The relationship between journalism and the state has been a central issue in studies of the news media in American society. It is, moreover, an issue that resonates with more general themes in social theory, in particular the concentration/diffusion of political power, and more broadly still the autonomy of societal institutions (Alexander 1981) or institutional fields (Benson and Nevu 2005; Bourdieu 1993, 1998). A frequent starting point for analyses is the model of the news media as an independent watchdog and counterweight to state power—a model that

When Does the Watchdog Bark?
Conditions of Aggressive Questioning in Presidential News Conferences

Steven E. Clayman
University of California, Los Angeles

John Heritage
University of California, Los Angeles

Marc N. Elliott
RAND Corporation

Laurie L. McDonald
RAND Corporation

In theories of the journalism-state relationship, the watchdog model of journalism competes with other models emphasizing either subservient or oppositional relations. Since actual journalistic practice is circumstantially variable, this study isolates the social conditions associated with aggressive journalism. Data are drawn from presidential news conferences from 1953 to 2000, and the focus is on the aggressiveness of the questions asked therein. Through multivariate models, four sets of explanatory conditions are explored: (1) the administration life cycle, (2) presidential popularity, (3) the state of the economy, and (4) foreign affairs. Results show (1) no evidence of a first-term honeymoon period, but significantly more aggressive questions during second terms, (2) the president’s Gallup job approval rating is not a significant independent predictor of aggressiveness, (3) both the unemployment rate and the prime interest rate are positively associated with aggressiveness; and (4) questions about foreign affairs are significantly less aggressive than questions about domestic affairs, and this differential has been stable for at least a half-century. We conclude by discussing the theoretical implications of these findings, which show that journalists modulate their conduct in complex ways that do not readily map onto any single model.

The relationship between journalism and the state has been a central issue in studies of the news media in American society. It is, moreover, an issue that resonates with more general themes in social theory, in particular the concentration/diffusion of political power, and more...
assumes a diffusion of power across highly differentiated political and journalistic institutions. There is widespread agreement that the watchdog model, although deeply engrained in journalistic culture as a normative ideal, is empirically inadequate. But there is disagreement over what should replace it.

One line of research casts U.S. journalism less as an independent watchdog and more as a subservient lapdog in its relationship to the state. Although reporters are formally autonomous, the practicalities of newsgathering make them dependent on government officials as authoritative sources of information and opinion (Cook 1994; Epstein 1973; Fishman 1980; Gans 1979; Molotch and Lester 1974; Sigal 1973; Tuchman 1978). Nonofficial sources and critical perspectives can be found within the news, but they are typically limited to the spectrum of opinion defined by the official mainstream (Bennett 1990) or are marginalized in various ways (Croteau and Hoynes 1994; Entman and Rojecki 1993; Gitlin 1980; Goldenberg 1975). The reliance on officials has become so routinized that their accounts remain prominent in news stories even when reporters are in a position to directly witness events for themselves (Livingston and Bennett 2003). This research challenges the capacity of journalism to serve as an adequate counterweight to official power.

More recent research has complicated this picture. Despite the continued prominence of official sources, news content in recent years has become more interpretive and more critical of officials and their policies. The growing prevalence of such content has been documented in studies of election campaign coverage (Hallin 1992; Patterson 1993) and political news more generally (Clayman et al. 2006; Clayman and Heritage 2002b; Entman 2003; Hart, Smith-Howell, and Llewellyn 1990; Robinson 1976, 1981; Rozell 1994; Sabato 1991; Smoller 1990; see also Cohen 2004). Although this could be interpreted as support for the watchdog model, some press critics argue that journalism has gone too far, presenting an indiscriminately critical and corrosively cynical view of officials and candidates. Here, the press is viewed as a relentless “attack dog” (Tannen 1998), an out-of-control “junkyard dog” (Sabato 1991), or—departing from the dominant canine metaphor—a “burglar alarm that just keeps ringing” (Bennett 2003).

How does the actual practice of journalism square with these models? Given nuanced phenomena such as Bennett’s (1990) indexing hypothesis in conjunction with recent work on event-driven news (e.g., Althaus 2003; Lawrence 2000; cf. Sparrow 1999), and the various phases of presidential coverage (Grossman and Kumar 1979; Manheim 1979; Smoller 1990), it seems increasingly clear that static models of journalism—whether of the watchdog, lapdog, or attack dog variety—are inadequate to capture the complexity and dynamism of actual journalistic practice. One way to advance understanding of the journalism-state relationship is to isolate the specific conditions under which news becomes more independent and critical. When, exactly, does the watchdog bark?

Our study explores the conditions associated with journalistic aggressiveness in the context of presidential news conferences. This project builds on previous research in several ways. First, unlike studies of single and often idiosyncratic news events, we encompass journalistic conduct over 48 years from Eisenhower through Clinton. This, in turn, facilitates multivariate analysis of a wider range of social conditions that might bear on journalistic practice. Finally, we examine journalistic conduct in a domain that has been subject to little formal quantitative analysis, namely presidential news conferences, where aggressiveness is embodied in the design of the questions that journalists put to presidents.

How vigorously do journalists discharge the task of questioning presidents, and what social conditions bear on this process? Four sets of conditions will be explored.

The administration life cycle. How does the life cycle of an administration bear on aggressive questioning? Presidential news coverage is hypothesized to follow discernable phases (Grossman and Kumar 1979; Manheim 1979; Smoller 1990), beginning with a congenial honeymoon period and then becoming more adversarial over the term in office, although later phases can bring about a reassessment of the administration and more favorable coverage. Second terms, by contrast, may lack a comparable honeymoon, and the prevalence of second-term scandals may make aggressive questioning more frequent.
A different chronological factor, but one of potential relevance, is the time lag since the previous news conference. Journalists may be inclined to punish presidents who have been inaccessible; alternatively, they may shrink from aggressive questioning for fear of having even less access to the president as a consequence.

