
This is the author’s final accepted version.

There may be differences between this version and the published version. You are advised to consult the publisher’s version if you wish to cite from it.

[http://eprints.gla.ac.uk/107109/](http://eprints.gla.ac.uk/107109/)

Deposited on: 08 June 2015
Are all policy decisions equal? Explaining the variation in media coverage of the UK Budget

Ana Ines Langer, University of Glasgow (Ana.Langer@glasgow.ac.uk)
Iñaki Sagarzazu, University of Glasgow (Inaki.Sagarzazu@glasgow.ac.uk)

Abstract

A country’s Budget is one of the most important public policy instruments, as it establishes the government’s policy priorities and has the potential to determine winners and losers. The Budget, however, is a mixture of different components and these get varying degrees of attention in the media. Drawing on sociology of news research, this paper seeks to explain this heterogeneous coverage of a Budget’s policy decisions. To do so, it uses a unique dataset of over 5,000 articles of press coverage of six UK Budgets (2008-2012). These articles are coded for the presence/absence of each of the Budget’s policy decision, via automated content analysis. Based on a multivariate negative binomial model we find that the salience of a policy decision in the coverage is determined by its cost; whether it is negative (i.e. tax hikes and spending cuts) or positive; the income group that is most affected by it; and the level of attention given to it by the government.

1 The order of the authors’ names is in alphabetical order. Both authors have contributed equally to all work. The authors would like to thank Yigal Harkavy for his invaluable research assistance. We are also grateful to Sarah Birch, Heike Klüver, Peter Van Aelst, Shaun Bevan, and the anonymous reviewers for their helpful comments and suggestions which helped improve the manuscript.
Introduction

The Budget presentation and approval process is a key element of the policy process. Budgets allow governments to establish and publicize their key priorities (Walgrave, Varone and Dumont 2006), even if they only have the ability to make marginal changes (Wildavsky 1964). Similarly, they can help the opposition in scoring political points, especially if the opposition is on the side of public opinion (Burstein 2003; Stimsom 2004). Furthermore, the public policy impact of the Budget cannot be denied, as well as its direct and indirect influence on voters (Whitten and Williams 2011). In the end, these political and policy dynamics make the act of presenting the Budget a highly mediatized event, even if—as we shall see—not all policy decisions receive the same degree of media coverage, or in some cases coverage at all.

So far, there has been considerable research into the Budget, including on how it is explained by paradigms of policy change and different causal mechanisms (Esser, Reinemann and Fan 2000; Waisbord 2011), the impact it has for different governing areas (John 2006; King 1998), the budgetary process and its outcomes (Molander 2001; Wagschal and Wenzelburger 2008) and on the strategic nature of the Budget itself (Alt and Chrystal 1981). However, given the high impact that budgets can have for a country and its politics, it is surprising that little has been devoted to the media coverage. This is the more important because, as research has consistently shown, the politics of attention are crucial (Baumgartner and Jones 2005), and the media—for whom, as for other actors, attention is scarce—often plays a key role in these dynamics. Moreover, previous research has found that the media is highly selective about which aspects of the policy process it reports (Van Aelst et al. n/d).

In the case of the Budget, the issues that it covers are evidently already on the policy agenda. Increases and decreases in budgetary spending, for instance, are tied to issue preferences of legislators (Mortensen 2009), even if these are constrained (Epp, Lovett and
Baumgartner 2014). Which of these policy initiatives, however, does the public get to know about? We argue that the issues that remain on the agenda, and those that disappear into the background, will partly depend on the degree of coverage. Moreover, this will be tied to how attractive different policy decisions are for the media. As such, it is important to understand: which items of the Budget do the media pay attention to? And how can this be explained?

In order to answer these questions, this paper draws on research on news values as well as commercial imperatives and media-source relations to explore the ‘selection bias’ of the media reporting of the Budget in the UK. Specifically, the paper will identify the key determinants that explain the difference in the amount of coverage allocated to different types of policy decisions presented in the Budget: from increases in direct taxes such as VAT to the reduction of corporation tax to changes in benefit indexation mechanisms. Using automated content analysis, this study analyses over 5,000 articles of press coverage of six Budgets, from 2008 to 2012 (with two in 2010); three during a Labour government and three during the Conservative-Liberal-Democrat coalition government. We then use media, policy and government communication variables in a multiple regression model to explain the variation in the coverage of the different components of the Budget.

This paper will be organized as follows: the first section will discuss the role of the media in the policy process and the factors that shape news coverage. This will be followed by the presentation of our hypotheses. The third section explains the data collected, the process of automated text coding, and our dependent and independent variables. This will be followed by section four, in which we present our results, and section five which summarizes our conclusions and suggests several avenues for further research.
The Media and Public Policy

What the media does (not) cover can matter a great deal in the policy process. For a long time, research emphasis has been on the media’s influence over public opinion; for instance, we know that the media affects perceptions of the economy (Duch, Sagarzazu and Meddaugh 2011; Goidel and Langley 1995; Harrington 1989; Holbrook and Garand 1996). Moreover, research on agenda setting and priming indicates not only that the media’s priorities affect those of the public, but also that parties and political representatives are more likely to be assessed on issues that are highlighted in the media (McCombs and Shaw 1972; Price and Tewksbury 1997; Scheufele and Tewksbury 2007; Weaver 1996; Weaver, McCombs and Shaw 2004). But recently it has also become clear that the media can more directly affect the policy process. Changes in salience, tone, and framing of coverage have been associated with policy change (Baumgartner and Jones 2005; Bonafont and Baumgartner 2013; John 2006; John et al. 2013; Rose and Baumgartner 2013; Soroka 2002; Walgrave, Nuytemans and Soroka 2008). In addition, research has shown not only that government and opposition often respond to the news agenda (Green-Pedersen and Stubager 2010; Vliegenthart and Walgrave 2010; Walgrave, Nuytemans and Soroka 2008) but also that news attention structures issue competition, influencing which issues are the subject of dialogue between opposition and government (Thesen 2011, 27). This is not just about symbolic reactions, however. Issues that the media pick up are more likely to create positive feedback into the policy cycle and hence be subject to more policy attention (Wolfe, Jones and Baumgartner 2013). Moreover, media attention to issues can help actors to politicize issues and, crucially, lack of media attention can help actors in trying to de-politicize them (Baumgartner and Jones 2005; Boydstun 2013).

