ecpr PRESS

New Perspectives on Negative Campaigning

Why Attack Politics Matters

Edited by Alessandro Nai and
Annemarie S. Walter

ecpr PRESS

ECPR Press Series Editors:
Dario Castiglione (University of Exeter)
Peter Kennealy (European University Institute)
Alexandra Segerberg (Stockholm University)

More from the ECPR Press Studies in Political Science Series

Please visit http://www.ecpr.eu/ecprpress for information about new publications.

Chapter Three

What is Negative about Negative Advertisements?

Barbara Allen and Daniel Stevens

Research on 'negativity' must first define the meaning of the term. Common definitions in political science and political communication, as reflected in this volume, include Lau and Pomper's (2000: 2) description of, 'talking about the opponent – his or her programmes, accomplishments, qualifications, associates, and so on – with the focus, usually, on the defects of these attributes,' Geer's (2006: 29) of a negative appeal as, 'any criticism or reason to vote against the opposition,' and Benoit's of 'attacks' as criticisms of the opponent (*see* Chapter Two in this volume). Although, each of these definitions pertains to research on American political advertising, they are – and have been – easily applied to other communication mediums such as debates or newspaper coverage (Ridout and Franz 2008), and to advertising and media in other countries such as Denmark (Elmelund-Præstekær 2010), the Netherlands (Walter and Vliegenthart 2010), Taiwan (Sullivan and Sapir 2012a), and the United Kingdom (UK) (Sanders and Norris 2005; van Heerde-Hudson 2011). They indicate widespread agreement on what we mean by negativity.[1]

Operationalising negativity is more complex, however. Whereas there is broad consensus on what negativity is, there is variation in the way negativity is identified and categorised. Some of this variation pertains to what is analysed: using the example of advertising that will be the focus of this chapter, some research examines a sample of advertisements – 'prominent' advertisements for example (West 2005), while other studies look at all the advertisements in an election; some research looks at advertisements made (Diamond and Bates 1992), others at advertisements aired (Goldstein and Freedman 2002). But additional variation stems from how exactly to code the 'tone' of an advertisement. For example, researchers have dealt differently with advertisements that combine positive and negative messages, some regarding advertisements with *any* negative content as 'negative,' while others view messages that combine negative and positive as 'comparative' or 'contrast.' Some research has recognised the distinction between negative and contrast advertisements but ultimately combined them on the grounds

---

1. Academic researchers now appear to agree that definitions of negativity should not include an evaluative dimension, i.e. that 'negativity' is also bad or undesirable, or that positivity is good or uplifting. These are considered to be empirical issues rather than matters of definition or judgement.

that they are 'functionally equivalent' (Goldstein and Freedman 2002), while other research argues that they are very different (Jamieson 2000). In addition, for some the unit of analysis is the advertisement (e.g. Goldstein and Freedman 2002), while others code discrete appeals (Geer 2006) or 'idea units' (Jamieson 2000), within an advertisement, and make further distinctions between issue and traits. Studying the tone of a newspaper article or a debate raises similar quandaries: does the researcher focus on sentences, paragraphs, the tone of an entire article in a newspaper or an entire answer in a debate?

It is this variation – in the operationalisation of tone – that is the subject of this chapter. Other studies have examined variation in tone *across* different media to see whether they provide different impressions of the negativity of a campaign, e.g. local news versus advertising (Franz and Ridout 2008). But this leaves a more fundamental question unanswered: does variation in the operationalisation of negativity *within* a medium lead us to different conclusions? We present the first study, to our knowledge, that takes the same universe of messages – political advertisements in the 2008 campaign – and pits different coding schemes against each other *using the same coders* (other comparisons have used secondary data) to see whether they provide different pictures of campaign negativity and its effects. The short answer is that while different coding schemes by and large converge on the same findings our study indicates that researchers need to be careful about the potential consequences of different choices.

Measuring 'negativity' in advertising

How has research conceived of and measured negativity in political advertising? The answer to these questions may matter for two reasons. First, the perception of trends in the use of negativity may depend on whether a single disparaging statement about one's opponent is seen as making an entire advertisement negative or whether the majority of the advertisement must be in this vein for it to be negative. Second, definitions of negativity make psychological assumptions about the individual voter, either explicit or implicit, that could be consequential for our understanding of the impact of tone.