**Presidential popularity.** Does the president’s popularity with the public condition the tenor of questioning? Given that opinion polling has long been a prominent fixture of presidential politics, and many news organizations administer their own polls, most White House reporters are presumably aware of the president’s standing with the public. Moreover, journalists’ self-understanding of their professional role—they see themselves as servants of the general public—may motivate them to attend to such information when designing their questions. As Helen Thomas, a fixture of presidential news conferences since the Kennedy era, put it, reporters act “as surrogates for all Americans who want to know what’s going on” (Thomas 1999:100). On the other hand, some research suggests that journalists are more sensitive to elite opinion than they are to public opinion (Bennett 1990). In any case, there are no systematic tests of the association between public support for the president and the tenor of White House journalism (but see Groeling and Kernell 1998).

**The economic context.** Is the business cycle associated with aggressive questioning? This is a plausible connection, given the post-New Deal tendency to view presidents as managers of economic affairs. Yet the relationship between the business cycle and the tenor of news, presidential or otherwise, has not previously been investigated. This neglect may result from a post-Watergate tendency to conceptualize the watchdog role of the press in terms of investigative journalism and the exposure of moral transgression, rather than more routine monitoring of presidential performance.

**Foreign affairs.** How do the well-known “rally ‘round the flag” syndrome and the maxim that “politics stops at the water’s edge” bear on the character of journalists’ questions? Journalistic deference under conditions of foreign conflict has been widely documented (e.g., Bennett and Paletz 1994; Zelizer and Allan 2002). However, some have suggested that such deference has eroded in recent decades, first because the Vietnam War opened up an official credibility gap (Hallin 1994), and more recently because the collapse of the Soviet Union and the end of the Cold War made it easier for reporters to question the official line (Entman 2003; cf. Holohan 2003).

A more general tendency to defer to officials on foreign and military affairs, regardless of the existence of active conflict, has not received equivalent attention. Since presidential news conferences are held infrequently, they are not well-suited to studying the impact of specific point-events such as military actions, but the broader phenomenon of deference on foreign affairs can be investigated through the vehicle of question content. Are questions about foreign affairs and military matters formulated less aggressively than domestic questions? Does journalistic aggressiveness “stop at the water’s edge” in the news conference environment? And how has such deference changed, if at all, in the post-Vietnam and post-Cold War eras?

With the exception of the administration life cycle, the emphasis here is on broad social conditions that might bear on journalistic practice. The overall objective in investigating such factors is to understand, for a key domain of journalistic professionalism, what makes the watchdog bark.

**Studying presidential news conferences**

In the long tradition of research on presidential news conferences and evolving president-press relations (e.g., Cornwell 1965; French 1982; Juergens 1981; Pollard 1947; Smith 1990; Tebbel and Watts 1985), quantitative studies of journalists’ behavior are rare and underdeveloped, limited mainly to the topical content of questions (Manheim 1979) and the prevalence of follow-up questions (McGuire 1967). More recent work, drawing on conversation analytic studies of talk in interaction, has gone beyond question content to explore more formal aspects of question design, especially those features that exert pressure and constraint on presidents. Using 10 features of question design encompassing both form and content, Clayman and
Heritage (2002b) developed a system for measuring the level of aggressiveness encoded in journalists’ questions. They applied that system in a comparative study of Dwight D. Eisenhower’s and Ronald Reagan’s news conferences, and later conducted a large-scale study of news conferences from Eisenhower through Clinton, documenting long-term historical trends toward more aggressive questioning (Clayman et al. 2006). The present study, using the same database, moves beyond the description of historical trends to isolate some of the major social conditions associated with variations in aggressiveness.

The presidential news conference is, in many respects, well-suited to the investigation of journalistic conduct toward the state. It is a key interface between the chief executive and the news media and, correspondingly, receives saturation media coverage. It has been a fixture of Washington journalism and presidential communication since the early decades of the twentieth century. Its emergence coincided with the growing stature and professionalization of journalism (Schudson 1978, 1982), as well as with presidents’ growing interest in and responsiveness to public opinion (Kernell 1986; Tulis 1987). Since Eisenhower, it has been a thoroughly public arena—an arena wherein White House journalists directly confront the president rather than writing or talking about him. It is thus not merely a conduit for information transfer from authoritative source to dependent journalist; it is also a ritual of political accountability in which presidents are obligated to explain and justify their policies before the public.¹ As an arena of accountability, the presidential news conference has been likened to prime minister’s question time in England (Cater 1959; French 1982), although the agents of accountability differ in these two systems. The British parliamentary system makes the prime minister answerable to the public through its elected representatives in Parliament; in the American system of divided power, the president is not answerable to Congress in the same way. The task of cross-questioning presidents thus falls to a cadre of journalistic professionals, where it has become incorporated into the normative watchdog role.

The interactional organization of news conferences offers both affordances and challenges for quantitative analyses. Because the role of questioner is shared by large numbers of participating journalists, some method of turn-taking is needed to determine which journalist is to ask each successive question. Although there has been occasional experimentation with different methods for selecting questioners (e.g., a prearranged order determined by lottery, used intermittently during the Reagan and current Bush administrations), by far the most common method has been for presidents to make the selection as the interaction unfolds. Specifically, as a given response by the president is winding down, journalists “bid” for the next question by raising their hands, calling out “Mr. President!”, and so forth, and the president then selects among the bidders. Consequently, with the exception of the occasional follow-up question, each successive question is asked by a different journalist. Since journalists are known to prepare and compose their questions well in advance, the questions tend to be topically disjunctive and are not contingent on prior question-answer sequences. These interactional conditions maximize the independence of questions from one another, thereby facilitating statistical analysis and the establishment of associations between question design and conditions other than the prior sequential context.

Some of these same conditions admittedly complicate the use of question design as a window into journalistic culture. The turn-taking system places discretion in the hands of the president, who may avoid journalists regarded as unfriendly or unduly aggressive. Insofar as such avoidance is accountable as such, it could inhibit others in the press corps from asking tough questions. These insights, coupled with the recent exposure of a partisan shill (Jeff Gannon) posing as a White House journalist, raise the possibility that question patterns may not be a direct reflection of journalistic dispositions. On the other hand, presidential discretion is not unlimited. Presidents almost never initiate questioning from a given reporter more than once per conference, which greatly restricts the capacity to exploit a few reliably friendly reporters. In any case, since journalistic culture cannot be pristine and independent of gov-

¹ For exemplary studies of the ritual aspect of journalism and the mass media, see Alexander (1988) and Dayan and Katz (1992).
ernment influence, it would be misguided to focus on journalistic dispositions to the exclusion of actual behavior. White House journalists are embedded within an evolving relationship with the president, and any valid analysis of journalistic vigorousness must track its context-sensitive expression in actual conduct.