So, clearly the media are important, but what do we know about what they do and why? The question of what shapes media coverage, and specifically why certain events and
issues receive more news attention than others, has been a key preoccupation of media and political communications scholars for decades (e.g. Gans 2004; Shoemaker and Reese 1996; Tuchman 1978; White 1950). However, despite the accumulated evidence of the importance of news values and other factors in affecting what (does not) get media attention, the determinants that explain the coverage of domestic policy—let alone fiscal policy and more specifically the Budget—have been generally understudied. This is surprising as the Budget is undoubtedly of key importance as the main instrument for setting—as well as giving publicity to—the country’s fiscal, and to an extent economic, policy for the next three to five years.

In public policy, there have been several studies that demonstrate the significance of ideas and discourse in the success or failure of policy measures (Cox 2001; Green-Pedersen 2002; Schmidt 2002), and even some that specifically emphasize the importance of political communication (Hood 2010; Wenzelburger 2011). But these mostly analyze government strategy and discourse, and in some cases public opinion data, where the media hardly figures. In studies of policy agendas and communication agenda setting (i.e. influence from the media’s to the public’s priorities), the media’s agenda is treated as independent variable, and hence generally left unexplored. On the other hand, studies of media and political agenda setting interactions explore whether media coverage is influenced by the political agenda (and vice versa), but generally do not try aim to uncover the influence of other factors in the media coverage (for overviews see Boydstun 2013; Van Aelst et al. 2014; Walgrave and Van Aelst 2006). Similarly, although there are, as mentioned above, studies (e.g. Baumgartner and Jones 2005; John 2006; John et al. 2013) that do incorporate dynamics of media attention as explanatory variables for policy change (including budgetary re-allocation), they do not try to explain which factors led to the coverage in the first place.
Within the media and sociological literatures, however, there is a strong body of literature which provides in-depth understanding of the different actors and dynamics that influence what appears in the news and how it is presented (e.g. Gandy 1982; Gans 2004; Golding and Elliott 1979; Schlesinger 1978; Tuchman 1978). These studies tend to analyze the news agenda in general and as such lack detail and specificity about the coverage of the policy process. They do, nonetheless, provide crucial insights into the influence of different factors—including news values—on the work of journalists, which will be discussed below. There have also been a number of studies that focus on particular events or issues, but they have predominantly been about international event news, from elections to terrorism to national disasters to human right violations (Caliendo, Gibney and Payne 1999; Golan and Wanta 2003; Jones, Van Aelst and Vliegenthart 2011; Koopmans and Vliegenthart 2011; Shoemaker, Chang and Brendlinger 1986; Wu 2007). On the other hand, the studies on domestic issues news have tended to concentrate heavily on protest (Oliver and Maney 2000; Oliver and Myers 1999) and crime (Buckler and Travis 2005; Pritchard and Hughes 1997) or on the relative visibility of specific political actors (e.g. Schoenbach, De Ridder and Lauf 2001; Squire 1988; Tresch 2009). The few that focus on domestic policy do so by comparing news about the economy vis-à-vis objective economic indicators in the US (Fogarty 2005; Harrington 1989; Soroka 2002), explaining aggregate patterns of attention across major policy issues (Boydstun 2013), analyzing the degree of media attention to two policy decisions about welfare cuts (Lindbom 2010), the evolution over time of the coverage of the UK Community Charge, known as the ‘poll tax’ (Deacon and Golding 1994), and the determinants of coverage of policy via analysis of legislative proposals (Van Aelst et al. n/d). While these studies have identified a range of key factors that shape coverage, including the importance of news values and sources and their communication strategies (most importantly
the government), they have left explanations of the coverage of fiscal policy—and most specifically the Budget—relatively unexplored.

In combination, however, what this research does tell us is that media coverage is generally fairly predictable and that it can be explained by a range of key determinants. Although initially media research placed a lot of emphasis on the influence of the individual preferences of journalists—and especially editors—in deciding what news was, over time it became clear that there are a range of influential factors, most of which are structural (both to do with media institutions and political and economic variables) rather than individual. Among them, one of the key considerations is shared professional understandings of newsworthiness and the news values that underpin them. ‘News values refer to common views, particularly among journalists, about what is believed to be intrinsically relevant and interesting for the public (McQuail, 1993)’ (Van Aelst et al. n/d, 4). To be clear, news values are not the only influences on news selection: economic, political and organizational factors, as well as ideology, also come into play (Golding and Elliott 1979; O’Neill and Harcup 2009; Shoemaker and Reese 2013). Moreover, news values are not devoid of subjectivity nor are they straightforward to operationalize. But, nonetheless, it is clear from the literature that news values do matter and that there are certain characteristics of events and issues, or the countries where they take place, that increase/decrease their potential for news coverage. In the next section, drawing on studies of news values as well as the broader literature on sociology of news, we discuss how known determinants of media coverage can be applied to the UK Budget and present our hypotheses.

Determinants of media coverage and hypotheses

Many studies, especially in international news, have tried to operationalize and test the concept of news values (e.g. Adams 1986; Chang, Shoemaker and Brendlinger 1987; Weber 2010). In fact there are several lists of criteria and many are rather extensive, which
can be problematic, especially for empirical testing (Harcup and O’Neill 2001; O’Neill and Harcup 2009). Here we focus on news values that have been repeatedly identified across key studies (Golding and Elliott 1979; Harcup and O’Neill 2001; O’Neill and Harcup 2009; Shoemaker and Reese 2013) and that for both theoretical relevance and matters of operationalization are most relevant to this paper, given its focus on fiscal policy coverage.

The first of these news criteria is magnitude or size. The idea is that the greater the magnitude of a ‘happening’, the more the media is likely to cover it. As studies on fiscal policy are lacking, there is no specific evidence of how this criteria applies in such a context. But this news value has indeed been found to be significant in a range of topics. Research has repeatedly shown that how much coverage an event abroad gets (e.g. natural disasters, elections, human right abuses) will be in part determined by the magnitude of what took place such as intensity of the earthquake or number protesters as well as—among other factors—the size of the country that the news refer to, measured by e.g. population, surface area, GDP or share of international trade (for a meta-analysis see Wu 2000). As applied to our study, we operationalize magnitude as a matter of cost. So we hypothesize that:

**H1 Magnitude:** the larger the cost of the policy decision for the Treasury, the more the media will cover it.

The second news value is timeliness, which refers to events that are ‘close to us in time’ (Shoemaker and Reese 2013, 107). Timeliness is considered to increase newsworthiness because ‘[w]e have limited attention and want to know what's happening

