The Effect of Salient Reputational Threats on the Pace of FDA Enforcement

MOSHE MAOR* and RAANAN SULITZEANU-KENAN*

Do reputational concerns affect the duration of enforcement decisions? We analyze “time to decision” in warning letter processes by two enforcement divisions within the Food and Drug Administration’s (FDA) Center for Drug Evaluation and Research. We find that nearly all criticism of these divisions revolves around the FDA’s primary consumer protection responsibilities (i.e., underenforcement), thus questioning the validity of the FDA’s unique reputation. We also found that as media coverage of the FDA’s consumer protection responsibilities becomes more positive, the agency takes enforcement decisions (warning letters) more slowly; in contrast, more critical media coverage leads to quicker action by the FDA. This effect is moderated by media salience; namely, it is found only for periods in which press coverage is relatively intense. An implication of this conditional relationship is an ability to assess the baseline role of reputation in the organization, namely, how concerned it is regarding its reputation in the absence of exogenous challenges.

Introduction

Previous research on bureaucratic politics has shown that “reputation matters” for the behavior of bureaucratic agencies. However, it has not provided a clear explanation as to when reputation matters, and how? This is not to say that previous research in this relatively new theoretical stream that has evolved alongside more traditional schools of thought—public interest theory and capture theory—asks narrow questions. On the contrary, studies that have placed the power of the agent at the center of analysis (Carpenter 2001, 2010a; Moe 2006) have generated several significant insights on the fundamental powers of bureaucratic agencies and on central issues of modern democracy. Paul Joskow (1974), for example, has argued that utility regulators worried more about nominal gas and electricity prices (the numbers that consumers and voters see) than about real prices (the observed numbers adjusted for inflation). Moffitt (2002) has suggested that public image considerations are important for the Food and Drug Administration (FDA) in determining whether and when it will call an advisory committee meeting. Krause and Douglas (2005) have demonstrated that reputation and audience considerations

*The Hebrew University of Jerusalem

© 2012 Wiley Periodicals, Inc.
shape government budget forecasts. Maor has shown that similar considerations influence the FDA’s strategy of jurisdictional claims (Maor 2010) and the FDA’s issuance of public warnings (Maor 2011). Maor, Gilad, and Ben-Nun Bloom (2012) have shown that reputation considerations are at the heart of a regulator’s attempt to strike a balance between strategic silence and regulatory talk. Gilad, Maor, and Ben-Nun Bloom (2012) have shown that reputation concerns determine regulatory differential reaction to audiences’ claims of under- and overregulation. Carpenter (2010a) has offered some generalized answers to issues regarding reputation and regulatory power, and has summed up the contributions of the aforementioned studies: “The lesson of this scholarship is that, when trying to account for a regulator’s behavior, look at the audience, and look at the threats” (Carpenter 2010b, 832, italics in original).

Yet what are the conditions under which reputation is more or less likely to matter, and what are the mechanisms that shape its effects? If the concept of “audiences” is so important to reputation-based explanations, how does the mechanism of visibility work? If the concept of “threat” to agency reputation is so crucial, what does negative coverage convey to the agency: punishment for past actions or threat guiding future conduct? These questions imply that the existing literature may benefit from a clearer approach to visibility. Such an approach should address the operational areas that will be strongly defended by a regulator facing threats to its reputation as compared to areas that will not: the policy instruments that will be employed by a regulator when facing threats to its reputation compared to instruments employed when no threats exist, and the regulator’s sensitivity to reputational challenges under different conditions of public visibility. This article addresses the latter question, namely, how do the mechanisms of visibility and reputational threats interact? It does so by focusing on the duration of agency decisions as an empirical instrument of testing claims about reputation.

Do reputational concerns affect the duration of enforcement decisions? If so, does an agency’s sensitivity to reputational challenges vary under different conditions of public visibility? We have developed an analytical framework that relates reputational threats to regulatory response duration, and have tested this empirically by estimating the effect of media valence (negative or positive tone of press coverage) regarding the FDA’s enforcement activity, on the “time to decision” in warning letter processes, by two enforcement divisions within the FDA’s Center for Drug Evaluation and Research (CDER). Our preliminary qualitative assessment of the press coverage reveals that nearly all criticism revolves around the FDA’s primary consumer protection responsibilities (i.e., underenforcement), thus questioning the validity of the FDA’s unique reputation. In the following stage, log-normal duration analyses show that such reputational threats are associated with shorter regulatory response duration. Specifically, as media coverage of the FDA’s primary consumer protection responsibilities becomes more positive, the FDA waits longer.
to issue warning letters; in contrast, more criticism leads to quicker action in issuing warning letters by the FDA. This effect is moderated by media salience; namely, it is found only for periods in which press coverage is relatively intense. These findings imply that an increase in the public visibility of the FDA’s enforcement activity increases its sensitivity to claims regarding underenforcement, thereby shifting the balance between overenforcement and underenforcement toward overenforcement.

The findings provide support for a reputation-based perspective (Carpenter 2001, 2010a; Maor 2007, 2010, 2011; Maor, Gilad, and Ben-Nun Bloom 2012), which allows us to shed useful light on the behavior of a regulatory agency once its enforcement policy is established. Moreover, it contributes to this literature in a number of ways. First, drawing on preliminary qualitative analysis, we identify the specific aspect of the agency’s reputation that is challenged—criticism over consumer protection versus criticism regarding overregulation. This identification enables us to select an appropriate testable hypothesis for the quantitative analysis from two theoretical alternatives. Second, by separately estimating the effect of both media valence (positive or negative) and media salience, we are able to show that only the combination of media content and intensity may affect enforcement decision duration, suggesting that the agency’s responsiveness to external evaluation of its performance is limited to conditions in which such challenges are likely to be publicly visible. These innovations in the measurement and analysis of the independent variable—reputational challenges—enable us to undertake a much more focused analysis of the role of reputation than previous research.

Furthermore, identifying a condition in which the agency is not sensitive to external evaluations of its performance extends the existing research agenda on organizational reputation, by allowing us to study not only the relationships between this concept and bureaucratic outcomes, but also the context in which these relationships operate. Importantly, it draws our attention to new questions regarding the baseline attitude of organizations to their reputation in the absence of exogenous challenges. Are they reputationally relaxed, only to become concerned about their reputation in the advent of external criticism (“reputationally relaxed baseline”)? Or are they reputationally concerned, and may relax only by visible public praise (“reputationally concerned baseline”)?

This article clearly demonstrates the importance of organizational time (Pierson 2004; Pollitt 2008) and how valuable a careful measurement of this concept may be for the study of bureaucratic decisions. Our findings imply that regulatory response duration is endogenous and plays a role in the range of ways in which agencies exercise responsiveness to external pressures. The idea that the duration of regulatory decisions might be endogenous to bureaucratic capacities and enforcement strategies is interesting not only for analyses of organizational reputation. It carries implications for understanding bureaucratic politics, the consequences of
public pressure on public institutions, and more generally for analyses of the gap between political information and action.

The layout of the article is as follows. The first section introduces the analytical framework and the derived hypotheses; the second section elaborates the methodology employed; the third section presents the empirical analyses and the findings; and the fourth section presents a discussion of the findings.

