

Issue Salience and Issue Ownership Online and Offline

Comparing Twitter and Survey Data

Bengü Hosch-Dayican (b.dayican@utwente.nl), University of Twente

Kees Aarts (c.w.a.m.aarts@utwente.nl), University of Twente

Chintan Amrit (c.amrit@utwente.nl), University of Twente

Adrie Dassen (a.dassen@utwente.nl), University of Twente

*Paper to be presented at the Annual Meeting of The American Political Science Association (APSA),
Chicago, August 29th - September 1st, 2013*

– Preliminary draft, please do not cite without permission by the authors –

Abstract

Social media, in particular Twitter, become an increasingly attractive tool for political scientists for the study of political attitudes and behavior because of its potential for tracking public opinion with minimal costs. Yet extracting reliable, valid and precise measurements of politically relevant concepts from these data sources still forms a major challenge. The present research is part of the project “Transforming Social Media Contents to Political Data” (funded by the Netherlands Organization for Scientific Research NWO), which aims at addressing this challenge by developing new tools and methods for harvesting, coding and analyzing messages from social media. Using data collected from Twitter during the last three weeks of the Dutch parliamentary election campaign 2012, we demonstrate in this paper how two prominent concepts of research on voting behavior – issue salience and issue ownership – can be measured online, and how these measures compare to traditional survey data. The preliminary results obtained by this study show that measures of issue salience and issue ownership using data collected from Twitter are comparable with the survey data to a considerable extent. However, we have also found some discrepancies between the two data sets on key issues, which suggest that our online measurements of the two concepts need to be further developed for a more sound validation.

Introduction

Electoral politics in Western European democracies has in recent decades been marked by a growing disenchantment with political parties as agents of representative democracy, which is reflected in declining turnout and party attachment, increasing voter volatility and success of populist parties (Dalton & Weldon 2005). Traditionally, voting behavior is thought to be affected by structural factors such as party identification and ideology, which in turn determine short-term evaluations of voters (Campbell et al. 1960). The alleged decline in the explanatory power of the social-structural model in explaining voting behavior (Franklin et al. 1992; Dalton 2006) has provided room for short-term explanatory factors like candidate and issue evaluations. Issue ownership and issue saliency have become prominent concepts for the explanation of voting behavior (a.o. Budge and Farlie 1983; Petrocik 1996; Van der Brug 2004; Bélanger and Meguid 2008). These concepts suggest that voters prefer parties whom they perceive to be particularly competent to resolve the issues that concern them.

While the competence of parties to handle specific issues appears to be relatively stable, the saliency of issues can vary in the short run – even in the course of an election campaign. For instance, actual developments in core issues like economy or welfare state might alter the saliency of established issues, and issues that may emerge shortly before elections. Over a longer period of time, the perceptions of which parties “own” which issues may change (Stubager and Slothuus 2012). Hence, for understanding the impact of political issues on electoral behavior it is essential to observe the short-term changes in issue saliency.

Short-term changes in saliency can be observed in various ways. Pre-election polls provide the most common instrument; other means include focus group interviews and content analyses of web pages, television broadcasts and the printed press. In this paper, we focus on the content analysis of new social media. We show that issue saliency can be measured by studying new social media, in particular Twitter. The advantage of studying issue saliency on Twitter is that short-term fluctuations can be observed continuously. As citizens are increasingly involved in political discussions on the Internet and particularly in social media, it seems important to explore social media as a data source for analyzing short-term influences on voting behavior.

This paper is part of a larger project, in which we aim to provide technical tools and empirical analyses to use information from Twitter in understanding short-term fluctuations in political preferences. The present paper focuses on some measurement issues. The central research question is: How do political issue salience and issue ownership manifest on Twitter?

We present an analysis of tweets mined during the Dutch 2012 election campaign. Rather than casting our net as wide as possible, we concentrate here on some issues that can be considered on the broader left-right dimension. These issues are: income differences, welfare state services, economy, and labour market. In addition, we also focus on an issue that was important in political debates right before the elections: Housing policies. Thus we will be able to contrast traditional bread-and-butter issues with an issue that does not easily fit the left-right dimension.

After a further introduction into the background of our study, we describe trends in issue salience on Twitter. We will then investigate the association between the selected issues and political parties. Previous survey research has looked into issue ownership. We will analyze whether or not such issue ownership is similarly present in the Twitter sphere. For this, we look at tweets that mention both a party/leader and an issue and investigate the extent to which these co-vary. From this point, we will make connections to survey data. We will compare the extent to which issue salience and ownership as manifested on Twitter space are associated with data collected by recent election surveys. By doing this, we attempt to validate our measures of issues using Twitter data. Questions that are of interest here are: Are the same issues mentioned in the survey data as in the Twitter data? Do parties and issues connect similarly in both data sets? And if not, can we explain any differences? Once we can find similarities, we can more strongly assume that our measures of the selected issues are valid.

Issue salience and issue ownership as determinants of voting behavior

Theoretical Background

Issue salience originally designated the importance individual voters attach to different issues when evaluating political candidates, i.e. the weight placed on an issue when making political decisions such as voting (Berelson et al. 1954, Wlezien 2005; Jennings and Wlezien 2011). It is also used to indicate how prominent an issue is in the mind of an individual voter, that is, how accessible the topic is when the individual is prompted to make a political decision (Wlezien 2005). Measures of issue salience are used widely in political science, particularly in agenda-setting research (Soroka 2002; Jones and Baumgartner 2004; Wlezien 2005), policy representation research (Soroka and Wlezien 2010; Lindeboom 2012), and voting behavior (Bélanger and Meguid 2008).

Issue salience is an important addition to models explaining how political issues can play a role on voters' party preferences. The directional and proximity models of issue voting hold that

voters make their party choice at elections based on the distance, or directional congruence, between the position of the party and their own position on issue dimensions (Enelow and Hinich 1984; Rabinowitz and Macdonald 1989; Macdonald, Listhaug and Rabinowitz 1991). While these models contribute greatly to the explanation of issue-based vote choice, they typically assume that the impact of issues on voting behavior is invariant over parties. But in practice issues may matter more to voters of some parties than to those of other parties. The salience model of issue voting complements other models by taking the relative weight of the issues for voters into account. Its basic assumption is that parties and candidates who are associated with (certain views or goals on) a specific issue are more likely to be elected by voters who currently perceive that issue to be salient (Borre 2001). Thus the more important an issue is for a voter, the more likely it is that he/she will favor a particular candidate or party on that issue, provided that he/she establishes a link between the party or candidate and the issue in question. This means that the degree to which a political issue is identified with a party is a decisive factor in the relationship between the relative importance of an issue and the voters' preference for that party.

The association between particular parties and issues is best conceptualized by the theory of issue ownership. The definition of issue ownership is however not straightforward. While earlier studies have defined issue ownership primarily as a party's perceived ability to deal with a specific problem or an issue (e.g. Budge and Farlie 1983; Petrocik 1996), more recent research stresses that the concept consists of two related dimensions: a competence dimension and an associative dimension (Petrocik, Benoit and Hansen 2003; Damore 2004; Holian 2004; Van der Brug 2004; Walgrave, Lefevere and Nuytemans 2009; Walgrave, Lefevere and Tresch 2012). Associative issue ownership implies that "[...] people, when hearing or reading about an issue, automatically and spontaneously start thinking about a certain party" (Walgrave et al. 2009: 156). Parties are thus considered to "own" an issue a) if they have a reputation for having a policy or program interest for the issue and are therefore being identified with it, and/or b) if their policy solutions are seen as credible and receive public support. For instance, in European party systems, social-democratic parties are typically considered to "own" welfare issues, while green parties are identified with environmental issues in the first place (Walgrave et al. 2009). Similarly, Republican and Democratic parties in the U.S. party system are associated with certain issues; while the Republican Party is thought to own issues such as fighting terrorism and crime, environment and social security issues are ascribed to the Democrats.

Both dimensions of issue ownership are thought to be determinants of voting behavior when combined with the salience of the issue in question (Bélanger and Meguid 2008; Walgrave et al.

2012). Factors such as a party's perceived competence on an issue or the likelihood of delivering particular policies in that issue area are crucial as "these factors can give one party an advantage with important implications for party differentiation" (Green and Hobolt 2008: 463). Yet their influence on voters' party preferences is not independent from the perceived salience of the issue. Bélanger and Meguid argue that "a party's competence on an issue should not influence voter behavior unless the issue is considered important" (2008: 477; see also Petrocik 1989, 1996). These assumptions have been confirmed empirically in a number of countries (see Walgrave et al. 2009 for an overview).