To illustrate, we will start with trends in presidential elections and the United States (US). For West (2005), an advertisement is negative if it contains *any* pejorative statement about the opponent. By this measure the vast majority of prominent advertisements made in the 1988 presidential election, 83 per cent, or about five in every six, were negative. While the year 1968 was very negative, it was followed by a steep decline in 1972 and 1976. The 1980s witnessed an upward trend, with 1984 and 1988 the most negative of all, but in the 1990s there was some decline (59 per cent). In earlier work, Jamieson calls an advertisement negative if the *majority* of the advertisement consists of such information. By her reckoning (1992: 270) presidential campaigns from 1952 to 1976 were very similar in terms of negativity, 1980 was notably negative, and 1984 and 1988 were closer to previous campaigns, though marginally

more negative. But in her later analysis, Jamieson (2000) divides the discourse contained in political spots into 'advocacy,' which promotes the sponsor's position, 'attack,' which focuses on criticism of the opponent, and 'contrast,' in which the two positions are compared. By this measure, 1952 becomes the most attack-oriented campaign. While there has been something of an upward trend since 1976, no recent election has been as negative. In an ostensibly similar measure to Jamieson's 1992 standard, Kaid and Johnston (1991) code by the positive or negative '*focus*' of the advertisement – sponsor or opponent. According to their analysis of presidential elections from 1960 to 1988, 1964 was the most negative campaign of all, and the trend has been slightly upward since 1968, without again reaching the level of 1964. They classify 37 per cent of advertisements in 1988 as negative compared to West's 83 per cent. Kaid and Johnston's (2001) later study, which includes 1992 and 1996, shows these two races to be the most negative campaigns of all. Finally, moving away from the advertisement or spot as the unit of analysis, Finkel and Geer (1998) and Geer (2006) break advertisements down into 'appeals.' For example, an advertisement might include appeals related to a candidate's record on inflation or defence, or be about moral values. They then code individual appeals as negative or positive, calculate the average proportion of negative and positive appeals from the candidates, and take the difference as the 'advertising tone' for that election. By this measure the 1964 election was also notably negative, and since 1976 there has been a steady decline in tone. The elections of 1984, 1988, 1992, and 1996 were all more negative in advertising tone than 1964.

In his work on negativity, Geer (2006) also compares many of these measures to his – West's, Jamieson's and Kaid and Johnston's – but adds Benoit's (1999). While Geer recognises the kinds of differences in the amounts of negativity *within* elections that these different coding schemes throw up, he points out that the trends in negativity they all show are similar and that the correlations between measures are mostly at 0.8 or above. Nevertheless, if one is interested in the effects of negativity within an election, as many analysts are, by West's reckoning prospective voters were barraged with negative advertising in 1988, at least in the 'prominent advertisements.' Yet according to Kaid and Johnston only one in three advertisements were negative that year. Jamieson finds no trend in the number of negative advertisements between 1952 and 1988 (1992, Appendix I Chart 4–3), whereas Geer indicates a steady decline in tone since 1976.

While they are illustrative of the potential consequences of duelling definitions of tone, in these examples it is not just the operationalisation of negativity that varies but also the advertisements that are coded from eras in which information about advertisements made versus advertisements aired is harder to obtain. We therefore do not know how much of the variation is due to different advertisements coded and how much is due to different ways of coding negativity. In the last few election cycles in the US, however, researchers have had access to Campaign Media Analysis Group (CMAG) data, which provide details of all the

presidential advertisements that were made and aired to the vast majority of the US population.[2] Electronic resources also make negativity in other media, such as transcripts of debates, newspaper content, even direct mail, much easier to obtain than in the past. The availability of such data reduces any problems related to coding of different samples of materials and means that most of the variation in researchers' findings about levels of negativity – if they were coding the same universe of messages – would stem different operationalisations of negativity.

By our reading, five methods of coding the tone of advertising messages have emerged over the last fifteen years in political science and political communication, two of which – the Geer and Jamieson's methods – appear in Geer's (2006) comparison:[3]

Goldstein and Freedman method

Goldstein (1997) and later Goldstein and Freedman (2000; 2002) pioneered the use of CMAG data. They combined these data on advertisements made and aired with the Wisconsin Advertising Project's method of coding the tone of advertisements (which has remained unchanged now that the research has moved to the Wesleyan Media Project). Coders are asked: 'In your judgment, is the primary purpose of the advertisement to promote a specific candidate, to attack a candidate or to contrast the candidates?' Possible responses are 'attack,' (negative), 'promote' (positive), and 'contrast.' Thus the unit of analysis in this method is the advertisement and advertisement tone has three categories.