Related Literature

Scholars of institutional political science and organization theory who study the behavior of bureaucratic agencies typically do not place the duration of agency decisions at the center of analysis. As Carpenter (2003, 25) notes, “It is as if political scientists have assumed that the timing of decisions has nothing to do with ‘politics’ and that the proverbial slowness of government agencies is, for all intents and purposes, neutral or trivial.” Among the relatively few studies dedicated to this phenomenon, Carpenter has offered some generalized answers by modeling the drug approval decision as an optimal stopping problem of a stochastic process (Carpenter 2001, 2002, 2003); Moffitt (2002) and Lavertu and Weimer (2009) considered the effect of FDA drug advisory committees on review times; Kosnik (2005) investigated the sources of regulatory delays in decisions by the Federal Energy Regulatory Commission regarding relicenses for hydroelectric dams; and Whitford (2005) has examined decision duration of the U.S. Environmental Protection Agency’s enforcement of hazard waste law for the acceleration and deceleration of policy implementation in response to sequential interventions by multiple principals. Recently, a few scholars have analyzed regulatory review deadlines and their influence upon regulatory decisions (Carpenter et al. 2009; Carpenter and Grimmer 2009; Carpenter, Zucker, and Avorn 2008; Gersen and O’Connell 2008; Yackee and Yackee 2010). While extensive insight has been generated by these studies, this article offers a new contribution by specifically assessing the particular aspect of the agency’s reputation that is challenged and by estimating the implications of reputational threats under various conditions of public visibility.

Another advantage of utilizing decision duration rather than decision content is its potential to offer a measure of unintended behavioral effects. Decision content is often amenable to strategic manipulation. Reputation protection and blame avoidance considerations may sometimes act to restrict the response to criticism in order not to tacitly acknowledge a problem (Sulitzeanu-Kenan 2010). Regulatory silence may be also employed to achieve this aim (Maor, Gilad, and Ben-Nun Bloom 2012). However, decision duration is rarely manipulated in such ways and thus offers a prism to behavioral effects that would otherwise be hidden. This logic follows the recent surge in employing response time measures in many fields of psychology and behavioral economics.
These measures have introduced new ways to gauge biases and attitudes respondents would not openly acknowledge, or even be conscious of (Bohner and Dickel 2011).

**Analytical Framework**

The fundamental assumption of this analytical framework is that a regulator safeguards its unique reputation for protecting citizen welfare (here “public safety”) (Carpenter 2001, 2010a; Heimann 1997; Krause and Douglas 2005; Maor 2010, 2011; Quirk 1980). “Organizational reputation” refers to a set of symbolic beliefs held by audience networks as to the actual performance of an organization, as well as its capacities, roles, and obligations to accomplish its primary organizational mission.³ Bureaucratic reputations “are valuable political assets—they can be used to generate public support, to achieve delegated autonomy and discretion from politicians, to protect the agency from political attack, and to recruit and retain valued employees” (Carpenter 2002, 491). “Reputation uniqueness,” according to Carpenter (2001, 5), refers to the demonstration by agencies that they can create solutions (e.g., expertise, efficiency) and provide services (e.g., moral protection) found nowhere else in the polity.

Reputation concerns are brought to the fore in view of the potentially catastrophic consequences of enforcement failures, and the potential reputational losses that they may entail. Studies in social and political psychology suggest that negative information carries greater weight in people’s judgment than an equivalent positive message, a finding known as “negativity bias” in the literature on impression formation (for a review, see Soroka 2006), or “loss aversion” in behavioral economics (Tversky and Kahneman 1981). Moreover, the sociological literature suggests that the ubiquity of risks is a dominant spirit of modern society (Beck 1992; Kasperson 1992). Under these conditions, officeholders are expected to do their utmost in order to avoid or diffuse blame (Hood 2002, 2011; Weaver 1986).

Yet resolving this concern is not as simple as it may seem to a “selectively fatalistic” audience (Sunstein 1999). In regulatory enforcement, the agency typically faces a trade-off between a set of risks that stems from underenforcement, and another set of risks that increases by overenforcement.

While this type of trade-off is known and established, it has mostly been studied in relation to the content of regulatory decisions (e.g., Bendor and Moe 1985; Brehm and Gates 1997; Hammond and Knott 1996; Huber and Shipan 2002; McCubbins and Schwartz 1984; Moe 1982, 1985; Weingast and Moran 1983; Wood 1988).⁴ However, this trade-off has an important bearing on decision duration. In order to minimize the risk of both over- and underenforcement, regulators strive to maximize enforcement accuracy—that is, where actual enforcement minimally deviates from truly appropriate enforcement for a given regulated activity. For an agency operating within a dynamic environment, it is very difficult to evaluate
whether a violation has occurred, to assess its risks, and to quickly act upon this assessment, while maintaining a perfect enforcement system whose decisions are all accurate. Enforcement accuracy is a function of expertise, resources, information, and time. Assuming that the first three factors are relatively stable, decision time provides the variability required to enhance accuracy under given conditions. Accurate decisions require time, and speedy response takes a toll in reduced accuracy. However, time is not independent from actual enforcement. A delay of enforcement decision is equivalent to (heretofore) underenforcement, thereby demoting the agency’s consumer protection reputation. However, expediting such decisions increases the risk of overenforcement. Thus, in situations like this, the regulator has to make a trade-off between decision duration and the risk of enforcement failures (either by omission or commission) in order to obtain a cost-effective system.

Of the works undertaken so far on decision duration analysis (e.g., Carpenter 2001, 2002, 2003; Carpenter et al. 2009; Carpenter and Grimmer 2009; Carpenter, Zucker, and Avorn 2008; Gersen and O’Connell 2008; Kosnik 2005; Whitford 2005; Yackee and Yackee 2010), this trade-off is manifested especially at the core of the works of Carpenter, as well as Carpenter et al., on drug approval. The challenge that this line of research faces is the difficulty of addressing the issue of waiting costs in regulatory areas wherein delays may have fatal consequences for citizens, and as a result, a significant reputational loss for the agency. Carpenter’s (2002) findings show that the FDA’s weighing of the harm resulting from unavailability of a drug in the course of drug approval appears to be influenced by interest group and media pressure. We, therefore, maintain that it is reasonable to expect that the agency’s weighing of the harm that results from an alleged regulatory violation in the course of regulatory enforcement will be at least as sensitive to such exogenous factors.

Against this premise, it is reasonable to consider the response of a reputation-sensitive regulator—insofar as its duration of enforcement decisions are concerned—to perceived threats to its reputation. As the analysis centers on the responsiveness of the agency to public agenda signals, the analytical framework draws on scholarship that informs answers to questions related to agency’s response to routine demands (Moe 1982, 1985; Weingast and Moran 1983; Wood 1988; Wood and Waterman 1991), the role of information flows and organizational routines in shaping bureaucratic behavior (Arrow 1974; Feldman and Pentland 2003; Hammond 1986; Hammond and Thomas 1989), and works related to agency response to new or exceptional demands revolving around the way organizational attention is organized in public bureaucracies (May, Workman, and Jones 2008).

Reputational considerations, in conjunction with the shorter time horizons of national governments operating within a blame culture (Hood 2002), drive regulators to adapt their responses to the numerous “moving parts” within and outside an agency, as well as to their strategic need
to limit the type of failures that are most threatening to their reputation (i.e., to produce a blame avoidance reaction). Faced with the aforementioned trade-offs, a reputation-sensitive agency is expected to take the maximum advantage of a given environment—that is, to scan the regulatory environment for reputational risks and opportunities—while satisfying the main needs of the regulator’s diverse organized interests and the media. Regulators may be motivated to select the desired enforcement “qualities” (e.g., decision speed) and adapt to reputational risks (e.g., negative media coverage regarding over- or underenforcement) and opportunities (e.g., positive media coverage) whenever they move from one “corrective order” (e.g., notice of inspectional observation, warning letter, administrative license action, recall, orders of retention, etc.) to another (Bendor and Moe 1985).