The remaining puzzle is how theories of issue salience and ownership could contribute to our understanding of the recent phenomenon of volatile voters. One plausible explanation is the dynamic nature of issue salience: While voters' policy positions remain rather stable over time, the importance they attach to certain issues are bound to change due to several factors such as media reporting or the general flow of events (Page and Shapiro 1992). For instance, security and fighting crime have traditionally been significant issues within the electoral competition, but after the terrorist attacks of September 11th, 2001, they became highly salient. Issue ownership, on the contrary, has been conceptualized as relatively stable. In line with classical cleavage theory (Lipset and Rokkan 1967), parties have been thought to build their issue agendas upon their traditional constituency's position on the dominant conflict dimensions (Klingemann, Hofferbert and Budge 1994; Petrocik 1996). Although the historical cleavages are losing their importance, core issues such as welfare, taxation or labor are still predominantly associated with social-democrat or liberal parties. Thus even if a change takes place in the ownership of such issues, it does so over a longer time period; short-time fluctuations would not come into question for such issues.

Yet newer accounts increasingly stress the dynamic character of issue ownership (Walgrave et al. 2009; Green and Jennings 2012; Stubager and Slothuus 2012; Walgrave et al. 2012). The perceptions of party reputation on a given issue can change over time due to several reasons. First and foremost among these are the parties' efforts in framing new issues and competing with each other over new as well as existing issues (Blomqvist and Green-Pedersen 2004; Holian 2004; Bélanger and Meguid 2008; Walgrave et al. 2009). In order to guarantee electoral success and government office, parties can strategically emphasize or manipulate certain issues over the course of the electoral campaign. Issue salience can be a decisive factor in determining parties to frame issues. If an issue becomes salient during the campaign, parties and candidates can respond to this by simply spending more time on discussing this issue (Ansolabehere and Iyengar 1994). Another strategy that parties can apply during a campaign is to address so-called

“wedge” issues (e.g. abortion, immigration etc.) in order to attract the cross-pressured voters (Hillygus and Shields 2008). Mass media coverage also has a significant influence on the perceived link between parties and certain issues and can contribute to changes in these perceptions, particularly for short-term issue ownership dynamics (Walgrave and De Swert 2007).

In spite of this potential for fluctuations in issue ownership, the changes are believed to take place under specific conditions (Walgrave et al. 2009). First of all, the issue type matters for stability and change of issue ownership. Core issues such as typical left-right concerns are solidly owned by particular parties and thus difficult for other parties to reframe, whereas peripheral or new issues can more easily be claimed by multiple parties. Second, the dimension of issue ownership matters: The competence dimension – i.e. the perceived ability to deal with issue – is more dependent on performance and therefore more prone to change than the associative dimension of ownership. Finally, the party system characteristics can determine the extent to which issue ownership can change: Multiparty systems are more vulnerable for fluctuations in perceptions of issue ownership while two-party systems appear to be more stable.

In short, both issue salience and issue ownership are dynamic phenomena that are likely to alter party competitions and thus cause fluctuations in the perceptions of voters. This, in turn, can lead to electoral volatility depending on the nature and intensity of the change in both phenomena. In the following, we introduce the issue environment and issue change in the Dutch context and discuss possible long- and short-term fluctuations in issue salience and ownership. Following upon this, we will make connections to the problem of electoral volatility that has been increasing in this context in the past decade.

Political issues in the Netherlands

Politics in the Netherlands has traditionally been dominated by two dimensions: a social class-related left-right dimension, and a religious or moral conflict dimension (Lijphart 1968). The left-right dimension refers to the general contrast between state intervention and free market economy. Political parties that are outspoken on this left-right dimension, include the traditional antagonists PvdA and VVD. The religious dimension for most of the 19th and 20th centuries referred to differences between Catholics, Dutch-Reformed, Calvinists and nonreligious persons. With the divides between Christian religions becoming less sharp after the 1950s, the religious dimension in Dutch politics increasingly resembled a general moral conflict dimension,

contrasting individual self-determination with collective and religious norms. Political parties with clear noncentrist positions on this religious/moral dimension include CDA, CU, and D66.

In more recent years, a third conflict dimension has presumably become more important in the politics of many Western-European countries, including the Netherlands (Kriesi 2007; Aarts and Thomassen 2008). This dimension is referred to as the libertarian-authoritarian dimension, of the globalization dimension. It distinguishes between voters and parties who promote an open economy and those who favor restrictions on globalization and European integration, for example in the labor market or through tariffs. The rise of several new political parties in the past two decades testifies to the electoral importance of this new dimension.

In sum, the issue dimensions that currently shape Dutch politics are:

1. Left-right ideological dimension (Income differences, social benefits and welfare state, in broader sense: also environment)
2. Religious or moral conflict dimension (Religion, ethical issues like euthanasia, abortion)
3. Libertarian-authoritarian dimension (Europe, globalization, Asylum seekers, ethnic minorities, crime and security, individualism and self-determination)

Citizens' priorities for specific issues have shown considerable fluctuations since the 1970s (Lindeboom 2012). In the late 1970s and 1980s, mass unemployment was generally considered to be the most important problem in the Netherlands. In the first years of the 1970s, and in the late 1980s, however, environmental pollution topped the priority list. During the 1990s and 2000s, issues related to immigration, crime, health care and the functioning of the political system became more important. Lindeboom furthermore shows that welfare state and economy issues are still prioritized by government, other issues such as health care are secondary. However, his analyses only cover the time space between 1971 and 2006, and it remains to be seen what the impact of the financial and economic crisis that started in 2007 is on the current issue priorities of parties and voters. Before the crisis, sociocultural issues were more prioritized in Dutch parliamentary debates. Since 2008, however, the socio-economic dimension is likely to be dominating the issue environment again (Van der Meer et al. 2012). The campaign for the 2012 parliamentary election appears to have been dominated by classical distributional issues indeed.

Given their relative importance in the campaign, we focus in this paper on traditional left-right issues as one of the core issue families that shape Dutch politics. From the other issues that were intensively debated during the campaign (education, housing, Europe, safety and security,

etcetera), we have picked housing policy. Housing policy in the Netherlands is dominated by the traditional 100% tax deductibility of mortgage interest payments. This deductibility is said to have caused artificially high property prices and extremely high mortgage debts, and makes it very difficult for younger persons to purchase a house. In addition, the market for rental real estate is characterized by a mismatch between rents and incomes. Housing policy has for a long time been a dormant issue in Dutch politics, but has now become an urgent priority with rising unemployment and dropping real estate prices. It became one of the dominant issues in the 2012 campaign debates. We expect that housing policy was visible in social media as well, in addition to the more traditional left-right based issues. We now turn to the measurement of issue saliency and issue ownership through social media.

Twitter as a data source for measuring political concepts

Since its launch in 2006, Twitter has already been exploited as a data source for studies on a variety of topics including politics. The growing interest in Twitter, among other social media, as a source of observing public opinion is due to several factors. First of all, Twitter is used by a large number of persons to express, distribute and discuss political opinions on a daily basis. The tweet posts are moreover available to a broad public which is not limited to the followers of the messenger. These characteristics make it possible to observe political expressions on Twitter continuously and within a large geography, while the cost of collecting this information is much lower when compared to surveys. A second reason to prefer Twitter to measure public opinion over other social network sites is that the messages on Twitter are more structured and organized. The use of hash tags to organize discussions around a particular topic enables researchers to track and quantify the information in tweet posts on various political actors and objects. The appropriateness of Twitter data for the study of political phenomena has been advocated by a number of scholars who utilized these data in their studies on political attitudes and behavior in recent years. For instance, Tumasjan et al. have compared party mentions on Twitter with the results of the 2009 German election and concluded that “the mere number of tweets reflects voters’ preferences and comes close to traditional election polls, while the sentiment of political twitter messages closely corresponds to the electorate’s sentiment” (2010: 13). Similarly, Tjong Kim Sang and Bos (2012) use tweets mined two weeks preceding the 2011 Dutch senate elections and show that a sentiment analysis of tweets mentioning parties performed equally well as election polls in predicting the number of seats per party. Another study on public sentiment toward U.S. presidential candidates has applied a time-series analysis to political public opinion polls and twitter messages that mentioned President Barack Obama

(O'Connor et al. 2010). Using software that measured the sentiment in Twitter messages, they were able to compare the public sentiment about Obama to traditionally collected public opinion polls, where they found a high correlation between the two collection methods. Based on their findings, the authors argued that Twitter provides accurate measurements of public opinion, and suggested that "expensive and time-intensive polling can be supplemented or supplanted with the simple-to-gather text data that is generated from online social networking" (O'Connor et al. 2010: 7-8.)