Freedman and Goldstein method

In another paper, Freedman and Goldstein (1999) describe a slightly different method, using a five-point scale. The variation is in how they deal with contrast advertisements (*see* their Appendix A). Advertisements with positive *and* negative appeals were split into three further categories of 'balanced' (given a '3' on the scale), 'predominantly negative, with a token mention of the sponsoring candidate' (a '4'), and, by implication – the third category is not explicitly mentioned in the article – more positive than negative ('2'). For the purposes of analysis, however, they took the average score given by multiple coders on their five-point scale, rounded to the nearest integer. They categorised only the middle category ('3') as a contrast advertisement, with a '1' or '2' on the five-point scale being a positive advertisement and a '4' or '5' a negative advertisement. Thus, the unit of analysis in this method is the advertisement and advertisement tone has five categories that are then collapsed into three.

2.   Jamieson's (2000) had access to television station logs, providing her with information on when and where advertisements aired.

3.   Benoit's functional analysis of advertisements, described in Chapter Two, has a longer pedigree. More importantly, it also correlates almost perfectly with Geer's measure (Geer 2006: 37). Thus, our findings with regard to Geer's method can be extrapolated to Benoit's.

Kahn and Kenney method

Kahn and Kenney (1999: 879–880) describe their approach to categorising tone as follows: 'we estimated the amount of negative information in each advertisement, including criticism of the opponent (i.e. negative information about issues or personality characteristics). We placed each commercial into one of three categories: no negative message (score=0), a minor emphasis on negativity (score=1), and a major emphasis on negative (score=2).'

Kahn and Kenney's method deals differently again with advertisements that mix positive and negative appeals: if the emphasis on negativity is 'minor' it is placed in the second category but if it is 'major' the advertisement is placed in the same category as advertisements that are purely negative. The unit of analysis in this method is the advertisement and advertisement tone has three categories.

Jamieson method

Jamieson (2000) breaks advertisements down into 'idea units' that pertain to issues or candidates' traits. She then codes the purpose of these idea units as 'attack' or 'advocacy.' Advertisements are given an 'attack score' based on the total number of words in the idea units categorised as attack as a proportion of the total number of words in the advertisement. Idea units that are about values or that are tag lines such as, 'Concerned about Barack Obama's naive foreign policy? You should be.' are excluded from the denominator. In Jamieson's (2000: 113) analysis, advertisements that contain more than 90 per cent attack words are attack advertisements, 30 per cent to 70 per cent attack words are contrast, and advertisements that are less than 30 per cent attack words are positive. The 71 per cent to 90 per cent range remains ambiguous in Jamieson's study – perhaps she found no advertisements with scores in this range. Thus, for Jamieson the unit of analysis is ultimately the advertisement and advertisement tone has three categories, but those categories are the result of the aggregation of idea units.[4]

Geer method

Geer's (2006) approach is similar to Jamieson's in that he divides advertisements into distinct 'appeals' that are directed at issues, traits or values that are positive or negative in tone. But Geer differs from Jamieson by not aggregating up from separate appeals to categorising the tone of the entire advertisement. He argues that there is so much variation within the category of contrast advertisements that

---

4. In her analysis of the impact of advertisements, Jamieson calculates an 'attack weight' for each advertisement, based on multiplying the attack score by the length of the advertisement in seconds and a measure of exposure or reach within each television market in the US calculated from Gross Ratings Points (GRPs). For our purposes, however, it is her operationalisation of tone that is important.

aggregation could mislead us about the overall tone of an ad campaign: 'These data underscore the advantage of a more precise measure for assessing the overall content of political advertising' (2006: 35). Thus, in the Geer method the unit of analysis is the appeal and issue, trait or value appeals are either positive or negative in tone.[5]

Not only are there subtle variations in these methods regarding what negativity is but also, our second point, in terms of how their protagonists think negativity in political advertising is likely to affect individuals. On the one hand, if the vast majority of an advertisement must be negative to be classified as a negative advertisement, as in Jamieson's scheme, the claim is that only when the scales tip very firmly toward negativity is an individual affected by 'negativity.' Advertisements that are predominantly negative in content are apparently experienced differently. Kahn and Kenney's most extreme category of advertisements that have a 'major emphasis' on negativity appears to make a similar assumption. Certainly, from Jamieson's method, the implication is that advertisements in which there is a mixture of positive and negative content, even if the negative messages outweigh the positive messages two to one, are not judged as harshly. This suggests that positive information is given more weight than negative information in that a little positive content dilutes a lot of negative content – 'Pundits and scholars who collapse all attack advertisements into the catchall category 'negative campaigning' are treating the complex world of political advertisements in a simplistic manner that is not shared by the citizenry at large' (Jamieson 2000: 79). Indeed, Jamieson's analysis of the 1996 presidential election indicates that contrast advertisements had a positive effect on turnout and negative or attack advertisements had a smaller negative effect. Geer's measure, on the other hand, gives negative and positive information equal weight, which implies that this is also what individuals do. Geer is also explicit in the quote above that treating advertisements as the unit of analysis rather than appeals could lead to misleading conclusions about the negativity of campaigns and thus, by implication, to the effects of negativity on individuals.

