STRENGTH IN NUMBERS (OF WORDS): EMPIRICAL ANALYSIS OF PREAMBLES AND PUBLIC COMMENTS

Anthony Moffa*

TABLE OF CONTENTS

* Associate Professor of Law, University of Maine School of Law. The author would like to thank Sarah Schindler, Jeff Maine, Thea Johnson, Richard Chen, Justin Pidot, Robert and Bruce Thompson for their thoughtful comments on earlier drafts. This paper also benefited from: the Colloquium on Environmental Scholarship at Vermont Law School, the Sabin Colloquium on Innovative Environmental Law Scholarship at Columbia Law School, and the Rocky Mountain Land Use Institute's Environmental Scholarly Workshop at the University of Denver's Sturm College of Law. Many thanks to Ashley Fagan for her helpful research assistance and to Madeleine Coles and Jessica Phipps, who, along with their team at the Nevada Law Journal, provided editorial wisdom. All errors are the author's alone.

INTRODUCTION

Almost a decade ago, Jason and Susan Webb Yackee foretold "a rich empirical literature that seeks to advance our understanding of whether, how, and why federal agencies are (or are not) able to satisfactorily achieve their regulatory responsibilities."[1] Their pioneering work, along with others, like Cary Coglianese and Cass Sunstein, exposed empirically some of the flaws of the so-called "ossification hypothesis," but left much still to be tested.[2] In the following decade, empirical administrative law scholarship has flourished, particularly in its examination of judicial review and, separately, of rulemaking processes. This work brings together those two veins of empirical scholarship, continuing in a quest to discern why rules have been getting longer and whether that reason makes sense in a world of limited resources.

The regulatory state invites much popular media critique and commentary from political pundits. Unfortunately, the great majority of that very public discussion takes place ungrounded in facts and empirical truth. The resulting debate pitting over- and under-regulation against one another misses a point more fundamental to the effective functioning of government—the form and content of regulations has been changing. In particular, as the data described below reveals, individual rules have steadily ballooned in length over time. That matters for a variety of reasons relevant to the core functioning of democratic governance. Citizens, in particular those who may be the subject of agency regulation, today must devote more time and energy than ever before to participate in any individual rulemaking. Regulators, likewise, today must devote more resources to the crafting of each of those rules. Those are real costs to the system and society. The questions that this Work sets out to answer empirically are first why costs increased on a per rule basis over time, and, relatedly, whether any discernible

---

[1] Jason Webb Yackee & Susan Webb Yackee, *Testing the Ossification Thesis: An Empirical Examination of Federal Regulatory Volume and Speed, 1950–1990*, 80 GEO. WASH. L. REV. 1414, 1482 (2012).

[2] *See id.* at 1479–80 ("This kind of rule-level comparative analysis might show, for example, that ossification is a reality for certain kinds of rules (perhaps rules of high salience, or rules that are economically significant.)").

benefit rationally explains that increase. The twin focuses of this study in particular are on the explanatory, or preambular, sections of individual rulemakings and the responses to public comments. The empirical analysis below investigates whether the former has any significant relationship to outcomes on judicial review, and whether the latter might explain the underlying trend of longer rules. In the end, the data were able to confirm, with a high degree of statistical significance, that the number of public comments contributes to the phenomenon of longer regulations.

In order to conduct the statistical analyses reported here, the necessary datasets were constructed with a focus specifically on rules promulgated by the Environmental Protection Agency (EPA) and judicial "arbitrary and capricious review" of the substance those rules. The relevant cases did not include those limiting review to the EPA's interpretation of statutory mandates pursuant to *Chevron v. Natural Resources Defense Council*.[3] Separating empirical analyses of opinions applying these two standards of review is consistent with the approach taken by other scholars, like Thomas Miles and Cass Sunstein.[4] Not only is it empirically sound, but it is also logically consistent with the hypotheses tested, which focus on the administrative process of rulemaking rather than the legislative process of delegation. The hypotheses described and tested below embody attempts to justify long and detailed rules, pointing alternatively to courts' increasingly searching inquiries into the scientific and economic rationale for those rules and to the commenting public's awareness of and argument with them after proposal. The Administrative Procedure Act's (APA) "arbitrary and capricious" clause provides the basis for the judicial review of interest.[5] In contrast, the *Chevron* deference doctrine concerns the statutory authority for a given rule, rather than the rationale underlying its substance.[6]

Under the APA, an agency action is unlawful if it is "arbitrary, capricious, an abuse of discretion, or otherwise not in accordance with law."[7] As the

---

[3] Chevron, U.S.A., Inc. v. Nat. Res. Def. Council, Inc., 467 U.S. 837, 842–43, 861 (1984) (laying out the boundaries of agencies' discretion in interpreting their statutory mandates and acting upon those interpretations).