Despite the potentials of using Twitter as a data source for measuring political phenomena, and the encouraging findings on its comparability with public opinion polls, there is still no consensus on the extent to which Twitter contents reflects general public opinion. Some scholars have provided counter evidence to the assumption that measurements extracted from Twitterspace can be used as reliable and valid indicators of offline political sentiments (e.g. Gayo-Avello et al. 2011; Jungherr et al. 2012). In addition to this, a well-known drawback of using Twitter data to measure political concepts is that the data is inherently biased due to the so-called digital divide (Norris 2001). The users tend to be younger and higher educated so that they do not match the demographics of the population of citizens at large, which means that the sample drawn from the Twitter stream – even if it is random – cannot be representative of the whole population (see Tjong Kim Sang and Bos 2012). For data consisting of conversations on political topics this bias is expected to be much greater; e.g. towards the more politically interested and knowledgeable. Finally, Twitter data is structurally different compared to survey data, which makes the validity of inferences derived from this data source questionable. Survey responses are generally less ambiguous to interpret. If a survey respondent mentions the economy as an answer to the question on the most important problem facing the country, it reflects – with less uncertainty – that he/she holds this view (see Mellon 2013). However, if a Twitter user posts a message about economy, there can be many motivations behind such an act, next to the odds that the user indeed finds this issue very important. Moreover, the search terms can have multiple meanings, so that the data collected using a specific term might end up being about a completely different topic than intended. Thus the data can potentially be afflicted with greater internal and external validity problems than survey data.

Twitter offers a rich and easily accessible source for tracking public opinion on a day-to-day basis, provided that the extracted data are valid and reliable indicators of public opinion. The promises and pitfalls of utilizing Twitter data in the study of public opinion described above call for a cautious approach to the usability of this data source to monitor issue salience and issue ownership. Political issues form a considerable portion of the topics discussed on Twitter.

Combined with the low costs and the continuous availability of information, Twitter offers a huge potential to capture trends in both issue salience and ownership. These concepts are surveyed regularly by the Dutch Parliamentary Election Studies, which usually takes place in two waves before and after each parliamentary election. However, the long intervals between the measurements makes the observation of fluctuations in issue priorities and party-issue associations over shorter time periods impossible. Using political discussions on Twitter to measure issue salience and issue ownership could potentially solve this problem. Yet the drawbacks outlined above indicate that the validation of the online measurements of these concepts is necessary. Without validating Twitter against survey data, we can thus not be sure if the aggregated term use in Tweets reflect the underlying attitudes. In this paper we will therefore contrast survey and Twitter measurements of both issue salience and ownership. In order to assess the criterion-related (or predictive) validity of Twitter data, we will compare the frequency of issue mentions and co-occurring references to parties and issues within tweet messages with the existing measures of issue salience in survey research. Similar attempts have been made for the measurement of issue salience with Google Insights for Search (e.g. Graefe and Armstrong 2012; Mellon 2012, 2013). Mellon (2012, 2013) has furthermore undertaken an attempt to validate his measures of issue salience using search indices. Such an attempt, to our knowledge, does not yet exist for the measurement of this concept on Twitter.

Data and Measurements

For the measurement and analysis of issue salience and issue ownership we utilize a data set consisting of Tweets on Dutch elections, collected during the parliamentary election campaign (Aug 25th - Sep 12th 2012)¹. In order to identify political discussions on Twitter, we mined tweets that have been tagged by a hash (#) symbol as being about political parties and candidates. As these so-called hash tags are central to organizing information on Twitter around particular events and topics (Small 2011), using these as search criteria seemed to be most relevant for our sampling purposes. However, when one defines the target population as “political discussions on Twitter” a case selection procedure based on hash tags poses a big challenge. Political messages can be about several different objects varying from political actors

¹ The data has been collected within the framework of the NWO-funded investment project “Transforming Social Media Contents to Political Data” (NWO grant 480-11-010), conducted at the University of Twente since July 2012. This project aims to build a research infrastructure for the study of social scientific research questions, by developing new tools and methods for harvesting, coding, and analysis of messages on social media. For more information see http://www.utwente.nl/igs/research/projects/social_media.

and institutions to actual political issues. This urges researchers to be very precise in defining the list of search terms prior to the data collection, whereas it is extremely difficult – if not impossible – to deductively decide which hash tags will be relevant indicators of a discussion thread for at least two reasons. First, the volume of tweets produced by any individual can be quite large and the contents can vary among a large number of topics, where politics forms a marginal part of all topics covered by messages on Twitter. Secondly and more importantly, the terms used to tag tweets are user-defined, emerge dynamically during the course of events, and thus are employed in a rather fluid way. Even in defining a relatively stable group of objects such as parties and candidates, the use of hash tags can vary to a great extent. To give an example, a party name can be used as a tag the way it is ('#partyX'), or it can be used in other ways to cheer for the party (e.g. '#gopartyX') or to express negative sentiment (e.g. '#neveragainpartyX') or even to explicitly persuade to vote for the party (e.g. '#voteforpartyX'). Therefore, mining on a fixed list of hash tags can mean that tweets tagged with derivatives of parties and leaders will be left out of the sample. Also, tweets which are organized around other tags related to several other aspects of the elections – such as party campaign slogans, news and candidate debates – will end up not being selected. In the case of political issues, any attempt to capture the online discussion by means of a fixed set of hash tags will be particularly difficult since issues evolve rather dynamically, especially over the course of an election campaign.

In short, case selection based on a fixed list of topics and hash tags is bound to be biased towards those tweets tagged with the researcher-defined key words. To limit such sampling bias, we decided to employ a “snowball method”. We started data collection with twenty hash tags, ten of which belonged to the political parties represented in the Second Chamber of the Dutch parliament, and ten for their leading candidates. Subsequently, we let a script extract other tags present in the mined tweets and had an algorithm assign relevance to each of those tags based on co-occurrence with the tags we decided to mine on. Once a tag’s relevance passed a certain threshold, it was added to the list of tags to mine on, thereby also becoming part of the set of tags used to identify new relevant tags. This procedure allowed us to expand the list of hash tags used for mining tweets in such a way that allowed us to overcome the above mentioned obstacles that result from purposive sampling based on a fixed hash tag list to a great extent. In this way we were able to retrieve alternative tags referring to elections, parties and candidates, as well as to identify several political issues by capturing the hash tags that emerged parallel to the unfolding events during the election campaign. The collection of tweets by means of this procedure resulted in a data set with over one million tweets.

Despite the advantages of the snowball sampling method in reducing the sampling bias, data collected by this method is flawed by being prone to contain a great deal of “noise”. Hash tags are not fixed in their meaning, i.e. a specific hash tag can be attributed to several different objects at the same time (e.g. #cda stands both for the Dutch Christian Democrat Party and for Canada). Similarly, a hash tag with a rather unequivocal meaning can potentially be used within the context of various discussions next to politics². This means that our raw data set of more than one million tweets contained irrelevant messages in other languages or on other topics than the Dutch elections to a significant extent: as we expanded the list of hash tags to mine on based on a co-occurrence of tags within tweets, tags that were falsely considered highly relevant by our algorithm would pick up a relatively large number of false positives compared to actually relevant tags. Therefore, the raw data needed extensive cleaning, which was performed for the time being in three stages. First, we automatically identified the language of tweets and discarded those identified as being non-Dutch. Second, we got rid of tweets which were identified as sent by spammers. Finally, tweets that were not related to Dutch politics were identified by using automated content analysis³ and excluded from the corpus. After this three-stage cleaning process, we were left with a data set containing 294,585 tweets, on which the measurements of issue salience and ownership will be based.

As a benchmark for these measures we will make use of provisional data from the 2012 Dutch Parliamentary Election Study (DPES). The DPES is a national election study which is conducted since 1971 around the parliamentary elections⁴. While data in most earlier studies have been collected in two rounds of interviews (pre- and post-election), the 2012 data was surveyed by combining these two rounds to one round of interviews (CAPI) with a total of 1,677 respondents between September 13th and October 31st 2012, i.e. in the first seven weeks after the elections.

² For instance, one of the main substantive issues in the campaign was the sustainability of the Dutch health care system. The hash tag #zorg (health care) therefore co-occurred rather often with party or candidate hash tags. Due to this frequent co-occurrence, #zorg was added to the snowball and tweets tagged with this particular hash tag collected. But tweets containing this tag more often than not referred to vacancies in health care institutions, or to people’s experiences with a particular health care provider, rather than to the substantive discussion of health care in the campaign, adding false positives to our data set.

³ Automated content analysis is a method that involves both human and machine coding of text data in two subsequent steps. In the first step, a randomly selected subsample of the text documents are coded by human coders based on a given coding scheme. This subsample is then used as a training set to code the rest of the documents automatically (for more information see Hopkins and King 2010).

⁴ Previous studies have been conducted around each parliamentary election since 1971. For more information see <http://www.dpes.nl>.

Operationalizations of concepts

In this section we introduce the online and offline measurements of issue salience and issue ownership. We will first discuss the measures of these concepts in our survey data. A long-established measure of issue salience is the open-ended question on “the most important problem facing the country at present”, which has been asked in several public opinion polls and academic studies (Wlezien 2005; Heath and Johns 2010). This measure is featured in the DPES 2012 as well, but the question wording is slightly different. Respondents have been asked about the most important problems in the country, thus allowing for multiple issues to be mentioned. The answers given by the respondents have been coded into nine categorical variables ranging from first most important national problem to the ninth, each of which consists of 16 issue categories (see appendix). The average frequency of mentions of an issue has been treated as the average importance of that issue among all respondents.