Studies of negativity in campaigns outside the US, such as some of those in this volume, face the same dilemmas about how to code the negativity of other media such as posters, press releases, or debate statements. For example, is a predominantly negative statement about an opponent in a debate that also includes some comparison categorically different from a statement that does not make such a comparison? The answers to such questions matter to our understanding of negativity. We ask two questions in this chapter: 1) Do differences in the operationalisation of the coding of negativity lead to different conclusions about the negativity of a campaign? 2) Do variations in the operationalisation of the coding of negativity lead to different conclusions about the relationship between the negativity of a campaign and dependent variables such as turnout?

---

5. Value appeals are similar to Benoit's (Chapter Two in this volume) category of messages about 'goals and ideals.'

Data

We trained a team of undergraduates to watch and code all of the presidential advertisements from the 2008 campaign in the US, along with advertisements for the US Senate and US congressional races in Minnesota. Our universe of advertisements included advertisements aired on behalf of the candidates and covered the networks and cable television as well as radio. These were provided to us by Video Monitoring Service (VMS[6]), a US company that offers media intelligence to its clients. Thus, all the advertisements we analyse here were aired in these races – we are not mixing 'advertisements aired' and 'advertisements made.' In total, we analysed the tone of 531 unique advertisements[7] that aired more than 700,000 times in 2008. Our team coded the negativity of all 531 advertisements using each of the five methods – Goldstein and Freedman, Freedman and Goldstein, Kahn and Kenney, Jamieson, and Geer – described above.[8] Coders were blind to the purpose of this study.[9]

As we outlined above, previous research that has focused on similar questions to ours has either compared negativity across different media with different coders (e.g. Franz and Ridout 2008) or has compared the coding of negativity for the same medium, advertising, but with different samples of advertisements and different coders (Geer 2006). This does not allow us to separate variation that is due to different media and coders from variation that derives from different methods of operationalising negativity. By keeping the medium constant and using the same coders our study allows us to isolate variation due to the operationalisation of negativity, which we have argued is a more fundamental question than variation by medium: before we compare negativity across different media we should be sure that researchers are not drawing very different conclusions about negativity and its effects because of different conceptions of negativity *within* the same medium.

Analysis

Figure 3.1 presents our first look at the levels of negativity in advertisements in 2008 according to the five measures. The unit of analysis is the advertisement,

---

6. http://www.vmsinfo.com/
7. The term 'unique' refers to the fact that some advertisements are essentially identical but may mention a different state, depending on where they air, or the text may be the same but in Spanish rather than English. Like Geer (2006), when we examine the tone of the advertisements made in the campaign, we do not include such advertisements more than once but count them as the same advertisement in order to avoid biasing our results. However, when we look at the campaign in terms of advertisements aired later in the chapter, we include unique and 'duplicate' advertisements from the presidential race – 654 in total.
8. An additional trained coder examined a subsample of forty-five of the advertisements and coded tone independently. Krippendorff's Alphas for intercoder reliability were above 0.8 for the Jamieson, Geer, Goldstein and Freedman, and Freedman and Goldstein measures, and 0.78 for the Kahn and Kenney method. These levels of agreement are more than sufficient for our purposes.
9. Indeed, the main purpose of the study was to assess the truth of the claims made in political advertising; coders had no motivation to try to make different coding methods 'agree' because they were unaware of how the data on tone would be used.