---

2 More specifically, for instance Oliver and Myers (1999) found in their study of media and protest that the size of the event (measured by the number of demonstrators) influenced the degree of coverage, with larger ones receiving greater attention. Similarly, Koopmans and Vliegenthart (2011, 645) found that both the number of deaths and the earthquake’s magnitude on the Richter scale had an effect on the degree of media attention.
now. Timely events are also more likely to require action’ (Shoemaker and Reese 2013, 171). Interestingly, although this is seemingly a rather self-evident criteria for influencing news coverage, it is not one that has much been empirically tested, perhaps precisely because of its apparent obviousness. The Budget is in this regard, however, rather atypical because although the event of unveiling it is clearly timely, policy decisions are scheduled to come into effect at a number of distinct time points (ranging from now to five years). So, whether or not this has an impact needs to be tested. Based on the news value of timeliness, we hypothesize that:

**H2 Timeliness**: the more immediate the impact of the policy decision, the more the media will cover it.

The third news value we incorporated in our model is *negativity*, which has long been considered a key criteria (Shoemaker and Reese 2013, 53-4) based on the idea that ‘bad news sells’ (Golding and Elliott 1979). It has also been particularly associated with economic news, where studies have found that ‘negative trends regularly get more coverage than positive trends’ (Soroka 2012, 521) and that the media overemphasizes the negative when the economy is doing poorly but not the positive when it is doing well (Fogarty 2005, 165). In relation to the Budget, negative news can be defined as policy decisions that reduce people’s disposable income and the benefits and services they receive from the state. So we hypothesise that:

**H3 Negativity**: policy decisions that announce cuts in spending and tax increases will be covered more than increases in spending and tax cuts.

Related to size or magnitude, theories of news values also suggest that journalists are more likely to write about ‘happenings’ that are of interest to many people (Strömbäck, Karlsson and Hopmann 2012). The implication is thus that the greater the number of people affected by it (a policy decision in this case), the more coverage it should receive. It might not be so simple, however. The public here can be defined by its size but also in general
terms (as in relevant to the ‘public interest’) or as specific readership profiles. It is well known that commercial imperatives shape journalists’ work, especially (assumed) audience interests and competences (Deacon and Golding 1994). Moreover, as newspapers are profit-seeking organizations, every potential reader might not have the same ‘value’, depending on the paper’s target readership as well as their marketability for advertisers (Hamilton 2003). Furthermore, as suggested by Gans (2004, 151) this judgement might also be affected by journalists’ background: more coverage might be given to decisions that affect the middle class, to which most journalists belong. Indeed, studies about crime coverage across countries have repeatedly found that crimes affecting the white middle-classes get more attention (for an overview see Lawrence and Mueller 2003). Similarly, Waisbord (2011, 153) found that issues affecting urban school systems used by the middle-classes are considered more newsworthy than those in deprived rural areas.

The issue here is that news values interact with commercial imperatives and other organizational factors, often producing a gap between what journalists normatively might think should be the most important event properties in deciding what is news (especially relevance to the ‘public interest’) and what they are in practice (Strömbäck, Karlsson and Hopmann 2012, 722). Following this, our hypothesis states that:

**H4 Middle class**: policy decisions that affect the middle class will be covered more than those covering other social classes or than administrative policy decisions that do not affect any particular group.

Following on from this, we also tested a second variable that combines news values with commercial imperatives: the impact on the degree of coverage of whether policy decisions affected mostly businesses or individual citizens. As explained above, the expectation is that the number of people directly affected and target readership of the paper will affect how much coverage an event (in our case the Budget measure) receives. This is
anticipated to be the case even though some policy decisions primarily affecting businesses might have stronger effects on the distribution of wealth and on the economy at large, and hence have a greater impact on the public interest. Applied to our study, given that the readership of all the papers bar the Financial Times is non-specialised, the hypothesis states that:

**H5 readership:** policy decisions affecting primarily businesses will receive less coverage than those primarily affecting individuals

The final news value we incorporated in our model is the involvement of elites, specifically here government officials, which has long been considered a crucial determinant of media coverage. There are two interrelated aspects to this. On the one hand, it is about elites as subjects: events associated with ‘important people’, regarded as such either because of their role in society and/or because of their celebrity status, have been shown to be considered more newsworthy (O’Neill and Harcup 2009; Shoemaker and Reese 2013; Strömbäck, Karlsson and Hopmann 2012). Specifically in relation to the policy process, ‘journalists report about what political elites are doing or planning to do, often ignoring other policy actors who are deemed to be less newsworthy than “officials”’ (Baumgartner and Bonafont 2015, 1). On the other hand, the involvement of elites is also important in their roles as sources. Due to organizational factors and professional norms, source considerations are crucial (Gans 2004) and, in this regard, the literature has consistently pointed at the influence of official sources, most especially government actors (for an overview see Bennett and Livingston 2003). Journalists often take the lead from government officials about what is most important. This is a result of both the fact that patterns of media attention are shaped by government attention (Bonafont and Baumgartner 2013; Boydston 2013) and the more direct influence of information subsidies from a source that has both high media capital and media-source affinity (Davis 2000). Specifically in relation to our study, it is the case that the
Budget announcement involves elites, specifically the Chancellor, in all its measures; hence in this regard involvement of elites can be considered a constant. But if a particular measure is highlighted by government communication, both due to news values and source considerations, it is expected to get more media attention. Thus, we hypothesize that:

**H6 official sources**: the more a policy decision is emphasised in government communications (Budget executive report and press notices), the more coverage the media will give to it.

**Identifying Newspaper Coverage of the Budget**

In the previous sections we presented our expectations regarding the visibility of the different policy decisions in media coverage of the Budget. In this section we describe the data collected to test our theoretical claims. First, we explain the newspaper data. Secondly, we describe how we used the policy decisions to generate a coding dictionary. Third, we explain how we coded the articles and performed reliability tests. Finally, we discuss our dependent and independent variables.

**Budget articles**

In order to identify how much the press covered each of the Budget measures, we downloaded articles from six British newspapers. We chose the print media not only due to ease of access but also because, given strict impartiality rules in broadcasting, the press is a more varied forum of national debate (Kuhn 2007). Moreover, the press has been found to

---

3 The Chancellor of the Exchequer is the top cabinet position in charge of financial matters. Sh/e has overall responsibility for the work of the Treasury, including the Budget.

4 As discussed in the literature review, the relationship has also been found to apply in the other direction: media coverage affecting the degree of attention paid to issues by both government and opposition. This is not something, however, that can be tested with the data analyzed here.
have intra-media agenda setting influence, including on that of TV news shows (Vliegenthart and Walgrave 2008, 862).

