Ruling by Repute: Agency Reputation on Judicial Affirmance of Agency Action

ELIZABETH GUO*

ABSTRACT

Courts rely on legal doctrine to decide cases, but attitudinal factors influence judicial review. Although agency reputation is not an explicit part of legal doctrine, judges have noted that agency reputation may have a subconscious hold on courts when deciding whether to rule in an agency's favor. To date, however, little empirical literature exists on actual differences in agency reputation or on how agency reputation may affect legal challenges to agency action.

This Article seeks to fill that empirical gap. This Article develops a series of tests to determine whether and how agency reputation affects judicial outcomes. Using data from the U.S. Food and Drug Administration (FDA) and the Environmental Protection Agency (EPA), the Article presents evidence that reputation before the judiciary differs between agencies. Courts view agencies differently regardless of whether a court decides to validate or invalidate agency action in a particular case. The evidence also suggests that an agency's long-term reputation is positively correlated with judicial validation rates; better reputations are associated with better judicial outcomes for an agency. This Article presents suggestive evidence that the relationship between agency reputation and judicial outcomes does not affect the frequency of judicial action.

INTRODUCTION

Federal agencies in the United States interact with and respond to external audiences, including Congress, the public, and other agency overseers and their constituents. One of the most important external audiences for any federal agency is the federal judiciary, which has the power to deny or uphold agency action. Because the judiciary can define the limits of agency authority, agencies attend to their relationships with the judiciary, mindful that judicial arbiters can dispose of a carefully crafted rule or overturn decades of agency policy.

The nature of the relationship between an agency and the judiciary rests on legal and attitudinal factors. Attitudinal factors include political affiliation and policy preferences, and scholars have studied the extent to which policy preferences form the basis of judicial decisions.[1] Other attitudinal factors, such as reputation, have received

* Ms. Guo is an associate at Covington & Burling LLP in the food, drug, and device group. The views presented are those of the author and are not those of Covington & Burling LLP or any of its clients. I am grateful to I. Glenn Cohen, Daniel Carpenter, Lewis Grossman, Richard Kingham, Rachel Sachs, Matthew Stephenson, and Crystal Yang for their helpful comments on earlier drafts.

[1] *See, e.g.*, JEFFREY A. SEGAL & HAROLD J. SPAETH, THE SUPREME COURT AND THE ATTITUDINAL MODEL REVISITED (2002); *see also* Jason J. Czarnezki & William K. Ford, *The Phantom Philosophy? An*

less attention, despite legal scholars suggesting that such factors can influence judicial decisions in ways that fall outside of traditional legal or political divides.

Agency reputation is not an explicit consideration in legal doctrine, but judges have hinted that it may have at least a subconscious hold on the judiciary.[2] Former D.C. Circuit Judge Patricia Wald once remarked:

> In my twelve years on the D.C. Circuit, I have reviewed hundreds of administrative actions by the [EPA], the [FERC], [OSHA], the Department of Health and Human Services (HHS), the [FCC], the National Labor Relations Board (NLRB), and other agencies . . . . I hold suspicions or impressions about who is on top or at the bottom of the regulatory honor roll.[3]

More recently, Justice Breyer, speaking about cases involving the Federal Trade Commission (FTC) stated, "You can't really understand these decisions very well unless you understand that the court just has confidence in certain agencies and it doesn't have confidence in certain others and we're not going to say which is which."[4] Former clerks to judges on the D.C. Circuit have echoed these sentiments by anecdotally noting that judges in the D.C. Circuit treat some agencies differently from others.[5]

Legal scholars support judges' anecdotal accounts. Christopher Edley writes that agencies with strong reputations fare better before the courts than agencies with weak reputations.[6] The Securities and Exchange Commission (SEC) holds a reputation as a "well-staffed professional agency," which predisposes courts to grant it greater deference in decisions related to agency action.[7] Elizabeth Garrett has posited that agencies such as the NLRB receive less deference in part because "its reputation

---

Empirical Investigation of Legal Interpretation, 65 MD. L. REV. 841, 847–55 (2006) (discussing the legal versus the attitudinal model of judicial review).

[2] *See* DANIEL CARPENTER, REPUTATION AND POWER: ORGANIZATIONAL IMAGE AND PHARMACEUTICAL REGULATION AT THE FDA 362 (2010) (noting that in issuing opinions in favor of FDA, "[Judge] Friendly was recognizing a different form of legal precedent, deferring not to the explicit language of other rulings (as in *stare decisis*), but to the flavor or temperature of those rulings and the implied portraits of the administrative agencies they embedded"). The relationship between reputation and agency outcomes extends outside the judiciary. *See, e.g.*, Jason MacDonald & William Franko, Jr., *Bureaucratic Capacity and Bureaucratic Discretion: Does Congress Tie Policy Authority to Performance?* 35 AM. POLITICS RESEARCH 790, 790–807 (2007) (finding empirical evidence that Congress is less likely to constrain administrative discretion for agencies with high agency grades in the Federal Performance Project).

[3] Patricia M. Wald, *The "New Administrative Law"—With the Same Old Judges In It?*, 1991 DUKE L.J. 647, 662 (1991).

[4] Erica Teichert, *Breyer Gives Antitrust Agencies Top Marks for EU Ties*, LAW 360 (Apr. 3, 2014, 7:31 PM), http://www.law360.com/articles/52485 l/print?section=appellate [https://perma.cc/XV6P-ZGA5].

[5] David Zaring, *Reasonable Agencies*, 96 VA. L. REV. 2317, 2360 (2010).

[6] *See* CHRISTOPHER EDLEY, ADMINISTRATIVE LAW: RETHINKING JUDICIAL CONTROL OF BUREAUCRACY 49 (1992).

[7] *Id.; see also* Bevis Longsreth, *The SEC after Fifty Years: An Assessment of Its Past and Future*, 83 COLUM. L. REV. 1592, 1608 (1983) (highlighting the SEC's "solidly based reputation for regulatory vigor and excellence in regulation").

makes it suspect in some quarters."[8] Likewise, agencies such as the Federal Election Commission and the Immigration and Naturalization Service receive less deference before the courts because of their "problematic reputations."[9]

To date, sparse empirical literature exists on actual differences in agency reputation before the courts or on how agency reputation may affect judicial challenges to agency action. As Edley notes, "Both the appraisal and decisional significance of reputation are unconfessed, unpoliced, and undisciplined through . . . effective empirical scrutiny."[10] No quantitative evidence exists on how courts perceive different agencies. Other than anecdotal commentary from certain judges, scholarly literature has not examined whether all agencies have similar reputations nor explored how agency reputation may affect judicial outcomes.

In this Article, I seek to move beyond anecdotal accounts to begin to fill the empirical gap. Borrowing from Nuno Garoupa and Tom Ginsburg's definition of reputation, I define reputation as "the aggregate of judgments used to predict future performance."[11] I use quantitative methods to assess whether agencies have different reputations before the courts and whether those reputations affect judicial outcomes. This question goes to the heart of judicial review. A reputational explanation of judicial review would suggest that the effects of prior agency actions could have belated effects on cases in the present. A reputation-based account of judicial decision-making would conclude that a court filters its assessment of the legality of an agency's actions through its subjective feelings about an agency's past actions.

I test this question by examining the reputations and judicial outcomes of two executive agencies in the United States: the U.S. Food and Drug Administration (FDA) and the Environmental Protection Agency (EPA). My empirical strategy is to treat a quantifiable measure—judicial attitude toward an agency in legal opinions—as a proxy for reputation. When courts rule on challenges to agency action, the opinions contain descriptive assessments of agency behavior that often convey the court's attitude, tone, and feelings towards an agency. As further explained below, I develop a quantifiable metric, "reputation score," to capture the court's attitude towards an agency in judicial opinions. One obvious difficulty with this approach is reverse causation. I measure reputation score by analyzing the court's attitude toward an agency in written opinions, so there is a high likelihood that judicial outcome influences the reputation score for each case. Whether a judge decides to rule for or against an agency in a particular case will likely affect how the judge describes the agency in an opinion. If I were to examine any particular opinion, it would be difficult to untangle how a judge's attitude towards an agency affects the outcome of the case and how the outcome of the case affects the judge's attitude towards the agency in a written opinion.

---

[8] Michael J. Hayes, *After "Hiding the Ball" is Over: How the NLRB Must Change Its Approach to Decision-Making*, 33 RUTGERS L. REV. 523, 565 (2002).