Figure 3.1: The tone of advertisements in the US in 2008 according to five different methods of coding negativity (no.=531)
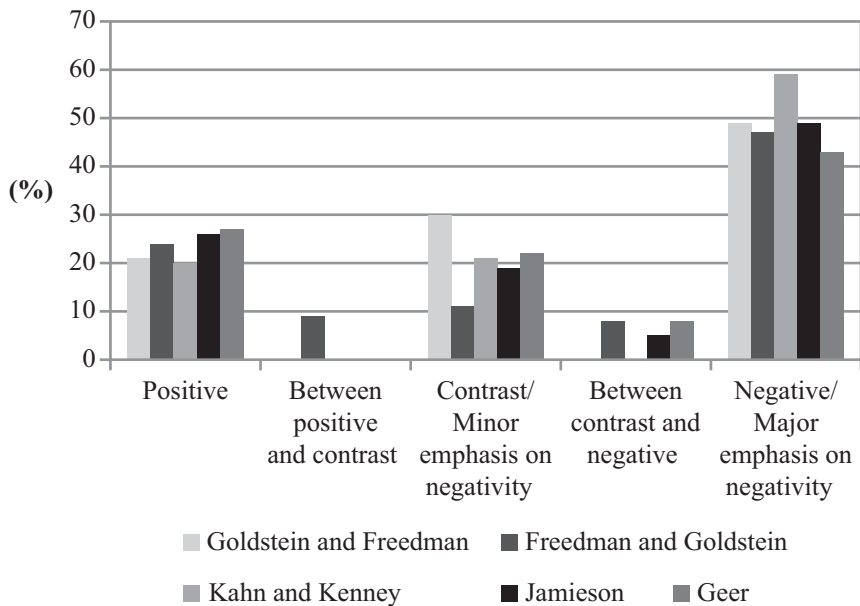
Notes: Goldstein and Freedman = Goldstein and Freedman's three category measure; Freedman and Goldstein = Freedman and Goldstein's five-category measure; Kahn and Kenney = Kahn and Kenney's three-category measure; Jamieson = Jamieson's 'attack score' approach, where advertisements with 0 per cent – 29 per cent attack words are positive or advocacy, 30 per cent – 70 per cent are contrast, 71 per cent – 90 per cent are between contrast and attack, and 91 per cent plus are negative. Geer = Geer's proportion of negative appeals in each advertisement, categorised using the same system as Jamieson's. Where there are no bars for a category it is because the category is excluded from an author's coding scheme.

which requires some additional explanation. First, we noted above that Jamieson's method leaves ambiguous classification of advertisements with 71 per cent to 90 per cent attack words; we put these advertisements into a fourth, 'Between contrast and negative' category in Figure 1, equivalent to Freedman and Goldstein's 'predominantly negative with token mention of the sponsor' category. Second, although Geer's unit of analysis is the appeal, for the purposes of Figure 3.1 we classified advertisements into four categories from Geer's method – positive, contrast, between contrast and negative, and negative – using the same thresholds as for Jamieson's method, e.g. 90+ per cent negativity constitutes a negative advertisement – but based on the proportion of negative appeals in the advertisement rather than the proportion of attack words. Third, for the Freedman and Goldstein method we present the full five categories in

Figure 3.1 for informational purposes, although we reiterate that they ultimately collapses these five categories into three.

Figure 3.1 shows that there is a great deal of overlap between the different methods. They converge, for example, on indicating that between 20 per cent and 27 per cent of the advertisements in 2008 were positive. There is more discrepancy at the other extreme, negative advertisements or advertisements with a major emphasis on negativity, but that is largely because the Kahn and Kenney method places more advertisements in this category; all the other measures put the proportion of negative advertisements between 43 per cent and 49 per cent. Kahn and Kenney's method, on the other hand, identifies 59 per cent of advertisements as belonging in this category. It seems that their criteria, in which negative advertisements are divided according to minor or major emphasis on negativity, places more advertisements in the extreme category than others. Indeed, if we count the 'between contrast and negative advertisement' category, which three of the methods include but Kahn and Kenney do not, as 'negative' the range across the five methods narrows to between 49 per cent and 59 per cent negative, with all but Goldstein and Freedman's indicating that a majority of the advertisements belonged in this category.

Given these results, it is not surprising that there is also a high level of agreement about contrast advertisements. If we put Freedman and Goldstein's 'between positive and contrast advertisements' – 9 per cent of the total – into the contrast advertisement category, on the grounds that they are not purely positive advertisements – four of the five methods indicate the proportion of contrast advertisements to be between 19 per cent and 22 per cent. The outlier here is the Goldstein and Freedman method, in which 30 per cent of advertisements appear in the contrast category. The reason seems to be that their definition of a positive advertisement is somewhat tighter than, for example, Jamieson's, for whom only when more than 30 per cent of the words in an advertisement are negative does the advertisement move out of the 'positive' category.

In summary, Figure 3.1 indicates high levels of agreement across different methods, with discrepancies in proportions of advertisements placed in the categories of 'positive,' 'contrast' or 'negative' advertisements no greater than about 10 per cent and often much less. All methods indicate that roughly 50 per cent of the advertising in 2008 was negative, with the remaining 50 per cent of advertisements being split fairly evenly between positive and contrast advertising. To put these kinds of differences into perspective, in Geer's (2006) comparison of different methods of coding negativity for the 1960–96 elections, even if we exclude West there were discrepancies of more than 10 per cent between methods for the 1964, 1968, 1972, 1976, 1980, 1984 and 1996 elections. Thus, maximum differences of 10 per cent do not appear large.