As argued above, the concept of issue ownership is thought to consist of an associative and a competence dimension (Walgrave et al. 2009). Accordingly, it is being measured in different ways. The most commonly used question to measure the competence dimension of issue ownership is the perceived competence of a party to deal with a specific issue (e.g. Van der Brug 2004); in most cases this item is combined with issue salience indicators (e.g. Bélanger and Meguid 2008; Green and Hobolt 2008). Next to these commonly used measures, the associative dimension can be measured using a more direct question. In previous DPES studies, the following questions were posed in the drop-off questionnaire: When you think of [party], what is the (first/second) issue that you first think of? And to what extent do you agree with the party on that issue? Unfortunately, these variables are not available in the most recent data. We used instead the following items to compute the issue ownership variable: 1) Party voted for in 2012 elections, and 2) Reason to vote for the party (string variable). 1,414 out of 1,677 persons have responded to the open-ended question on the reason to elect the party they voted for, and a total of 585 respondents have mentioned specific issues as the main reason why they voted for the particular party. These responses have been coded into corresponding categories (see appendix). Since some respondents have associated more than one issue with the party voted, individual dummy variables have been computed for each issue depending on whether it has been mentioned or not. The ranking of party votes within each mentioned issue is our measure of issue ownership (see table in the appendix).

The operationalization of these concepts with Twitter data is quite similar to the measurements with DPES data, yet the procedure is far less straightforward. In DPES, the data consists of responses to survey questions which are of rather general nature. For instance, the question on

most important problem is answered in most cases by clear-cut expressions, e.g. economy, health, education, etc. Contrarily, tweet messages are not responses to questions but rather freely formulated thoughts and opinions of the tweeter on a particular topic. As a consequence, the number of words and expressions used to define a political party, leader or an issue can be far more than those used to answer open-ended questions in surveys. Due to the high volume of tweets, however, detecting all terms that define a particular issue by hand-coding is highly time-consuming – in bigger data sets than ours, it is nearly impossible.

This urges us to develop different strategies to *automatically* categorize tweets according to the issues they address. One can adopt either a deductive (automated search for words and coding based on a pre-defined list of terms) or an inductive (automated detection of clusters of terms and exploring the underlying constructs) procedure in doing this. In this paper we use the deductive method. Political parties and leaders have been extracted from the data on the basis of lists of objects consisting of political party names and name/surname of party leaders (see coding scheme in the appendix). Also, party campaign slogans (e.g. nu vooruit by D66) have been featured on each respective list. In order to make sure that as many tweets as possible can be correctly coded, we made use of “word stems” instead of words. This means that we have defined both party and leader names as word stems in the search algorithm, so that tweets which contain derivatives of party and leader names – such as “pvvfail” or “teamroemer” – can also be correctly classified into the party categories⁵.

We applied a similar procedure for the coding of issues within the Tweet corpus. As mentioned in the theoretical section, the main groups of political issues we are interested in are labor market, economy & finances, income differences and welfare state, health care, and housing policy issues. Based on party manifestos for the Dutch election campaign 2012, we first defined the scope of each issue group: labor market, for instance, does not only consist of issues of unemployment but also covers other related subjects like work conditions, employer and employee relations, taxes on income, and retirement. Second, we identified a list of terms and of synonyms for each term – again to be used as word stems – for the automated extraction of these issues. One problem that might result from this approach is that the usage of common words such as care (zorg) can cause many tweets that are not actually related to that policy area (e.g. tweets about job vacancies in health care sector) to be falsely coded into the respective

⁵ For some of the word stems defining a party or a leader, hash tags have been preferred to word stems. For instance, Sap refers to the leader of Groen Links, but at the same time also means juice, or it can easily appear as a letter combination within words. Therefore we used #Sap instead of the word stem, based on the assumption that hash tags organize discussions and therefore the chance for an accurate classification is higher.

issue category. As we have reduced the corpus of collected tweets to the ones that are election-related, however, we expect a high accuracy in classification of tweets into issue groups.

Having identified categories of tweets about a particular party, leader or issue, we have measured issue salience as the number of tweets mentioning a particular issue category and issue ownership as co-occurring mentions of parties/leaders and issues within the same tweet.

Analyses

Issue salience

We begin with the analysis of issue salience by focusing on descriptive statistics of the selected issues in Twitter and DPES data. We first display how much attention has been paid to which issue in the political discussions taking place on Twitter during the election campaign. Following upon this, we compare the amount and ranking of the mentioned issues to the results obtained by the DPES. Finally, using Twitter data we will demonstrate how attention for a particular issue developed during the campaign.

Examining the frequency of mentions of the core left-right issues and housing issue, we can establish that these issues have not been paid much attention within all tweet messages about elections. While 234,976 tweets (79.8% of the analyzed corpus) have mentioned at least one party, only 43,254 tweets (20% of the analyzed corpus) have mentioned at least one of our five selected issues⁶. As shown in figure 1, labor market related issues are the most mentioned group of issues among all, followed by health care and economy issues. Housing and social welfare have been cited least by Tweeters posting about the elections. Figure 1 furthermore contrasts the frequencies and rankings of these issues on Twitter to the findings of DPES. The left-hand side of the figure presents the percentages of respondents that have mentioned the respective issue among all respondents of the CAPI interview (N=1,677). The percentages on the right side of the figure, on the other hand, refer to the shares of individual issues among the tweet posts on the selected issues (N=43,254). Since survey and Twitter data are different in nature, it makes little sense to make a blunt comparison of the percentages of mentions of the issue groups as compared to the total number of observations of the both data sets. As discussed above, surveys

⁶ This might suggest at a first glance that candidates and parties get more attention from the electorate than the central issues of the campaign; yet considering that the Twitter discussion on elections might contain several other issues which are not observed in the present analysis, we refrain from drawing such conclusions for the time being.

provide a stimulus for the respondents to mention an issue by asking questions, while such stimulus does not exist in Twitter messages. It can be therefore expected that the observed percentage of mentioned issues will turn out much higher in DPES data, which should not necessarily mean that the two data sets produce different results. In order to provide for comparability, we thus decided to treat the tweet messages on issues as responses to stimuli and calculated the displayed shares of issues on the basis of these observations.

[Figure 1 about here]

As can be seen in the figure, almost all selected issues are mentioned to a comparable extent and with a similar ranking on both Twitter and survey data. Only with respect to economic and financial issues we can observe significant differences between the results provided by the two data sources. In DPES this issue has been addressed by far as the most important problem in country by 87.1% of the respondents. Since the measurements of other issues seem to be equivalent, how can we explain this deviation? One possible factor is that the word “economy” has been used in a rather general way by the DPES respondents. Since the problems in labor market or the radical changes in the health care system mostly go hand in hand with economic problems, there is reason to assume that the respondents have mentioned economy as a substituting or complementing term for other issues. Another likely cause is the measurement of economy and finance issues on Twitter data. As our operationalization of issues was only based on key words selected from party manifestos, there is the possibility that specific words belonging to this broad issue category have been excluded from the search term list. Adding a wider list of words to the search term list might have resulted in economic problems being the most mentioned issue and thus provide comparable results to those obtained by the Dutch election study.

On the other hand, a glance at the trends in issue salience on Twitter throughout the election campaign suggests that the above displayed similarities cannot be straightforwardly be interpreted as validating the online measurement of this concept. As figure 2 shows, the daily frequencies of issues mentioned in Twitter discussions fluctuate heavily during the whole campaign period. The most mentioned issue on average on Twitter, the labor market issue, has been a dominating discussion only in the last two days of the election campaign and on the day of election. The second most mentioned issue, health care policies, followed a similar pattern during that period, yet there are two other spikes on August 29th and September 4th, which correspond to the TV debates broadcasted on the respective days. In these debates, health care was one of the main issues on which the leaders had to take positions. This seems to have mobilized a simultaneous discussion on Twitter on this issue, which resulted in health care

being the most important issue of the day. These fluctuations suggest two things: that the discussions on Twitter not only do reflect independent opinions of users, but also mass media influences to a considerable extent; and that the data collection period matters for the results obtained with these measurements.

[Figure 2 about here]

In sum, the Twitter data informs us that attention paid to both core issues and the peripheral issue housing has been subject to alterations in the last three weeks of the election campaign. Particularly housing seems to have gained importance in the final days before the election, though on average it appeared to be a secondary issue compared to the traditional left-right issues. As we have mentioned before, changes in the salience of peripheral issues are particularly important since they can play a decisive role for the outcome of the elections in case they are emphasized by a specific party or group of parties in the run-up to the elections. One needs therefore complementary information on issue ownership and its trends in order to tell whether the increasing salience of this issue has stimulated a particular political party to become the owner of this issue.

Issue ownership

Out of the tweet posts on electorally relevant issues, 37,645 (87%) of the tweets mentioning issues have mentioned at least one party at the same time; whereas 5,609 (13%) have mentioned these issues independently, i.e. not together with a party or leader. In this section we will first compare the results of the online issue ownership measure with those obtained by the DPES measurement of issue ownership. Moreover, we will analyze whether and to what extent there are fluctuations in the associations between parties and issues on Twitter.