Drawing on this analysis, the current article explicitly assesses the facet of reputation under threat—over- or underregulation—by qualitative content analysis of media coverage of the FDA’s regulatory activity, and draws two clearly distinct hypotheses based on this assessment. Moreover, drawing on recent advances in the study of blame avoidance (Hood et al. 2009; Sulitzeanu-Kenan 2010), we provide separate measurements for media valence and media salience, thereby estimating their independent and interactive effects on regulatory decision duration. In doing so, we contribute to the existing literature on the role of organizational reputation in bureaucratic decision making by studying an agency’s sensitivity to reputational threats (valence) in various levels of public attention (salience).

It is reasonable to expect that for a reputation-sensitive agency, regulatory response duration will be relatively short under conditions of intense media coverage that is negative in tone and focuses on underenforcement on the part of the agency, as compared to less intense media coverage, or intense coverage with a positive tone. This is because such an agency will have an incentive to counter criticism and to avoid enforcement failures when its activities are in the media limelight by demonstrating visible and easily understandable evidence for a decline in its enforcement response duration (or by avoiding exposure of delays).

\[\text{Hypothesis 1: If threat to reputation pertains to underenforcement, decision duration will be positively associated with media valence (i.e., negative valence pertaining to underenforcement will be associated with shorter decision duration).}\]

If, however, critical media coverage is focused on overenforcement, an inverse relationship is expected. The more critical such coverage is, the longer the duration for taking an act of enforcement will be, as the risk of overenforcement bears a greater threat to the agency’s reputation.

\[\text{Hypothesis 2: If threat to reputation pertains to overenforcement, decision duration will be negatively associated with media valence (i.e., negative valence pertaining to overenforcement will be associated with longer decision duration).}\]
However, if media criticism has a causal effect on the agency’s decision duration, this relationship should be conditional upon media salience. Media salience in this context is a measure of the extent to which our main independent variable—media valence regarding the agency’s enforcement activity—actually comes to bear. Under minimal media coverage (just enough to establish media valence), media valence is expected to have an insignificant effect on decision duration, while under intense media coverage media valence is expected to exert its greatest effect. Moreover, since reputational concerns are not expected to solely dominate the activity of the agency, it is expected that media criticism, either directed at under- or overenforcement, will have an impact on the agency only when overall press coverage of the regulator’s enforcement activity is intense. More formally, we hypothesize that media salience moderates the effect of media valence.

Hypothesis 3: In either of the above relationships, the effect of media valence will be moderated by media salience. The effect of media valence will be stronger under high media salience, and weaker under low media salience.

Ultimately, we hypothesize that reputational concerns create a responsive mechanism that produces institutional outcomes (decision durations) over and above the content of rules, guidance, procedures, structures, or statutes. Our attention turns now to the methodological section.

Research Design

This research analyzes the effect of reputational threats on enforcement activity by gauging the duration for issuing warning letters by the FDA following on-site inspections. All the warning letters issued between 1998 and 2008 by two divisions of the CEDR—the Division of Manufacturing and Product Quality (DMPQ) and the Division of Scientific Investigations (DSI)—are included in the analysis (136 cases). Warning letters are issued when there are serious violations (such as repetitive misconduct) or a potential for serious health risks to the public. They require the regulatee (manufacturers, clinical research facilities, review boards, etc.) to satisfy the regulator that it is on the road to compliance, and they warn that should the regulatee fail to respond, it may be subject to additional action. Warning letters reflect significant institutional investment on the regulator’s part and are approved at a high organizational level. Therefore, they are serious and threatening documents (Cooper and Fleder 2005, 479). In the context of a competitive market, the issuance of a warning letter is also a de facto sanction for the regulatee. The publication of a warning letter may entail harsh implications for the survival of medium and small firms, and may constitute a significant blow to large ones.

“Warning letters” replaced “regulatory letters” in 1991, a practice created in the mid-1970s by which virtually all letters were reviewed by the Office of Chief Counsel. An attempt to empower field districts and
headquarters with more latitude was initiated in 1991 by the then-commissioner, David Kessler, who introduced warning letters, which did not require review by the Office of Chief Counsel. District offices had the authority to issue warning letters directly to firms when inspection outcomes did not meet FDA regulation. Only letters regarding enforcement activity overseas remain within the ambit of the Office of Chief Counsel. Yet, on November 29, 2001, the FDA commissioner announced that all warning letters should be preapproved by the Office of Chief Counsel. The idea was to review the letters for legal adequacy and consistency with agency policy.

The warning letters analyzed here cover six established regulatory domains, namely, U.S. clinical investigations; foreign clinical investigations (FCI); institutional review boards (IRB); sponsors, monitors, or contract research; good manufacturing/laboratory practices (GMP/GLP); and vivo bioequivalence. Clinical investigations are experiments in which a drug, device, or biologic is administered or dispensed to, or used by, human subjects. IRB are volunteer committees that oversee the ethics of human studies and monitor their scientific value. They comprise at least five people that include at least one nonscientist and one person with no formal ties to the research institution. “Sponsors, monitors or contract research” refers to a person or a corporation that provides financial, managerial, and/or technical or scientific support to the pharmaceutical/biotech industry in areas such as product development, management of clinical trials, selection or monitoring of investigations, evaluation of reports, preparation of material to be submitted to the FDA, and so on. “Good manufacturing/laboratory practices” (GMP/GLP) are guidelines that outline the aspects of production that would affect product quality. And finally, “vivo bioequivalence” refers to experiments that examine whether a drug is therapeutically equivalent to an existing patented drug.

Why focus on warning letters of all enforcement actions available? Warning letters are tools for providing firms with an opportunity to voluntarily make appropriate corrections without harsher enforcement action. For this reason, they are among the first corrective actions to be undertaken and therefore a seismograph of changing considerations by the organization initiating the corrective action. Why focus solely on on-site inspections? A regulatory agency operating in an environment that gives much weight to cost-effective considerations may reorient its monitoring activities toward low-cost actions, such as learning through Web site products, reviewing flashcards, reviewing retail sale sheets, reviewing professional sales aid and journal advertisement, and so on. Estimating the extent to which decision duration is affected by reputation concerns requires scholars to neutralize cost-effectiveness considerations. This can be done by concentrating our analysis on on-site inspections—some of which are low-cost actions (e.g., inspections of clinical trials undertaken in the United States) while others require substantial investment (e.g., inspections of foreign clinical trials).
Much of the previous research on the topic of warning letters addresses the frequency, rather than the decision duration, of warning letters. These studies have been conducted by economists (Stewart and Neumann 2002), public health specialists (Salas et al. 2008), bioethicists (Bramstedt 2004), and practitioners (Wilshir 2002). This research employs duration analysis to ascertain the effect of reputational considerations on FDA regulatory enforcement activity. CDER’s decisions were investigated by a number of scholars (e.g., Carpenter 2002, 2004; Olson 1997), yet to the best of our knowledge, no study has ever accounted for variation at the subunit (divisions) level of this center. The FDA warning letters database contains all warning letters sent by the FDA’s centers and the Office of Enforcement, and can be accessed via the FDA’s Electronic Freedom of Information Reading Room.6

Data and Measurement

The data for this article were gathered from three sources: (1) FDA warning letters database (1998–2008), (2) LexisNexis media database, and (3) U.S. Senate official Web site.7 Summary statistics are provided in Table 1.

Decision Duration

The dependent variable in our study is the time from the end of an inspection until a warning letter is issued. The analysis is based on the entire set of 136 warning letters issued by the DSI and DMPQ divisions of CDER over a period of 11 years, between 1998 and 2008. The time between the end of the inspection and the issuance of a warning letter was recorded for all the cases in days.8 The relevant dates were extracted from each of the warning letters. Figure 1 presents the Kaplan–Meier estimate of the survival curve for the time-to-warning letter variable in our study.