Table 3.1 presents additional evidence based on the correlations between measures of levels of negativity using the different methods. In Table 3.1, rather than collapsing the Jamieson attack score and Geer's appeal measures into three categories we allow them to vary over their full range from zero to 100 for each advertisement, i.e. Jamieson's attack score and Geer's proportion of negative appeals

Table 3.1: Correlations between five different methods of measuring negativity (no.=531)

	Goldstein and Freedman	Freedman and Goldstein	Kahn and Kenney	Jamieson	Geer
Goldstein and Freedman	1				
Freedman and Goldstein	0.93	1			
Kahn and Kenney	0.89	0.93	1		
Jamieson	0.88	0.91	0.88	1	
Geer	0.87	0.91	0.88	0.92	1

for each advertisement. The correlations are universally high, with the *lowest* being the 0.87 correlation between the Goldstein and Freedman and Geer methods, which is still a very healthy correlation by any standard. Similarly, a factor analysis of the five methods shows only one very strong factor, with the factor loadings for all five methods greater than 0.9. We also examined different subsets of the data, such as just presidential advertisements, Republican versus Democrat advertisements, radio versus television advertisements, television advertisements that used more visual 'cuts' (to see whether more visual activity in an advertisement might open up greater disagreement between coders), and advertisements with more traits in them (perhaps the tone of issue oriented advertisements is easier to code). No matter how we sliced the data, the correlations between methods remained in excess of 0.8.

Thus, all of the evidence we have presented so far indicates that conclusions about the prevalence of negativity in a campaign would be only marginally affected by decisions like when the mixture of positive and negative appeals makes an ad 'negative' rather than 'contrast' and by the unit of analysis. This still does not quite answer the question of whether and why there might be disagreement about the tone of particular advertisements, however. We therefore examined the coding of specific advertisements among the 531. We would not expect to find instances in which one method describes an advertisement as positive and another as negative – and we do not. As Figure 1 suggested, where we see most instances of disagreement is with advertisements that mix negative and positive messages. For example, an advertisement from the Senate race in Minnesota received an attack score of twenty-one by the Jamieson method and about 13 per cent of the appeals were negative according to the Geer method, putting it in the advocacy or positive advertisement category, where it also belongs according to the Freedman and Goldstein method. Coding by the Kahn and Kenney method, however, places the same advertisement in their middle category of 'minor emphasis on negativity' and using the Goldstein and Freedman method made it a contrast advertisement according to our coders. In a few other instances, differences appear to arise between the methods that treat the advertisement as the unit of analysis and those that break them down into ideas or appeals. For example,

when our coders broke down a National Rifle Association advertisement critical of Barack Obama on gun control into idea units for the Jamieson method, or into appeals for the Geer method, they gave it a Jamieson attack score of fifty, while the proportion of negative appeals by the Geer method was 62.5. This renders these advertisements contrast. But using the other three methods, where the unit of analysis is the advertisement, our coders put the same advertisement in the most negative category. The reason appears to be because the first half of the advertisement told a story about gun ownership, culminating in the line that, 'You might call it an heirloom, or just a wall hanger,' before going on to excoriate Obama for apparently calling 'it' an assault weapon. When the unit of analysis was the advertisement, coders did not see the first half of the advertisement as positive, but by the Jamieson and Geer methods it diluted the negativity of the ideas or appeals in the advertisement such as to put it in the contrast category.

Such differences are, however, both fairly limited and rare in our data. Nevertheless, so far we have only looked at the first of our two questions, differences in the coding of the negativity of advertisements made. Researchers are also usually interested in our second question, the effects of advertisements aired. It may be that when we look at negativity in this way small differences in the coding of individual advertisements become larger differences when we are trying to account for advertisements that often air thousands of times. Similarly, content analysis of negativity in debates or media coverage will code hundreds of cases, meaning that small but systematic differences in the coding of individual cases may become more consequential.

We examine this by looking at the relationship between negativity in advertisements aired in the presidential race and turnout in the 2008 presidential election according to the five methods. VMS provided us with data on how many times each presidential advertisement aired in 103 television markets in the US – the largest one hundred and three others – which cover more than 80 per cent of the population (the US is divided into 210 television markets, or 'Designated Market Area(s)' (DMA)). DMAs are smaller than states – Texas alone has seventeen – but they are larger than counties. We used census data on turnout by county and then aggregated up the counties in each DMA to obtain turnout in each DMA. It would be a mistake to simply use this turnout figure, however, because some areas will usually have higher turnout at elections than other areas, regardless of advertising; what we want to examine is the relationship between negativity and deviations from 'normal turnout' in a presidential election. Our dependent variable is therefore the per cent deviation from average turnout in the DMA over the five previous presidential elections (1988–2004).