[9] Elizabeth Garrett, *Legislating* Chevron, 101 MICH. L. REV. 2637, 2651 (2003); *see also* Samuel Êstreicher, *Policy Oscillation at the Labor Board: A Plea for Rulemaking*, 37 ADMIN. L. REV. 163, 170–75 (1985) (discussing how the perception of the NLRB as an unstable agency means that "courts are reluctant to pay little more than lip service to the doctrine of deference to agency policymaking").

[10] EDLEY, *supra* note 6, at 49.

[11] Nuno Garoupa & Tom Ginsburg, *Reputation, Information and the Organization of the Judiciary*, 4 J. COMPARATIVE L. 228, 228 (2009).

I tackle the collinearity problem in two ways. When examining differences in reputation, I control for outcomes by measuring differences in reputation score within judicial outcomes. First, I compare FDA and EPA reputation scores from favorable decisions separately from those in unfavorable decisions. Second, when determining whether a relationship exists between judicial outcomes and reputation, I examine the relationship between outcomes and lagged measures of reputation, using lagged measures of judicial outcomes as control. For example, I compare judicial outcomes decided in each four-year interval (e.g. from 2006 to 2009) with reputation scores for cases decided in the preceding four-year interval (e.g. between 2002 and 2005), using as control judicial outcomes for cases decided in the preceding four-year interval (e.g. 2002 to 2005). The lagged measures of reputation capture an agency's pre-existing reputation and would not change based on a court's final decision in the case.

I use the quantifiable measures of reputation to conduct three analyses. First, I examine whether FDA and EPA reputations differ before federal district and circuit courts. I then examine whether any difference in reputation is associated with differences in judicial outcomes in cases concerning agency action. I consider the possibility that the relationship between reputation and agency outcomes may differ before and after 1984, the year the Supreme Court decided *Chevron U.S.A. Inc v. NRDC*.[12] Finally, I examine the relationship between agency reputation and the relative frequency of cases decided by the court. Although this association does not directly measure the relationship between agency reputation and judicial outcomes, the results of the analysis help interpret the results of the first two analyses. If there is a relationship between agency reputation and the frequency of cases decided by the courts, the results of the first two analyses could have been affected by the changing strength of challenges against the agency over time.

My findings are summarized as follows. First, the evidence suggests that agency reputation before the courts can differ between agencies, and that courts may view agencies differently despite similar judicial outcomes. EPA has a lower reputation score than FDA, even after controlling for judicial outcome. I also find evidence of a positive relationship between lagged measures of agency reputation and judicial outcomes. Long-term reputation is positively correlated with judicial validation rates for FDA and EPA, after controlling for the rate of affirmance in prior agency decisions. This empirical finding is consistent with the anecdotal observations of judges and judicial observers. The higher the long-term reputation score, the higher the rate of affirmance. Third, the evidence suggests that agency reputation may not be associated with the frequency of decisions. A change in lagged measures of reputation for FDA and EPA has little effect on the rate that different types of cases are decided before the courts. This result does not necessarily mean that litigant behavior is unreceptive to changes in agency reputation. As discussed further below, the result could signify that both agencies and litigants are highly attuned to the attitude and tone of judicial decisions and adjust their litigation and regulatory strategies accordingly.

The rest of the Article proceeds as follows. In Section I, I provide a background of reputation and the literature around reputation and judicial outcomes. In Section II, I describe the methodology used to measure reputation. I then describe the results of each analysis in Section III. I conclude in Section IV by describing the legal ramifications of my findings. The results in the Article support anecdotal accounts

[12] Chevron U.S.A., Inc. v. NRDC, 467 U.S. 837 (1984).

describing the role of agency reputation in judicial review. Agency reputation may be a factor that agency officials and litigants should consider when deciding whether to take regulatory action or when litigating challenges to agency action.

I. BACKGROUND ON REPUTATION

Agency reputation consists of a set of beliefs about an agency's future actions based on its past actions.[13] Reputation encompasses information about an agency's "capacities, intentions, history, and mission that are embedded in a network of multiple audiences."[14] Reputation encapsulates an agency's past performance and captures signals provided by the agency in the present.[15] For a federal agency, reputation is simultaneously developed and reinforced by external audiences (e.g., elected officials, industry, consumer groups, ordinary citizens, and media) and internal audiences (e.g., employees and executive leadership).[16] Agency reputation can also change over time.[17] Changes to agency reputation could occur gradually or suddenly, due to highly publicized events.[18]

Agency reputation serves multiple functions. To an agency's external audiences, reputation performs an important informational role, allowing them to predict an agency's future actions based on the agency's past actions.[19] Agency reputation can communicate what an agency intends to do and the quality with which it can accomplish its mission. To internal audiences, an agency's reputation is central to its members' identities and self-esteem and can affect whether an agency's members feel attached to the agency. These internal perceptions may be further reinforced by the members' interactions with external audiences.[20]

All federal agencies have reputations to maintain among several audiences, but this Article focuses on agency reputation before the courts. Federal regulatory agencies are repeat players before federal courts and must interact with them on a regular basis. The courts can serve as a major source of power for agencies, by affirming their actions and allowing them to seize property, prohibit activities, or impose monetary penalties

---

[13] *See* Rachel Brewster, *Unpacking State's Reputation*, 50 HARV. INT'L L.J. 231, 235 (2009) (defining reputation as "a belief about the state's future actions based on its past actions"). As Brewster notes, this definition of reputation is one used in economic and political science literature; *see also* Reinhard Selten, *The Chain Store Paradox*, 9 THEORY & DECISION 127 (1978) (economics; using reputation as the monopolist's history of driving entrants out of the market as a deterrent for future entrants into the market); JONATHAN MERCER, REPUTATION AND INTERNATIONAL POLITICS 6 (1996) (political science; defining reputation as "a judgment of someone's character (or disposition) that is then used to predict or explain future behavior").

[14] Daniel Carpenter & George Krause, *Reputation and Public Administration*, PUB. ADMIN. REV., Feb. 2012, at 26.

[15] *See* Garoupa & Ginsburg, *supra* note 11, at 231.

[16] CARPENTER, *supra* note 2, at 26.

[17] Jennifer Nou, *Interagency Design*, 129 HARV. L. REV. 421, 458 (2016) (noting that EPA's reputation changed during Anne Gorsuch's tenure as head of the agency. Gorsuch had a reputation as a foe of robust environmental protection policies.).

[18] *See* CARPENTER, *supra* note 2, at 261 (arguing that the thalidomide scandal in Europe and Frances Kelsey's role in refusing to approve thalidomide in the United States "symbolically multiplied" FDA's reputation as a "cop, gatekeeper, and protector").

[19] Brewster, *supra* note 13, at 235.

[20] *See* William Kovacic, *Building an Agency Brand*, 22 GEO. MASON L. REV. 238, 255 (2015).

against regulated entities. The judiciary can also serve as a shield, protecting the agency from external actors seeking to access the agency's confidential communications or validating the agency's refusal to grant requests. Conversely, courts can reduce an agency's power by restricting its authority or jurisdiction, invalidating its actions, or preventing the agency from enforcing its decisions.[21]

For the most part, judicial power over an agency lies in whether courts will decide cases in favor of or against the agency. Although legal doctrine plays a primary role in determining whether judges will affirm the decisions of administrative agencies,[22] judges have significant discretion. Judicial discretion may lead to some judges favoring interest groups that align with the judges' policy preferences.[23] For example, Cass Sunstein and Thomas Miles observe that ideological preferences play a role in judicial review of agency decisions for arbitrariness.[24] Their empirical examination of arbitrariness review for the NLRB and the EPA found that Democratic appointees show liberal voting patterns and Republican appointees show conservative voting patterns, with patterns amplified when appointees sit on all-Democratic or all-Republican panels.

Judicial discretion could also allow courts to favor groups with better reputations. Scholars often assume that agencies have different reputations before the judiciary and that these differences affect judicial outcomes, particularly in cases where judges have discretion to defer to an agency's decisions. Daniel Carpenter suggests, for example, that FDA's strong reputation has affected the judiciary, leading courts to adopt a "multidecade record of deference" to FDA's decisions.[25] As Judge Fred Parker asserts, the potential effect of institutional reputation on judicial outcomes means that federal agencies are careful to avoid actions before the judiciary that would harm their reputation because "the institutional credibility of these offices with the courts is largely their stock in trade."[26]

---

[21] In this Article, I treat the judiciary as one entity, though I recognize that an agency may have different reputations among different members of the judiciary. As Garoupa and Ginsburg have noted, the judiciary functions collectively and as a result of the decisions of its individual judges. While an individual judge may act in ways that affect his personal preferences, a judge, as a member of the judiciary, also reflects the attributes and nature of the judiciary as a whole. For example, an individual judge's decision could flow from his personal ideological preferences but may also be tempered by the image that a judiciary as a whole hopes to project on its external audiences. Nuno Garoupa & Tom Ginsburg, *Hybrid Judicial Career Structures: Reputation Versus Legal Tradition*, 3 J. LEGAL ANALYSIS 411, 416 (2011).