[Table 1 about here]

Table 1 shows that parties owning the selected issues are almost identical in both data sources. The ranking of parties within each issue is also quite similar. In line with the theoretical expectations, labor party was predominantly associated with labor market issues whereas the economic and financial issues were associated with VVD, the liberal party, both in survey responses and on Twitter. With respect to health care discussion, the socialist party (SP) has been assessed as the owner of this issue by both survey respondents and Twitter users. Again here, the measurements on survey and Twitter diverge on one issue: social inequality and welfare state. In DPES, the majority of the respondents associated this issue with the labor party PvdA (40.4%), followed by the SP and the VVD. These results support the theoretical

assumptions that the social democrat or socialist parties are the typical owner of redistribution issues. However, the Twitter posts reflect a different picture. Here, the liberal party appears to be the party that is most linked to social inequality and welfare (39.2%), followed by the SP and the PvdA. This discrepancy might again have been caused by measurement issues. The measurement of issue ownership in DPES is based on a positive relationship between the party and the issue in question, i.e. the performance of party in that issue area has been stated as the reason to vote for the party. On the contrary, the measurement of this concept in Twitter reflects a rather straightforward association between issues and parties. For instance, in the Tweet posts about social redistribution and welfare, the liberal party's positions have been supported by some users, but also often criticized to be deepening the gap between the income classes and thus being anti-social. Therefore, the measurement of this issue is likely to reflect the controversial discussion about the position of the liberal party on this issue rather than its perceived competence to deal with the issue.

In the theoretical section we discussed that the owners of core issues are mostly fixed and do not fluctuate much, but the peripheral issues are rarely owned by a particular party their ownership can be crucial for election outcomes, especially in case they have a high salience. We have demonstrated in the previous section that the average salience of housing policy was not very high, but the issue gained weight in the final days of the election campaign. Therefore we observe the development of traditional left-right issues and housing issue for three parties over the course of campaign period.

[Figure 3 and 4 about here]

As can be taken from table 1, housing issue has been associated with the liberal party VVD in both survey and Twitter data with a high frequency. A longitudinal analysis of the Twitter posts on this issue further suggests, however, that its ownership has been through serious alterations (see figure 4). Figure 3 demonstrates that ownership of core issues fluctuates as well, though to a lesser extent. Salience of left-right issues, especially of labor market issue, has increased during the last days before the elections. Parallel to this, the association between all three parties and these issues has shifted as well; especially the linkage between the social-democrat PvdA and these core issues seems to have increased dramatically in the last three days before the elections. Returning to figure 4, we can establish that an even stronger volatility in the link between the housing issue and the three parties has taken place. This could be interpreted as a reflection of the increasing party competition on the ownership of these issues in the perceptions of the electorate as measured on Twitter. However, it is questionable whether it is an indicator of a real competition since both figures show that the increase in the frequency of

associations of an issue with a specific party rarely occurs at the expense of the other parties – rather, all three parties display the same trend in the ownership of issues. This once again demonstrates that the measurement of issue ownership on Twitter is clearly not related to competence ownership, and is highly influenced by the trends in perceived salience of issues. Therefore, it is not possible to make any direct connections to the electoral success of these parties by relying solely upon Twitter data.

Conclusions

In this paper we took a preliminary step towards providing sound measurements for two prominent concepts in electoral research by utilizing social media contents. Our primary goal was to explore the extent to which Twitter discussions on political issues can serve as a data source for the observation of short-term changes in issue salience and issue ownership. Researchers increasingly pay attention to this new medium as a data source for the study of political opinions and behavior since it offers a huge potential for tracking public opinion on a longitudinal basis with minimal costs. But because of the unique characteristics of Twitter posts which differ notably from traditional surveys, dealing with these data in terms of extracting reliable, valid and precise measurements of politically relevant concepts still forms a major challenge. The present research makes a first move to address this challenge by contrasting measurements of issue salience and ownership on Twitter to the measures in traditional survey data. We made use of tweets on political discussions collected during three weeks before the Dutch parliamentary election on September 12th, 2012, and the Dutch Parliamentary Election Study (DPES) which has been fielded right after the election. Our analyses have yielded the following insights:

First of all we have established that measures of issue salience and issue ownership using data collected from Twitter are comparable with the survey data to a considerable extent. For a majority of the issues we selected for our analyses, we found that both the shares and rankings of issues obtained by the online and offline measures of issue salience are highly similar. This was also the case for issue ownership. The parties that were identified with particular issues turned out to be identical in both data sets for the bulk of the issues. The findings thus indicate that there seems to be a link between the public political opinion as measured by survey data and the political discourse on Twitter. However, the measurements from the two data sources diverge from each other on two important issues, economy and social security, which suggests that the findings should be approached with caution.

As we have discussed above, the found discrepancies in measurements of both issue salience and issue ownership might have been caused by the daily fluctuations in priorities given to specific issues as well as in party-issue associations in Twitter data (see figures 2 to 4). The results obtained from Twitter data can thus be quite dependent on the measurement period, which suggests that more solid information is necessary in order to validate the measurements. Our analyses are flawed to some extent by the fact that the collection of both data sets have not taken place at the same time. Twitter reflects the pre-election period whereas DPES has surveyed the post-election attitudes. More reliable results can be obtained in case survey data covers the same period as the Twitter data; e.g. daily polls conducted simultaneously with data collection from Twitter would provide a better tool for the validation of online measurements.

Different operationalizations of the same concept in two data sets could provide another explanation for the discrepancy between online and offline measurements on certain issues. Leaving out or including particular words in the coding scheme of issues on Twitter data might result in completely different findings. This means that the measurements of issue-related concepts should be carefully revised before drawing conclusions on comparability of online and offline data. The used methods for measuring these concepts also need to be reflected upon. In the present study we used a deductive method to code and measure issues, i.e. for each issue we extracted relevant terms from election manifestos of parties and then used the list of these terms as a basis for automated coding of tweets. Other alternatives are semi-automated coding, which combines deductive and inductive procedures, and fully automated detection and coding of issues using semantic analysis on a purely inductive basis. These coding procedures might produce different measurements and thus different results; thus it is essential to compare these methods with each other and use surveys again as a benchmark to decide which procedure provides the most accurate measurements.

A final step to be taken in the future to improve the measurement of issue ownership on Twitter is to identify the sentiments in tweets which mention political parties and issues simultaneously. What we have measured in this study was associative ownership since we based our operationalization on the co-occurrence of parties and issues in the same tweet. Competence ownership, which is another important dimension of the issue reputation of parties, implies that the association between the party and issue is a positive one, yet we cannot observe this dimension with the current measurement. For instance, in Twitter data we found the liberal party VVD to be the owner of welfare issues, yet this could have been caused by the high number of negative tweets criticizing VVD's redistribution and welfare policies. Identifying the sentiment in tweets would make it possible to distinguish between critical and supportive contents about

the position of a party on a specific issue, and thus enable the measurement of the perceived competence of parties.

With our preliminary findings we have demonstrated that Twitter is highly eligible as a data source of for measurements of key political concepts, and that the measurements can be possibly developed to be exploited in studies of voting behavior. As a matter of fact, data extracted from Twitter cannot be utilized in the same way as survey data, e.g. for making causal inferences on motivational backgrounds of voting behavior. Yet once reliable and valid measurements are obtained, Twitter contents can potentially supplement survey data, and help to improve surveys in a way that the motivational backgrounds of volatile voters can be better studied.