For our estimates of negativity, knowing the tone of the advertisement according to each method and how often it aired allows us to calculate the five measures of the proportion of negative, contrast, and positive advertisements aired in each DMA. The Goldstein and Freedman method uses these exact categories. For the Freedman and Goldstein method, we categorised advertisements 'between positive and contrast' (*see* Figure 1) as positive and advertisements 'between contrast and negative' as contrast; for the Kahn and Kenney method, advertisements with no negativity were categorised as positive, advertisements with minor emphasis on negativity as contrast, and advertisements with major

Table 3.2: Estimated relationships between negativity and turnout according to five different methods (no.=103)

Negative advertisements

	Goldstein and Freedman	Freedman and Goldstein	Kahn and Kenney	Jamieson	Geer
Estimated relationship with turnout	0.15 (0.06)*	0.13 (0.06)*	−0.09 (0.07)	0.13 (0.06)*	0.16 (0.07)*

Contrast advertisements

	Goldstein and Freedman	Freedman and Goldstein	Kahn and Kenney	Jamieson	Geer
Estimated relationship with turnout	−0.24 (0.07)*	−0.19 (0.08)*	0.26 (0.12)*	−0.26 (0.08)*	−0.26 (0.08)*

Notes: *p<0.05 #p<0.10 (two-tailed test). Standard errors in parentheses. Models include controls for per cent black, per cent Hispanic, per cent in poverty, per cent in the labour force, and per cent with a higher education degree, according to census data, as well as the average age of adults, and the 'normal' Democrat and Republican votes in the DMA (average vote in presidential elections from 1988–2004).

emphasis on negativity as negative; with the Jamieson and Geer methods we also categorised advertisements 'between positive and contrast' as contrast. Thus for purposes of comparison we use the Geer method, in which the unit of analysis is the appeal, to categorise advertisements (employing the same criteria described in Figure 3.1).

Table 3.2 presents the results of separate regression models in which we first examined the relationship between the proportion of negative advertisements aired and turnout and then the relationship between the proportion of contrast advertisements aired and turnout according to each measure. The models control for the proportion of voting age adults in the DMA that are black, Hispanic, in poverty, in the labour force, and have a higher education degree, according to census data, as well as the average age of adults, and the 'normal' Democrat and Republican vote, also based on the previous five elections. We do not present those estimates here; our focus is solely on the estimated relationships between negativity and turnout.[10]

Four of the five methods provide very similar estimates of the effects of negativity, showing a statistically significant positive relationship between the proportion of negative advertisements and turnout in a DMA, ranging from 0.13 to 0.16, and a statistically significant negative relationship between the

---

10. The adjusted R^2 in the models is between 0.41 and 0.48.

proportion of contrast advertisements and turnout in a DMA, ranging from −0.19 to −0.26. The outlier is Kahn and Kenney's measure, which indicates no relationship between negative advertising and turnout and a positive relationship between the proportion of contrast advertisements and turnout. On closer inspection, this appears to be because, as Figure 3.1 showed, Kahn and Kenney's method tends to put more advertisements with negative content into their most extreme category. In the Philadelphia market, for example, where all the other methods suggest that 34 per cent to 41 per cent of the advertisements were negative, the Kahn and Kenney method indicates 51 per cent. In the Charlotte market, while the other four measures indicate between 34 per cent and 44 per cent of the advertisements were negative the Kahn and Kenney method suggests 56 per cent. Indeed, in more than 75 per cent of the 103 markets the Kahn and Kenney method indicates higher levels of negative advertising than any other.

These differences can become particularly pronounced in markets where there was relatively little advertising. Thus, we repeated the analysis of Table 3.2 but divided the sample of DMAs into two – the fifty-two with the fewest advertisements aired and the fifty-one with the most advertisements aired – our reasoning being that if the different estimated relationships using the Kahn and Kenney method are due to large discrepancies in markets where few advertisements were aired, we should see similar estimated relationships when we focus on the markets with the most airings. Table 3.3 presents the results. Dividing the sample in this way limits sample size, as well as the variance in dependent and independent variables, and we are thus less likely to see statistically significant relationships, but our interest is also in the size and signs of the estimated relationships. Table 3.3 confirms that the discrepancy between the estimates using the Kahn and Kenney measure versus the other measures is particularly pronounced in markets in which fewer advertisements aired. The coefficients for the relationship between the proportion of negative advertisements and turnout in these markets are between 0.00 and −0.03 for the other four methods but it is −0.19 and statistically significant using the Kahn and Kenney method. For markets with more airings there is also some variation, with all the other methods indicating a negative relationship while the coefficient using the Kahn and Kenney is 0.01; nevertheless, none of the relationships is statistically significant, meaning we would not be led to different conclusions by using any of the five methods in this case. The results for contrast advertisements are similar in that all the measures other than Kahn and Kenney's indicate a negative relationship with turnout in markets with fewer airings of advertisements, with three of the four statistically significant at p<0.10 while the relationship using the Kahn and Kenney method is positive. In addition, none of the methods suggests a statistically significant relationship with turnout in markets where more advertisements aired – all the coefficients are smaller than their standard errors.