[22] *See, e.g.* E. Donald Elliot & Peter H. Schuck, *To the Chevron Station: An Empirical Study of Federal Administrative Law*, 1990 DUKE L.J. 984, 1057 (1990) (finding that courts validated agency decisions at a high rate in the time period just after the Supreme Court issued *Chevron U.S.A. v. NRDC.*); *c.f.* Zaring, *supra* note 5, at 2353 (finding that courts validated agency action at similar rates regardless of the degree of deference a court a court was supposed to accord to the agency).

[23] R. SHEP MELNICK, REGULATION AND THE COURT: THE CASE OF THE CLEAN AIR ACT 343 (1983).

[24] Thomas Miles & Cass Sunstein, *The Real World of Arbitrariness Review*, 75 U. CHI. L. REV. 761, 768 (2008).

[25] CARPENTER, *supra* note 2, at 45–60; *see also* David Zaring, *Regulating by Repute*, 110 MICH. L. REV. 1003, 1006 (2012) (stating that "Carpenter is right to recognize that reputation gets agencies breathing room from courts and Congress").

[26] Fred I. Parker, *Foreword: Appellate Advocacy and Practice in the Second Circuit*, 64 BROOK. L. REV. 457, 462 (1998); *see also* Melissa Wasserman, *Deference Asymmetries: Distortions in the Evolution of Regulatory Law*, 93 TEX. L. REV. 625, 670 n. 141 (noting that an agency's reversal by the court "may also cause reputational harms" and "diminish its credibility before the court").

Scholars have proposed a number of frameworks to capture how agency reputation may affect judicial outcomes. William Kovacic argues that agencies have different "brands" before the courts, which signal institutional quality. Courts use agency "brands" as a proxy for an agency's ability to do "strong substantive work and us[e] sound methods to make policy."[27] Agencies with good reputations before the courts have "substantive programs, sound procedures (e.g., meaningful disclosure of information, rigorous testing of evidence, and regular assessment of outcomes), strong capabilities . . . and a healthy culture."[28] A positive reputation translates into judicial deference for the agency; an agency with a "good reputation" is more likely to gain the benefit of the doubt from a reviewing court than an agency with a "weak reputation."[29]

James O'Reilly takes a similar, agency-wide view of agency reputation but believes that agency reputation may hinge on high-profile cases before the judiciary. O'Reilly argues that FDA's behavior during a number of high-profile, highly-politicized decisions has injured its reputation and decreased the level of deference it receives from courts.[30] O'Reilly argues that courts traditionally "have been quick to give deference to FDA" because of their confidence in FDA's "reputation for superior science and expertise."[31] The deference granted to FDA, and its "reputation for careful preparation of cases[,]" meant that Department of Justice prosecutors routinely concurred with FDA's enforcement discretion.[32] This deference declined in the early 2000s, O'Reilly argues, as a result of several "politically motivated" decisions that "tarnished [FDA's] reputation for scientific impartiality."[33] These cases gained broad media coverage.[34] In the years since the early 2000s, FDA "frequently lost cases where deference previously would have aided the Agency."[35]

This Article intends to provide quantifiable evidence to determine whether any of the theoretical frameworks on agency reputation have empirical validity. The Article uses a new data set to first examine whether there is evidence to support the underlying assumption that FDA and EPA have different reputations before the courts. The Article then attempts to provide estimates of the impact of FDA's and EPA's on judicial affirmance before the courts. Although this Article cannot provide irrefutable proof of the accuracy of any of the frameworks presented previously, the results of each analysis may help scholars restructure or refine their theoretical frameworks analyzing agency reputation.

---

27 Kovacic, *supra* note 20, at 238.

28 *Id.*

29 *Id.* at 239.

30 James T. O'Reilly, *Losing Deference in the FDA's Second Century: Judicial Review, Politics, and a Diminished Legacy of Expertise*, 93 CORNELL L. REV. 939, 943 (2008).

31 *Id.* at 949.

32 *Id.* at 952.

33 *Id.* at 964.

34 *Id.* at 954.

35 *Id.* at 973.

II. METHODOLOGY

I use a new data set to empirically examine whether reputation may have an effect on judicial decisions evaluating agency action. By distinguishing between cases involving FDA and EPA and between cases related to questions of law and findings of fact, I try to understand whether and how agency reputation may affect judicial outcomes.

A. Data Set

To produce the data set, I examine, analyze, and organize judicial opinions for administrative law cases challenging FDA and EPA action[36] in federal district and circuit courts between 1972 and 2014.[37] I examine all cases challenging FDA action between 1972 and 2014. Because there were thousands of cases challenging EPA action between 1972 and 2014, I randomly selected ten cases from each year that met the inclusion criteria to include in the data set. Collecting ten random cases per year ensured that I had enough variety between cases to complete each analysis.

The final data set includes cases challenging FDA or EPA actions that meet certain inclusion criteria. Cases in the data set must have been decided on the merits. As I discuss later in this Article, it is possible that agency reputation might have some bearing on whether a judge might dismiss a case on procedural grounds (e.g., lack of jurisdiction and no standing).[38] However, because the outcomes of dismissed cases generally do not rely on judicial deference to agency action, they did not pertain to the outcome I want to measure.

If a court decided a case on the merits, I determine whether the case primarily concerned administrative law. I generally include in my final data set cases that hinge on administrative law but exclude two groups of administrative law cases. First, I exclude cases deciding issues peripheral to an agency's core mandate. For example, I

---

[36] I chose FDA and EPA for my analysis because both are repeat players before a range of federal courts. To minimize the effect of confounders, I wanted to examine two agencies with similar organizational and hierarchical structures and a similar range of administrative law cases (i.e., questions of law vs. findings of fact) before the court. There is no indication that judges would view reputation differently for executive and independent agencies, and anecdotal scholarly accounts of agency reputation seem to treat executive and independent agencies interchangeably. However, it is possible that the results from this analysis may not apply as well to agencies that do not directly report to the President.

[37] I focus on cases before the federal district and circuit courts. Studies have examined agency validation rates before the Supreme Court, looking at the effect of agency ideology, judge ideology, and the nature of the agency. *See, e.g.*, Bradley Canon & Micheal Giles, *Recurring Litigants: Federal Agencies Before the Supreme Court*, 25 W. POL. Q. 183 (1972). However, district and circuit courts are more suitable for three reasons. First, agencies are likely to be repeat players before district and circuit courts, and may only argue a case before the Supreme Court once every few years. This is especially true for an agency like FDA, which fields a low number of cases every year. Examining Supreme Court decisions would not provide enough observations to draw conclusions. Second, because Supreme Court cases are rare, the cases that reach the Supreme Court are often affected by a number of factors that would bias the sample. These factors include whether the solicitor general allows the agency to appeal the decision, and whether the Supreme Court decides to grant certiorari. These issues also exist on the circuit level, but are exacerbated in the Supreme Court. Finally, although the Supreme Court has the final say, district and circuit courts hold most sway in expanding or constraining agency action because they are more likely to rule on challenges to agency action. The decisions of district and circuit courts may have less impact per case than the Supreme Court, but the large volume of cases heard in district and circuit courts mean that in practice, these courts have greater influence on agency decisions.

[38] For example, one could argue that courts might be more willing to dismiss cases challenging agencies with a good reputation before the courts.

exclude cases based on employment law and the Freedom of Information Act (FOIA). Although FOIA is also part of administrative law, judicial determinations related to FOIA are not core to an agency's mandate. Because FOIA issues are not agency-specific, it would be difficult to distinguish whether a change in reputation is due to the courts' views about a particular agency or the courts' views about agency action in general. Second, I exclude agency enforcement actions. Although enforcement actions can also implicate the extent of an agency's discretion under a statute, the validity of an agency's enforcement action implicates concepts of guilt and liability in civil and criminal law, questions that rely on different doctrinal assumptions than assumptions underlying judicial deference in administrative law.