Reputation-Related Measurements and Variables

The data used for estimating the effect of the level of threat to the agency’s unique reputation rely on the quantity and content of press coverage, and their variations over time. Using the LexisNexis media database, we have gathered all the articles published by U.S. newspapers (including the Herald Tribune) in the relevant years (1997–2007) that mention the keywords “FDA” and any of the defined inspection topics (or their shorthand names, e.g., “good manufacturing practices” or “GMP”)—a total of 164 articles.9 The articles from the automated searches were then assessed for their actual relevance.10

Before moving on to the quantitative analysis, a qualitative evaluation of the content of press coverage was utilized to assess whether it was indeed directed at the dominant unique reputation of the FDA. In other words, this analysis was used to assess whether critical press
coverage related to underenforcement (i.e., threatens the agency’s unique reputation), overenforcement (e.g., as overly aggressive to the industry), or both. This analysis has shown that apart from very few articles, all critical press articles regarding these activities of the FDA direct their criticism at what appears to be failings of the agency to live up to its unique reputation as protector of public safety. Appendix A provides some examples of these critical claims. Given this finding, Hypothesis 2 is no longer relevant, and the following analysis is intended to assess the validity of Hypotheses 1 and 3.

Next, the press coverage data were coded to create two variables: yearly press valence and yearly press coverage. For the valence measure, all the articles were coded as neutral, positive, or negative in valence on the basis of the title, subtitle, and content of each piece. A random sample of 20 articles was coded independently by three coders. Treating the

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Descriptive Statistics</th>
<th>Mean (SD)</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warning letter decision time</td>
<td>230.26 (187.87)</td>
<td>29</td>
<td>1,171</td>
<td></td>
</tr>
<tr>
<td>Media valence</td>
<td>0.482 (0.348)</td>
<td>0.021</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Dummy: High media salience (= 1)</td>
<td>0.404 (0.493)</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Dummy: DSI division (= 1)</td>
<td>0.640 (0.482)</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Division staff: DSI</td>
<td>39.5 (4.7)</td>
<td>36</td>
<td>52</td>
<td></td>
</tr>
<tr>
<td>DMPQ</td>
<td>45.3 (14.2)</td>
<td>29</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td>Dummy: Centralized legal approval (= 1)</td>
<td>0.728 (0.447)</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Dummy: U.S. clinical investigations (= 1)</td>
<td>0.346 (0.477)</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Dummy: Foreign clinical investigations (= 1)</td>
<td>0.015 (0.121)</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Dummy: IRB (= 1)</td>
<td>0.140 (0.348)</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Dummy: Sponsors, monitors or contract research (= 1)</td>
<td>0.0221 (0.147)</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Dummy: Good manufacturing/laboratory practices (= 1)</td>
<td>0.434 (0.497)</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Dummy: Letter addressed to a company</td>
<td>0.654 (0.477)</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Dummy: Letter addressed to individual medical practitioners (= 1)</td>
<td>0.125 (0.332)</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Dummy: Letter addressed to nonprofit organizations (= 1)</td>
<td>0.206 (0.406)</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Dummy: Letter addressed to IRB (= 1)</td>
<td>0.015 (0.121)</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Number of violations</td>
<td>4.788 (3.952)</td>
<td>1</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td>Number of interactions</td>
<td>2.147 (1.358)</td>
<td>1</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Democratic representatives</td>
<td>214.257 (11.391)</td>
<td>202</td>
<td>233</td>
<td></td>
</tr>
<tr>
<td>Share of Democrats in Senate</td>
<td>0.468 (0.023)</td>
<td>0.440</td>
<td>0.500</td>
<td></td>
</tr>
<tr>
<td>Share of Democrats in Senate in the Health, Education, Labor, and Pensions Committee</td>
<td>0.462 (0.033)</td>
<td>0.420</td>
<td>0.500</td>
<td></td>
</tr>
</tbody>
</table>

SD, standard deviation; DSI, Division of Scientific Investigations; DMPQ, Division of Manufacturing and Product Quality; IRB, institutional review boards.
coding data as ordinal showed acceptable levels of intercoder reliability (Krippendorff’s $\alpha = 0.743$). The average press valence for each year was then calculated based on the proportion between the difference in the number of positive and negative articles and the total number of articles in that year. The resulting valence scale was adjusted to vary in the positive range between 0 and 1. Each case (warning letter) was assigned the value of the yearly average press valence for the calendar year that preceded the inspection to which it referred ($MVpy$).

It is hypothesized that media salience moderates the effect of media valence on the FDA’s decision duration. For this reason, a dummy variable ($High_{MS}$) represents calendar years with above median number of press articles, and the interaction term ($High_{MS} \times MVpy$) is included in the analysis.

Organizational Variables

A dummy variable measured 1 for DSI and 0 for DMPQ accounted for the division that issued the warning letter. Enforcement decisions duration is expected to vary between the two divisions due to the different nature of the activities these divisions regulate. When the DMPQ monitors methods, facilities, and technical procedures (e.g., production and packaging), it is typically done against clearly defined regulations containing
minimum requirements (e.g., current good manufacturing practices). By contrast, when the DSI monitors the treatment of human subjects (e.g., their safety, welfare, and rights), it relies on statutory and ethical requirements that require broader discretion in their application in particular cases (e.g., clinical trials involving pregnant HIV-infected women). Moreover, monitoring the treatment of human subjects may require the probing of human-run institutions (e.g., committees) in order to gauge their capacity and effectiveness. All in all, it is hypothesized that enforcement decision duration involving the mechanics of a production processes is more likely to be shorter than the duration of decisions that follows noncompliance involving human research subjects.

Since the two divisions’ resources are expected to have an impact on the time required to take an enforcement decision, the analyses control for the (current) number of staff in each division in each case. The yearly number of inspections conducted by CDER serves as a proxy for the division workload. In order to account for the November 2001 (re)centralization of legal approval of all warning letters, a dummy variable, measured 1 for centralized legal approval (pre-2002 foreign inspections and all post-2002 inspections) and 0 otherwise, was included in the analysis. Finally, in order to control for organizational memory regarding previous enforcement activity, the analysis includes a dummy variable, measured 1 for cases with a reference to previous violation by the regulated organization, and 0 for cases with no such reference.

**Task Variables**

In order to control for inspection topic, our coding follows the six categories employed by CDER for this purpose: (1) U.S. clinical investigations; (2) FCI; (3) IRB; (4) sponsors, monitors, or contract research; (5) GMP/GLP; and (6) vivo bioequivalence. Five dummy variables account for these categories in the analysis. The type of addressees of the warning letters were also controlled for based on the following categories: (1) companies, (2) individual medical practitioners, (3) nonprofit organizations, and (4) IRBs. Three dummy variables account for these categories in the analysis. A proxy for the danger parameter is the number of violations mentioned in each of the warning letters, based on the number of headings describing observed violations. Additionally, the number of interactions between the regulated body and the CDER division—measured by the number of times a firm formally approaches the agency in an attempt to satisfy it that the firm is on the road to compliance.

**Oversight Variables**

In order to account for political oversight, two types of variables were employed. The first gauges the level of congressional attention to the FDA and relies on the annual number of testimonies of FDA officials before the
House and Senate Committees and subcommittees in the preceding year for each case. The second set of variables accounts for the partisan element of oversight institutions. It includes the relative proportion of Democrats in the House of Representatives, in the Senate, and in the Senate Health, Education, Labor, and Pensions Committee, which oversees the FDA. Another dummy variable scored 1 if a Democrat was president. Due to strong correlations among these four partisan variables, they were included separately in the analysis.