CONVERSATION ANALYSIS AND QUANTIFICATION

Since the framework for analyzing aggressiveness in questioning is based on prior conversation analytic research, some discussion of conversation analysis and its relationship to quantification is in order.

Conversation analysis (CA) is the dominant approach to the study of language use and talk in interaction across the social sciences. It investigates patterns of social interaction for evidence of practices of conduct that exhibit systematic design—design associated with the production of intelligible social actions and organized sequences of actions. To be identified as a systematic practice, particular elements of conduct must be recurrent, contextually situated, and attract responses that discriminate them from related or similar practices. A central feature of this methodology is that the analysis of a given practice (e.g., negatively formulated yes/no questions such as “Isn’t it true that . . .”) in terms of the action that it performs (e.g., it invites or favors a yes-type response) can be validated by examining how others respond. These sequentially unfolding practices are a core part of the verstehende resources with which actors manage their engagement with others in the social world. To study them requires close analyses of individual cases and collections of cases.

Integrating analyses of verstehende practices with the environmental circumstances of action or its outcomes may, however, require formal quantification and statistical analysis. Questions such as whether talk has changed over time, or whether it varies systematically with characteristics of the social environment, typically cannot be answered in any other way. Most quantitative extensions of CA have focused on practices that are associated with the design of individual turns at talk, if only because the “statistics of counts are less problematic than those of series” (Inui and Carter 1985:536). These extensions have become increasingly common (Boyd 1998; Heritage, Boyd, and Kleinman 2001; Maynard et al. 2002; Perakyla 1998; Robinson and Heritage 2005; Stivers et al. 2003). When quantitative measures of talk are derived from prior CA findings, they are internally validated in an “emic” (Pike 1967) sense—that is, grounded in the understandings of interactants themselves—and are proving to have strong predictive and explanatory value. It is in this spirit that the present study of aggressive questioning in news conferences, which employs a coding scheme grounded in extensive prior work on question design in broadcast news interviews and other environments (Clayman et al. 2006; Clayman and Heritage 2002b), was developed.

METHODOLOGY

HISTORICAL TIME FRAME AND SAMPLING PROCEDURE

Our study encompasses the administrations of Eisenhower through Clinton (1953 to 2000), a time frame that roughly spans the era of the public news conference. Earlier news conferences were essentially private encounters between presidents and journalists, with strict rules governing the use of quotations and the manner in which they could be attributed to the president (Cornwell 1965; Smith 1990). These rules were relaxed significantly during the Truman administration, but it was not until Eisenhower that the news conference became fully public and “for the record,” with most content available for verbatim quotation, full attribution, and subsequent broadcast.

Using transcripts reprinted in Public Papers of the Presidents, four news conferences were sampled per year from 1953 to 2000. The conferences were staggered quarterly over the course of each year, selecting the first conference held after February 1, April 1, July 1, and October 1. Although the frequency of conferences varies (mean/year = 17.7, standard deviation = 11.5), we chose a quarterly sample to maximize the power to detect associations with time.² Conferences held beyond White House

² A proportional sampling approach would have better supported design-based inference. If the regression models are not correctly specified, the estimates
grounds, and those involving other officials in addition to the president, were excluded from the sample. Twenty-eight of 192 quarters (15 percent) contained no presidential news conferences, so this sampling procedure yielded a database of 164 conferences and 4,608 distinct questions.

**INDEPENDENT VARIABLES**

The administration life cycle. We investigated four variables concerning the chronological position of a given news conference within the president's term in office: (1) a contrast between the first conference and later conferences in the first term, (2) a linear trend variable over the course of the first term, (3) for those presidents who served at least part of two terms, a contrast between the first and second term, and (4) the time lag since the previous conference. The first three variables are designed to test for the existence of phases in president-press relations, whereas the fourth tests whether presidential avoidance of regular conferences is associated with questioning practices.

Presidential popularity. We operationalized popularity as the president's Gallup job approval rating that most closely preceded the conference in question. This measure is standardized across the data sample and is based on the response to the question, "Do you approve or disapprove of the way [name] is handling his job as president?"

The economic context. The business cycle was indexed via four variables: (1) the unemployment rate (using Bureau of Labor Statistics data), (2) the inflation rate (using retrospectively standardized consumer price index data), (3) the prime interest rate, and (4) the stock market (using the Dow Jones index, retrospectively standardized). For each measure, the figures that most closely preceded each news conference were used. Neither the unemployment rate nor the prime interest rate had consistent trends in mean from 1953 to 2000. They were standardized traditionally, as Z-scores based on means and standard deviations of the entire 1953 to 2000 time period. In contrast, the consumer price index and Dow Jones average increased strongly over time, rendering levels in 1953 not directly comparable in 2000. To put these figures in their historical context, we transformed the variables through what we call "retrospective standardization," so that for a given news conference they are Z-scores based on the mean and standard deviation of only those values between January 1, 1949 and the date of the news conference. The measures may be distinguished in terms of their salience to either the Main Street or Wall Street economy. The first variable (unemployment rate) captures economic conditions on Main Street, while the third and fourth variables (interest rate and stock market) are more relevant to the Wall Street economy. The second variable (inflation) has a broader salience that spans the Main Street and Wall Street domains.

Foreign affairs. Question content was operationalized as a binary variable, so that questions concerning domestic affairs could be compared with those concerning foreign/military affairs. To assess the stability of the impact of question content, we also investigated interaction effects in relation to time and to significant economic variables. This variable differs from others in that it reflects partially endogenous choices of journalists, rather than entirely exogenous conditions.

Control variables: historical trends. Our previous research on the same database (Clayman et al. 2006) documented long-term historical trends for all dimensions of aggressiveness, all involving increasing aggressiveness over time. These rising trends are linear for initiative, assertiveness, and adversarialness, and level off over time for directness and accountability. Accordingly, all of the present models include a linear control variable for time (1 unit
= 4 years), and the directness and accountability models have an additional (centered) quadratic time variable. Since these historical trends have already been reported, they serve as control variables in the present study.