All major daily national newspapers were included, hence covering both tabloids and broadsheets and different partisan and ideological positions. It thus enabled us to compare newspaper types, which are assumed, and indeed have been found by some studies, to differ in their criteria of newsworthiness (Deacon and Golding 1994; Harcup and O’Neill 2001). Specifically we included the broadsheets: The Guardian, The Daily Telegraph, The Times, and the Financial Times; and the tabloids: The Daily Mail, The Daily Mirror, and The Sun. In terms of ownership, two of these newspapers (The Sun and The Times) are owned by Rupert Murdoch’s News International – organization considered to be of center-right ideological orientation (Esser, Reinemann and Fan 2000)-; none of the others are owned by the same conglomerate.

The articles were identified using Newsbank with the key word ‘budget’\(^5\). The samples covered one full working week (Sundays excluded) starting the day after the Budget was announced. The short timeframe allows us to focus on Budget-specific stories. We focused on the London editions because this is the only one that all newspapers have. After manually cleaning for duplicates, we identified a total of 5,501 unique newspaper articles for the six years. As shown in table 1, there are—on average—about 900 articles per Budget; tabloids have a significantly lower number of articles than broadsheets, which is expected due to their weaker focus on public affairs and smaller pagination.

---

**Coding policy decisions**

The UK Budget is generally structured around two reports: (1) an economic and fiscal strategy report and (2) a financial statement and Budget report. While the former focuses on

---

\(^5\) To obtain the Daily Mail articles we used the same search but through the Lexis Nexis database.
how the measures and decisions announced in the Budget help advance the Government’s long-term goals, the latter ‘brings together in summary form all the measures and decisions that affect the Budget arithmetic’ either in the form of revenues or spending (Budget 2008, p. 109). 6 These ‘policy decisions’ are then the main feature of the UK Budget and the focus of our paper.

For the purpose of this analysis, we used the policy decisions listed as such in the Budget’s official documents. However, we only included decisions: i) made in the Budget under study and ii) with costed budgetary impact; hence we excluded policy plans without monetary quantification and decisions carried on from a previous Budget. Following these criteria, we identified 299 policy decisions over the six Budgets under study. 7

In order to code how many of the 5,000+ articles in our sample covered the policy decisions of the respective Budgets, we took advantage of dictionary coding. Dictionary coding is a particular type of quantitative text analysis. It links specific words to predetermined categories, be they policy positions (Laver and Garry 2000), issue areas (Pardos-Prado and Sagarzazu n/d), or more complex measures such as assuredness and tentativeness (Hart and Childers 2005). The main advantage of this tool is that it enables researchers to code a larger number of documents and do so with a stronger degree of reliability than hand-coding provides (Mikhaylov, Laver and Benoit 2012).

6 ‘These include tax measures, national insurance contributions (NICs) measures, measures that affect Annually Managed Expenditure (AME), and additions to Departmental Expenditure Limits (DEL)’ (Budget 2008, 109).

7 Policy decision #23 of the 2008 budget was removed due to its high overlap with a policy decision emanating from 2007. Similarly, policy decision #19 of the 2008 budget was removed due to the limited information provided for it in the budget together with the fact that, in abstract, it duplicated other PDs from the same budget (#17, 18 and 20).
To build our dictionary we took the tokens (both words and numbers) that were mentioned in the different policy decisions. These tokens were stemmed (words) and normalized (for different representations of the same numerical concept, i.e. 2%, 2pct). Once this was done our dictionary had 1,279 unique terms of which 1,146 were word stems and 133 were numbers. Based on these terms we classified the different articles. Because these computerized methods, in this case in particular a dictionary, can suffer of lack of validity (Grimmer and Stewart 2013) we tested its validity by comparing the results with a hand-coded random sample of 235 articles. These articles were manually coded for the presence/absence of references to each policy decision, and then we compared our two measures to check their validity. In our test we correctly coded 98.3% of our sample. Furthermore, we also estimated the correlation ($r = .61$, $p < .001$) of our dictionary compared to the hand coded measure; the largely cross validate each other. The online appendix describes in further detail the process we used to build our dictionary and classify the texts, as well as the validity test performed.

Figure 1 shows the total number of articles reporting at least one policy decision as a percentage of all the articles in the dataset. As can be seen, on average 30% of tabloid articles and about 40% of broadsheet articles in our sample (returned with the search string ‘budget’) referred to specific policy decisions.

<FIGURE 1>

As expected, coverage of policy decisions is not homogenous: while some receive zero mentions others are mentioned in 200 articles. On average an individual policy decision is mentioned in 21 articles. This of course also varies by Budgets; for instance, the 2008 Budget has the lowest average with 13 articles per policy decision, while the first coalition Budget has the most with 32. On average tax and transfer policy decisions are mentioned similarly (23 and 21 articles respectively), while policy decisions regarding other
administrative measures receive the least attention (6 articles). This, however, also varies: the 2008 Budget has the highest average mentions of transfers and the 2010 coalition Budget has the highest average mentions for taxes (See figure 3). This variation in attention makes this study of the causes for this heterogeneity of attention to the different policy decisions highly significant.

<FIGURE 3>

**Variables**

The dependent variable was the number of articles that mentioned each of the policy decisions (PDs). Data was then aggregated simultaneously by policy decision and newspaper, resulting in a pooled dataset of 2,000 observations of the dependent variable. We considered the possibility of also trying to establish prominence, based on the placement of the piece (e.g. front page) and number and location of the references to the PDs. However, given the limited information available in the databases about page numbers and the characteristics of our methodology, this would not have been achievable. This is not a major limitation, however, because other studies have found that using both types of measure produced very similar results (Boomgaarden et al. 2010; Sulitzeanu-Kenan 2007).

Each policy decision was coded for several independent variables, following our hypotheses. Firstly, we used estimated policy costings based on the governments' figures. Whilst both Labour and the coalition government used the National Accounts basis, consistency across all six Budgets required us to make two choices. First, static costings were chosen, rather than post-behavioural ones. Static costings do not generally take into account projected changes in the behaviour of taxpayers or social security claimants, in response to the policy decisions in question (e.g. changes in how much such individuals work, as a results of changes in minimum wage or welfare payments, etc.). Although limited in some
regards, it is more reliable than trying to predict behavioural changes. Second, indexed costings were used, rather than non-indexed ones. This means that, in calculating a policy decision’s net Exchequer impact, it was assumed that the relevant pre-policy decision fixed-amount tax or benefit regime (e.g. the personal allowance, benefit amounts, etc.) would have kept pace with inflation, even without the policy decision in question.⁸

These policy costings were used for three variables. For the variable magnitude (hypothesis 1), we added all years and transformed the sum into absolute numbers. Following standard practice, as these figures were highly skewed, they were log-transformed so that they had a normal distribution in the regression analysis. For timeliness (hypothesis 2), we coded for whether the financial impact started in the fiscal year of the Budget (0), in the following year (1), or beyond (2). For hypothesis 3, we used the sign of the costing (i.e. whether it entailed savings or spending for the Exchequer) combined with type (described below) to code for negative (-1), neutral (0), or positive (1) news.