B. Analyzing and Measuring Reputation

My research design for measuring reputation relies on written judicial opinions in cases challenging agency action. Judges rarely speak openly about their feelings toward specific agencies, but they convey their attitudes within written opinions describing or ruling on agency action. Judges make broad assumptions about the importance of an agency's mission, how an agency ordinarily conducts its business, or about the integrity of agency staff. The courts may speak glowingly of an agency's objectives and expertise. For example, in response to a challenge to FDA policies on mercury in fish, the court ruled for FDA and stated, "Due to its expertise, the Administration must be permitted flexibility in navigating the tough choices that come along with its expansive safety docket."[39]

The courts may also be highly critical or dismissive of an agency's character or modus operandi. In an opinion denying FDA's efforts to regulate electronic cigarettes as drugs or devices under the Federal Food, Drug, and Cosmetic Act, the court concluded that the case was " . . . yet another example of FDA's aggressive efforts to regulate recreational tobacco products . . . . Unfortunately, its tenacious drive to maximize its regulatory power has resulted in its advocacy of an interpretation of the relevant law that I find, at first blush, to be unreasonable and unacceptable."[40] Each opinion addressing a specific agency action may contain sentences that convey the court's general attitude toward an agency at that time.

Reputation is difficult to measure; I quantify agency reputation through a proxy coding system.[41] I review each written opinion for sentences or phrases that convey

---

[39] Ctr. for Sci. in the Pub. Interest v. FDA, 74 F.Supp.3d 295, 305 (D.D.C. 2014).

[40] Smoking Everywhere vs. FDA, 680 F.Supp.2d 62, 78 (D.D.C. 2010).

[41] Reputation is complex and difficult to quantify, and I did not find validated measures of organizational reputation that would be appropriate for this study. Scholars use different measures to quantify organizational reputation depending on the organization, data available, and research question. For example, scholars have attempted to quantify reputation for corporate and non-profit organizations using a multi-dimensional modeling approach. Manfred Schwaiger developed a model to measure corporate reputation that encompassed six dimensions: likeability, competence, quality, performance, social responsibility, and attractiveness. These dimensions derived from qualitative surveys and interviews conducted with a sample of the general population. Manfred Schwaiger, *Components and Parameters of Corporate Reputation—An Empirical Study*, 56 SCHMALENBACH BUS. REV. 46 (2004). On the other hand, scholars who have attempted to measure the reputation of government agencies have used fewer external inputs to determine agency reputation. These measures of agency reputation have differed depending on the topic in question. For example, Lee and Whitford measured the public reputation of federal agencies by using a combined index of the number of denied requests and time period to respond to requests from the agency's annual Freedom of Information Act (FOIA) report. Soo-Young Lee & Andrew Whitford, *Assessing the Effects of Organizational Resources on Public Agency Performance: Evidence from the US Federal Government*, 23 J. PUB. ADMIN. RES. & THEORY 687 (2013). Moshe Maor, however, measured the

the overall attitude of the judiciary towards each agency. For each case, I try to analyze the opinion separately from my analysis of the court's decision in the case. I assign a "reputation score" based on the overall attitude and tone of the opinion. Each opinion receives a score based on a three-point rating system.[42] Opinions that express a negative image of an agency receive a "1." Judges openly chastise agencies in "1" opinions. For example, *Beaty v. FDA* received a "1" because the opinion was highly critical of FDA, with statements including, "FDA appears to be simply wrapping itself in the flag of law enforcement discretion to justify its authority and masquerade an otherwise seemingly callous indifference to the health consequences of those imminently facing the executioner's needle. How utterly disappointing!"[43] Opinions that express neither a negative nor positive image of an agency (or both) receive a "2," or a neutral rating. Opinions where judges openly praise the agency receive a "3," or a positive rating. For example, *American Public Health Ass'n v. Veneman* received a "3" with statements including, "Because the FDA has responsibility in matters which directly and literally affect the nation's health and welfare, it is one of the most important of all Federal regulatory agencies."[44]

I do not calculate sub-reputation scores for each issue within each opinion. Although I recognize that judges could express different attitudes towards the agency between different arguments within an opinion, it was difficult to parse out the judge's tone to produce an accurate assessment of judicial reputation based on each issue litigated in the case. In cases with concurring or dissenting opinions, I analyze the controlling opinion, since these opinions speak for the views of the courts.

Figure 1 shows the distribution of reputation scores for cases challenging FDA and EPA action.

---

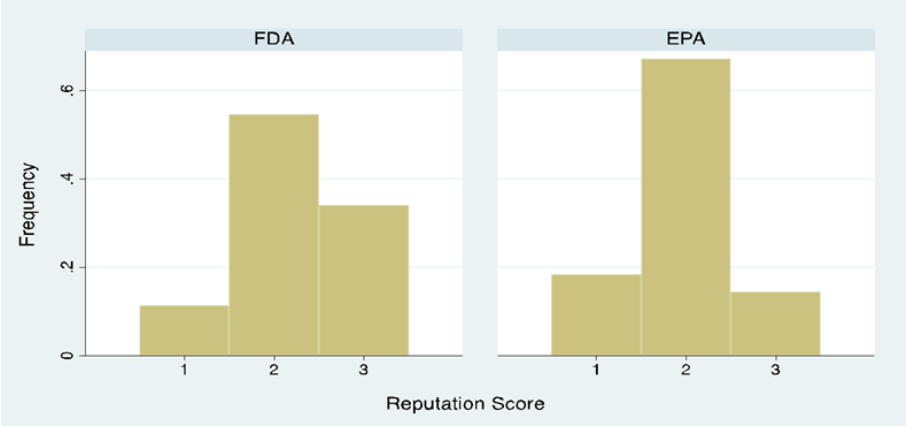
reputation of the Israeli Banking Supervision Department (BSD) between 1998 and 2009 through coding articles involving opinions—whether negative or positive—about the BSD and its performance. Moshe Maor et al., *Organizational Reputation, Regulatory Talk, and Strategic Silence*, 23 J. PUB. ADMIN. RES. & THEORY 581 (2013). While Lee's measure of agency reputation might be appropriate to serve as a proxy for "reputational resources," an agency's response to FOIA reports cannot capture the court's response to the full range of an agency's administrative decisions. Maor's measure of agency reputation provides a more general and comprehensive view of agency reputation that encompasses all aspects of an agency's administrative authority. Maor's measure is similar to the measure of agency reputation that I will use for this study.

[42] I used a three-point system to better ensure reproducibility and reliability. A five-point system may be able to identify and differentiate the harshest opinions from slightly negative critiques (e.g., "highly negative," "negative," "neutral," "positive," and "highly positive"). However, nuances in attitudes towards different agencies are hard to identify, and I did not attempt to subdivide each decision using a five-point system. Using a three-point system would likely lead to higher inter-coder reliability for an admittedly subjective characteristic.

[43] 853 F.Supp.2d 30, 43 (D.D.C. 2012).

[44] 349 F.Supp. 1311, 1317 (D.D.C. 1972).

Figure 1. Distribution of Reputation Score



C. Other Measures Related to Outcomes and Reputation

Aside from reputation, I collect a number of characteristics for each case. I categorize each opinion based on the arguments presented. Looking at each issue decided in the opinion, I determine whether litigants presented arguments involving questions of law or findings of fact. I find far fewer cases litigated on findings of fact than on questions of law. I characterize cases as involving "findings of fact" when the opinion decided any issue concerning a finding of fact, even if the opinion also discussed questions of law.[45] I characterize cases as "question of law" when the case involved questions of law and do not concern issues related to findings of fact. Even though some cases litigated more issues than others, each opinion had equal weight; I was more interested in whether a litigant brought forth a specific type of issue than the number of issues litigated before the courts.

I also collect information about the ideology of the judges hearing each case, the ideology of the agency (whether the case was litigated by a Republican or Democratic administration),[46] and whether the opinion was issued before or during and after 1984—the year when major doctrinal shifts occurred in administrative law. As a proxy for judicial ideology, I use judicial ideological scores compiled by Epstein, Martin, Segal, and Westerland.[47] For circuit court decisions issued by panels of judges, the

---

[45] Under this characterization, I could produce two mutually exclusive sets of data and avoid duplicating certain cases. The nature of challenges before the court meant that I could not invert the characterization by labeling cases as "questions of law" when the case involved both a question of law and a finding of fact. There were too few cases challenging agency action that involved only a finding of fact.

[46] I labeled the opinion "1" if it was decided during a Republican administration and "0" if decided within a Democratic administration, regardless of whether another administration initiated the case.