A summary of the social conditions examined, and the specific independent variables through which they were indexed, are reported in Table 1.

**DEPENDENT VARIABLES: THE QUESTION ANALYSIS SYSTEM**

The question analysis system decomposes the phenomenon of aggressive questioning into five outcome measures:

1. **Initiative**: the extent to which questions are enterprising rather than passive in their aims.
2. **Directness**: the extent to which questions are blunt rather than cautious in raising issues.
3. **Assertiveness**: the extent to which questions invite a particular answer and are in that sense opinionated rather than neutral.
4. **Adversarialness**: the extent to which questions pursue an agenda in opposition to the president or his administration.
5. **Accountability**: the extent to which questions explicitly ask the president to justify his policies or actions.

Each measure is operationalized in terms of various features of question design, and these serve as our key indicators (see Table 2).

Granting the difficulty of either distinguishing or fully separating the substantive “content” of a question from its linguistic or discursive “form,” the first three measures are predominantly concerned with matters of form, while the fourth and fifth are predominantly concerned with content. Below is a brief sketch of the coding system (for a fuller discussion, see Clayman et al. 2006; Clayman and Heritage 2002b).

**INITIATIVE**. Journalists exercise initiative when they (1) preface their questions with statements that construct a context for the question to follow, (2) ask more than one question within a single turn at talk, or (3) ask a follow-up question. Each of these practices embodies a more vigorous posture by the journalist.

**DIRECTNESS**. Directness is measured by the absence of various practices that embody an indirect or cautious stance toward the question. Journalists are markedly indirect when they frame their questions with *self-referencing* phrases (e.g., “I wonder whether . . .,” “I want/would like to ask . . .,” “Can I/Could I/May I ask . . .”) invoking their own intentions or desires before launching into the question proper. With the more indirect variants (“Can I/Could I/May I ask . . .?”), the journalist is virtually requesting permission to proceed with the ensu-

---

**Table 1. Independent Variables**

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Independent Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Administration Life Cycle</td>
<td>1st conference versus later 1st-term conferences</td>
</tr>
<tr>
<td></td>
<td>Linear trend across first term</td>
</tr>
<tr>
<td></td>
<td>1st term versus 2nd term</td>
</tr>
<tr>
<td></td>
<td>Time lag since previous conference</td>
</tr>
<tr>
<td>Presidential Popularity</td>
<td>Gallup job approval rating</td>
</tr>
<tr>
<td>The Economic Context</td>
<td>Unemployment rate</td>
</tr>
<tr>
<td></td>
<td>Consumer price index</td>
</tr>
<tr>
<td></td>
<td>Prime interest rate</td>
</tr>
<tr>
<td></td>
<td>Dow Jones</td>
</tr>
<tr>
<td>Foreign Affairs</td>
<td>Domestic versus foreign/military questions</td>
</tr>
<tr>
<td></td>
<td>Foreign × time interaction</td>
</tr>
<tr>
<td></td>
<td>Foreign × prime interest rate interaction</td>
</tr>
<tr>
<td></td>
<td>Foreign × unemployment interaction</td>
</tr>
<tr>
<td>Historical Trends</td>
<td>Year</td>
</tr>
<tr>
<td></td>
<td>Year squared</td>
</tr>
<tr>
<td>Measure</td>
<td>Item (Indicator)</td>
</tr>
<tr>
<td>------------------</td>
<td>----------------------------------------</td>
</tr>
<tr>
<td>Initiative</td>
<td>Statement Prefaces</td>
</tr>
<tr>
<td></td>
<td>Multiple Questions</td>
</tr>
<tr>
<td></td>
<td>Follow-Up Questions</td>
</tr>
<tr>
<td>Directness</td>
<td>Absence of Other-Referencing Frames</td>
</tr>
<tr>
<td></td>
<td>Absence of Self-Referencing Frames</td>
</tr>
<tr>
<td>Assertiveness</td>
<td>Preface Tilt</td>
</tr>
<tr>
<td></td>
<td>Negative Questions</td>
</tr>
<tr>
<td>Adversarialness</td>
<td>Preface Adversarialness</td>
</tr>
<tr>
<td></td>
<td>Global Adversariality</td>
</tr>
<tr>
<td>Accountability</td>
<td>Accountability Questions</td>
</tr>
</tbody>
</table>
ing question. Indirectness is also manifest through the use of other-referencing frames that invoke the president’s ability (e.g., “Can you/Could you tell us . . . ?”) or willingness (“Will you/Would you tell us . . . ?”) to answer the question, and hence allow for the possibility that he may be unable or unwilling to answer. Both self- and other-referencing frames reduce the level of coercion encoded in the question. Conversely, the absence of such frames is indicative of directness and represents a more forceful way of putting issues before the president.

**Assertiveness.** Assertiveness is measured only for yes/no questions (n = 2,519), for which the phenomenon is most easily assessed. Yes/no questions can be designed to invite or favor either a yes- or no-type response in two distinct ways: (1) through a prefatory statement (e.g., “Unemployment rose sharply last month. Are we in an economic downturn?”), or (2) through the linguistic form of the question itself, which can be negatively formulated and thus tilted in favor of yes (e.g., “Aren’t we in an economic downturn?”).

**Adversarialness.** An oppositional stance can be encoded (1) in the preface to the question only, or (2) in the design of the question as a whole. Prefaces were coded as adversarial if they disagreed with the president or were explicitly and strongly critical of the administration. We also noted whether the subsequent question focused on the preface (e.g., “What is your response to that?”), thereby treating it as debatable, or whether the question presupposed the truth of the preface, the latter being more adversarial. Questions as a whole were coded as adversarial when an oppositional or critical posture ran through the question in its entirety.

**Accountability.** Accountability is operationalized as questions that explicitly ask the president to explain and justify his policies. Because such questions decline to accept policy at face value, they are to some extent aggressive, although the degree of aggressiveness depends on the linguistic form of the question. *Why did you*-type questions invite a justification without prejudice, whereas *How could you*-type questions are accusatory, implying an attitude of doubt or skepticism regarding the president’s capacity to adequately defend his actions. Note that accountability, unlike the other measures, has a single indicator—the occurrence of *why did you/how could you*-type questions.