In order to identify which income groups were most affected (hypothesis 4) by each policy decision (i.e. according to income groups, and businesses vs. individuals), we coded policy decisions based on whether they affected the poor (specifically the two lower income deciles as defined by the Office of National Statistics)⁹, the middle class (median-income households), the rich (the highest income decile), or none of the three in particular (for instance, administrative policy decisions). Similarly, to identify which policy decisions target

---


businesses (*hypothesis 5*) it was determined whether the policy decision directly affects mostly businesses (including sole traders).\(^\text{10}\) The variable was coded according to the actual, rather than formal, economic incidence of each policy decision; in the case of taxes, for instance, this is to say, anyone who is out of pocket as a result of the imposition of the tax, rather than only those who have the actual legal liability for paying the tax (Kay and King 1990, 6).

For *government attention* (*hypothesis 6*), we coded, using the same dictionary and criteria for inclusion, the presence/absence of each policy decision in the Treasury’s press notices and in the Chancellor’s speech in the House of Commons. If present in one of them, we coded as 1, if in both as 2, and 0 if not present. Given the different strategies for specific years, we proceed to normalize this variable for the year by estimating how differently policy decisions were mentioned versus the mean mention for the year.

In addition, we created a set of control variables. First, we differentiated policy decisions as to whether they announced *taxes, transfers, or other* measures (essentially administrative and miscellaneous); we set *taxes* as our control category. Second, we generated a dummy variable to identify a *Labour* government (1) versus a *Conservative* government (0) and another one to identify ideology, i.e. whether the source was *left leaning* (1) for the *Daily Mirror* and *The Guardian* or not (0) for the other four titles. Finally, in the full model we controlled for differences in newspaper types with a dichotomous variable identifying whether a source was a *tabloid* (1) or a *broadsheet* (0), and with dummies for the different budgets. Table 2 shows the summary statistics for all variables.

\[\text{<TABLE 2}>\]

Having defined all variables, we can proceed to our analysis. As we discussed earlier, the dependent variable is a count of the number of articles in a newspaper mentioning a

\(^{\text{10}}\) A research assistant, with expertise in Tax law, performed the coding.
policy decision. As such, given the non-negative integer nature of our variable, a count model needs to be used (Cameron and Trivedi 2013). Within this family of models, we specifically chose to use a Negative Binomial regression given that—since the sample variance exceeds the sample mean—this model is more efficient compared to the Poisson regression (Long 1997).

Results

For the multiple regressions we ran three separate models: one for the whole sample and two models for the sample split by type of newspaper\(^{11}\). As can be seen in Table 3, the full model shows—as expected—that tabloids have given the Budget less coverage overall than broadsheets. Regardless of this disparity, the main point to note is that, although there are some differences between the models that will be discussed below, the coverage in both types of newspapers is largely driven by the same explanatory variables. In other words, both types of newspapers appear to be affected by the same factors in their news selection. Reinforcing this indication of commonality across newspaper types, neither the party in government nor ideological leaning were statistically significant. Given our methodology, we obviously cannot speak to the tone of the coverage; but in terms of salience this result indicates that, despite the marked partisanship of British newspapers, they all focus their attention on the same policy decisions, regardless of editorial ideological leaning and which government is in charge (ceteris paribus).

\(^{11}\) Given patterns of ownership (see p. 13), we ran a separate model to test whether our results were different for newspapers owned by News International versus the other newspapers, and we compared this to our full model. The results (available on request) were consistent with the models presented here, except that the coefficient of the variable for policy decisions targeting businesses (hypothesis 5) loses significance. Nonetheless, the sign is the same (negative) and it is similar to the results of our only tabloids model where the coefficient is also non-significant.
In regard to our hypotheses, most of them have been confirmed, although there are also some interesting exceptions. Regarding hypothesis 1, which refers to magnitude, it is clear that the monetary value of the policy greatly matters. The relationship is significant and in the expected direction: the greater the cost (whether saving or spending for the Exchequer), the greater the number of articles that discussed it. For example, an additional £3,792.95 (million) in the cost of a policy decision (one standard deviation) is associated with an increase in coverage of 5 articles.

The second hypothesis referred to timeliness, the assumption being that policy decisions that have more immediate effect would receive more coverage than those with deferred implementation. Interestingly, this variable is significant for all three models but in the opposite direction to what we expected: the later a policy decision comes into effect (i.e. starting in year 2 or 3), the more coverage it receives. We do not have a clear explanation for this yet but the fact that policy decisions resulting in more structural change, which do not necessarily correlate with cost, have deferred implementation is—we suspect—at least part of the answer.

The third hypothesis referred to negativity, which has long been considered a key news value and indeed the data provides support for hypothesis 3. This variable, coded to include tax increases and spending cuts, was significant for all three models. This is to say that, even when controlling for other variables including economic costing, negative policy measures defined as above received more coverage than positive ones (i.e. tax cuts and spending increases) in both tabloids and broadsheets.

<TABLE 3>

The next hypotheses combined news values with commercial imperatives. We hypothesize that not all income groups matter the same for newspapers. The data shows that there is indeed a stronger focus on what matters for the middle classes, especially by
broadsheets. On average, policy decisions affecting them received more coverage. However, interestingly, in tabloids the policy decisions affecting mostly the poor received almost as much attention as those for the middle classes. This is significant while controlling for type of policy decision, so it is not a result of tabloids’ ‘obsession’ with welfare ‘hand-outs’, but an indication of their target readership, and what is assumed to be their (self) interest, and hence of the importance of market considerations in shaping the coverage. We also tested to see whether policy decisions focusing on businesses, and hence directly affecting a smaller target audience, would receive less coverage. This was indeed confirmed by the data: corporate policy decisions are, ceteris paribus, considered less newsworthy.