[47] Lee Epstein et al., *The Judicial Common Space*, 23 J.L. ECON. & ORG. 303, 303–25 (2007). The authors calculate judicial ideology based on measure using methodologies based on the NOMINATE Common Space scores and Martin-Quinn scores. NOMINATE Common space scores measure the preferences of judges that takes into consideration measures for Representatives, Senators, and Presidents who matter in judicial appointments. *See* KEITH POOLE & HOWARD ROSENTHAL, CONGRESS: A POLITICAL-ECONOMIC HISTORY OF ROLL-CALL VOTING (1997). Martin-Quinn scores provide a vote-based measure of ideology derived from justices' voting patterns. *See* Andrew Martin & Kevin Quinn, *Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953-1999*, 10 POL. ANALYSIS 134, 134–53 (2002).

judicial ideology associated with each case is the average of the ideological score for each of the judges on the panel. Judicial ideological scores range from -0.502 for the most liberal judges to 0.538 for the most conservative judges.

D. Summary Statistics

Table 1 displays summary statistics for each agency. I include all of the available cases for each outcome and further subdivide each outcome into cases categorized as a question or law or a finding of fact. Cases involving FDA have a higher validation rate than cases involving EPA. This holds true when cases are subdivided into cases involving a question of law or a finding of fact. FDA has a higher overall reputation score than EPA. As expected, cases involving FDA and EPA have similar overall agency ideology and judicial ideology scores and when subdivided by case type. I expected similar ideology scores for FDA and EPA because judges are randomly assigned to cases.

Table 1. Summary Statistics of FDA and EPA Cases

	FDA (1)	EPA (2)
Question of Law (Freq)	0.67	0.56
Finding of Fact (Freq)	0.33	0.44
Validation Rate	0.74 (0.03)	0.55 (0.02)
Question of Law	0.69 (0.04)	0.49 (0.03)
Finding of Fact	0.83 (0.04)	0.62 (0.04)
Reputation Score	2.23 (0.04)	1.96 (0.03)
Question of Law	2.10 (0.05)	1.91 (0.03)
Finding of Fact	2.49 (0.07)	2.02 (0.05)
Agency Ideology	0.53 (0.03)	0.56 (0.02)
Question of Law	0.54 (0.04)	0.6 (0.03)
Finding of Fact	0.5 (0.05)	0.51 (0.04)
Judicial Ideology	-0.04 (0.02)	0.04 (0.01)
Question of Law	-0.02 (0.02)	0.02 (0.02)
Finding of Fact	-0.07 (0.03)	0.06 (0.02)
Total Number	258	402

Means, standard errors in parentheses. Note: This table reports summary statistics of FDA and EPA cases decided in federal district and circuit courts. The cases involving FDA comprise the complete dataset of cases that meet the inclusion criteria between 1972 and 2014. The cases involving EPA include a random subset of ten cases each year between 1972 and 2014 that meet the inclusion criteria.

III. RESEARCH DESIGN AND RESULTS

I first study whether agencies have different reputation scores before the courts and then determine how the changes in reputation score are associated with agency validation rates and frequency of challenges to agency action over time.

A. Do Agencies Have Different Reputations Before the Courts?

My first step is to investigate whether different agencies have different reputations before the courts. If agencies have different reputations, I would expect judicial behavior to reflect those differing reputations, regardless of the actual outcome of a given case. I expect that a court, whether validating or invalidating agency action, to speak better of an agency with a strong reputation than an agency with a weak reputation.

I examine whether agencies have different reputations by comparing the reputation scores between FDA and EPA in cases where a court affirmed agency action and in cases where a court rejected agency action. I make the same comparison after controlling for the type of case involved (questions of law or findings of fact).

Table 2 reports average reputation scores when courts validate FDA and EPA decisions. It also reports average scores when cases are subdivided based on the predominance of questions of law and findings of fact. FDA has a higher reputation score than EPA in cases where courts validate agency action. These differences persist even after controlling for case type; differences in reputation score between EPA and FDA are similarly pronounced for cases involving questions of law and for those focused on findings of fact.

Table 3 reports average reputation scores when courts invalidate FDA and EPA decisions. Overall, the differences between reputation scores for FDA and EPA are not significant in cases where the courts invalidate agency action. The differences are also statistically insignificant regarding cases focused on questions of law. The difference for cases concerning findings of fact is significant at the one percent level. The sizeable gap between the number of cases related to questions of law and the number of cases related to findings of fact may account for this result.

Table 2. Reputation Scores of FDA and EPA in Cases Validating Agency Action

	FDA (1)	EPA (2)	Difference of (1) – (2)
Overall	2.36 (0.04)	2.15 (0.03)	0.21 (0.05)***
	[190]	[220]	
Question of Law	2.26 (0.05)	2.06 (0.04)	0.21 (0.06)***
	[119]	[111]	
Findings of Fact	2.51 (0.07)	2.24 (0.05)	0.27 (0.08)***
	[71]	[109]	

*Means, standard errors in parentheses, and number of observations in brackets. Note: * denotes difference significant at ten percent level, ** denotes difference significant at five percent level, and *** denotes difference significant at one percent level. Bonferroni correction applied.*

Table 3. Reputation Scores of FDA and EPA in Cases Invalidating Agency Action

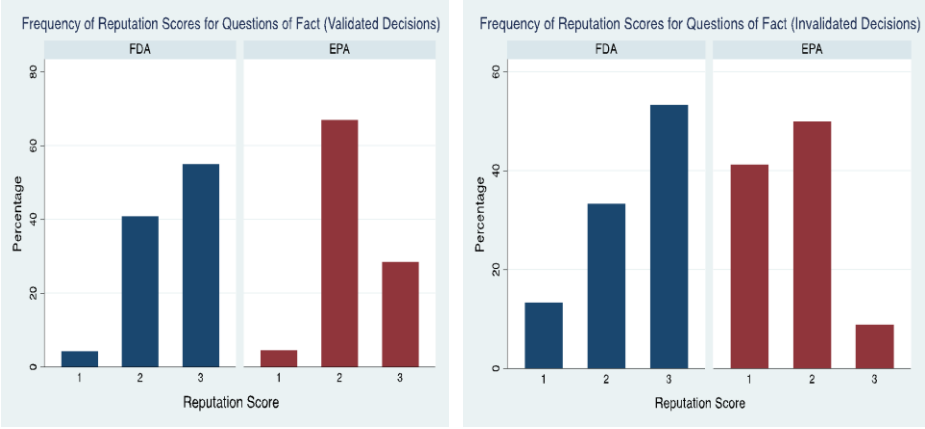
	FDA (1)	EPA (2)	Difference of (1) – (2)

Overall	1.87 (0.09) [68]	1.73 (0.04) [182]	0.14 (0.09)
Question of Law	1.72 (0.09) [53]	1.76 (0.05) [114]	-0.05 (0.1)
Findings of Fact	2.4 (0.19) [15]	1.68 (0.08) [68]	0.72 (0.19)***

*Means, standard errors in parentheses, and number of observations in brackets. Note: * denotes difference significant at ten percent level, ** denotes difference significant at five percent level, and *** denotes difference significant at one percent level. Bonferroni correction applied.*

The differences in reputation score may be attributed to the likelihood that judges may be more hesitant to write glowing EPA opinions than glowing FDA opinions, regardless of the outcome of the case. For example, Figure 2 shows the percentage of cases with a reputation score of 1, 2, or 3 for cases involving findings of fact for validating and invalidating decisions. FDA enjoys a higher percentage of opinions with the highest reputation score across outcomes.

Figure 2. Distribution of Reputation Score for FDA and EPA



Courts are also slightly more willing to write opinions with the lowest reputation score for EPA than for FDA. In cases invalidating agency action based on findings of fact, EPA receives the lowest reputation score more often than FDA. The same trend does not present for opinions that validated FDA or EPA decisions on findings of fact. This finding might be attributable to the high degree of deference that courts are supposed to grant to an agency's decisions for a finding of fact. Even if an agency has a weak reputation before the courts, the positive language used to support the high degree of deference accorded to the agency for a finding of fact may balance out any negative language the courts may use on agencies with weak reputations.

B. Are Changes in Agency Reputation Associated with Changes in Agency Validation Rates?

The second step is to investigate the relationship between agency reputation before the courts and the agency's validation rate. If such a relationship exists, I would expect a positive correlation between agency reputation and validation rate. There are a

number of ways that agency reputation, as measured by reputation score, could affect validation rate. First, short-term reputation could have a stronger influence on validation rates than long-term reputation. Under this hypothesis, agency reputation in its most recent cases will have the largest impact on the courts. Alternatively, agency reputation could affect its validation rate through a slower process. Under this framework, the agency's validation rate before the courts would not be affected by one or two strongly critical, recent decisions before the courts. However, a long line of critical decisions, accumulated over a decade, would seriously undermine the agency's reputation and could result in lower validation rates before the courts.