In sum, differences between the measures when we look at their relationships with turnout are largely confined to the Kahn and Kenney method; all the others

Table 3.3: Estimated relationships between negativity and turnout controlling for the number of advertisements aired

Negative advertisements

	Goldstein and Freedman	Freedman and Goldstein	Kahn and Kenney	Jamieson	Geer
Estimated relationship with turnout in 52 markets with lowest number of airings	0.00 (0.09)	−0.03 (0.08)	−.019 (0.07)*	−0.02 (0.09)	−0.01 (0.10)
Estimated relationship with turnout in 51 markets with highest number of airings	−0.17 (0.20)	−0.21 (0.19)	0.01 (0.24)	−0.15 (0.20)	−0.06 (0.22)

Contrast advertisements

	Goldstein and Freedman	Freedman and Goldstein	Kahn and Kenney	Jamieson	Geer
Estimated relationship with turnout in 52 markets with lowest number of airings	−0.17 (0.08)#	−0.08 (0.10)	0.20 (0.15)	−0.17 (0.10)#	−0.17 (0.10)#
Estimated relationship with turnout in 51 markets with highest number of airings	0.09 (0.25)	0.24 (0.30)	−0.23 (0.31)	0.13 (0.29)	−0.12 (0.30)

Notes: *p<0.05 #p<0.10 (two-tailed test). Standard errors in parentheses. Models include controls for per cent black, per cent Hispanic, per cent in poverty, per cent in the labour force, and per cent with a higher education degree, according to census data, as well as the average age of adults, the 'normal' Democrat and Republican votes in the DMA (average vote in presidential elections from 1988–2004).

tell a similar story. The reason appears to be because the tendency for the Kahn and Kenney method to code more advertisements in the most negative category can be particularly exaggerated in markets where few advertisements aired. In markets with more advertisements, we see some variation in the sign of the relationships across the five methods, but none of these relationships approaches statistical significance, i.e. all five methods lead to the same substantive inferences about the relationship between negativity and turnout.

Discussion and conclusion

This chapter has provided a unique examination of different measures of negativity in a campaign. We asked two questions at the outset: Do differences in the operationalisation of the coding of negativity lead to different conclusions about the negativity of a campaign; and do variations in the operationalisation of the coding of negativity lead to different conclusions about the relationship between the negativity of a campaign and dependent variables such as turnout? Our results are reassuring on the first question and suggest the need for caution with respect to the second. On the one hand, while we pointed to the fact that the five different methods we examined – what we have termed the Goldstein and Freedman, Freedman and Goldstein, Kahn and Kenney, Jamieson, and Geer measures – make slightly different assumptions about how negativity affects individuals, they correlate very highly – at 0.87 and above – when our interest is in levels of negativity in a campaign. Thus, if researchers' focus is on the tone of a campaign, or when and whether a candidate or party 'goes negative,' our study suggests that variation in what constitutes a 'negative' versus a mixed or 'contrast' message, or variation in the unit of analysis makes little difference. On the other hand, when we turned to an examination of the *effects* of negativity – in our case to the relationships the five methods imply between advertisements aired and turnout – these relatively minor differences in the coding of negativity in individual advertisements can become much greater. We observed that while four methods provided similar answers about the relationship between negativity and turnout, the Kahn and Kenney method, which lead to more advertisements being coded in the most extreme negative category, suggested quite different relationships, particularly in markets where few advertisements aired.

Such differences in measures of negativity may be less consequential in countries like the UK where very few advertisements are aired compared to the US, but when the medium is negativity in newspapers or debates we may see similar patterns: small differences in the coding of the negativity of individual articles, for example, may become more substantial if the researcher then creates a measure that multiplies the negativity of hundreds of articles by likely exposure to them. What is the solution? We offer two possibilities in concluding this chapter. One is that researchers adopt more than one method by which to code negativity and check to see whether analysis using different measures converges on the same answers. It seems particularly important to be cognisant of different ways of categorising messages that mix positive and negative information. A second way forward is that researchers are simply more aware of the assumptions that lie behind their coding of negativity and, at a minimum, justify them, though it would be preferable to actually test for the robustness of different categories of tone by conducting experiments.