between two countries was (alongside the presence or absence of news agency bureaus) the most significant variable driving coverage. This finding was replicated in a later study which also encompassed news websites (Wu, 2007). A country’s economic development can also be an influential factor: Golan’s (2008) analysis of US television found that, as well as trade, national GDP can predict coverage of African countries, while Kim and Barnett’s (1996) analysis of international newspaper trade data found economic development to be the most significant of a host of contextual factors.

Other studies have investigated the role of demographic and geographic factors. When population is used as a variable it is usually found to have some predictive power (Dupree, 1971; Ishii, 1996; Charles et al., 1979) though this typically accounts for less influence than other factors. The distance between two countries has also been shown to partially drive coverage; a 1987 literature review of research in this area suggested ‘a relatively stable pattern of foreign news coverage [which] above all is characterized by “regionalism”, i.e. a preference for news from nearby countries’ (Wilke, 1987: 150). Analysis of the *New York Times*’ coverage of foreign disasters similarly found that distance from the USA was the only pertinent contextual factor (van Belle, 2000), and

the regionalism effect was also detected in Oceania (Nnaemeka and Richstad, 1980). This finding was moderated in Chang and Lee (2009), which found that geographic proximity was significant only for television coverage, not newspapers.

Finally, political factors have sometimes been found to play a role. The 'relevance' of a country to the USA was found to be significant in Chang et al. (1987) (albeit this was measured in a somewhat limited fashion as a dichotomous variable). More systematic analysis of six African newspapers in 1981 found that former colonial ties still exert some influence over coverage, although this is related more to coverage of fellow former colonies in Africa than ex-colonial powers in Western Europe (Skurnik, 1981). Meyer (1989) also found 'neimperial' effects in the flow of news, in relation to former French and British colonies in Africa and the sphere of US influence in Latin America.

The volume of previous research that has tested for the effects of *event-oriented* variables is slight compared to contextual factors. Nonetheless, some studies have found these types of variables to be significant. Chang et al. (1987) built on the concept of event 'deviance', developed in Shoemaker et al. (1986), as a factor that could explain news coverage. Of the seven contextual and event-oriented variables they tested for in a content analysis of American newspapers and television broadcasts, they found that both the normative deviance of an event in relation to the norms of the USA, and the potential for social change in the country in which the event took place, were among the most statistically significant factors explaining coverage. Thus events which *would* threaten domestic norms were they to occur at home, and events which *did* pose the possibility of social change within the foreign country, were both more likely to be covered. However, a later study which sought to replicate this earlier work with data from 1994 found that only the latter factor was significant in newspaper and television coverage ten years on (Chang and Lee, 2009). ('Loss of lives or property', a variable added to the 1994 data analysis, was found to be statistically significant for newspaper but not television coverage.)

Other studies have assessed the effect of other event-oriented factors. Van Belle (2000) found that the number of people killed in natural disasters is a statistically significant factor in the volume of coverage of that event. Golan and Wanta (2003) found that, in election coverage, elections were significantly more likely to be covered in regions where conflict was taking place. Golan (2008: 53) showed that the majority of American coverage of African stories 'focused on negative and highly deviant issues such as conflict and disasters both natural and human

caused', although overall rates of death among a country's population was not significant in relative levels of coverage between countries.

In this study, we test for the effects of many of the contextual and event-oriented factors outlined here. However, we also expect that these factors do not explain everything there is to know about online hyper-linking. Rather, we also expect these linking practices to obey certain logics of their own, within the overall structures conditioning international news reporting. Several aspects are worth considering here.

First and most obvious is the language of the website being linked to, with news websites likely to favour external sites that have the same language as they do. While obvious and mechanical, this hypothesis nevertheless has significant implications, as it means news readers are much more likely to learn more about countries with which they share a language through this mechanism. The second factor is the number of websites available relevant to the country of interest: countries with a larger digital presence are more likely to attract web links. This is again significant as larger and more developed countries inevitably have more of a web presence.