The final hypothesis about the role of official sources was also confirmed: policy decisions highlighted in government communications via the Chancellor’s speech and press notices received greater coverage in the media. This is of course while controlling for all other relevant factors, so it is not simply that these policy decisions are ‘objectively’ (e.g. cost) more important; instead, it suggests that the government’s communication influences which policy decisions the media give more salience to, because of the public interest value allocated to what the government says, as well as organizational and routine factors that encourage reliance on official sources,

An interesting point to make regarding the attention the government gives to the different policy decisions is that while it matters, it is also conditional on the government’s selectiveness about which PDs to highlight. When we model government attention simply as the number of places where a policy decision is mentioned: none (0), both speech and press release (2), or in one of the two (1), we find not only a similar significance for the attention variable but a corrective negative impact on coverage during the Labour years. This is

12 The coefficients for rich, poor, and middle classes are statistically significantly different from each other. Results will be made available upon request.
because Labour's press notices were substantially longer and touched on a much larger number of policy decisions, suggesting that while government information subsidies work on shaping coverage, they do so only as long as they are focused.\textsuperscript{13}

**Discussion and conclusions**

This paper analyzed the amount of press coverage for each of the policy decisions in six UK Budgets (2008-1012), providing not only data but also an original and useful framework for further research on a subject that has been clearly understudied. The analysis offers three main insights: firstly, it is clear that there is a huge disparity in the degree of coverage that different policy decisions receive; some are simply of no interest to the press, receiving no attention at all. Secondly, the limited differences between the models and the lack of significance of two of the control variables shows that there is a remarkable commonality about the factors that explain the attention given by the press to the Budget’s policy decisions. This indicates that news values to a large extent transcend differences in type of newspaper, their ideological orientation and party/ies in power. Thirdly, the analysis confirmed most of our hypotheses about the factors that shape the salience given to each of the policy measures. This clearly suggests that the coverage of the Budget is highly predictable. In terms of news values directly associated with the policy decisions themselves, both magnitude (based on economic costing) and negativity (defined as tax hikes and spending cuts) were statistically significant and in the expected direction (i.e. more coverage). The fact that cost matters is perhaps not very surprising: the more money at stake, the more media interest it is likely to get. However, as Lindbom (2010) and Adams (1986) have argued, there are also other factors at play as well. Findings on negativity are consistent with previous literature, which have found negative developments to be more newsworthy (e.g. Golan and Wanta 2003), especially in economic news (Fogarty 2005; Soroka 2012).

\textsuperscript{13} Results will be made available upon request.
Immediacy, on the other hand, challenged our expectations: PDs with delayed implementation received more coverage.

The relative salience of different measures in the coverage, however, was not explained just by properties of the policy decisions and their ‘inherent’ news value. In line with the literature, government influence also proved to be an important factor in explaining the attention to different measures. And so were commercial imperatives, and it is here that we see differences across newspaper types, because of differences in their target readerships. Moreover, it is these findings that raise normative issues about what the press covers and what it omits, and hence what citizens do (not) get to know about, most especially in relation to the low attention to what affects the rich and businesses, as well the poor. There was a general emphasis on policy decisions most directly affecting individuals, rather than businesses, which makes sense from a commercial point of view. From a normative point of view, however, it raises some questions about how well the general public is being informed about what are in many cases highly significant decisions, given their potential impact on the economy as a whole and for the redistribution of wealth (or rather lack thereof). A similar point can be made about the findings regarding target group. Focusing, as broadsheets did, on measures that affect the middle class is a sound commercial strategy, given the size of the group and their disposable income—and hence marketability for advertisers—compared to the poor. However, 'quality' newspapers overlooking the impact of the Budget on the poor raises normative questions about how well they serve the public interest. It also provides an interesting angle of support to arguments about the importance of universality in sustaining support for welfare provision.

These ‘selection biases’ are key because of the potential impact they might have on both the political and the policy process. In terms of the former, this heterogeneity in coverage—and the fact it can be affected by the government—can have significant
repercussions for accountability. This is the case because availability of information—or at the very least visibility—is crucial for attribution of responsibility, and hence credit claiming and blame avoidance (Powell and Whitten 1993). Moreover, in terms of the policy process, we know that policies that are ‘invisible’ to citizens (or that citizens are unaware of) are on the one hand tempting if the aim is to legislate by stealth (Pierson 1994)—such as in the case of cuts—and on the other, in the case of expenditure, gives elected representatives less incentive to implement them (Alt and Chrystral 1981). Altogether, this highlights the importance of studying the media coverage of policy and its determinants.

The results also suggest several fruitful avenues for further research. The government is clearly a key actor but it would be interesting to explore the role of interest groups, opposition parties and think tanks, such as the Institute of Fiscal Studies (IFS), in shaping media attention, especially if the timeframe was extended beyond the period immediately after the Budget announcements. Another avenue to expand upon would be to include different types of media, especially television, to uncover similarities and differences in the main determinants of attention to different policy decisions. Moreover, the study could be replicated in a comparative setting hence enabling us to test the validity of the conclusions in different policy, political and media systems.