**Statistical Analysis**

We have tested our hypotheses using maximum likelihood duration analysis. The decision whether to issue a warning letter offers serious implications for regulated institutions, and bears serious consequences to the supply of medical products. Thus, the predicted form of the hazard function entails only a few cases of short duration before issuance of a warning letter, followed by a rising propensity, and then diminishing, as most inspections do not end in an issuance of a warning letter. Figure 2, indeed, demonstrates that the natural log transformation of warning letter

**FIGURE 2**

Logged Duration Distribution

![Log Duratioan Distribution](image.png)

*Note:* The bars represent actual distribution of logged duration, and a smoothed kernel density plot is shown against a normal plot.
decision durations included in this study follows a normal distribution. For these reasons, a log-normal distribution is used as a functional form (Kleinbaum and Klein 2005).

An important aspect of the analysis is the need to account for potential variability among the cases due to unobservable variables. Our main data source—the warning letters—are censored by the FDA before becoming publicly available. Important information, such as the subject of the inspection, the nature of the flaws that were found, their health implications, etc., is deleted from the text. In order to account for this case-specific heterogeneity, the analysis includes a “frailty” random component (Kleinbaum and Klein 2005). The inverse Gaussian distribution was selected as the frailty distribution.

Findings
Specification of the various log-normal regression models had to account for multicollinearity in the data. Strong and significant correlations were found between organizational and task variables, reputation and partisan oversight variables, and between reputation and time trend (valence becomes more negative over time).

Hypotheses 1 and 3 were tested in nine log-normal decision duration models, reported in Tables 2 and 3. Table 2 presents four log-normal duration analyses with reputational, organizational, and task control variables. Model 1 presents a barebones specification of the associations between press valence and salience, and decision duration. Model 2 adds the interaction between press valence and press salience. The coefficient of the High-Salience ¥ Valence-Prv-Year interaction term in model 2 is positive and statistically significant ($P < 0.001$), suggesting that the association between media valence and decision duration varies significantly across media salience, and thus provides support for Hypotheses 1 and 3. Controls for organizational variables and for task variables are reported in models 3 and 4, respectively. These controls appear to contribute to the predictive ability of the model. Yet the coefficients for the High-Salience ¥ Valence-Prv-Year interaction term remain substantively similar and statistically significant across these model specifications, indicating a positive association between media valence and decision duration when media coverage is high. The estimated association between media valence and decision duration when media coverage is low is less consistent. Models 2 and 4 suggest that when media coverage of FDA enforcement activity is low, decision duration increases as media valence grows more negative. However, this association is not statistically significant in model 3. In fact, in the six models that control for organization factors (3, and 5–9), no association was found between media valence and decision duration when media coverage is low. This is important, as these models provide the best fit to the data.
TABLE 2
Log-Normal Decision Duration Analyses with Reputational, Organizational, and Task Variables

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reputation variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Valence-Prv-Year</td>
<td>−0.603* (0.239)</td>
<td>−0.898*** (0.233)</td>
<td>0.217 (0.260)</td>
<td>−0.986*** (0.200)</td>
</tr>
<tr>
<td>High-Salience</td>
<td>−0.016 (0.184)</td>
<td>−0.710** (0.223)</td>
<td>−0.622*** (0.159)</td>
<td>−0.954*** (0.205)</td>
</tr>
<tr>
<td>High-Salience × Valence-Prv-Year</td>
<td>−</td>
<td>2.461*** (0.608)</td>
<td>2.373*** (0.400)</td>
<td>2.701*** (0.583)</td>
</tr>
<tr>
<td>Organizational variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Division (DSI)</td>
<td>−</td>
<td>−</td>
<td>1.233*** (0.126)</td>
<td>−</td>
</tr>
<tr>
<td>Division staff</td>
<td>−</td>
<td>−</td>
<td>0.016* (0.008)</td>
<td>−</td>
</tr>
<tr>
<td>Inspection intensity</td>
<td>−</td>
<td>−</td>
<td>−0.003** (0.001)</td>
<td>−</td>
</tr>
<tr>
<td>Legal approval centralization</td>
<td>−</td>
<td>−</td>
<td>0.901*** (0.168)</td>
<td>−</td>
</tr>
<tr>
<td>Previous violation</td>
<td>−</td>
<td>−</td>
<td>−0.112 (0.132)</td>
<td>−</td>
</tr>
<tr>
<td>Task variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U.S. clinical investigation</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>0.727* (0.333)</td>
</tr>
<tr>
<td>IRB</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>0.443 (0.338)</td>
</tr>
<tr>
<td>Foreign clinical investigation</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−0.393 (0.383)</td>
</tr>
<tr>
<td>Sponsors</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>1.100** (0.328)</td>
</tr>
<tr>
<td>GMP</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−0.001 (0.325)</td>
</tr>
<tr>
<td>To MD</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>0.135 (0.205)</td>
</tr>
<tr>
<td>To NPO</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>0.212 (0.153)</td>
</tr>
<tr>
<td>To IRB</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>0.479** (0.176)</td>
</tr>
<tr>
<td>Number of violations</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>0.002 (0.009)</td>
</tr>
<tr>
<td>Interactions</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>0.035 (0.037)</td>
</tr>
<tr>
<td>Constant</td>
<td>5.461*** (0.191)</td>
<td>5.671*** (0.175)</td>
<td>5.216*** (0.596)</td>
<td>5.278*** (0.323)</td>
</tr>
<tr>
<td>ln_p</td>
<td>−0.337** (0.108)</td>
<td>−0.356*** (0.052)</td>
<td>−0.594*** (0.063)</td>
<td>−0.527** (0.063)</td>
</tr>
<tr>
<td>ln_theta</td>
<td>−2.897 (2.641)</td>
<td>−16.397*** (0.394)</td>
<td>−16.811*** (0.563)</td>
<td>−</td>
</tr>
<tr>
<td>Log pseudolikelihood</td>
<td>−151.882 (2.641)</td>
<td>−144.576 (2.621)</td>
<td>−112.242 (2.641)</td>
<td>−117.777 (2.641)</td>
</tr>
<tr>
<td>Cases</td>
<td>136</td>
<td>136</td>
<td>136</td>
<td>136</td>
</tr>
<tr>
<td>Observations (days)</td>
<td>31,316</td>
<td>31,316</td>
<td>31,316</td>
<td>31,316</td>
</tr>
</tbody>
</table>

*P < 0.1, **P < 0.05, ***P < 0.01, ****P < 0.001; robust standard errors in parentheses.
GMP, good manufacturing practices; MD, individual medical practitioners; NPO, nonprofit organizations; IRB, institutional review boards.
### TABLE 3
Log-Normal Decision Duration Analyses with Organizational and Oversight Variables