Third, a variety of other more subtle factors about the perceived trustworthiness of the content being linked to may come into play. This may relate to background knowledge the journalists themselves have about the country in question, or perceptions generated by reading websites related to any given country. Finally, it is important to note that we restrict our study here to the coverage of only one news organization, BBC News. The outlinking decisions taken by BBC journalists are undoubtedly also shaped to some degree by the characteristics of the organization: as an established, esteemed, publicly funded broadcaster, the priorities of and pressures on BBC reportage are likely to diverge from, say, up-and-coming and/or commercially funded news outlets. One strength of our single-organization approach is that these factors are controlled for across different national domains, but one trade-off is that we are not able to generalize fully to a wider array of broadcasting organizations.

## Data, methods and descriptive statistics

In this chapter, we test these propositions by focusing on the case of BBC News Online. Before describing the dataset collected in more detail, it is worth reflecting a little on this organization. The migration of traditional news media organizations on to the internet has typically been

uneven and often inchoate. This is due in part to confusion – on the part of governments as well as news organizations themselves – over the increasingly sophisticated affordances of the web, in relation to existing broadcasting technology. Traditional broadcasting organizations are licensed by governments, a policy which has its roots in the original ‘scarcity’ of broadcast frequencies (Moe, 2003). In this context, the web’s increasing support for audio and video playback led to tortured definitions of what constitutes ‘broadcasting’ on the internet, as in the case of Australia’s state broadcaster ABC (Martin, 2005).

The experience of bringing the BBC online was also somewhat uneven. The Conservative government’s original aim in the mid-1990s was for the BBC’s web presence to be commercial (Born, 2003). The last minute decision of the then-BBC director general John Birt to pull out of a commercial deal in 1996 was described by a BBC executive as ‘the most important thing [Birt] ever did’ (Connor, 2007). In 1999, this shift towards a public service provision was solidified with the BBC’s submission to the licence fee review panel; significantly, the first core element of the online provision was ‘the provision of news and information’ (Graf, 2004: 69). In practice, too, the technical development of the BBC’s online public service offering was driven largely by real-world news events: the 1996 budget, the 1997 and 2001 general elections, the terrorist attacks of September 2001 in New York and July 2005 in London and the Indian Ocean tsunami of 2004 all yielded new capacities and approaches for the BBC website (Thorsen, 2010).

To date, the BBC has continued to innovate and iterate its online services, now firmly under the rubric of public service delivery. The BBC’s digital services were grouped under the Future Media division in 2011 after a restructure, and innovation efforts continue in the BBC News Labs project. As part of measures aimed at cutting the online budget by 25% by 2013, many subsections of the BBC’s website were taken down, yet ‘high quality news’ remained the top of the list of the corporation’s revised online strategy in 2011 (Huggers, 2011). The BBC’s continued investment and innovation has been vindicated by its consistent popularity among UK web users: at the time of writing, it was the seventh most visited site in the UK, and the only British organization represented among the top ten most visited sites in the UK.<sup>2</sup> As a large and prominent media organization, the BBC has navigated initial confusion over the status of public broadcasters online – as well as recent budget cutbacks – to sustain a popular, resourceful web presence over the course of 20 years, with the reporting of domestic and international news as its flagship function.

The size and prominence of the BBC makes it an excellent case study with which to test our hypotheses. However, it should also be noted that this case does come with certain compromises. First, as the BBC's newsgathering activity must meet with stringent editorial standards, its hyperlinking should as well – suggesting that material which the BBC links to should not be objectionable. (Although it is noteworthy that under the list of 'Related Internet Links' common to BBC news stories in our period of investigation, the phrase 'The BBC is not responsible for the content of external internet sites' appears as a disclaimer, suggesting that these standards are perhaps not as complete as for the content actually published by the BBC.) These standards will naturally differ in different organizations. Second, the BBC operates an automatic external link generation system which contributes some of the external links found on its web pages, especially those relating to foreign news organizations.<sup>3</sup> Again, this system is rather unique to the BBC. Both of these factors decrease the generalizability of our findings.