For the measures involving multiple indicators (initiative, directness, assertiveness, adversarialness), we combined discrete indicators into a single composite measure with higher values corresponding to more aggressive practices or multiple practices used in combination (see Clayman et al. 2006). We treated these composite measures as ordinal variables, not assuming interval scale properties or a normal distribution. To test whether a single underlying construct is being measured ordinarily throughout each scale, we predicted the outcomes from the time the conference was held using ordinal logistic regression and examined the test of the assumption of proportional odds. The tests revealed only a single violation of the assumption of proportional odds (p < .05), which was rectified by collapsing two adjacent levels of one scale.

As a framework for coding aggressiveness, the question analysis system is highly reliable, in part because most indicators are relatively formal aspects of question design. Unlike content-based coding categories, which tend to be more interpretive and require considerable judgment to apply (Krippendorff 1980:62–63), categories based on formal design features are relatively concrete and straightforward to apply. A team of 14 coders working in pairs, whose decisions required consensus, coded the data. Problem cases were resolved in weekly meetings involving the entire research team. Reliability was assessed by a joint recoding of a subsample of 10 conferences, and evaluated using Cohen’s kappa. Kappa scores exceed .90 for 4 of the 10 indicators and exceed .75 for 8 of 10 (see Table 2). Note that the reliability of the composite measures (far right column), each composed of two to three indicators, tends to be somewhat higher than that of the discrete indicators, with three out of four exceeding .80 and one just shy of that level at .78. Since the statistical analysis is based mainly on composite measures, and since Kappa scores above .75 are generally understood to indicate at least 90 percent agreement (and even greater agreement for coding categories with few codes; see Bakeman et al. 1997), the system is demonstrably reliable.
As for validity, formal features of question design have been the subject of substantial prior research, both on journalistic questioning per se and on questioning practices in interaction generally. This research demonstrates that specific design features are indeed understood and treated by interactants as embodying aggressiveness in various forms, and thus provides grounds for the use of such features as indicators of aggressiveness. Further, as noted above, each composite measure appears to tap into a single underlying construct.

The various measures are also moderately correlated. At the level of the individual press conference, the five outcomes have a mean correlation of .30, ranging from a minimum of 0 for assertiveness and directness to a maximum of .55 for accountability and hostility, with the other 8 of 10 correlations ranging from .21 to .44. A factor analysis found only one eigenvalue greater than 1. Together, these results suggest moderate correlations of these five measures of press behavior for a given press conference.

**Statistical Analysis**

Although these data have a time series structure, autocorrelation does not appear to be a significant issue. Because the outcome variables are ordinal, standard tests and corrections for autocorrelation don’t apply directly. Nevertheless, we estimated and tested for autocorrelation in the scores that would have resulted if we treated the scales as having interval properties with normally distributed residuals. The Durbin-Watson test of serial correlation found no evidence of autocorrelation ($p > .2$) for three of the five outcomes, although assertiveness and adversarialness exhibit small negative autocorrelations ($p < .05$). This suggests a mild compensatory or “pendulum” effect in aggressiveness—if the press corps has been more assertive/adversarial than the model predicts in a given conference, they tend to be slightly less assertive/adversarial in the next sampled conference, and vice versa. Uncorrected standard errors and p-values are actually conservative in the presence of negative autocorrelation; for this reason we present slightly conservative unadjusted results that preserve the ordinal nature of the outcomes.

To assess the impact of various social condition variables on aggressive questioning, we ran 14 sets of staged ordinal logistic regression models. Within each set, we ran models for each of the five outcomes in Table 2, for a total of 70 individual regression models. Factors that were not significant at $p < .05$ after correcting for multiple testing across outcomes were removed from subsequent models. Table 3 summarizes the parameterization of this series of 14 sets of models.

The first set of models test for historical trends by examining the impact of linear and quadratic time variables (periodized by year). As reported in a previous article (Clayman et al. 2006), linear trends are significant for all dimensions, but quadratic trends are significant for only two outcomes—directness and accountability. Therefore, we retain the quadratic variable only for models involving these two outcomes. In the present analysis, these serve as base models for the investigation of social condition variables.

The second through fifth sets of models assess the impact of the administration life cycle on aggressive questioning by adding a single term each to the base model set. These variables are time since last conference, a first conference indicator, time since start of administration, and a first term indicator for models two through five, respectively. The third and fourth model sets are restricted to first term data in order to focus on effects within first terms. Of these four variables, only the dichotomous variable for first/second terms yields significant results. Accordingly, only this variable is retained for subsequent models.

The sixth through ninth sets of models add economic variables, one at a time, to the fifth set of models: (1) the unemployment rate, (2) the prime interest rate, (3) the consumer price index, and (4) the Dow Jones average, respec-
tively. Of these four economic variables, two—unemployment rate and prime interest rate—emerge as significant predictors. They are both added to the fifth set of models to form a tenth set of models.

An eleventh set of models add Gallup job approval ratings to the tenth set of models and finds no significant results for the Gallup ratings. Side analyses show that when added to the base set of models, Gallup ratings are significantly negatively associated with initiative and adversarialness, and that they have strongly negative and significant bivariate correlations with both unemployment and the prime interest rate.

The twelfth set adds a dichotomous indicator of foreign/military content to the tenth set of models. This variable proves statistically significant. The thirteenth set of models adds interactions between the foreign indicator and linear time to the twelfth set of models, while the fourteenth set of models adds interactions of the foreign indicator with both unemployment and prime interest rate to the tenth set of models. After corrections for multiple comparisons, no interaction terms are statistically significant.

RESULTS

Table 4 displays the results of the final (twelfth) set of models. Note that odds ratios (ORs) significantly greater than one reflect positive associations and ORs significantly less than one reflect negative associations. Concerning the magnitude of the association, ORs denote the amount by which the odds of a higher rather than lower level of the outcome occurring are multiplied for each unit of a given predictor, after controlling for other predictors, at any cut-point on the ordinal outcome scale. For example, the OR of .52 for foreign affairs with respect to the assertiveness outcome means that the odds of an assertiveness score of 3 versus 0–2, 2–3 versus 0–1, or 1–3 versus 0 are approximately halved (multiplied by .52) for foreign affairs questions, as compared to domestic questions, after controlling for term, unemployment, interest rate, and time trends.