<table>
<thead>
<tr>
<th></th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reputation variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Valence-Prv-Year</td>
<td>0.374</td>
<td>0.421</td>
<td>0.230</td>
<td>0.206</td>
<td>0.453</td>
</tr>
<tr>
<td></td>
<td>(0.266)</td>
<td>(0.285)</td>
<td>(-0.274)</td>
<td>(0.275)</td>
<td>(0.642)</td>
</tr>
<tr>
<td>High-Salience</td>
<td>-0.731***</td>
<td>-0.945*</td>
<td>-0.479*</td>
<td>-1.517*</td>
<td>-0.720***</td>
</tr>
<tr>
<td></td>
<td>(0.155)</td>
<td>(0.400)</td>
<td>(0.272)</td>
<td>(0.628)</td>
<td>(0.181)</td>
</tr>
<tr>
<td>High-Salience × Valence-Prv-Year</td>
<td>2.814***</td>
<td>3.087***</td>
<td>2.595***</td>
<td>4.199***</td>
<td>2.747***</td>
</tr>
<tr>
<td></td>
<td>(0.423)</td>
<td>(0.660)</td>
<td>(0.458)</td>
<td>(1.187)</td>
<td>(0.649)</td>
</tr>
<tr>
<td>Organizational controls</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Oversight variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Congressional attention</td>
<td>0.016+</td>
<td>0.018+</td>
<td>0.014</td>
<td>-0.016+</td>
<td>0.017</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.010)</td>
<td>(0.010)</td>
<td>(0.009)</td>
<td>(0.013)</td>
</tr>
<tr>
<td>Senate committee–Democrats share</td>
<td>-</td>
<td>2.350</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Senate–Democrats share</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Congress–Democrats share</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Democrat president</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-0.065 (0.487)</td>
</tr>
<tr>
<td>Constant</td>
<td>4.653***</td>
<td>3.461+</td>
<td>7.171**</td>
<td>1.175</td>
<td>4.607***</td>
</tr>
<tr>
<td></td>
<td>(0.702)</td>
<td>(2.054)</td>
<td>(2.149)</td>
<td>(2.656)</td>
<td>(0.797)</td>
</tr>
<tr>
<td>Ln_p</td>
<td>-0.602***</td>
<td>-0.604***</td>
<td>-0.606***</td>
<td>-0.609***</td>
<td>-0.603***</td>
</tr>
<tr>
<td></td>
<td>(0.062)</td>
<td>(0.062)</td>
<td>(0.062)</td>
<td>(0.061)</td>
<td>(0.062)</td>
</tr>
<tr>
<td></td>
<td>(0.511)</td>
<td>(0.515)</td>
<td>(0.515)</td>
<td>(0.496)</td>
<td>(0.538)</td>
</tr>
<tr>
<td>Log pseudolikelihood</td>
<td>-111.048</td>
<td>-110.858</td>
<td>-110.541</td>
<td>-110.153</td>
<td>-111.037</td>
</tr>
<tr>
<td>VIF (multicollinearity)</td>
<td>2.95</td>
<td>4.97</td>
<td>4.25</td>
<td>9.18</td>
<td>6.97</td>
</tr>
<tr>
<td>Cases</td>
<td>136</td>
<td>136</td>
<td>136</td>
<td>136</td>
<td>136</td>
</tr>
<tr>
<td>Observations (days)</td>
<td>31,316</td>
<td>31,316</td>
<td>31,316</td>
<td>31,316</td>
<td>31,316</td>
</tr>
</tbody>
</table>

*P < 0.1, *P < 0.05, **P < 0.01, ***P < 0.001; robust standard errors in parentheses.

VIF, variance inflation factor.
Table 3 presents a series of analyses that include controls for oversight and partisan factors. These models extend model 3, which provided the best fit to the data. Most of the oversight variables correlated and thus were controlled for in separate regressions. Some of the oversight variables are also associated with reputation variables, and therefore introduce problems of multicollinearity. Our approach here is to present the results while indicating the extent of this problem for each analysis (variance inflation factor).

Model 5 introduces congressional attention to the decision duration analyses. Models 6–9 add a measure for the proportion of Democrats on the Senate Health, Education, Labor, and Pensions Committee, the share of Democrats in the Senate, the share of Democrats in the Congress, and a dummy variable for democratic president. None of the oversight variables present a clear association with the agency’s enforcement decision duration. These apparently inconsistent and statistically weak results are not conducive to a clear conclusion regarding the effect of partisan oversight factors on the FDA’s enforcement activity. However, the High-Salience × Valence-Prv-Year interaction term remains substantively stable and statistically significant across these model specifications.

As noted above, model 3 provides the best fit to the data and therefore is used here for estimating the relationships between media valence and salience, and enforcement decision duration. The significant interaction of these variables attests to the theoretical and methodological importance of separating the measurements of valence and salience. The estimates of model 3 were selected for plotting Figure 3, which graphically depicts the relationship between media valence and enforcement decision duration across low and high media salience. The distinctly different slopes of the two plots represent the diverging effect of media valence in different levels of media salience, as suggested by Hypothesis 3. A rise of one standard deviation in yearly media valence (from 0.308 to 0.656) under conditions of high media salience is associated with an average increase of 543 days (from 372 to 915) in enforcement decision duration. However, under low media salience such a shift in yearly media valence is associated with an average increase of merely 26 days (from 333 to 359), a statistically insignificant difference.

The findings suggest that some of organizational variables have an effect on decision duration. As expected, decision duration for issuing warning letters is (on average) longer at the DSI compared to the DMPQ. Somewhat surprisingly, the number of division staff was found to have a positive association with decision duration, although this association was not statistically significant across all the specifications. Importantly, while media valence was found to predict the total number of staff in the two divisions over the duration of this study ($R^2 = 0.61, \ P = 0.004, \ N = 11$), the associations between media valence and decision duration remain stable and statistically significant when controlling for division staff, suggesting that resource allocation alone does not account for the association between
media valence and decision duration. The measure of workload represented by the number of inspections per year conducted by CDER was found to be positively associated with decision duration. The application of centralized legal approval increases decision duration, as could be expected. However, previous violations by the regulated organization were not found to be associated with decision duration.

Most of the task variables were not found to be associated with decision duration. Still, the findings suggest that enforcement decisions in the domain of “US clinical investigations”; “sponsors, monitors, or contract research”; nonprofit organizations (marginally significant); and IRB have relatively longer durations.

Robustness Tests

In order to further assess the validity of the findings regarding the effect of media valence on decision duration, four sets of robustness tests were conducted, and are reported in detail in Appendix B. The first is a cure (split-population) model, which is a special type of survival analysis model where it is assumed that there is a proportion of cases that will never experience the event. In this study, most inspections do not end in

![Figure 3: The Relationship between Media Valence and the Duration of Enforcement Decisions across Low and High Media Salience](image-url)
the issuance of a warning letter, and thus accounting for them in the
analysis is important. However, information on inspections that do not
end in enforcement is sparse and is limited to the number of inspections
per year. However, based on the year, these cases can be attributed with
preceding media valence and salience, thus allowing us to estimate model
2 (Table 2) with a cure model, as reported in the left column of Append-
ix B. The substantive results are similar to those found in models 2
through 9 (Tables 2 and 3), although somewhat stronger.

Next are three alternative estimations of model 3 (Table 2). The first uses
a Cox proportional hazard model. While overall model fit is inferior com-
pared to the log-normal model (log likelihood = −491.680), the associa-
tions between the independent variables and decision duration are
substantively similar in both models, as reported in the second column of
Appendix B.

The final two regressions are estimates of model 3 with alternative
measures of decision duration. Since our measure of decision duration
is in days, while all other independent variables are given in yearly or
time-fixed values, these analyses assess whether our findings are consist-
tent when measuring decision duration in larger time units. Therefore,
decision duration was converted to months (rounded) and quarters
(rounded). As reported in the two right-hand columns of Appendix B,
all the substantive findings were found to be statistically significant
in these analyses, attesting to the robustness of the decision duration
measurement.

Discussion

This research provides evidence that reputational concerns play a variable
role in determining the duration of enforcement decision by the FDA,
depending on the level of visibility of reputational challenges. Our pre-
liminary qualitative analysis has shown that nearly all media criticism
(negative valence) refers to underenforcement (thus rendering Hypoth-
esis 2 irrelevant). The log-normal duration analyses provide support for
Hypotheses 1 and 3, suggesting that the agency’s sensitivity to reputa-
tional threats is limited to occasions when such threats are more likely to
be noticed. This finding appears to be robust in various model specifica-
tions, and when controlling for various organizational, task, and oversight
variables. Organizational and task variables were generally found to have
an association with decision duration, while partisan oversight variables
yielded, at best, mixed results. To conclude, as media coverage of the
FDA’s primary consumer protection responsibilities becomes more posi-
tive, the FDA waits longer to issue warning letters; in contrast, more
critical coverage leads to quicker action in issuing warning letters by the
FDA. Our empirical analysis provides support for the notion that the
admittedly vague concept of reputation may translate into concrete
regulatory outcomes. It does so by revealing, more selectively, the
conditions under which waiting becomes costly to the reputation-sensitive regulator facing negative media coverage.