To test my hypothesis, I create a model that controls for a number of endogenous variables. The first is doctrine. In 1983 and 1984, the Supreme Court issued two opinions that had major implications for administrative law. *Motor Vehicles Manufacturers Ass'n v. State Farm Mutual Automobile Insurance Co*.[48] laid out the court's interpretation of arbitrary and capricious review. The opinion instructed courts to require detailed justifications for agency action and examined the reasonableness of the agency's conclusions under the appropriate standard of arbitrary and capricious review.[49] In 1984, the Supreme Court decided *Chevron U.S.A. Inc v. NRDC,*[50] which provided a two-step framework for judicial review of an agency's interpretation of a statute that it is charged with administering. Because changes in judicial review could affect both how a court describes an agency's action and the validation rate before the courts, I included only cases decided after 1984 for the analyses in this section.[51]

I employ two levels of stratification to account for other endogenous variables. The first variable divided the data by "agency." An agency with a stronger reputation may have stronger validation rates due to agency-specific effects that affect both the agency's reputation before the courts and the agency's validation rate. If there is a relationship between agency reputation and validation rate, the relationship should exist within the agency. Thus, I should see changes in validation rate that correlate with changes in reputation even though the FDA may have an overall higher reputation and validation rate than EPA for all data samples.

The second level of stratification divides the data by "case type," whether the cases involve only questions of law or whether the cases also involve findings of fact. The same logic for excluding pre-1984 cases supports stratification by case type. The level of judicial review a court gives to an agency's decisions differs when the case involves a question of law or a finding of fact. Agency decisions are subject to arbitrary and capricious review or "substantial evidence" review for cases that involve a finding of fact. "Arbitrary and capricious" review and "substantial evidence" review are two separate standards on their face, but in practice, scholars have suggested that,

---

[48] Motor Vehicles Ass'n v. State Farm Mutual Automobile Insurance Co., 463 U.S. 29, 43 (1983) (holding that agency action is "arbitrary and capricious" if the agency does not examine the data and articulate a satisfactory explanation, including "a rational connection between the facts found and the choice made").

[49] *See* Miles & Sunstein, *supra* note 24, at 771.

[50] Chevron U.S.A. Inc v. NRDC, 467 U.S. 837 (1984).

[51] Empirical assessments have come to different conclusions about whether *Chevron* had an effect on validation rates. *Compare* Elliot & Schuck, *supra* note 22, at 1057 (finding, in the years following *Chevron*, validation rates for agencies changed), *with* Zaring, *supra* note 5, at 2353. *See also* Kent Barnett, Christina Boyd & Christopher Walker, *Administrative Law's Political Dynamics*, 71 VAND. L. REV. 1463, 1463 (2018).

especially after *State Farm*, "review under the substantial evidence standard is essentially the same as review under the arbitrary and capricious standard."[52] The same might not be said of the standard of review for cases involving questions of law, which often require *Chevron* review.[53]

The final data set for evaluating the relationship between validation rate and reputation consisted of judicial opinions decided after 1984, stratified by agency and by case type. For this analysis, I grouped observations into four-year intervals so that I would have enough observations within each interval to make representative estimates for the average reputation score and the validation rate.[54]

My basic model of the relationship between validation rate and reputation consists of the following framework:

$$\text{Outcome}_{it} = \alpha_0 + \alpha_1\text{ReputationScore}_{i,t-1} + \alpha_2\text{LaggedOutcome}_{i,t-1} + \alpha_3\text{Agency}_i + \alpha_4\text{CaseType}_i + \beta_{i,t} + \varepsilon_{i,t}$$

In this equation, *Outcome* is the rate of judicial affirmance of agency action for each stratified case observation *i* during the four-year time interval *t*. Each stratified case observation *i* equals the average of values calculated from between three and twenty-one cases during each time interval *t*, after stratification by agency and case type. I was unable to control for fixed effects of judicial circuits due to too few observations in certain circuits, precluding an accurate assessment of the effect that a regional circuit could have on judicial outcomes.

ReputationScore is the lagged short-term or long-term reputation score for each analysis. *LaggedOutcome* is the lagged short-term or long-term validation rate. The lagged regression analysis intends to protect against endogeneity. Using lagged variables, I ensure that I calculate agency reputation before the courts decide a case—that reputation was not predetermined by the realization of the outcome variable. *Agency* is a categorical variable, whether the agency is the FDA or EPA, and *CaseType* denotes the question in the case—whether it relates to a question of law or finding of fact. β denotes time-varying and agency-specific legal attributes, such as judicial ideology and agency ideology. The error term ε captures observation- and time-variant unobserved heterogeneity.

I use simple linear regression to determine the short-term and long-term effects of reputation score on validation rate. To measure the short-term effects of reputation score on validation rate, I conduct a simple linear regression for validation rate over a four-year period against a lagged variable of reputation score for the prior four years, using standard errors clustered by agency. To determine long-term effects, I conduct a simple linear regression of validation rate over a four-year period against lagged

---

[52] Miles & Sunstein, *supra* note 24, at 764.

[53] *Cf.* Matthew Stephenson & Adrian Vermeule, *Chevron Has Only One Step*, 95 VA. L. REV. 597, 597 (2009) (arguing that *Chevron*'s inquiry, properly understood, consists of only one step).

[54] It would also be interesting to calculate the frequency that plaintiffs filed each type of case in each four-year interval. However, this is difficult to calculate because judges dismiss some cases before they are adjudicated on the merits and because parties decide to settle cases instead of litigating or appealing cases to circuit courts.

average reputation score for the past eight years, using standard errors clustered by agency.[55]

Table 4 reports estimates of the relationship between agency validation rate and short-term reputation score in four models that build on each other. In Model 1, I examine only the variable of interest with control variables for the lagged validation rate. In Model 2, I add the fixed effects variable for agency, and in Model 3, I add the fixed effects variable for case type. In Model 4, I include time-related and agency-specific variables that account for the ideological characteristics of the legal environment associated with the cases in each period.

Although there is some association between short-term reputation and agency validation rates in Models 1, 2, and 4, this relationship does not exist for all the models. Short-term reputation is significantly associated with validation rate at the five percent level in Model 1. The relationship between validation rate and short-term reputation is negligible if I limit the model to my main variables of interest; it remains weak once I control for potentially endogenous variables. The association between short-term reputation and validation rate is no longer significant in Models 2 and 3. In Model 4, a one-unit increase in lagged short-term reputation is associated with a 0.31 unit increase in validation rate. The evidence suggests that short-term reputation may not affect court validation rates and that an association between short-term reputation scores and validation rates may be attributed to agency and case-type fixed effects.

Table 4. Short-Term Reputation and Judicial Validation Rate

Variables	(1) Validation Rate	(2) Validation Rate	(3) Validation Rate	(4) Validation Rate
Lagged Short-Term Reputation	0.650**	0.459	0.280	0.310*
	(0.020)	(0.082)	(0.076)	(0.033)
Lagged Validation Rate	-0.125	-0.238	-0.395**	-0.447
	(0.022)	(0.020)	(0.015)	(0.301)
Agency		-0.132	-0.223*	-0.255*
		(0.021)	(0.027)	(0.021)
Case Type			0.160	0.173
			(0.062)	(0.038)
Judicial Ideology				0.131
				(0.279)
Agency Ideology				0.270
				(0.194)
Constant	-0.479	-0.064	0.394	0.176
	(0.093)	(0.172)	(0.155)	(0.137)
Observations	28	28	28	28
R-squared	0.443	0.509	0.633	0.688

**p < 0.10; **p<0.05; ***p < 0.01. Standard errors are clustered at the agency level. Model (4) also controls for time period.*

[55] I included weighted and unweighted averages for these reputation scores. I calculated the unweighted average as $(\text{ReputationScore}_{t-2} + \text{Reputation}_{t-1})/2$. The weighted average assigned greater weight to the more recent reputation score $(\text{ReputationScore}_{t-2} + 2*\text{Reputation}_{t-1})/3$.