THE ADMINISTRATION LIFE CYCLE. Most of the intra-administration chronological variables turn out to be insignificant. However, second terms (following re-election) are significantly different from first terms, with the questioning more aggressive on all dimensions except directness during the second term in office. The effect is strongest for assertiveness, whereas directness actually decreased in second terms.

PRESIDENTIAL POPULARITY. Popularity as measured by the Gallup job approval rating is not a significant independent predictor of any aspect of aggressive questioning. As noted earlier, side analyses show that when added to the base models, Gallup ratings are significantly

<table>
<thead>
<tr>
<th>Model Set</th>
<th>Independent Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (Base model)</td>
<td>Year, Year squared^a</td>
</tr>
<tr>
<td>2</td>
<td>Model Set 1 + time since last conference</td>
</tr>
<tr>
<td>3</td>
<td>Model Set 1 + first conference indicator^b</td>
</tr>
<tr>
<td>4</td>
<td>Model Set 1 + time in administration^b</td>
</tr>
<tr>
<td>5</td>
<td>Model Set 1 + 2nd term indicator</td>
</tr>
<tr>
<td>6</td>
<td>Model Set 5 + unemployment</td>
</tr>
<tr>
<td>7</td>
<td>Model Set 5 + prime interest rate</td>
</tr>
<tr>
<td>8</td>
<td>Model Set 5 + Consumer Price Index</td>
</tr>
<tr>
<td>9</td>
<td>Model Set 5 + Dow Jones average</td>
</tr>
<tr>
<td>10</td>
<td>Model Set 5 + unemployment, prime interest rate</td>
</tr>
<tr>
<td>11</td>
<td>Model Set 10 + Gallup Poll</td>
</tr>
<tr>
<td>12</td>
<td>Model Set 10 + foreign indicator</td>
</tr>
<tr>
<td>13</td>
<td>Model Set 12 + foreign × time</td>
</tr>
<tr>
<td>14</td>
<td>Model Set 12 + foreign × unemployment, foreign × prime interest rate</td>
</tr>
</tbody>
</table>

^a Only directness and accountability outcomes retained quadratic terms.
^b First terms only.
Table 4. Predictors of Aggressive Questioning

<table>
<thead>
<tr>
<th>Condition</th>
<th>Predictor</th>
<th>Initiative</th>
<th>Directness</th>
<th>Assertiveness</th>
<th>Adversarialness</th>
<th>Accountability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Administration Life Cycle</td>
<td>2nd terms odds ratio</td>
<td>1.79***</td>
<td>.82*</td>
<td>2.04***</td>
<td>1.68***</td>
<td>1.44*</td>
</tr>
<tr>
<td></td>
<td>p value</td>
<td>&lt;.001</td>
<td>&lt;.019</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Economic Context</td>
<td>unemployment rate odds ratio</td>
<td>1.15***</td>
<td>1.05</td>
<td>1.17***</td>
<td>1.22***</td>
<td>1.12*</td>
</tr>
<tr>
<td></td>
<td>p value</td>
<td>&lt;.001</td>
<td>&lt;.082</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>prime interest rate odds ratio</td>
<td>1.02</td>
<td>1.98</td>
<td>1.06***</td>
<td>1.05**</td>
<td>97</td>
</tr>
<tr>
<td></td>
<td>p value</td>
<td>97</td>
<td>159</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Foreign Affairs</td>
<td>foreign/military Qs odds ratio</td>
<td>.96</td>
<td>1.03</td>
<td>.52***</td>
<td>.40***</td>
<td>.58***</td>
</tr>
<tr>
<td></td>
<td>p value</td>
<td>428</td>
<td>.635</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Secular Trends</td>
<td>time odds ratio</td>
<td>1.11***</td>
<td>1.23***</td>
<td>.99</td>
<td>1.08***</td>
<td>1.18***</td>
</tr>
<tr>
<td></td>
<td>p value</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>time squared odds ratio</td>
<td>—</td>
<td>.98***</td>
<td>—</td>
<td>—</td>
<td>.97**</td>
</tr>
<tr>
<td></td>
<td>p value</td>
<td>&lt;.001</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>&lt;.003</td>
</tr>
</tbody>
</table>

Note: For the unemployment and prime interest rates, odds ratios are standardized.
negatively associated with two dimensions of aggressiveness (initiative and adversarialness), but because Gallup ratings also have strongly negative and significant bivariate correlations with economic condition variables (discussed below), these ratings lose their predictive significance in subsequent models incorporating economic variables.

**THE ECONOMIC CONTEXT.** Of the economic indicators, two emerge as significant predictors. The prime interest rate is mildly associated with increased assertiveness and adversarialness. The unemployment rate is moderately associated with increased aggressiveness on every dimension but directness.

**FOREIGN AFFAIRS.** Questions concerning foreign affairs and military matters are less assertive, less adversarial, and embody less accountability than domestic affairs questions. In each case these effects are very strong, and their magnitude remains essentially stable over time and under different economic circumstances, as evidenced by the general absence of interaction effects for question content in relation to both temporal and economic variables.

**HISTORICAL TRENDS.** It should be noted that secular trends remain after controlling for all these factors, with all aspects of aggressiveness other than assertiveness increasing mildly to moderately over time.

Summarizing these results, the strongest and most consistent exogenous predictor of journalistic aggressiveness is the president’s term in office, followed by unemployment and unexplained secular trends, and finally the prime interest rate. Question content is also a strong predictor, with foreign affairs questions markedly less aggressive than domestic questions. The one outcome that is relatively poorly explained by these contextual factors is directness, which displays an increasing secular trend and a second-term trend contrary to all other dimensions. These results suggest that directness may not be linked to the same construct of aggressiveness targeted by the other four outcomes.