In order to test our hypotheses, particularly those relating to sporadic and infrequent events across multiple countries, a dataset which covers as wide a time period as possible is required. For this reason, we chose to collect our data from the Internet Archive (IA), an organization which has been capturing and archiving web pages since 1996 (Kahle, 1997). The IA made available a large set of data on web pages specifically emerging from the .uk country level domain, which constitutes the 'JISC UK Web Domain Dataset'.<sup>4</sup> From this dataset, a set of hyperlinks was extracted during the course of a separate project (see Hale et al., 2014), together with the text to which the hyperlink was attached. These data were then filtered out to include only links emerging from the BBC itself. The web archive dataset is considerable, containing data from almost 17.5 million BBC news pages. It has excellent coverage for the period 2002–2010, when the BBC was visited and archived on average 354 days per year, and reasonable coverage for the period 1999–2001, when on average 205 days per year were captured (not much was archived before 1999). It is difficult to estimate, however, the absolute coverage of the hyperlink dataset, as we do not know to what extent archival visits to the BBC were complete (i.e. the IA may have saved some of the pages but not all of them). However, we have no reason to suspect that the IA's visits were biased to including coverage of one country more than another.

These data are used to create the two major variables used in the dataset. First, we count the number of links made from the BBC website to other country specific 'top-level domains' (TLDs) (across the entire time period of the archive).<sup>5</sup> Such links frequently appear to provide

extra background and context to ongoing news stories. For example, the BBC often links to the government page of a particular country if it is reporting on a news story from that country; or it might link to the website of a particular organization, if the story is about an organization. A top-level domain, in a general sense, is the last part of a hyperlink which indicates the top level of the website in question. For example, the ‘.fr’ in [www.lemonde.fr](http://www.lemonde.fr) indicates that the website has a French top-level domain. In our analysis, a TLD is taken to include all second-level content (such as [www.edu.au](http://www.edu.au), where [.au](http://www.edu.au) indicates the TLD for Australia and [.edu](http://www.edu.au) indicates Australia’s academic SLD). We focus our analysis solely on ‘country code top-level domains’ (ccTLDs) – that is, TLDs which are reserved for countries and other recognized territories. As such, for practical reasons we exclude generic TLDs such as [.com](http://www.com), which are typically country-neutral (although other research has suggested ways to incorporate the [.com](http://www.com) domain into studies of the international hyperlink network: cf Barnett et al., 2011). Moreover, we restrict our analysis to ccTLDs which can be unequivocally linked to one country, removing ccTLDs which have come to be used for non-country specific purposes. For example, the ‘.tv’ domain is partially owned by the island nation of Tuvalu, but since the government’s leasing of the TLD in 1999, it has frequently been used for websites which aim to broadcast television and video content.<sup>6</sup> In total that left us with 222 ccTLDs which had at least one outlink from the BBC website in our dataset.