[4] *Compare* Thomas J. Miles & Cass R. Sunstein, *The Real World of Arbitrariness Review*, 75 U. CHI. L. REV. 761, 766 (2008) (analyzing arbitrariness review), *with* Thomas J. Miles & Cass R. Sunstein, *Do Judges Make Regulatory Policy? An Empirical Investigation of* Chevron, 73 U. CHI. L. REV. 823, 825 (2006) (analyzing the application of *Chevron* deference).

[5] 5 U.S.C. § 706 ("The reviewing court shall . . . hold unlawful and set aside agency action, findings, and conclusions found to be . . . arbitrary, capricious, an abuse of discretion, or otherwise not in accordance with law.").

[6] *See Chevron*, 467 U.S. at 843 ("The power of an administrative agency to administer a congressionally created . . . program necessarily requires the formulation of policy and the making of rules to fill any gap left, implicitly or explicitly, by Congress." (omission in original) (quoting Morton v. Ruiz, 415 U.S. 199, 231 (1974))).

[7] 5 U.S.C. § 706(2)(A). The APA, as interpreted by the courts, also imposes a number of other requirements on rulemaking in § 553, including statements about legal authority, data supporting the rulemaking, an opportunity for public comment, responses to material comments, and a defense of the final rulemaking as rationale and as a logical outgrowth of the proposed rule.

Supreme Court has noted, the reviewing court must determine whether the agency based its decision on a consideration of "the relevant factors" or whether it made "a clear error of judgment."[8] Such a determination necessarily entails a "searching and careful" fact-specific inquiry, but "the ultimate standard of review is a narrow one."[9] "The court is not empowered to substitute its judgment for that of the agency."[10]

The subject of this Work—EPA rulemaking—was not selected solely because of the urgent need in the current moment for effective regulatory action on the most significant environmental challenge in human history. That would surely be reason enough. But EPA rulemaking is also worthy of study because it is emblematic of agency activity across various subjects and the frequent target of regulatory reform advocates. The public perception, and the scientific and bureaucratic reality of our time, hold the EPA responsible for a large amount of new federal regulations. In the eyes of administrative law observers, and its own employees, the EPA stands out as an agency committed to its mission and to achieving it through the process of good governance.[11]

To lay the foundation for the statistical analyses that constitute the primary contribution of this Work, Part I presents a picture of the trend of increasing rule length over time that is grounded in the initial, high-level data. Parts II and III identify gaps in the existing qualitative and quantitative administrative law scholarship, particularly with respect to the previous study of the judicial review and public comment processes. Part IV presents the results of statistical analyses of the preamble text, which sought, and failed, to identify a statistically significant correlation between outcomes on judicial review and the length of rules' preamble sections. Part V presents the results of statistical analyses of the number of public comments on proposed rules, which demonstrated a significant positive

See id. § 553(b)–(c); Weyerhaeuser Co. v. Costle, 590 F.2d 1011, 1031 (D.C. Cir. 1978); United States v. Nova Scotia Food Prods. Corp., 568 F.2d 240, 251 (2d Cir. 1977).

[8] Citizens to Pres. Overton Park, Inc. v. Volpe, 401 U.S. 402, 416 (1971) (citations omitted).

[9] *Id.*

[10] *Id.* The Supreme Court has provided some much-needed context to this rather amorphous standard over the years. The most cited definition, from the Court's opinion in *Motor Vehicle Manufacturers Ass'n of the United States v. State Farm Mutual Automobile Insurance Co.*, holds that the "arbitrary and capricious" standard is violated when

> the agency has relied on factors which Congress has not intended it to consider, entirely failed to consider an important aspect of the problem, offered an explanation for its decision that runs counter to the evidence before the agency, or is so implausible that it could not be ascribed to a difference in view or the product of agency expertise.

Motor Vehicle Mfrs. Ass'n of the U.S. v. State Farm Mutual Auto. Ins. Co., 463 U.S. 29, 43 (1983).

[11] *See, e.g.*, Sally Katzen, *A Reality Check on an Empirical Study: Comments on "Inside the Administrative State,"* 105 MICH. L. REV. 1497, 1499 (2007) ("[I]n my experience senior political appointees at EPA clearly stand out from their colleagues at other agencies for both the intensity of their enthusiasm for their agency's mission and their faith in regulatory solutions.").