Table 5. Long-Term Reputation and Judicial Validation Rate

Variables	(1) Validation Rate	(2) Validation Rate	(3) Validation Rate	(4) Validation Rate
Lagged Long-Term Reputation Score	0.785**	0.682	0.453***	0.560***
	(0.053)	(0.185)	(0.006)	(0.003)
Lagged Validation Rate	-0.227	-0.436	-0.914**	-0.943*
	(0.142)	(0.174)	(0.047)	(0.082)
Agency		-0.142*	-0.320**	-0.341**
		(0.017)	(0.012)	(0.009)
Case Type			0.235	0.227*
			(0.039)	(0.0061)
Judicial Ideology				0.317
				(0.139)
Agency Ideology				0.318
				(0.243)
Constant	-0.838	-0.403	0.388	0.156
	(0.010)	(0.284)	(0.031)	(0.161)
Observations	24	24	24	24
R-squared	0.547	0.605	0.800	0.862

**p < 0.10; **p<0.05; ***p < 0.01. Standard errors are clustered at the agency level. Model (4) also controls for time period. I used weighted lagged reputation scores in this analysis.*

In Table 5, I present estimates of the relationship between agency validation rate and reputation score for long-term reputation, using the same models as in Table 4. In three of the four estimations, reputation is positively and significantly associated with agency validation rates. The small sample size urges caution in interpreting the results, but reputation score remains positively and significantly associated with agency validation rate after controlling for agency and case-type fixed effects. The evidence supports the view that reputation may affect court validation rates in cases involving agency action. My methodology in Model 4, the most robust of the models, implies that a one-unit increase in reputation score is associated with a 0.56 unit increase in validation rates.

Agency, but not case type, is significantly associated with validation rates. In Table 5, FDA is the reference group for the "Agency" variable, so the negative coefficient on "Agency" indicates that FDA has higher validation rates than EPA, even taking into account the effect of agency reputation. It is unsurprising that different agencies would have different validation rates. Agency-specific factors, such as agency organization, internal policies and procedures, litigation strategies, and the reach of agency action, may play a role in both agency reputation and validation rates. The regression also suggests that opinions involving findings of fact may not be associated with different validation rates than questions of law. Other studies have reported similar results.[56]

[56] *See* Zaring, *supra* note 5, at 2353.

These results deserve some qualification. They do not mean that a boost in reputation score would flip a judge's assessment of a case so that the judge would decide all cases in an agency's favor for an agency with a strong reputation or would decide cases against an agency for an agency with a weak reputation. Legal doctrine and ideological preferences may moderate the practical effect of agency reputation on judicial outcomes. Nevertheless, the results suggest that after controlling for external variables, reputation may have a non-trivial role in cases evaluating agency action.

C. Are Changes in Agency Reputation Associated with Changes in the Frequency of Challenges Decided Before the Courts?

I attempt to qualify the findings detailed above by analyzing whether the quantifiable association reflects the real association between reputation score and validation rate. As I have suggested, agency reputation may also affect the actions of other actors before courts. If litigants believe in the power of agency reputation and can distinguish between agencies with strong and weak reputations, the rate of challenges to agency action could be affected by whether an agency's reputation is strong or weak. If the agency has a weak reputation, I would expect to see more challenges because litigants may believe that the likelihood of success is higher. If the agencies have a strong reputation, I might hypothesize that outside litigants would be more cautious before bringing an action before the courts.

Agency decisions could also be affected by the agency's perceived reputations. It is possible that an agency with a strong reputation may be more comfortable making less legally defensible decisions, knowing that their reputations may subject their decisions to less judicial scrutiny. Likewise, an agency with a weak reputation may be more careful about the decisions they make lest the courts invalidate their decisions. Reputation may play a similar role in how willing an agency will be to settle or appeal a negative decision before the courts.

Assuming that litigants are rational, the frequency of challenges decided before the courts serves as an imperfect proxy for the nature of challenges against agency action. Fewer challenges could mean that litigants are less aggressive, would less likely litigate, and would more likely choose to settle. The challengers that do litigate may bring stronger legal arguments against agency action. The analysis here is analogous to that of the effect of judicial doctrine on the rate of challenges to agency action. Miles and Sunstein posit that changes in the intensity of arbitrariness review can lead to two effects that cut in different decisions; as review becomes less intense, "litigants will challenge fewer decisions and agencies will be more likely to make decisions that aggressive courts would have struck down as arbitrary."[57] The ultimate consequences to changes in the intensity of arbitrariness review will depend not only on the direction, but also on the magnitude of the effects of the behavior of the litigant and agency.[58]

As with changes to the intensity of judicial review, the consequences of changes to agency reputation will depend on how comparatively responsive litigants and agencies

---

[57] Miles & Sunstein, *supra* note 24, at 804.

[58] *See id.* (stating that "[a] reduction in the intensity of review should first tend to raise the validation rate. If neither litigants nor agencies were responsive to the intensity of review, an increase in the validation rate would be the sole consequence of less intense review. But if litigants were highly responsive to intensity shifts and if agencies were not, the content of agency decisions would remain the same while litigants would decline to bring the more marginal challenges. The volume of arbitrariness challenges would decline, and the validation rate could remain fairly constant.").

are to changes in agency reputation. Assume that agency reputation weakens. If the agency is less responsive but litigants are more responsive to an agency's weakening reputation, we may see litigants bringing more marginal challenges against agency action, though the content of an agency's decisions remains the same. If the number of litigant challenges increases significantly, a decrease in agency reputation could result in an increase in judicial validation of agency action due to the dramatic increase in frequency of challenges to agency action.

To examine the association between frequency of challenges and reputation, I measure the relationship between lagged short-term and long-term reputation rates and the frequency of challenges to agency action. I assume that the frequency of decisions for challenges to agency action is driven more by the behavior of litigants than by the behavior of agencies. Challenges to agency action do not begin until a litigant decides to bring a suit against the agency. Although an agency could still affect the frequency of cases through its willingness to settle or willingness to appeal a negative decision, I assume that the magnitude of the effect would be less than a litigant's willingness to begin a case before the courts.

For each four-year interval, I calculate a simple statistic for frequency of challenges decided before the courts. The frequency statistic represents the percentage of cases in each four-year period involving findings of fact and the percentage of cases in each four-year period involving questions of law for each agency. In Table 6, I examine the relationship between short-term and long-term reputation score and the frequency of cases in each four-year period. Models 1 and 2 examine short-term reputation, and Models 3 and 4 examine long-term reputation. In Models 2 and 4, I add control variables to account for heterogeneity in previous validation score and agency, case type, and year fixed effects.

I do not find a statistically significant association between short-term and long-term reputation and the frequency of cases brought before the courts. The association between frequency and lagged short-term reputation score is not statistically significant in Model 1. The association between frequency and lagged long-term reputation score is not statistically significant in Model 3. The results are the same in Models 2 and 4, respectively, when controlling for agency and case-type fixed effects.

Table 6. Reputation and Frequency of Challenges to Agency Action

Variables	(1) Frequency	(2) Frequency	(3) Frequency	(4) Frequency
Lagged Short-Term Reputation Score	-0.437 (0.228)	-0.560 (0.151)		
Lagged Long-Term Reputation Score			-0.501 (0.295)	-0.691 (0.167)
Lagged Validation Rate		-0.227 (0.389)		-0.523 (0.799)
Agency		-0.241 (0.060)		-0.346 (0151)
Case Type		-0.121 (0.081)		-0.052 (0.011)
Agency Ideology		-0.409 (0.547)		-0.499 (0.639)
Constant	1.43 (0.557)	1.66 (0.459)	1.569 (0.707)	2.46** (0.193)

Observations	28	28	24	24
R-squared	0.321	0.763	0.334	0.799

**p < 0.10; **p<0.05; ***p < 0.01. Standard errors are clustered at the agency level, reported in parentheses. I used weighted lagged long-term reputation scores in this analysis.*

IV. DISCUSSION

In this Article, I investigate the effect of agency reputation on an agency's validation rate before the courts for FDA and EPA. A cautious interpretation of the results suggests that the anecdotal accounts of agency reputation affecting judicial decisions have some basis in reality. The EPA has a significantly lower reputation score than FDA, suggesting that agencies have different reputations before the courts. Further, agency reputation before the courts may have an effect on the courts' decisions and litigant behavior. Long-term agency reputation is positively correlated with judicial validation rates, suggesting that an increase in agency reputation may mean that courts are more likely to favor agency decisions. This result should be viewed in light of how agency reputation affects litigant behavior. I did not find a relationship between agency reputation and the frequency of decided challenges to agency action, meaning that if agency reputation leads to a change in litigant behavior, the relationship is not so reliable as to affect the relationship between reputation and judicial outcomes.