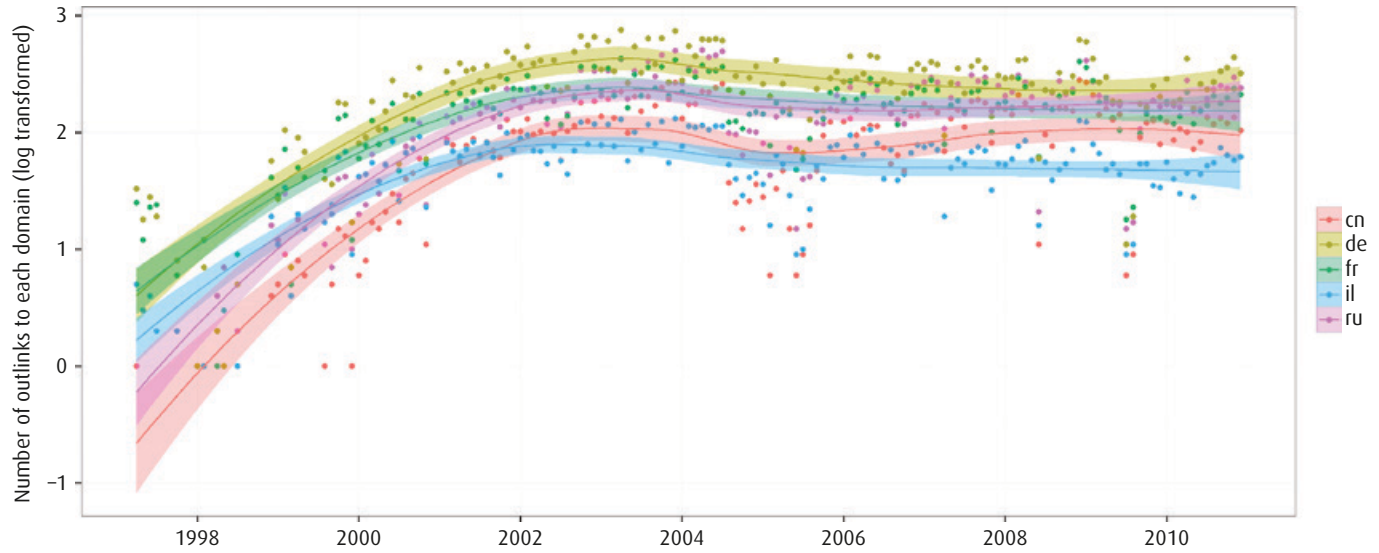
Second, we counted the number of times each country was mentioned in the text of links to news articles found on the BBC News pages. News articles themselves were identified on the basis of a previously developed schema used in other research (Bright and Nicholls, 2014; Bright, 2015). The text of the link is most frequently the title of the news article, and hence can be used as a means of identifying what the article is about. Based on a list of country names, and common abbreviations for those countries, we checked each title to see how many times a country had been mentioned. This provides an indication of the level of coverage that country receives.

Of course, this method is not a perfect proxy: in particular, it is likely to understate the total amount of coverage each country receives, because not every article about a country will have the name of the country within it (for example, it might refer instead to the capital of the country, or that country’s prime minister). However, we do not expect this understatement to be uneven across different countries, hence as a measure of the relative distribution of coverage between countries we still expect this to be valid.

One issue to highlight with the dataset is that the IA's method of archiving pages is quite ad-hoc, based on a web 'spider' which crawls over the internet following hyperlinks from one page to another. There is therefore no guarantee that the same page will be archived consistently over time. Furthermore, as we highlighted above, the volume of pages captured is also not constant over time. However, we do not expect these sampling issues to affect one country disproportionately, since our analysis relies solely on the presence of BBC News pages in the IA over time. We need not assume that the BBC News website was captured in its entirety throughout the period, because we see no reason for the IA to have 'over-captured' BBC News pages covering a particular country compared to any other. Therefore, we still believe that these measures can be used as effective proxies. This is a contention supported by Figure 5.1 below, which shows how absolute counts of outlinks to selected country domains fluctuate over time, but the relative order of countries remains largely unaffected (in Figure 5.1, each point represents the total number of outlinks observed during a one month period in the archive).

We created the following independent variables for this study.<sup>7</sup> Beginning with contextual factors: first, to investigate whether the sheer size of a country influences coverage, we collected data on the total population for each country from World Bank statistics.<sup>8</sup> Second, to assess whether trade flows with the UK affect news coverage, we collected the total combined trade between the UK and other countries from official UK trade data.<sup>9</sup> Third, as a gauge to measure the importance of a country's overall wealth we collected GDP per capita (in current US dollars).<sup>10</sup> Fourth, we collected data on the geographic distance, in kilometres, from London to every other capital city.<sup>11</sup> Fifth, we created the dichotomous variable of whether a country was a member of the Commonwealth of Nations – an intergovernmental organization of member states, most of which were formerly territories of the British Empire – to assess whether the historical legacy of colonization affects modern news coverage or outlinking.