References


Table 1. Number of articles by newspaper and budget

<table>
<thead>
<tr>
<th>Newspaper</th>
<th>2008</th>
<th>2009</th>
<th>March 2010</th>
<th>June 2010</th>
<th>2011</th>
<th>2012</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily Mirror</td>
<td>35</td>
<td>30</td>
<td>16</td>
<td>38</td>
<td>30</td>
<td>48</td>
<td>197</td>
</tr>
<tr>
<td>Daily Mail</td>
<td>136</td>
<td>63</td>
<td>125</td>
<td>101</td>
<td>115</td>
<td>100</td>
<td>640</td>
</tr>
<tr>
<td>The Sun</td>
<td>59</td>
<td>50</td>
<td>42</td>
<td>43</td>
<td>44</td>
<td>43</td>
<td>281</td>
</tr>
<tr>
<td>Telegraph</td>
<td>141</td>
<td>128</td>
<td>181</td>
<td>227</td>
<td>218</td>
<td>198</td>
<td>1,093</td>
</tr>
<tr>
<td>Financial Times</td>
<td>151</td>
<td>201</td>
<td>158</td>
<td>183</td>
<td>174</td>
<td>155</td>
<td>1,022</td>
</tr>
<tr>
<td>Guardian</td>
<td>174</td>
<td>265</td>
<td>225</td>
<td>196</td>
<td>262</td>
<td>244</td>
<td>1,366</td>
</tr>
<tr>
<td>The Times</td>
<td>129</td>
<td>160</td>
<td>160</td>
<td>175</td>
<td>123</td>
<td>155</td>
<td>902</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>825</strong></td>
<td><strong>897</strong></td>
<td><strong>907</strong></td>
<td><strong>963</strong></td>
<td><strong>966</strong></td>
<td><strong>943</strong></td>
<td><strong>5,501</strong></td>
</tr>
</tbody>
</table>
Table 2.- Summary statistics of dependent and independent variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Articles coded</td>
<td>2,000</td>
<td>3.483</td>
<td>6.293</td>
<td>0</td>
<td>66</td>
</tr>
<tr>
<td>Cost (log)</td>
<td>2,000</td>
<td>4.666</td>
<td>2.794</td>
<td>-4.605</td>
<td>10.903</td>
</tr>
<tr>
<td>Timeliness</td>
<td>2,000</td>
<td>0.645</td>
<td>0.759</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Negativity</td>
<td>2,000</td>
<td>-0.004</td>
<td>0.975</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>Social class affected: None</td>
<td>2,000</td>
<td>0.437</td>
<td>0.496</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Social class affected: Poor</td>
<td>2,000</td>
<td>0.118</td>
<td>0.323</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Social class affected: Middle</td>
<td>2,000</td>
<td>0.359</td>
<td>0.479</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Social class affected: Rich</td>
<td>2,000</td>
<td>0.084</td>
<td>0.278</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Business affected</td>
<td>2,000</td>
<td>0.617</td>
<td>0.486</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Government attention</td>
<td>2,000</td>
<td>0.000</td>
<td>0.517</td>
<td>-1.864</td>
<td>0.825</td>
</tr>
<tr>
<td>Type: Tax</td>
<td>2,000</td>
<td>0.717</td>
<td>0.451</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Type: Transfer</td>
<td>2,000</td>
<td>0.209</td>
<td>0.407</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Type: Other</td>
<td>2,000</td>
<td>0.073</td>
<td>0.261</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Left leaning</td>
<td>2,000</td>
<td>0.299</td>
<td>0.457</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Labour government</td>
<td>2,000</td>
<td>0.516</td>
<td>0.499</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Tabloid</td>
<td>2,000</td>
<td>0.402</td>
<td>0.490</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Budget 1</td>
<td>2,000</td>
<td>0.175</td>
<td>0.380</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Budget 2</td>
<td>2,000</td>
<td>0.206</td>
<td>0.404</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Budget 3</td>
<td>2,000</td>
<td>0.135</td>
<td>0.341</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Budget 4</td>
<td>2,000</td>
<td>0.144</td>
<td>0.351</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Budget 5</td>
<td>2,000</td>
<td>0.199</td>
<td>0.399</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Budget 6</td>
<td>2,000</td>
<td>0.140</td>
<td>0.347</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 3.- Results from Negative Binomial Regression on Number of Articles Coded

<table>
<thead>
<tr>
<th></th>
<th>ALL</th>
<th>Tabloid</th>
<th>Broadsheet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost (log)</td>
<td>0.118***</td>
<td>0.129***</td>
<td>0.116***</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.024)</td>
<td>(0.015)</td>
</tr>
<tr>
<td>Timeliness</td>
<td>0.222***</td>
<td>0.164*</td>
<td>0.243***</td>
</tr>
<tr>
<td></td>
<td>(0.051)</td>
<td>(0.089)</td>
<td>(0.060)</td>
</tr>
<tr>
<td>Negativity</td>
<td>0.153***</td>
<td>0.114*</td>
<td>0.169***</td>
</tr>
<tr>
<td></td>
<td>(0.034)</td>
<td>(0.060)</td>
<td>(0.040)</td>
</tr>
<tr>
<td>Social class affected [base = none]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poor</td>
<td>0.543***</td>
<td>0.983***</td>
<td>0.377**</td>
</tr>
<tr>
<td></td>
<td>(0.137)</td>
<td>(0.249)</td>
<td>(0.160)</td>
</tr>
<tr>
<td>Middle</td>
<td>0.758***</td>
<td>1.158***</td>
<td>0.599***</td>
</tr>
<tr>
<td></td>
<td>(0.081)</td>
<td>(0.143)</td>
<td>(0.097)</td>
</tr>
<tr>
<td>Rich</td>
<td>0.415***</td>
<td>0.502**</td>
<td>0.408***</td>
</tr>
<tr>
<td></td>
<td>(0.123)</td>
<td>(0.221)</td>
<td>(0.146)</td>
</tr>
<tr>
<td>Business affected</td>
<td>-0.242***</td>
<td>-0.144</td>
<td>-0.296***</td>
</tr>
<tr>
<td></td>
<td>(0.086)</td>
<td>(0.146)</td>
<td>(0.103)</td>
</tr>
<tr>
<td>Government attention</td>
<td>1.313***</td>
<td>1.124***</td>
<td>1.387***</td>
</tr>
<tr>
<td></td>
<td>(0.073)</td>
<td>(0.133)</td>
<td>(0.086)</td>
</tr>
<tr>
<td>Control variables:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type [base category=Tax]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transfer</td>
<td>-0.371***</td>
<td>-0.201</td>
<td>-0.453***</td>
</tr>
<tr>
<td></td>
<td>(0.104)</td>
<td>(0.182)</td>
<td>(0.122)</td>
</tr>
<tr>
<td>Other</td>
<td>-1.114***</td>
<td>-1.600***</td>
<td>-1.022***</td>
</tr>
<tr>
<td></td>
<td>(0.142)</td>
<td>(0.345)</td>
<td>(0.161)</td>
</tr>
<tr>
<td>Left leaning</td>
<td>-0.186***</td>
<td>-0.819***</td>
<td>0.108</td>
</tr>
<tr>
<td></td>
<td>(0.066)</td>
<td>(0.121)</td>
<td>(0.081)</td>
</tr>
<tr>
<td>Labour government</td>
<td>-0.384***</td>
<td>-0.235</td>
<td>-0.444***</td>
</tr>
<tr>
<td></td>
<td>(0.115)</td>
<td>(0.192)</td>
<td>(0.138)</td>
</tr>
<tr>
<td>Tabloid</td>
<td>-1.415***</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.066)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Budget – 2009</td>
<td>0.435***</td>
<td>0.160</td>
<td>0.494***</td>
</tr>
<tr>
<td></td>
<td>(0.105)</td>
<td>(0.176)</td>
<td>(0.125)</td>
</tr>
<tr>
<td>Budget – 2010a</td>
<td>0.506***</td>
<td>-0.217</td>
<td>0.751***</td>
</tr>
<tr>
<td></td>
<td>(0.115)</td>
<td>(0.227)</td>
<td>(0.132)</td>
</tr>
<tr>
<td>Budget – 2010b</td>
<td>0.084</td>
<td>-0.229</td>
<td>0.266*</td>
</tr>
<tr>
<td></td>
<td>(0.122)</td>
<td>(0.222)</td>
<td>(0.144)</td>
</tr>
<tr>
<td>Budget – 2011</td>
<td>-0.269**</td>
<td>-0.522***</td>
<td>-0.166</td>
</tr>
<tr>
<td></td>
<td>(0.109)</td>
<td>(0.185)</td>
<td>(0.131)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.560***</td>
<td>-0.836***</td>
<td>0.523***</td>
</tr>
<tr>
<td></td>
<td>(0.156)</td>
<td>(0.269)</td>
<td>(0.185)</td>
</tr>
<tr>
<td>Parameter α</td>
<td>0.116**</td>
<td>-0.068</td>
<td>0.080</td>
</tr>
<tr>
<td></td>
<td>(0.053)</td>
<td>(0.125)</td>
<td>(0.060)</td>
</tr>
<tr>
<td>N</td>
<td>2000</td>
<td>804</td>
<td>1196</td>
</tr>
<tr>
<td>Log Likelihood</td>
<td>-3914.794</td>
<td>-1016.541</td>
<td>-2855.100</td>
</tr>
<tr>
<td>AIC</td>
<td>7867.588</td>
<td>2069.083</td>
<td>5746.200</td>
</tr>
</tbody>
</table>
Figure 1.- Percentage of articles coded with at least one policy decision, by budget and output type
Figure 2.- Mean number of articles covering policy decisions based on type
### Appendix 1.- Example of Policy Decisions and Dictionary terms