The evidence does not definitely establish a causal relationship between agency reputation and judicial validation rates. It is difficult to eliminate the possibility that the observed positive correlation between long-term reputation and validation rates is better explained by overall trends in the judiciary that would affect both agency reputation and validation rates. For example, it is possible that courts may have generally grown less deferential and respectful of agencies in the last few decades and thus have overturned agency action at higher rates while simultaneously revealing lower esteem towards agencies.

While the observational results themselves cannot prove causation, the results do support a causal inference when viewed alongside anecdotal accounts. I cannot completely rule out endogeneity, omitted variables, false-positives, or other explanations that could account for the findings in this study. Nevertheless, the results of this empirical investigation do support anecdotal accounts from judges, clerks, and legal observers of the role that reputation plays in judicial decisions on FDA or EPA action.

This Article examines the role of reputation in agency action decided on the merits in a subset of cases involving the agency. It is, however, possible that agency reputation has a stronger effect on agency proceedings before the courts because of the many decisions that courts must consider before and after litigants argue a case before the merits. Before a court hears a case on the merits, it must examine and decide questions of jurisdiction, standing, and ripeness, among other threshold issues. Courts must also adjudicate challenges to agency action after a court makes a decision on the merits. For example, litigants may challenge an agency's action after a court remands the case back to the agency. An agency with a positive reputation could see more challenges dismissed on threshold issues so that courts hear stronger arguments on the merits. Even if a court does rule against an agency with a positive reputation, the court could be more likely to uphold an agency's decision on remand. This Article does not examine how agency reputation can affect court decisions before and after a

determination on the merits nor how those decisions affect the likelihood that a court will decide a case in favor of the agency.

Two recently decided cases involving FDA and EPA illustrate how reputation could affect a case outside of an opinion on the merits. In *Ferring Pharmaceuticals v. Azar*, the court had initially granted the plaintiff's motion for summary judgment, concluding that FDA's decision to deny five-year exclusivity to Ferring's drug was arbitrary and capricious and remanded the action to FDA "for further proceedings not inconsistent with [its] opinion."[59] Eight months after the remand, FDA changed its decision regarding the identity of the active moiety of Ferring's drug. As a result of the change in identity, FDA concluded on remand that Ferring's drug was not entitled to five-year exclusivity. Ferring argued that FDA's switch in position about its drug is barred by judicial estoppel. The court denied Ferring's challenge. Although there "is no doubt that Ferring is inconvenienced by the FDA's new position . . . , [a]t the very least, this revelation is a setback."[60] Instead, the court adopted FDA's reasoning to conclude that "the FDA's late change in position may have been due to inadvertence or mistake" and did not grant Ferring the relief it sought on the basis of judicial estoppel. In this case, FDA's positive reputation on issues of fact might have helped the judge side with the agency, both on the scientific issue at hand and FDA's consideration of how it arrived at its scientific conclusion.

In contrast, a court decided not to accept EPA's arguments in revising carbon monoxide emission limits for boilers. In *Sierra Club v. EPA*, the court refused to accept the Administration's carbon monoxide emission limit where the Administration had not provided substantial evidence to support its conclusion.[61] The court refused to accept EPA's "change of direction on the carbon monoxide limits," describing EPA's conclusions as a "hunch."[62] The court continued by saying that "it would be particularly inappropriate to give EPA a pass on backing up its apparent hunch here, where EPA was operating against the backdrop of its own prior reasoned judgment . . ." [63] Although courts generally defer to an agency's scientific judgment, the court in this case refused to accept EPA's apparent change in position based on EPA's less than full accounting of its change. EPA's reputation may have been a factor in how the judge decided whether or not EPA's scientific rationale for the change constituted substantial evidence to support the change.

Courts ostensibly review agency decisions in light of legal standards and precedent, but judicial review inevitably involves subjective decisions. Extensive empirical research suggests that the subjectivity in judicial decisions may reflect a court's ideological leanings. In determining whether or not a court should trust an agency's decisions and its explanation for its decisions, the court may subconsciously factor agency reputation into the legal calculus. Although the effect of an agency reputation on judicial validation may be smaller than other subjective factors, reputation affects judicial decision-making at the margins.

---

[59] Ferring Pharm., Inc. v. Burwell, No.: 15-0802(RC), 2016 WL 4734333, at *11 (D.D.C. Sept. 9, 2016), *enforcement denied*, 296 F.Supp.3d 166, 167 (D.D.C. 2018).

[60] Ferring Pharm. Inc. v. Azar, 296 F.Supp.3d 166, 167 (D.D.C. 2018).

[61] Sierra Club v. EPA, 884 F.3d 1185, 1189 (D.D.C. 2018).

[62] *Id*. at 1198.

[63] *Id*. at 1198.

Courts decide cases involving hundreds of federal departments, agencies, and sub-agencies; reputation may provide a valuable signal to courts trying to determine the quality or significance of agency action or the justification for a challenge to agency action. Where there is uncertainty about agency action, reputation may provide courts with some indication about an agency's general effectiveness, reliability, or impartiality, information that may be relevant to a decision but may not be readily available from the facts of a particular case. The information may also be helpful to courts deciding a judicial remedy, including whether to vacate a decision or remand to an agency without vacatur.

On the other hand, overreliance on agency reputation may be problematic where reputational analysis replaces legal or factual analysis or outweighs legitimate legal arguments or the facts presented in a particular case. Where a court relies on agency reputation to place more or less weight on one side's arguments or evidence, reputation may compromise a court's impartiality and fairness.

Overall, the results have three main implications for those seeking to improve an agency's reputation before the courts. First, the findings indicate that courts might not view an agency differently based on whether the case concerns questions of law or questions of fact. Courts might not view agencies differently depending on the type of deference an agency receives. An agency with a strong reputation before the courts for a case involving a question of law may find that the reputational effects for one type of case could transfer to other types of cases before the courts. That said, the analysis did not differentiate cases by subject matter, so it is possible that an agency could have a strong reputation litigating on certain substantive issues and a weak reputation when litigating on other substantive issues.

Second, the findings indicate that agency reputation may change slowly, and long-term shifts in agency reputation may affect validation rates before the courts. For agencies headed by a commissioner or general counsel serving terms shorter than eight years, short-term reform of agency reputation may not lead to immediate changes in validation rates before the courts. Instead, only longer-term efforts to change an agency's decision-making processes could lead to an observable shift in judicial validation rates. On the other hand, one or two damaging decisions to agency reputation may not lead to long-term shifts in validation rates before the agency.

Third, the findings question whether the effect of an agency's perceived reputation would have a reliable effect on litigants deciding to bring challenges before the courts. Any intentional shift by the agency to improve agency reputation by the courts may not be countered by a shift in litigant behavior. This conclusion is constrained by the limits of the Article's data set. First, it is possible that a change in agency reputation changes the frequency of complaints before the courts. I assume for the purposes of the analysis that litigant behavior has a larger effect on the frequency of opinions decided by the courts. This assumption could be wrong; agency action (e.g., the likelihood that the agency would settle a case) could equally counterbalance any shift in litigant behavior due to changes in agency reputation. Second, it is also possible that litigants recognize signals in agency reputation but that any shift in litigant behavior is offset by more significant predictors of the frequency of challenges before the courts. Whether changes in agency reputation affect the absolute frequency of judicial challenges to agency action is an area for further study.

Whether the results of this Article have bearing across agencies is an issue for further study. The results of this Article focus on FDA and EPA reputation and do not examine reputations from other agencies. It is possible that the reputational effects

found in this Article are confined to FDA and EPA or to agencies with a heavy scientific focus; courts may defer to those agencies' scientific judgments more regularly. It is also possible that the reputational effects may be similar across executive agencies but not for independent agencies with less executive oversight. The evidence does not, however, suggest that agency reputation only applies to executive agencies. In fact, the anecdotal accounts of reputation provided by judges at the beginning of this Article, describing the reputation of the NLRB, for example, seems to suggest that agency reputation spans across executive and independent agencies. It is possible that independent agencies might have reputation effects that linger longer than executive agencies, which could be more directly impacted by changes in administration.

CONCLUSION

Legal doctrine and attitudinal factors shape judicial decisions. Reputation adds a complicating dimension to how attitudinal factors influence the courts, as agency reputation falls outside of ideological lines. This Article supports anecdotal accounts of the relationship between agency reputation and judicial outcomes by demonstrating that different agencies can have different reputations before the courts and that positive agency reputations can correspond to trends in favorable judicial decisions for an agency. The Article's results offer a more comprehensive understanding about how courts consider agency action and provide an empirical basis for agencies and litigants to assess how reputation affects the viability of their actions before the courts.