As we highlight above, alongside these general, context-oriented indicators, we also expect factors that are related more specifically to the 'newsworthiness' of a country to have an effect on both the amount of coverage it receives and the amount of links it receives. It is worth noting here that – in contrast to earlier research in this area, which typically gauged the impact of events on a qualitative case-by-case basis – since the dataset we have is so large, we operationalize the 'eventful-ness' of countries by using summary statistics. As such, we introduce three variables which measure a country's newsworthiness, as a proxy for event-oriented



**Figure 5.1** Evolution of outlinks to top five country domains over time

hypotheses. First, we look at a country’s ‘disaster risk’, based on data from the World Risk Report.<sup>12</sup> This report measures both the potential for natural disasters such as earthquakes to occur in a country, and the extent to which the country in question is prepared to deal with such an event (see World Risk Report: p. 46). Higher scores on this scale mean a country is at greater risk of disasters. Second, we look at the extent to which a country is ‘peaceful’, using data from the Global Peace Index.<sup>13</sup> This index measures internal safety and security within a society (taking into account factors such as violent demonstrations), the extent to which it is involved in domestic and international conflicts, and the extent of its militarization (Global Peace Index: p. 2). Higher scores on this scale mean a country is less peaceful. Finally, we measure the crime rate of a country, looking in particular at murder rate statistics provided by the UNODC.<sup>14</sup> This statistic, it should be noted, is also taken into account in the peace index; but it is more specific, focusing solely on internal crime rather than also taking war into account. Higher scores on this scale mean more homicides per 100,000 people within a given country.

Initial descriptive statistics are provided in Table 5.1. As will be apparent from comparing the mean and median values, many of the variables in the dataset (including our key independent variables) are highly skewed. This means both news coverage and outlinking patterns are distributed unevenly, with a small amount of countries receiving a large proportion of the attention. It also suggests that transformations of these variables is appropriate to improve the fit of our statistical models; these transformations are discussed below in the analytical section.

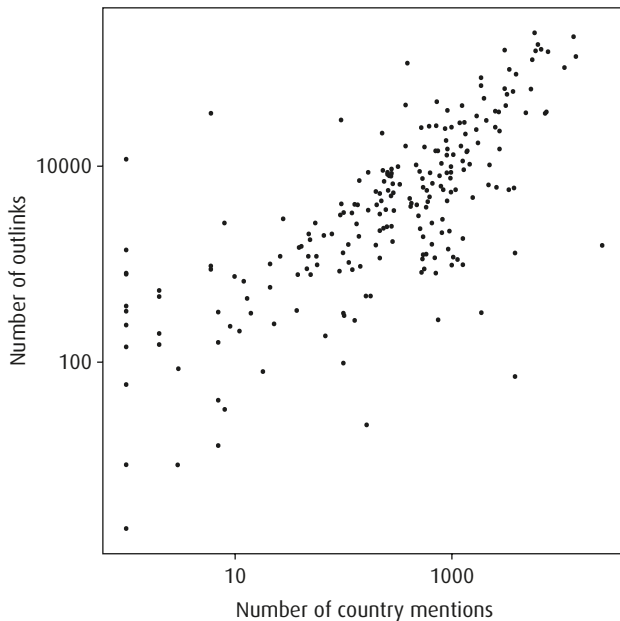
**Table 5.1** Descriptive statistics

	Mean	Median
<i>Context-oriented variables</i>		
Outlinks (whole period)	17,147	4,096
Mentions (whole period)	1,213	418
Population (2005)	31,410,000	5,904,000
GDP per capita (\$) (2005)	11,375	3,172
Distance from London	6,600	6,600
Trade with UK (2005)	2,053,000,000	78,550,000
<i>Event-oriented variables</i>		
Disaster risk (2015)	0.07	0.07
Peace Index (2015)	2.02	1.98
Homicide rate (per 100,000, 2015)	8.80	4.8

## Analysis

The main aim of this chapter is to explain outlinking patterns from the BBC to different country top-level domains. In this section, we will explore this question using a series of regression models. As highlighted above, the nature of BBC outlinks means that we expect country coverage to have a significant impact on outlinks themselves, as outlinks are prevalent on news articles, and are themed to the article in question. In fact, a major aim of the chapter is to explain how outlinks vary when taking these differential levels of coverage into account.