References

- Aarts, K. and J. Thomassen (2008). "Dutch voters and the changing party space 1989-2006", *Acta Politica* 43(1/2), pp. 203-234.
- Ansolabehere, S. and S. Iyengar (1994). "Riding the Wave and Claiming Ownership Over Issues: The Joint Effects of Advertising and News Coverage in Campaigns", *Public Opinion Quarterly* 58, pp. 335-357.
- Bélanger, É. and B. M. Meguid (2008). "Issue salience, issue ownership, and issue-based vote choice", *Electoral Studies* 27(3), pp. 477-491.
- Berelson, B. R., P. F. Lazarsfeld, and W. N. McPhee (1954). *Voting: A Study of Opinion Formation in a Presidential Campaign*. Chicago: University of Chicago Press.
- Blomqvist, P. and C. Green-Pedersen (2004). "Defeat at home? Issue ownership and social democratic support in Scandinavia", *Government and Opposition* 39(4), pp.587-613.
- Borre, O. (2001). *Issue Voting: An Introduction*. Aarhus: Aarhus University Press.
- Budge, I. and D. J. Farlie (1983). *Explaining and Predicting Elections: Issue Effects and Party Strategies in Twenty-Three Democracies*. London: Unwin Hyman.
- Campbell, A., P. E. Converse, W. Miller and D. Stokes (1960). *The American voter*. New York: Wiley.
- Dalton, R.J. and S. A. Weldon (2005). "Public Images of Political Parties: A Necessary Evil?", *West European Politics* 28(5), pp. 931-951.
- Dalton, R. J. (2006). *Citizen Politics: Public Opinion and Political Parties in Advanced Industrial Democracies*. Chatham: Chatham House.
- Damore, D. F. (2004). "The dynamics of issue ownership in presidential campaigns", *Political Research Quarterly* 57, pp.391-397.
- Enelow, J. M. and M. J. Hinich (1984). *The Spatial Theory of Voting: An Introduction*. Cambridge: Cambridge University Press.
- Franklin, M., T. Mackie and H. Valen (eds.) (1992). *Electoral Change: Responses to evolving Social and Attitudinal Structures in Western Europe*. Cambridge: Cambridge University Press.
- Gayo-Avello, D., P. Metaxas and E. Mustafaraj (2011). "Limits of electoral predictions using social media data". Proceedings of the International AAAI Conference on Weblogs and Social Media, Barcelona.
- Graefe, A. and J. S. Armstrong (2012). "Predicting elections from the most important issue: A test of the take-the-best heuristic", *Journal of Behavioral Decision Making* 25(1), pp. 41-48.
- Green, J. and S. B. Hobolt (2008). "Owning the issue agenda: Party strategies and vote choices in British elections", *Electoral Studies* 27(3), pp. 460-476.

- Green, J. and W. Jennings (2012). "The dynamics of issue competence and vote for parties in and out of power: An analysis of valence in Britain, 1979–1997", *European Journal of Political Research* 51(4), pp. 469-503.
- Heath, O. and R. Johns (2010). "Measuring Political Behavior and Attitudes", in: Martin Bulmer, Julie Gibbs, and Laura Hyman (eds.) *Social Measurement through Social Surveys. An Applied Approach*. Farnham: Ashgate, pp. 47-68.
- Hillygus, D. S., and T. G. Shields (2008). *The Persuadable Voter: Wedge Issues in Presidential Campaigns*. Princeton, New Jersey: Princeton University Press.
- Holian, D. B. (2004). "He's stealing my issues! Clinton's crime rhetoric and the dynamics of issue ownership", *Political Behavior* 26(2), pp. 95-124.
- Hopkins, D.J. and G. King (2010). "A Method of Automated Nonparametric Content Analysis for Social Science", *American Journal of Political Science* 54(1), pp. 229-247.
- Jennings, W. and C. Wlezien (2011). "Distinguishing between most important problems and issues?", *Public Opinion Quarterly* 75(3), pp. 545-555.
- Jones, B.D. and F.R. Baumgartner (2005). *The Politics of Attention: How Government Prioritizes Problems*. Chicago: The University of Chicago Press.
- Jungherr, Andreas, Pascal Jürgens and Harald Schoen (2012). "Why the Pirate Party Won the German Election of 2009 or The Trouble With Predictions: A Response to Tumasjan, A., Sprenger, T. O., Sander, P. G., & Welpe, I. M. "Predicting Elections With Twitter: What 140 Characters Reveal About Political Sentiment", *Social Science Computer Review* 30(2), pp. 229-234.
- Kriesi, H. (2007). "The Role of European Integration in National Election Campaigns", *European Union Politics* 8(1), pp. 83-108.
- Klingemann, H., R. Hofferbert and I. Budge (1994). *Parties, policies and democracy*. Oxford, UK: Westview.
- Lijphart, A. (1968). *The politics of accommodation: Pluralism and democracy in the Netherlands*. Berkeley: University of California Press.
- Lindeboom, G.-J. (2012). "Public priorities in government's hands: Corresponding policy agendas in the Netherlands?", *Acta Politica* 47(4), pp. 443-467.
- Lipset, S. M. and S. Rokkan (Eds.) (1967). *Party systems and voter alignments: Cross-national perspectives*. London: Macmillan.
- Macdonald, S. E., O. Listhaug and G. Rabinowitz (1991). "Issues and Party Support in Multiparty Systems", *American Political Science Review* 85(4), pp. 1107-1131.
- Mellon, J. (2012). "Search Indices and Issue Salience: The Properties of Google Trends as a Measure of Issue Salience", Paper presented at the Elections, Public Opinion and Parties Conference, Oxford, 7-9 September 2012.

- Mellon, J. (2013). "Where and When Can We Use Google Trends to Measure issue Salience?", *PS Political Science & Politics* 46(2), pp. 280-290.
- Norris, P. (2001). *Digital Divide: Civic Engagement, Information Poverty, and the Internet Worldwide*. New York: Cambridge University Press.
- O'Connor, B., R. Balasubramanyan, B. R. Routledge, and N. A. Smith (2010). "From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series", Proceedings of the International AAAI Conference on Weblogs and Social Media, Washington, DC, May 2010.
- Page, B. and R. Shapiro (1992). *The rational public: Fifty years of trends in Americans' policy preferences*. Chicago: University of Chicago Press.
- Petrocik, J. R. (1989). "The theory of issue ownership: Issues, agendas, and electoral coalitions in the 1988 elections", Paper presented at the annual meeting of the American Political Science Association, Atlanta, GA.
- Petrocik, J. R. (1996). "Issue Ownership in Presidential Elections, with a 1980 Case Study." *American Journal of Political Science* 40(3), pp. 825-850.
- Petrocik, J. R., W. L. Benoit, and G. J. Hansen (2003). "Issue Ownership and Presidential Campaigning, 1952-2000", *Political Science Quarterly* 118(4), pp. 599-626.
- Rabinowitz, G. and S. E. MacDonald (1989). "A Directional Theory of Issue Voting", *American Political Science Review* 83(1), pp. 93-121.
- Small, T.A. (2011). "What the Hashtag? A content analysis of Canadian politics on Twitter", *Information, Communication & Society* 14(6), pp. 872-895.
- Soroka, S.N. (2002). *Agenda-Setting Dynamics in Canada*. Vancouver, BC: UBC Press.
- Soroka, S.N. and C. Wlezien (2010). *Degrees of Democracy: Politics, Public Opinion, and Policy*. New York: Cambridge University Press.
- Stubager, R. and R. Slothuus (2012). "What Are the Sources of Political Parties' Issue Ownership? Testing Four Explanations at the Individual Level", *Political Behavior*, doi. 10.1007/s11109-012-9204-2.
- Tjong Kim Sang, E. and J. Bos (2012). "Predicting the 2011 Dutch Senate Election Results with Twitter", Proceedings of SASN 2012.
- Tumasjan A., T.O. Sprenger, P.G. Sandner and I.M. Welpel (2010). "Predicting elections with Twitter: What 140 characters reveal about political sentiment", Paper presented at the Fourth International AAAI Conference on Weblogs and Social Media. Washington, DC: George Washington University.
- Van der Brug, W. (2004). "Issue Ownership and Party Choice", *Electoral Studies* 23(2), pp. 209-233.

Van der Meer, T., R. Lubbe, E. van Elsas, M. Elff and W. van der Brug (2012). "Bounded volatility in the Dutch electoral battlefield: A panel study on the structure of changing vote intentions in the Netherlands during 2006-2010", *Acta Politica* 47(4), pp. 333-355.

Walgrave, S. and K. De Swert (2007). "Where does issue ownership come from? From the party or from the media? Issue-party identifications in Belgium, 1991-2005", *The Harvard International Journal of Press/Politics* 12(1), pp. 37-67.

Walgrave, S., J. Lefevere and M. Nuytemans (2009). "Issue Ownership Stability and Change: How Political Parties Claim and Maintain Issues Through Media Appearances", *Political Communication* 26(2), pp. 153-172.

Walgrave, S., J. Lefevere and A. Tresch (2012). "The Associative Dimension of Issue Ownership", *Public Opinion Quarterly*, doi: 10.1093/poq/nfs023.

Wlezien, C. (2005). "On the salience of political issues: The problem with 'most important problem'", *Electoral Studies* 24(4), pp. 555-579.