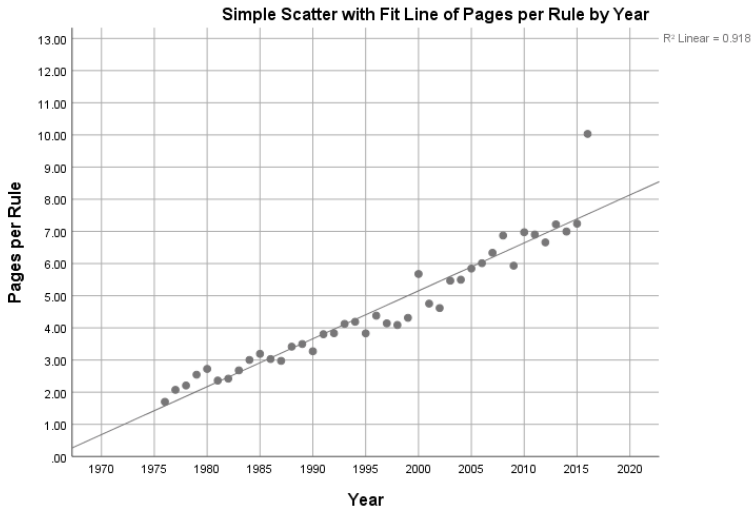
correlation between quantity of comments and rule length. Part VI describes the implications of the study's findings for policymakers, advocates, and scholars.

I. LONGER RULES

A. Documenting the Trend

In *Word Limited*, I demonstrated empirically that over the course of the last four decades, the length of new rules (measured in Federal Register pages of final rule entries) has steadily increased, while the number of separate, individual rule-makings has decreased. Combining the two sets of data yielded a linear trend of increasing Federal Register pages per rule over time.[12]

FIGURE 1[13]



As the above figure demonstrates, the social regulation and environmental protection era of the 1970s produced a lot of new regulations, but each of them accounted for only about two Federal Register pages.[14] In the decades since, that pages-per-rule figure has steadily increased, ballooning to almost ten Federal Register pages for each new rule.[15]

The foundational regression model in *Word Limited* quantified the strength of the statistical relationship between pages per rule and year of the rule's promulgation.[16] That model found that over 90% of the variance in the number of

---

[12] Anthony Moffa, *Word Limited: An Empirical Analysis of the Relationship Between the Length, Resiliency, and Impact of Federal Regulations*, 20 NEV. L.J. 733 (2020).

[13] This figure is reprinted exactly as it appeared in *id.* at 744 fig.3.

[14] *See supra* Figure 1.

[15] *See supra* Figure 1.

[16] Moffa, *supra* note 12, at 744.

Federal Register pages a rule comprises can be explained by the year that the rule was promulgated, with newer rules comprising more pages than older rules.[17] This regression model empirically reaffirmed the contention of Cary Coglianese and other scholars who pushed back on the popular narrative that the regulatory state had not ossified beginning in the 1980s.[18]

MODEL SUMMARY[19]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.958[a]	.918	.916	.54152

a. Predictors: (Constant), Year

The foundational data published in *Word Limited* demonstrated quite clearly that individual regulations published in the Federal Register have become wordier over time—a noteworthy, but perhaps unsurprising, phenomenon.[20] Many threads of debate and analysis could flow from just the documentation of that initial observation—both critical and supportive of the maturing administrative state. My own project has been an attempt to explain the reason for the trend in an effort to improve the efficiency of government agency work. This Work continues that important effort.

---

[17] *Id.*

[18] *See* Cary Coglianese, *The Rhetoric and Reality of Regulatory Reform*, 25 YALE J. ON REGUL. 85 (2008).

[19] This summary is reprinted exactly as it appeared in Moffa, *supra* note 12, at 744 fig.3.