Figure 5.2 is a scatter plot of the relationship between country mentions and observed outlinks for the entire time period. It provides strong support for the idea that coverage is a major underlying driver of outlinks, as we might expect. A strong positive correlation between mentions and outlinks can be observed when points are plotted on a log 10 scale ( $R = 0.72$ ). In other words, as the amount of times a country is mentioned by the BBC goes up, so does the number of outlinks to domains linked to that country. With this in mind, an initial analytical task is to explain news mentions themselves. This is something we tackle



**Figure 5.2** Correlation between outlinks and mentions of a country in BBC News Online

**Table 5.2** Linear regression model explaining amount of country news mentions on BBC online

Variable	Coefficient	Standard error
Population (log transformed)	0.63***	−0.11
Trade with UK (log transformed)	−0.09	−0.1
GDP per capita (log transformed)	0.32	−0.17
Distance from UK (log transformed)	0.08	−0.14
Homicide rate	−0.02*	−0.01
Peace Index	0.53*	−0.23
Disaster risk	−3.13	−2.59
Commonwealth ember	−0.34	−0.36
Internet penetration	0	−0.01
English as an official or primary language	0.89*	−0.34
adj. R-squared	0.43	
N	148	

in Table 5.2, with a linear regression model, which enables us to assess how multiple explanatory variables might relate to the BBC increasing or decreasing its coverage of a certain country. The dependent variable in this model is the log transformed mentions variable.<sup>15</sup>

Three main drivers of international news coverage can be observed. First, population size is strongly and positively associated with news coverage, with more populous countries receiving more mentions in the news; replicating some previous findings (Dupree, 1971; Charles et al., 1979; Ishii 1996). GDP is found to have a positive correlation, however the p-value is above the conventional cut-off for statistical significance, thus providing no real support for previous findings (Kim and Barnett, 1996; Golan, 2008). Volume of trade with the UK and distance from the UK were also not found to have any effect, despite strong findings in previous research.

In terms of event-driven factors, the Peace Index also shows a strong positive correlation, with less peaceful countries receiving more coverage. This supports earlier research suggesting greater coverage of less peaceful regions (Golan and Wanta, 2003). The homicide rate has a negative correlation which is also statistically significant, although we have no theory to explain why a higher level of homicides might lead to

less coverage. Finally, despite good theoretical cause to believe disasters should increase coverage, we find no evidence for this in our data. Overall, the relative scarcity of evidence for the effects of event-specific coverage here chimes with the fairly limited findings of significance in earlier research.

We also include three variables which we considered of importance in explaining outlinks: membership of the Commonwealth, internet penetration and use of English as a primary or official language. Using these factors in the mentions model is useful in order to provide a basis for comparison in the next model. Commonwealth membership and internet penetration are shown to have no impact on news coverage. However, use of English is shown to be significant and positively associated with increasing news coverage. Of final importance in this model is the adjusted  $R^2$  of 0.41. This indicates that the model explains a reasonable amount of the variance in coverage, but also that a significant portion of it goes unexplained.

We will now move on to look at factors driving outlinking itself. We fit the same model as in Table 5.2, but with news coverage now included as an independent variable. This model, which will allow us to assess factors which seem to drive outlinking whilst controlling for news mentions, is presented in Table 5.3. The dependent variable, outlinks, is again log transformed.<sup>16</sup> It is worth highlighting that the adjusted  $R^2$  of this model is 0.69, meaning that it explains a considerable amount of the total variance in observed outlinks.