We should note that this analysis relies on observational data, in which the main dependant variable—media coverage—cannot be assumed to be exogenous, and thus our findings can only provide support for hypothetical causal relationships but not establish them conclusively. While the possibility of reversed causation has been addressed by using press coverage data that preceded the enforcement process, other confounding relationships and omitted variable bias cannot be eliminated in such a research design. Furthermore, the period covered by this study (1998–2008) was characterized by increasing criticism of the FDA for failures in its consumer protection mission. It is possible that inclusion of the 1980s and early 1990s would have revealed more media criticism of overregulation, and therefore may have resulted in different results.

The findings of this research have bearing on the prevailing view of bureaucratic passivity found in congressional dominance and bureaucratic politics literature. Our analyses suggest that press coverage of the FDA appears to alter regulatory activity directly, while influencing congressional attention to the regulator. These findings suggest that what might sometimes appear to be congressional dominance may actually be driven by media agenda that shapes both policy agenda—congressional attention to the FDA—and regulatory activity as a result of reputation protection motivation.

Empirically, our analyses of the decision duration of the FDA’s corrective actions provide us with insight into how regulatory agencies select their decision speed and adapt to reputational risks (e.g., negative media coverage regarding over- or underenforcement) and opportunities (e.g., positive media coverage). Importantly, reputational challenges appear to influence both the allocation of agency resources (staff variations) and decision duration, while the latter does not appear to be mediated by the former (decision duration was found to be strongly associated with media valence even when controlling for division staff figures).

This article also points to an additional and more subtle way in which an agency can handle conflicting demands, goals, and expectations from within the organization as well as from its environment. An agency’s enforcement policy is composed of inspections and enforcement activity at the conclusion of inspections. However, another important policy tool often ignored by scholars and practitioners alike is the temporal gap between these two. Although less visible than inspections and corrective actions at the conclusion of inspections, this gap offers an effective way to handle conflicting demands from within the organization, from the political arena, and from other social actors by varying its organizational output in a responsive way. For example, a long-term increase in the number of inspections as a result of pressures by politicians and the public for more corrective actions can mask the strategic delays in the issuance of warning
letters. Inasmuch as public pressure is expressed in the press, the agency demonstrates its responsiveness by altering its decision duration.

An alternative explanation for these findings may rely on shifts in regulatory attention to performance shortfalls as a result of processing press coverage information, rather than simply being the product of reputation-protecting behavior. Let us first identify the central difference between these two explanations. While both theoretical accounts have much in common, they are distinct from one another in one important way. Information processing relies on the availability of performance information, while assuming the motivation to act upon it is constant. Reputation protecting behavior, on the other hand, also relies on the availability of performance information but is variably motivated to act upon it based on the visibility of this information to the agency’s relevant audiences. Our finding that media valence is only influential when media coverage is high provides more credence to the reputation protection hypothesis; that is, performance information is acted upon only when it is publicly visible.

Moreover, and perhaps more importantly, the variance in our dependent variable—decision duration—more appropriately represents a reputation-protecting behavior, rather than resulting from decision content adjustment on the basis of new information. Since nearly all underenforcement criticism is directed at actual activity (lack of sufficient inspections) or decision content, its effect on enforcement duration is not a likely outcome if this media input is merely treated as an additional source of performance information.

By providing separate measurements for media valence and media salience, our analyses contribute to the existing literature on the role of organizational reputation in bureaucratic decision making in various levels of public attention and visibility. The findings, indeed, show that the agency’s sensitivity to reputational threats is limited to occasions when such threats are more likely to be noticed. Yet identifying a condition in which the agency is not sensitive to external evaluation of its performance provides an insight into the role of reputation in regulatory activity that, to the best of our knowledge, has not been addressed in previous studies.

Our point of departure was that reputational threats directed at the FDA’s primary consumer protection responsibilities result in shortening the duration of enforcement decisions when these threats receive high media attention. Although this appears to be so, an implicit assumption in this article pertains to the duration of such decisions in the absence of external reputational challenges. Under this condition, it was (latently) assumed that the agency is “reputationally relaxed”; hence, enforcement decision durations are typically long. Yet this expectation is not consistent with our findings, as depicted by Figure 3, which suggest that enforcement decision duration is typically short for periods in which media coverage of the agency regulatory activity is low. This pattern suggests that FDA enforcement divisions are basically “reputationally concerned” and
thus do not shorten their enforcement decision duration under reputational challenges but rather extend the time for taking action when their reputation is visibly praised.

For a graphic demonstration of this distinction, consider Figures 4A and 4B. Figure 4A presents a conditional relationship between media content and salience, and enforcement decision duration in a schematic way. The solid line represents the relationship between media valence and enforcement decision duration when media salience is high, and the dotted line represents the absence of an effect when media salience is low. It is clear that when media salience is low, decision duration is as long as
when media salience is high and its content positive. The agency then shortens its time until enforcement only if press coverage is salient, and increasingly so the more critical it becomes. Such a situation suggests that the organization’s baseline attitude is “reputationally relaxed.” On the other hand, Figure 4B presents a different situation. While the conditional relationships are similar, it is clear in this case that when media salience is low, decision duration is as short as when media salience is high and its content critical. The agency then extends its time until enforcement only if press coverage is salient, and increasingly so, the more positive it becomes. Such a situation suggests that the organization’s baseline attitude is “reputationally concerned.”

This distinction extends the existing research interest regarding the relationships between reputation and bureaucratic outcomes, and offers a closer understanding of the context in which these relationships operate. It opens the possibility of studying the variance in “baseline reputational sensitivity” from which exogenous positive or negative shocks may shift outcomes. Such studies require varying this baseline by comparing across agencies, within agencies across time, or both.

At a more practical level, our findings imply that interest groups and other actors wishing speedy decisions of corrective actions should keep feeding negative stories to the press so that the agency never receives disproportionately or unqualifiedly positive press. In contrast, firms, organizations, and individuals privy to impending bad news following an agency inspection may be well advised to actively plant seeds for positive stories in the media, as well as for allegations of overenforcement, to gain more time to control damages.14

While this article has explicitly analyzed the moderating effect of public visibility on the responsiveness of an agency to reputational challenges, it should be noted that the FDA is a single case of a regulator characterized by an exceptionally powerful reputation. This raises the question of whether a focused study on one unique agency is of any value for understanding other organizations. We contend that the international regulatory arena in the field of drug safety does not offer “a larger population sufficiently homogenous in so many respects that we would care generalize about . . .” (Carpenter 2010a, 20). Given the unique FDA’s role in global health, we share Carpenter’s view that “there is a value in studying a singular process not because it stands in for so many others, but because it differs so radically and starkly from others, so much so that the act of comparison is itself problematic,” and . . . because it influences so many others” (Carpenter 2010a, 20). Still, the analysis presented here can be modified and applied to other national regulators in order to investigate similar hypotheses.

In this respect, a comparative consideration can be added. The reputational considerations of the FDA should be similar to other types of agencies operating in liberal democracies that are in the business of making government decisions involving uncertainty and risks.
Some agencies may develop their desired basis of reputation, some may “shadow” expert regulators, and others may be inclined to develop a basis of reputation that is more suitable to the particular environment within which they are operating (Maor 2011). But when the media is centrally controlled, the various symbols that represent the organization may be manipulated. When the structure of the organization’s relationships to its different audiences is closed, the organization will not be able to shape symbolic beliefs and embed them in its various audiences. The more open and competitive the interface of subject and audience, the interface of regulator and regulated, the more intertwined reputation and bureaucratic politics are. Under such conditions, each agency may defend its preferred operational areas that are determined by the nature of the symbolic beliefs held by the audiences and the structure of the organization’s relationships to its various overlapping audiences.