**DISCUSSION**

The idea that president-press relations during the first term unfold in standardized phases—a congenial honeymoon period followed by more contentious relations—is not supported by our analyses. First sampled news conferences are not significantly different from other conferences, and there are no significant linear trends that extend across first terms. Of course, if the honeymoon period is brief, it could be that the sampling procedure used here (four conferences per year, staggered quarterly) is simply too sparse to detect it. Granting such limitations, our results do not support the existence of recurrent phases within first terms.

Our findings do, however, support the existence of phases extending across terms in office, with second terms substantially more contentious than first terms. This difference is probably not a product of the honeymoon effect, which as noted above is not significant in these data. It could very well be driven by independent events, such as the scandals that have marred presidents’ second terms (Nixon and Watergate, Reagan and Iran-Contra, Clinton and Lewinsky), although this begs the question as to how these scandals came to be and whether their association with second terms is coincidental or systematic. One president-centered explanation—that presidents become more reckless during second terms because they do not have to face the voters again in an election—is ruled out by the fact that each of the above scandals can be traced to conduct initiated during the first term. A more likely explanation is that a combination of factors—the persistent efforts of administration critics to expose misconduct, the time lag required for such efforts to bear fruit, the public’s greater willingness to take allegations of misconduct seriously when they persist beyond election campaigns (Lang and Lang 1983), which in turn may license journalists to be more vigorous in the second term—converge in fostering the emergence and escalation of second-term scandals.

As for broader circumstantial factors that bear on aggressive questioning, it is useful to consider these results against the backdrop of the watchdog model of journalism and its alternatives. White House reporters are plainly discriminating in their conduct toward presidents, avoiding the extremes of both total passivity and relentless aggression. Moreover, at least
some of the ways in which they discriminate are not inconsistent with an elaborated version of the watchdog model. This is perhaps most clear with respect to the relative impact of opinion polls versus economic factors. Journalists do not appear to be influenced by public perceptions of presidential performance. Although unpopular presidents are questioned somewhat more aggressively than popular presidents, objective economic conditions appear to be the driving force behind this association. Journalists, it would seem, are attentive to the real state of the nation, growing more aggressive as economic conditions worsen. Moreover, their sensitivity to the economy is multidimensional, encompassing both conditions on Main Street (as indexed by the unemployment rate) and to a lesser extent conditions on Wall Street (as indexed by the prime interest rate). The somewhat greater sensitivity to the Main Street economy imparts a mildly populist tilt to the questioning, although unfavorable conditions on either front are consequential for both adversarial question content and for modes of questioning that exert greater pressure on presidents.

This finding highlights an aspect of journalism that can be subsumed within the independent watchdog model but has not been adequately appreciated in previous discussions of the subject. Both journalistic and scholarly treatments of the watchdog role follow the Watergate prototype in focusing primarily on investigative journalism and the impetus to uncover official transgressions and wrongdoing (e.g., Ettenla and Glasser 1998; Gans 2003; Protess et al. 1992; Waisbord 2000). But another aspect of this role, one that does not require elaborate investigative methods and appears to be a much more routine journalistic practice, is the impetus to monitor presidential performance with respect to the domestic economy. Such an impetus is suggested by the robust association between poor economic conditions and aggressive questioning and would in turn explain that association.

While sensitivity to the domestic economy is broadly consistent with an elaborated version of the watchdog model, the scope of such sensitivity further complicates the model. As we have seen, a downturn in the business cycle leads to more aggressive questioning not only on domestic affairs, but also on foreign affairs and military matters. White House journalists, it would seem, are not immune to gestalt processes in their assessment of the president. As they monitor the president’s stewardship of the economy, poor presidential performance in this area appears to contaminate the president’s image in other areas, leading journalists to become generally more aggressive.

The foreign affairs arena remains, however, an area of relative journalistic deference. Compared to the domestic arena, presidents are to some degree shielded from vigorous questioning on the international scene, and this armor remains in place even in the absence of active foreign threats or other “rally ‘round the flag” events. Journalists thus treat domestic and foreign news as distinct domains to which different journalistic norms apply (cf. Hallin 1984). This does not mean, however, that aggressive questioning “stops at the water’s edge,” for the foreign shield is not impermeable. Whenever domestic questions have grown more aggressive (i.e., in recent decades and during economic hard times), foreign/military questions have as well and at a proportionally similar rate. In other words, whenever White House journalists have generally been inclined toward a more vigorous or adversarial posture, the inclination has extended to foreign/military subjects. In the news conference environment, then, aggressiveness does not stop at the water’s edge; but it does become measurably more restrained.

The salience of the economy provides a new way of explaining the rise of a more vigorous and at times adversarial journalism during the 1970s, as documented in studies of presidential news conferences (Clayman et al. 2006; Clayman and Heritage 2002b) and in studies of presidential news more generally (Hallin 1992; Hart et al. 1990; Patterson 1993; Robinson 1976, 1981; Smoller 1990; see also Cohen 2004). Most explanations focus on declining trust in the presidency resulting from the decep tions associated with the Vietnam War and the Watergate scandal. But given that the 1970s were also a period of persistent stagflation, with unemployment and interest rates reaching sustained high levels, economic factors also seem to have played a role in the contentious press relations that characterized this era. The fact that aggressive questioning persisted into the recession-plagued years of the early 1980s, but then declined during the better economic times...
of the later 1980s, adds support to this analysis (Clayman et al. 2006).

This explanatory framework also brings a fresh perspective to more recent historical trends in the foreign news arena. Previous research (Entman 2003; Hallin 1994; Holohan 2003) has documented more independent and vigorous coverage of foreign affairs news and has attributed the trend to events specific to the foreign arena—the fracturing of the Cold War consensus in the aftermath of Vietnam, and more recently, the collapse of the Soviet Union and the end of the Cold War. But when foreign news trends are lined up against domestic news trends, as in the present study, it becomes clear that the decline of journalistic deference toward officials has been a broader and more encompassing development. This, in turn, suggests the need for a more encompassing explanatory framework, one that includes the domestic economic conditions to which journalists are broadly responsive.

In understanding the postwar history of U.S. journalism, the economy is plainly not the whole story. When economic and other factors are controlled, there remains a secular trend toward more aggressive questioning that has persisted at least through the Clinton era. It may be that other unexamined conditions might account for this trend, which could in turn suggest that journalistic norms have remained essentially constant over time. Alternatively, it may be that Vietnam, Watergate, and related events have indeed brought about a sea change in the culture of journalism itself. Only further research can adjudicate between these possibilities.