As we would expect given Figure 5.2, news mentions are strongly correlated with news outlinks. Population and GDP also continue to be important factors. Interestingly, however, the Peace Index is now negatively correlated with outlinks. Internet penetration in the country of destination also appears as an important factor, as does the use of English as either a primary or official language of the country. Being a member of the Commonwealth has, however, no effect. This analysis underlines our hypothesis that outlinks on news articles are understood to serve a different function for the user than the coverage itself. Where the country being covered has a robust internet infrastructure, with content likely to be in English, journalists and editors may see greater utility in linking to content local to those domains. Where this is not the case – and where non- and inter-governmental organization websites might provide more authoritative sources of additional information, in the case of conflict-ridden countries – we may have detected less of an urge to link to ‘native’ content.

**Table 5.3** Linear regression model explaining amount of country outlinks on BBC online

Variable	Coefficient	Standard error
Population (log transformed)	0.52***	(0.10)
Trade with UK (log transformed)	0.02	(0.08)
GDP per capita (log transformed)	0.11	(0.14)
Distance from UK (log transformed)	0.04	(0.11)
News mentions (log transformed)	0.16*	(0.07)
Homicide rate	0.01	(0.01)
Peace Index	−0.44*	(0.19)
Disaster risk	−3.34	(2.05)
Commonwealth member	−0.41	(0.29)
Internet penetration	0.02***	(0.01)
English as an official or primary language	0.66*	(0.27)
adj. R-squared	0.69	
N	148	

## Discussion

This chapter has sought to explain variation in the rate that the news media link outwards to websites in different countries, even when taking into account variations in the news coverage of those countries. Based on a long-term temporal dataset of articles from the BBC News Online website, it has shown that more populous countries and less peaceful countries receive greater levels of news coverage. However, it has also shown that, with the level of news coverage taken into account, less peaceful countries receive fewer outlinks. Countries with greater internet penetration and which use English as a primary or official language are also more likely to receive links.

In closing, we offer a few remarks on the significance of these findings. As we argued in the introduction, the disjuncture between ‘the world outside and the picture in our heads’ is a key topic in research on international news coverage, and it is also a topic which the emergence of the internet as a crucial venue for news consumption has the potential to revolutionize. Even if major news outlets continue to dominate in the online environment, selective outlinking from these sites has the

potential to inform audiences like never before, as they can be taken to media portals directly within the country they are reading about. Given that the international hyperlink network was perfectly interconnected by 2009, the potential of the hyperlinked web to take users 'closer' to the geographic source of a given news story is clearly possible in principle (Park et al., 2011). However, as we have shown (and as the aforementioned study would suggest), this outlinking is not evenly distributed, or simply a function of how much a given country is in the news. Rather, certain countries do appear to receive more links than others. This means that the impact of these links will be skewed, and the 'picture in our heads' only partially corrected.

Further, it is important to situate this study in the far longer-running stream of analysis of international news coverage. For example, the fact that we have observed here significant differences between the factors explaining coverage, on one hand, and those which explain outlinks, on the other, highlights the value of analysing phenomena specific to the internet as a location of news coverage. Journalists and editors have – perhaps unconsciously – adopted forms of practice relating to the affordances of web-specific phenomena such as hyperlinks. This sort of finding is nothing new to the study of communications in general, which has long demonstrated the importance of the form and affordances of a given medium on how it is used. But it is nonetheless important to note the extension of this phenomenon here, in a study of news coverage on the internet.

It is appropriate to conclude, however, by highlighting again the limitations of this study. Chief among these is the focus only on the BBC, which is an organization with a specific set of values and practices. Furthermore, the identification of 'country relevant' websites is also potentially problematic: not every country uses its top-level domain equally, and many nationally specific websites nevertheless use the generic .com TLD. The study could also be improved by studying temporal variation in outlinks, and rather than just the level observed over the entire period. Finally, we have done little to differentiate between different types of websites within these TLDs, something which may obscure important secondary patterns (for example, how many governments around the world are linked to from the BBC). Further work could usefully pursue these questions, and thus offer us a fuller explanation of the determinants of international hyperlinking practices in online news media.