Another caveat of our article stems from the fact that other reputational threats (e.g., allegations of overenforcement or of being too harsh/unfair toward the industry) may be leveled at the agency through channels other than the press. It is conceivable that such threats indeed exist, yet we cannot account for them in this article. Further research should be conducted to assess the relative weight, and possible interactions, of multiple sources of reputational appraisals under varying degrees of visibility. To conclude, it has been demonstrated that reputational concerns create a responsive mechanism that produces institutional outcomes (decision durations) over and above the content of rules, guidance, procedures, structures, or statutes.

Acknowledgments

The authors would like to thank Sharon Gilad, Micha Mandel, Dan Miodownik, Tamir Sheafer, Shlomi Segall, Mike Ting, the editors of Governance, and three anonymous reviewers for their excellent insights and thoughtful comments.

Notes

1. I thank Sharon Gilad for raising this question.
2. “Valence” refers to whether a message/signal is negative or positive. It is commonly used in psychology (Forgas 1995), behavioral economics (Elster 1998), and communication studies (Sheafer 2007).
3. This definition is based on Ruef and Scott (1998) and Carpenter (2001, 2010a).
4. For a review of the studies on the content of regulatory decisions, see Carpenter (2003).
5. Moreover, only warning letters that are issued following on-site inspections contain the precise dates of the inspection, that along with the dates the letters were issued provide the value of decision duration.
6. This Web site was revised during 2009, and a number of warning letters from earlier years that were not included in the Web site prior to this revision were added to it. We assume that this database is now complete.
7. See https://www.senate.gov/index.htm (September 28).
8. The maximum delay for the issuance of a warning letter in our data is 1,171 days (3 years and 76 days) from its respective inspection.
9. For more details on these categories, see subsection “Task Variables.” Although the discussion regarding the impact of print media versus audio-visual media still continues in the literature, most studies support the primacy of newspaper agenda-setting effect on political institutions (Walgrave, Soroka, and Nuytemans 2008).
10. Nonrelevant articles are those wherein the object of search does not appear fully in the text, that is, one or more word/s describing the object of search appear/s in isolation from the rest of the term’s words and/or in different contexts throughout the text. The proportion of relevant articles from automated searches varies between 92% in the case of IRB and 15% in the case of FCI.
11. \[ MV = \frac{\sum \text{positive} - \sum \text{negative}}{\sum \text{positive}, \text{neutral}, \text{negative}} \]. This scale fully accounts for the three-level (positive, negative, neutral) coding of the press articles. The resulting scale varies between −0.39 and 0.07, hence occupying both the positive and negative domains of the range, yet exhibiting asymmetric press responsiveness to negative events (Soroka 2006). However, for convenience, it was adjusted to vary between 0 and 1.
13. It is possible that this association reflects an alternative causal path. For example, it is possible that a more effective division at a particular period may have lower decision durations, and thus have more time to increase the number of inspections conducted. It should be noted that the use of temporally preceding measures of press valence eliminates the potential for inverse causation.
14. Such steps may complement existing ones, such as attempts to persuade the agency that the company is, or soon will be, in compliance, and that a finding of past noncompliance should not lead the agency to expect future noncompliance by the company.

References


Appendix A

Notable Examples of Critical Press Coverage of the FDA

“In a sharply critical report, federal investigators say the Department of Health and Human Services has made only ‘minimal progress’ in protecting the many thousands of participants in clinical trials, despite prior warnings that the system for safeguarding patients was seriously flawed.” ...”The Food and Drug Administration and the National
Institutes of Health, the HHS arms responsible for oversight of clinical trials . . .” (USA Today April 12, 2000).

“The FDA’s paper review of foreign trials is ‘very late in the pipeline to worry about the consumer or the patients in those trials, . . . And when the FDA reviews a drug this way, its essentially saying ‘caveat emptor’ ['buyer beware'] when the drug is approved for sale” (Washington Post December 18, 2000).

“Half the FDA’s budget for the evaluation of new drugs now comes from drug company users’ fees, making the agency dependent on the industry it regulates—an obvious conflict of interest” (Washington Post June 20, 2001).

“Congressional investigators are looking into whether the FDA erred in not issuing warnings sooner” (Washington Post October 16, 2004).


“The withdrawal of the popular painkiller Vioxx from market in September because of safety concerns shows that federal regulators need a better way to pinpoint potentially hazardous drugs, said the lead author of a journal article that describes a complementary private drug-monitoring system” (The Boston Globe May 4, 2005).

“The Food and Drug Administration’s office responsible for policing the safety of prescription drugs is underfunded, understaffed, and lacks a ‘clear and effective’ process for deciding whether and how to act when it determines a drug is unsafe . . .” (The Boston Globe April 25, 2006).

“The FDA, charged with protecting the public, acknowledges that it seldom checks up on doctors doing trials” (St. Petersburg Times [Florida] September 2, 2007).

“The Food and Drug Administration does very little to ensure the safety of the millions of people who participate in clinical trials, a federal investigator has found” (The New York Times September 28, 2007).
Appendix B

Estimating the Association between Media Valence and Enforcement Decision Duration—Robustness Tests

<table>
<thead>
<tr>
<th></th>
<th>Cure Model Estimation (Model 2)</th>
<th>Estimating Model 3 with Decision Duration in Months</th>
<th>Estimating Model 3 with Decision Duration in Quarters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cox Proportional Hazard Model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reputation variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Valence-prv-year</td>
<td>−0.765* (0.345)</td>
<td>0.343* (0.172)</td>
<td>0.235 (0.268)</td>
</tr>
<tr>
<td>High-valence</td>
<td>−1.648*** (0.340)</td>
<td>2.83** (1.031)</td>
<td>−0.608*** (0.157)</td>
</tr>
<tr>
<td>Valence-prv-year</td>
<td>4.358*** (0.792)</td>
<td>0.013*** (0.015)</td>
<td>2.322*** (0.400)</td>
</tr>
<tr>
<td>High-salience</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High-salience × valence-prv-year</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organizational variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Division (DSI)</td>
<td>0.090*** (0.031)</td>
<td>1.241*** (0.129)</td>
<td>1.150*** (0.116)</td>
</tr>
<tr>
<td>Division staff</td>
<td>0.961* (0.017)</td>
<td>0.015* (0.008)</td>
<td>0.012* (0.007)</td>
</tr>
<tr>
<td>Inspection intensity</td>
<td>1.007*** (0.002)</td>
<td>−0.003** (0.001)</td>
<td>−0.003** (0.001)</td>
</tr>
<tr>
<td>Legal approval</td>
<td>0.154*** (0.055)</td>
<td>0.930*** (0.178)</td>
<td>0.810*** (0.152)</td>
</tr>
<tr>
<td>centralization</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Previous violation</td>
<td>1.450* (0.296)</td>
<td>−0.120 (0.138)</td>
<td>−0.107 (0.105)</td>
</tr>
<tr>
<td>Constant</td>
<td>4.411*** (0.246)</td>
<td>1.825*** (0.618)</td>
<td>0.841** (0.498)</td>
</tr>
<tr>
<td>Ln_p</td>
<td>−0.566*** (0.070)</td>
<td>−0.736*** (0.060)</td>
<td></td>
</tr>
<tr>
<td>Ln_theta</td>
<td>−16.425*** (0.441)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log pseudolikelihood</td>
<td>−1518.753</td>
<td>−491.680</td>
<td>−116.058</td>
</tr>
<tr>
<td>Cases</td>
<td>7,472</td>
<td>136</td>
<td>136</td>
</tr>
</tbody>
</table>

*p < 0.1, *p < 0.05, **p < 0.01, ***p < 0.001; robust standard errors in parentheses.