<table>
<thead>
<tr>
<th>Budget Year: 2008</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Policy Decision:</strong> 1 Charities: gift aid transitional rate</td>
<td></td>
</tr>
<tr>
<td><strong>Additional information:</strong></td>
<td>Gift Aid A.39 With effect from 6 April 2008, charities will be able to claim Gift Aid at a transitional rate, consistent with a basic rate of income tax of 22 per cent, for three years.</td>
</tr>
<tr>
<td><strong>Dictionary words:</strong></td>
<td>able, aid, april, basic, business, charities, claim, consistent, effective, gift, rating, taxes, transitional, 22%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Budget Year: 2011</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Policy Decision:</strong> 15 Stamp Duty Land Tax: bulk purchasing</td>
<td></td>
</tr>
<tr>
<td><strong>Additional Information:</strong></td>
<td>2.155 Stamp Duty Land Tax (SDLT): reform of the rules for bulk purchases – The Government will introduce changes to the SDLT rules for bulk purchases of residential properties. If the buyer chooses, the rate of SDLT on purchases of multiple residential properties will be determined by the mean value of the dwellings purchased (subject to a minimum rate of 1 per cent), rather than their aggregate value as is currently the case. (Finance Bill 2011)</td>
</tr>
<tr>
<td><strong>Dictionary words:</strong></td>
<td>aggregate, bulk, buyers, dutifully, dwellings, land, means, multiples, property, purchase, residential, stamped, taxes, 1%</td>
</tr>
</tbody>
</table>
Online appendix: Coding of Policy Decisions

Our dictionary was created by first taking the words from the Budget, specifically the tables listing the policy decisions as well as the text describing each of these in more detail. Second, following standard quantitative text analysis practices (e.g. Klüver 2009; Proksch and Slapin 2010) we excluded stop-words and stemmed all terms (Porter 1980) to make sure we covered a wide range of uses of the words. Because of the importance of numbers to the policy decisions, we made sure to include them in our dictionary. While it is not a common approach to include numbers in quantitative text analyses, we did so because they can provide information unique to the different policy decisions. As such we treated numbers as tokens just as if they were a word. This, however, posed a challenge as they do not have a stem. Nonetheless, upon inspection of the texts we noticed different representations for the same concept; for instance two percent was found as 0.02, 2pct, and 2%. So, for each numerical term we used all these variations to increase the accuracy of our dictionary.

Once this initial list of terms for each policy decision was made, we extended each to include synonyms of some of the key words. This is particularly important given the differences in languages between tabloids and broadsheets. In total our dictionary contained 1,279 unique terms of which 1,146 were word stems and 133 were numbers. It is important to note that our dictionary does not place exclusivity between a term and a policy decision, and as such one term can be included in several policy decisions (for more details see Appendix).

Coding the articles

Once our dictionary was defined, we proceeded to code each newspaper article based on the number of terms associated with each policy decision that was present. This coding resulted in three specific measures for each article: 1) the simple count of unique terms in the
article that were in the dictionary for each policy decision (repeated terms where counted as one); 2) the percentage of total words in the article that represented the terms from the policy decision; and 3) the average of the uniqueness of the terms from the policy decision present in the article. This latter measure is created from a uniqueness index of each term $t$ which indicates the extent $t$ is specific to a single policy decision $p$ compared to the other policy decisions in the same Budget (i.e. if a term is unique to one PD in a given year uniqueness=1 if it belongs to all PDs of the same year uniqueness=0).

After coding the articles using the dictionary, we established a threshold for our three indicators. As such, for a policy decision to be coded present in the article it needed to include: 1) at least six unique words from the dictionary category; this means that almost 50% of the unique dictionary words for a policy decision need to be present; 2) the total count of words of a policy decision present in the article needs to represent at least 2 per cent of the valid words of the article; 3) the average uniqueness score has to be higher than .2. Figure 1 shows the distribution of articles by policy decisions for the three variables and the cut-off points. As can be seen, this threshold allowed us to be highly inclusive while at the same time controlling the presence of false positives and false negatives.14

<FIGURE 1>

In order to validate our dictionary, we proceeded to do a manual coding of a random sample of at least 5% of the articles for each newspaper-Budget combination. However, given the small number of articles for the tabloids, we oversampled these sources. This resulted in a sample of 385 articles. These articles were manually coded for the presence/absence of references to each policy decision, and then we compared our two measures to check their validity. Out of a total of 20,378 observations, we correctly coded

---
^{14} We tested alternative thresholds; however, we found this was more accurate and we corroborated this against our hand-coded sample.
20,037, i.e. 98.3% of our sample.\textsuperscript{15} Despite the high number of correctly predicted observations, we also estimated the correlation of our dictionary compared to the hand coded measure. The estimates correlate well ($r = .61, p < .001$) and therefore largely cross validate each other. As such, we are confident that our dictionary is properly identifying the presence and absence of policy decisions.

\textsuperscript{15} We correctly coded the absence of a PD in 19,760 observations and the presence in 277. Our dictionary identified 172 observations which our manual coding didn’t and missed 169 observations.
Figure 1. Distribution of articles based on coding scores

a) Count

b) Percentage

c) Uniqueness