Figures

Figure 1: Issue salience on Twitter and DPES

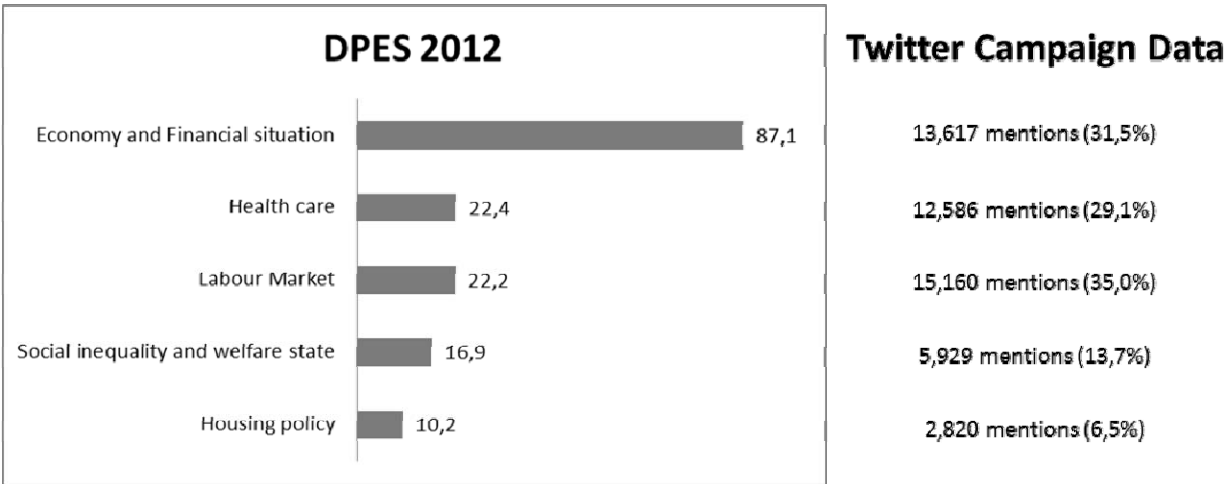


Figure 2: Trends in issue salience on Twitter during the election campaign

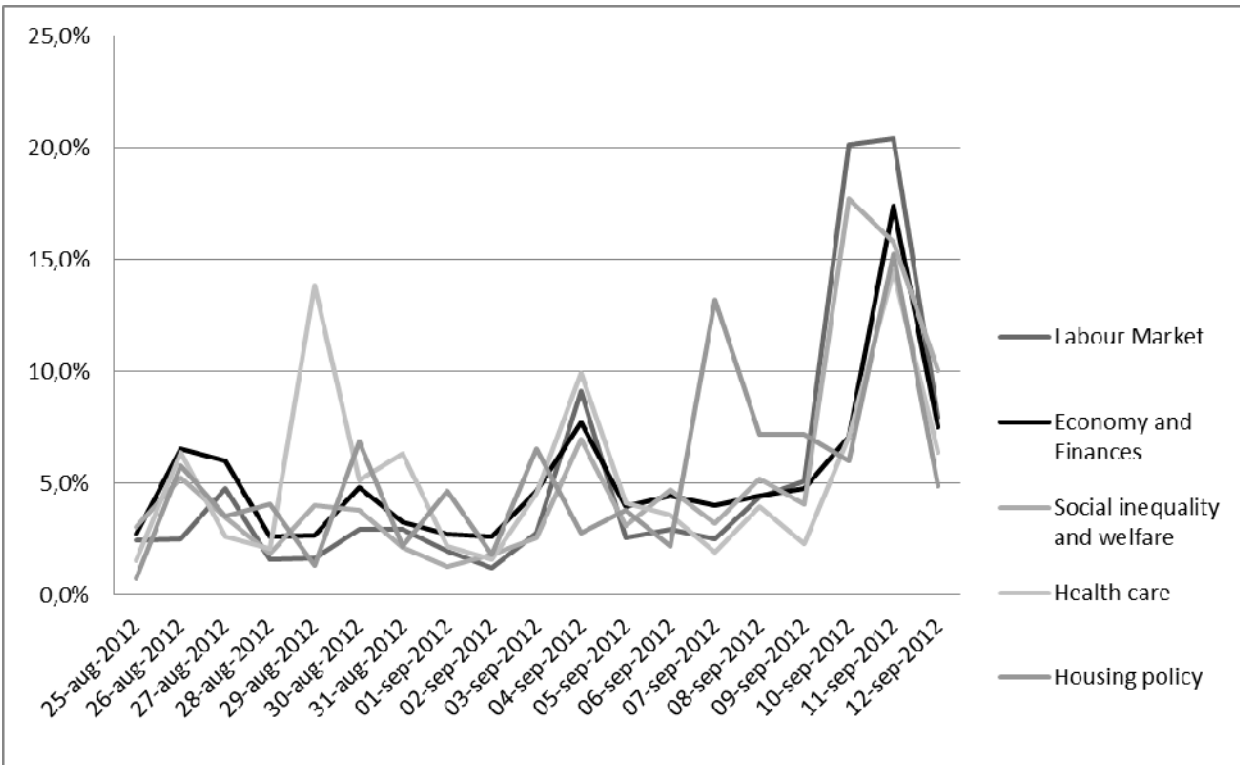


Figure 3: Trends in ownership of traditional left-right issues among three biggest parties

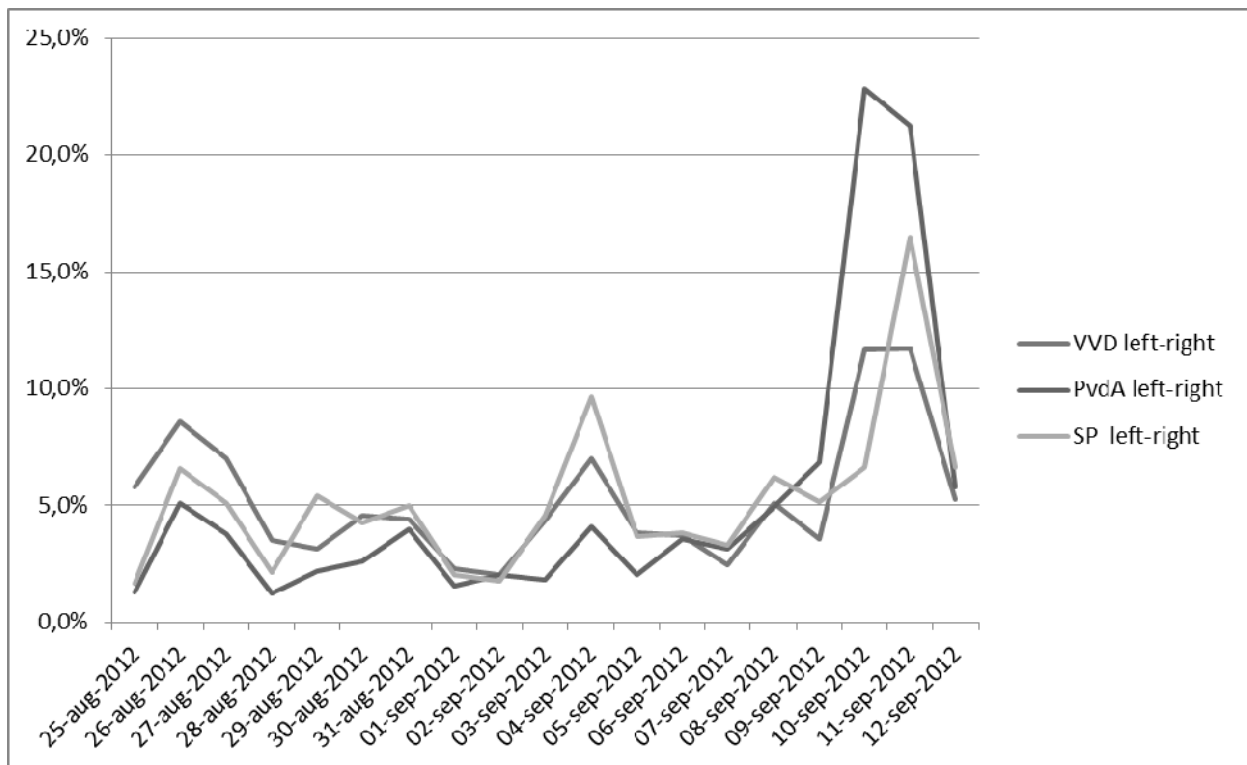
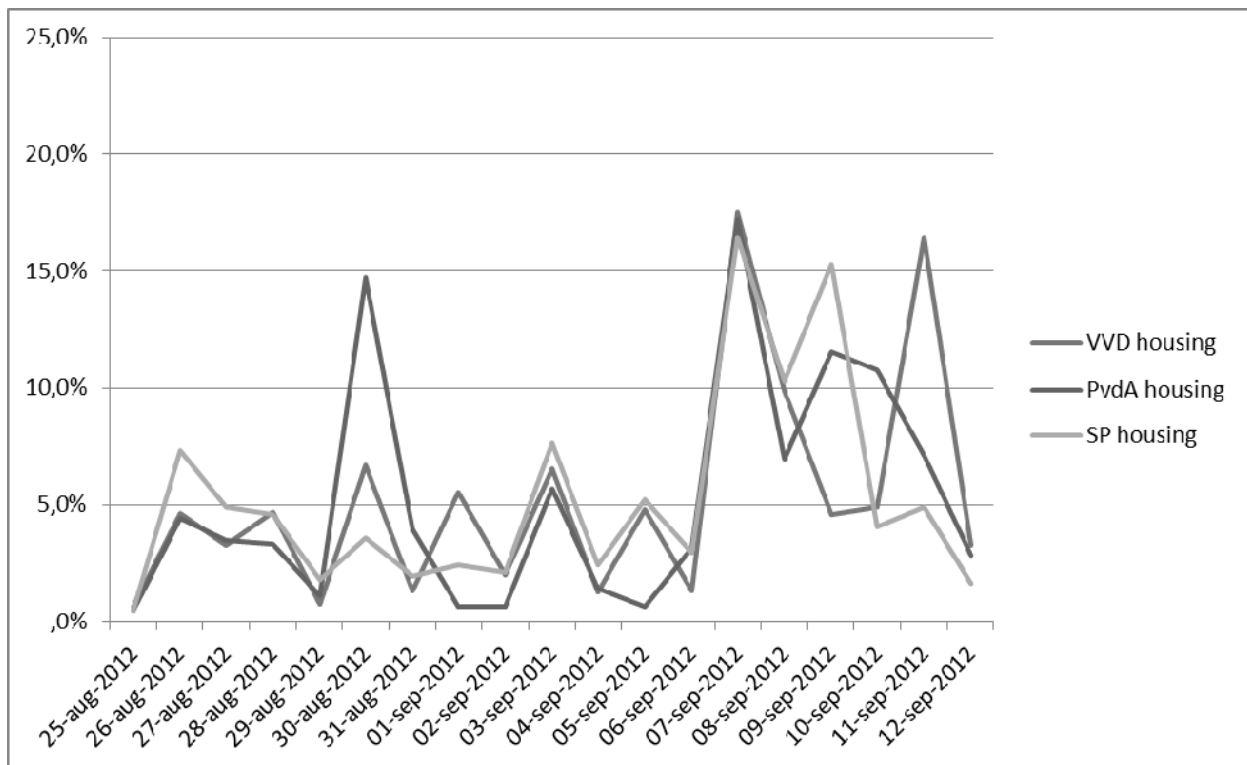