Among the various dimensions of aggressive questioning examined in this study, directness stands out as most resistant to explanation. The fact that the array of social conditions analyzed here explain very little of the variation in directness dovetails with previous research (Clayman et al. 2006) demonstrating that directness exhibits a smooth historical trendline of steady growth over time. Both the trendline and the multivariate results indicate that directness is neither driven by nor sensitive to local events or conditions, and that its growth over time is a thoroughly secular trend. This adds further support for the suggestion (in Clayman et al. 2006) that rising directness may not be a journalistic trend per se, but rather one manifestation of a broader cultural change involving the decline of formality in American life and the coarsening of public discourse (Ferris 2002; Tannen 1998; see also Maynard 2003:55).

The focus on historical variation should not obscure more robust and persistent regularities of journalistic conduct, and the differential between domestic and foreign questioning is a previously-unrecognized instance that emerges from this study. While aggregate levels of aggressiveness have risen and fallen substantially over time, the domestic/foreign differential has been remarkably stable. For at least a half-century, White House journalists have been more cautious and deferential in the foreign news arena, and their relative cautiousness has remained substantially unchanged through periods of war and peace, recession and prosperity. Both structural and cultural factors may contribute to this pattern. Journalists may have less access to independent information in the foreign affairs arena and are thus disproportionately dependent on the administration’s framing of events. Moreover, most White House journalists are also U.S. citizens, and the latter identity is apt to be foregrounded and rendered more salient when raising foreign and military matters. Thus, when Dan Rather said on CNN, Look, I’m an American. . . . And when my country is at war, I want my country to win, whatever the definition of “win” may be. Now, I can’t and don’t argue that that coverage is without a prejudice. About that I am prejudiced. (April 14, 2003) he was capturing a fundamental dynamic that appears to condition the news not only during wartime, but whenever news content breaches the water’s edge.

Journalists also exercise restraint when it comes to enforcing the ritual of the news conference itself. When presidents decline to hold regular news conferences—as they sometimes do, especially in recent decades and in particular during times of scandal—they do not experience any measurable consequence in subsequent news conferences (although inaccessibility may of course be consequential for other aspects of news coverage). Journalists, it would seem, are not inclined to punish presidents for their inaccessibility, although they do not appear to be particularly intimidated by it either. In any case, the lack of consequence may be one factor, among many others, contributing to the relatively infrequent and at times
sporadic character of what was once a more regular ritual of political communication.

Various directions for future research are suggested by this study. It would be interesting to explore the historical scope of the association between aggressive journalism and the business cycle. While journalists have monitored presidential performance with respect to the economy for at least five decades, this aspect of the watchdog role may be limited to the post-New Deal era when presidents’ accountability for the economy reached new levels (Lynch 1999). Other conditions that might bear on aggressive questioning remain unexplored, including the characteristics of journalists and presidents. And shifting from causes to consequences, future studies might examine how the forms of questioning identified here impact presidents’ responses, as well as subsequent media coverage of news conferences. The latter would shed light on whether the exercise of journalistic vigorousness inside the news conference can have ripple effects extending beyond the confines of the conference itself, and what those effects might be. More generally, it would be illuminating to investigate whether the conditions of aggression that operate in news conference questioning also operate in other modes of journalistic practice. The fact that the historical trend toward more vigorous questioning roughly parallels trends toward more interpretive and critical news stories (Clayman et al. 2006) suggests that the factors identified here may indeed be more broadly consequential for how journalism is done.

**CONCLUSION**

This is the first multivariate analysis of journalism as it is practiced in presidential news conferences. We use reporters’ questions—drawn from a 48-year sample of news conferences—as a window into the social conditions that bear on aggressive journalism. We employ a multidimensional system for measuring the aggressiveness of questions that is both reliable and validated by previous conversation analytic research on question design. Our study demonstrates that journalists modulate their aggressiveness in complex ways that do not readily map onto any single model of the journalism-state relationship. Some patterns (e.g., aggressive questioning is conditioned by objective economic circumstances rather than presidential popularity) are consistent with an elaborated version of the watchdog model. Other patterns (e.g., aggressiveness is reduced for foreign and military questions) identify domains of journalistic deference toward the president, although even foreign/military questions become more aggressive when the economy is in decline. Linking journalistic vigorousness to the business cycle suggests a new way of understanding the watchdog role and its limits. It also provides a new explanatory framework for previously-documented historical trends in news coverage, although a secular trend of rising aggressiveness remains even when economic and other factors are controlled. Finally, our study documents a robust regularity in aggressive questioning, namely the domestic/foreign differential, the magnitude of which has remained stable for at least a half-century.

Steven E. Clayman is Professor of Sociology and affiliated with the Communication Studies Department at UCLA. His research concerns the intersection of talk, interaction, and mass communication in various contexts, with an emphasis on broadcast news interviews and presidential news conferences. He is the author of The News Interview: Journalists and Public Figures On the Air (with John Heritage).

John Heritage is Professor of Sociology at UCLA. He is the author of Garfinkel and Ethnomethodology, and The News Interview: Journalists and Public Figures On the Air (with Steven Clayman), and the editor of Structures of Social Action (with Max Atkinson), Talk at Work (with Paul Drew), and Communication in Medical Care (with Douglas Maynard). He is currently working on a range of topics in physician-patient interaction and on presidential news conferences (with Steven Clayman).

Marc N. Elliott is a RAND Senior Statistician with more than 80 peer-reviewed publications and expertise in survey sampling, experimental design, causal inference, and case-mix adjustment. Elliott leads an Office of Minority Health project developing methods to improve the precision of national health estimates for small racial/ethnic subgroups. He has been RAND’s lead statistician on the AHRQ CAHPS® project assessing consumer healthcare experiences since 1996. As the Center Statistician for the RAND/UCLA CDC Center for Adolescent Health Promotion, Elliott provides cross-site leadership of all design/analytic work on the longitudinal Healthy Passages project.
REFERENCES


Completion Relevance of Current Symptoms.” 
Social Science and Medicine 61:481–93.