Figure 4: Trends in ownership of housing issues among three biggest parties



Tables

Table 1: Issue ownership in DPES 2012 and Twitter data

	Labour market		Economy & Finance		Social inequality & welfare state		Health Care		Housing policy	
	DPES	Twitter	DPES	Twitter	DPES	Twitter	DPES	Twitter	DPES	Twitter
CDA(Christian democrat party)	1,9	6,7	2,3	9,6	1,1	7,1	0	11,0	0	7,4
PvdA (Labour party)	34,6	41,4	12,8	26,7	40,4	28,2	29,2	19,2	3,6	22,4
VVD (Liberal party)	21,2	31,6	70,9	33,3	14,9	39,2	4,2	23,9	67,9	56,5
GL (Green party)	0	5,3	1,2	9,7	4,3	10,3	2,1	6,3	0	4,4
SP (Socialist party)	23,1	25,1	2,3	24,4	30,9	29,7	37,5	26,1	3,6	21,8
D66 (Social-liberal party)	5,8	10,8	5,8	12,2	5,3	7,3	18,8	10,1	17,9	11,8
CU (Christian Union party)	0	2,8	2,3	4,0	2,1	2,8	0	5,2	3,6	4,1
SGP (Christian reformed party)	0	0,7	1,2	1,6	0	1,0	0	1,2	0	0,9
PVV (Wilders' freedom party)	11,5	6,0	0	8,5	1,1	4,9	6,2	10,5	3,6	7,3
PvdD (Animal welfare party)	0	0,7	1,2	1,6	0	1,5	2,1	1,2	0	0,9
50Plus (Fifty plus party)	1,9	0,4	0	0,3	0	0,2	0	0,3	0	0,2
Total N	52	15160	86	13617	94	5929	48	12586	28	2820

DPES: Entries are percentages of party votes as reported by respondents within the category of issues as reason to vote for party

Twitter: Entries are percentages of party mentions within the group of tweet posts on each respective issue

Appendix A: Measures of issue salience and issue ownership in DPES 2012

Issue salience is measured by the following open-ended question: “What do you think are the most important problems in our country?”. The answers have been coded into nine categorical variables including first to ninth answer, each featuring the issue categories displayed at the left-hand side of the table. The corresponding categories in the paper are indicated at the right-hand side of the table.

1 Economy/Financial situation	Economy and financial situation
2 Social security	Social inequality and welfare state
3 Politics	
4 Crime	
5 Defense	
6 Health care	Health care
7 Education	
8 Income/Price levels/Taxes	Economy and financial situation
9 Employment	Labour Market
10 Traffic/mobility	
11 Housing	Housing
12 Environment	
13 Population	
14 Minorities	
15 Norms and Values	
16 Media	
99 There are no problems	
999 DK/NA/Cannot be coded	

Issue ownership is measured by combining the following variables:

a) Party voted for in 2012 parliamentary elections

1 CDA	10 Partij voor de Dieren
2 PvdA	12 50Plus
3 VVD	13 Other party
4 GroenLinks	14 Blanco
5 SP	995 INAP
6 D66	998 NA/Refusal
7 ChristenUnie	999 Don't know
8 SGP	
9 Partij voor de Vrijheid	

b) Reason for party choice (string variable) – the answers have been coded manually into the issue categories labour market, economy and financial situation, social inequality and welfare state, health care, and housing.

Appendix B: Coding Scheme for Quantifying Issues and Parties in Twitter Data

Issue Family	List of word stems
Labour market	AOW, Arbeid, arboregel, arbowet, Baan, Banen, #CAO, flexwerk, forensen, inkomstenbelasting, loonbelasting, minimumloon, ontslagrecht, pensioen, reiskosten, salaris, uvw, vakantiegeld, vakbond, verlof, wajong, wao, werkelo, werkend, werkgelegen, werkgever, werklo, werknemer, werktijd, wia, ww-er, ww-uitkering, zelfstandigen, ziekwet, ZZP
Economy and Finance	Banken, bankier, bankrun, belasting, beurs,bedrijf, bedrijven, begroting, bezuinig, ondernemer, economi, recessie, staatsschuld, fiscale, heffing, krediet, btw, spaard, spaargeld, koopkracht, accijn, mkb, bonuscultuur, #BNP
Income differences and welfare state	Sociaal, inkomen, bijstand, uitkering, welvaart, draagkracht, armoede, #eerlijkdelen, herverdeling, kinderbijslag, bestaansminimum, bestaanszekerheid, middenstand
(Health) Care	Zorg, ouderen, patiënt, patient, eigenrisico, eigenbijdrag, marktwerk, awbz, huisarts, verpleeg, verpleging, medicijn, medicatie, apotheker, geneesmiddel, zieken, ziekte
Housing	woning, hypotheek, hypotheken, huurhuis, huurprij, huurmarkt, huurverhoging, huurtarie, huurstijging, huursubsidie, huurtoeslag, huursector, huurder, overdrachtsbelasting, makelaar, huisvesting, bouwsparen, bouwbedrij, bouwsector, bouwcorporatie, huizenkoper, huizenmarkt, huizenbezitter, huizenprij, huisbezitter, huiseigena, huisprijs
Parties and Leaders	List of word stems
VVD	VVD, rutte , zekernu, nietdoorschui
PvdA	PvdA, partijvande arbeid, Samsom, Samson, Diederik, nlsterk
PVV	Pvv, partijvoordevrijheid, geert, wilders
CDA	Cda, sybrand, buma, haersma, samenkunn, nieuwemoraal, bouwenaande
SP	Sp, emile, roemer, emiel, stemsociaal, armoedewerktniet
D66	d66, alexander, pecht, nuvooruit
GroenLinks	GL, Groenlinks, jolande, jolanda, #sap , detijdisnu, #4dw, 4dagenwakker, zinindetoe kom

Christen Unie	#CU, Christenunie, slob, voordeverand
SGP	Sgp, staaij, staay, daadbijhet
PvdD	pvdd, partijvddieren, partijvoordedier, thieme, houvastanjeide
50 Plus	50plus, #krol, #henkkrol

Appendix C: Issue salience in DPES 2012

	First issue	Second issue	Third issue	Fourth issue	Fifth issue	Sixth issue	Seventh issue	Eighth issue	Ninth issue	Total (%)
Labour Market	181 (10,9%)	115 (8,9%)	53 (7,1%)	22 (8,1%)	1 (1,2%)	0	0	0	0	372
Economy and Finances*	712 (42,9%)	418 (32,5%)	227 (30,3%)	74 (27,2%)	16 (19,5%)	10 (47,6%)	1 (14,3%)	2 (66,7%)	1 (50,0%)	1461
Social inequality and welfare state	80 (4,8%)	112 (8,7%)	66 (8,8%)	18 (6,6%)	4 (4,9%)	2 (9,5%)	1 (14,3%)	0	0	283
Health care	184 (11,1%)	104 (8,1%)	60 (8,0%)	14 (5,1%)	12 (14,6%)	2 (9,5%)	0	0	0	376
Housing policy	43 (2,6%)	57 (4,4%)	44 (5,9%)	22 (8,1%)	4 (4,9%)	1 (4,8%)	0	0	0	171
N	1659	1286	748	272	82	21	7	3	2	

*Economy and Finances consists of the addition of two different categories in the DPES: Economy/financial situation and income/price levels/taxes