[20] As I acknowledged in *Word Limited*, some scholars might point to their own older studies as evidence against the trend I document. Most notably, Anne Joseph O'Connell's seminal work, *Political Cycles of Rulemaking: An Empirical Portrait of the Modern Administrative State*, 94 VA. L. REV. 889 (2008), compiles a truly remarkable amount of data on administrative rulemaking, including the number of Notices of Proposed Rulemaking (NPRMs) over time. O'Connell uses that figure as one indicator of regulatory activity. *See id* at 940, chart 6. O'Connell's data does not show a decline in NPRMs over time, which contrasts with the decline in rules over time reported by the Office of the Federal Register. *Id.* There are a few potential explanations for the discrepancy, the first of which being that NPRMs have not historically been issued for every rulemaking. The second explanation has to do with the composition of O'Connell's dataset, which she constructed using "federal agency reports in the Unified Agenda, which is published twice a year in the Federal Register, from 1983 to 2003." *Id.* at 924. These reports included agency-provided information on rulemakings, including the date of the NPRM, "the date(s) of the comment period(s), [and] the date . . . [of] the final rule" or withdrawal. *Id.* As O'Connell herself concedes, the database "has some disadvantages," the most significant of which being that the information included in the Unified Agenda is all self-reported by the agencies. *Id.* at 927. Jerry Mashaw contends that "[t]he EPA . . . does not report any rulemaking activity that it considers insignificant [to be included in the Unified Agenda]." Jerry L. Mashaw, *Improving the Environment of Agency Rulemaking: An Essay on Management, Games, and Accountability*, 57 L. & CONTEMP. PROBS. 185, 198 n.41 (1994). O'Connell explicitly disputes Mashaw's account but does ultimately agree with his contention "that it is not feasible 'for the untutored eye to discern from the reporting in the Unified Agenda . . . whether activity levels are primarily in a regulatory or deregulatory direction.' " O'Connell, *supra*, at 928 (omission in original) (quoting Mashaw).

B. Previously Tested Hypotheses

Over at least the last three decades,[21] scholars and policymakers have become increasingly concerned with judicial scrutiny contributing to the perceived "ossification" of regulatory activity.[22] Numerous scholars have empirically debunked the ossification thesis, without acknowledging the trend of declining individual rules published in the Federal Register described above.[23] This Work is the second in a series seeking, among other things, to fill that noteworthy gap in the empirical scholarship.

The insulation hypothesis, which was tested in the *Word Limited* study, surmised that policymakers have included more detailed legal and scientific support in new regulations, and thereby increased their length relative to previous regulations, because the additional detail provided more insulation from increasingly searching judicial review.[24] The empirical test of that hypothesis provided an initial, rather crude metric for the concern of judicial intervention translating to longer rules in the hope that they would better stand up to scrutiny. The data, however, failed to confirm the existence of a statistically significant[25] correlation between the length of a rule and its ability to withstand judicial scrutiny.[26]

That finding, while important, cannot put the argument to bed; any serious observer of the Federal Register recognizes that different sections of a rulemaking entry serve different purposes. And agencies have historically utilized the preamble sections of Federal Register entries to insulate their rules from judicial vacation or remand—particularly the "concise general statement of . . . basis and purpose."[27] Thus, the second-order study herein will test two "preambular insulation hypotheses." Those hypotheses can be stated as follows:

---

[21] *See* Yackee & Yackee, *supra* note 1, at 1418 n.18 (finding the term "ossification" produced over 1,000 hits in a search of documents in Westlaw's Journals and Law Reviews database).

[22] *See* Cary Coglianese, *Empirical Analysis and Administrative Law*, 2002 U. ILL. L. REV. 1111, 1130 (2002) (describing the "ossification hypothesis, which supposes that agencies have to work harder to produce rules that will withstand judicial scrutiny").

[23] *See, e.g.*, Yackee & Yackee, *supra* note 1, at 1421, 1436–38; Coglianese, *supra* note 22, at 1127, 1129, 1131, 1133 (2002).

[24] *See* Moffa, *supra* note 12, at 747–62.

[25] *Id.* at 745 n.49 ("Statistical significance refers to the confidence that a correlation (positive or negative) found in the data is not the product of random variation. This work will utilize a confidence level of 95 percent to define statistical significance.").

[26] *Id.* at 747–62.

[27] 5 U.S.C. § 553(c); *see* Sidney A. Shapiro & Richard E. Levy, *Heightened Scrutiny of the Fourth Branch: Separation of Powers and the Requirement of Adequate Reasons for Agency Decisions*, 1987 DUKE L.J. 387, 412 (1987) (analogizing the statement of basis and reasons provided by an agency in a rulemaking to judicial opinion writing, "giv[ing] a 'reasoned elaboration' for . . . actions according to norms of consistent, neutral and candid decisional processes" (quoting G. Edward White, *The Evolution of Reasoned Elaboration: Jurisprudential Criticism and Social Change*, 59 VA. L. REV 279, 286 (1973)).

> Preambular Hypothesis One: the length of a rule's preamble in the Federal Register, measured in number of words, correlates positively with that rule's ability to withstand judicial review.
>
> Preambular Hypothesis Two: the relative length of a rule's preamble sections in the Federal Register as opposed to the text of the rule itself, measured by the percentage of words in the Federal Register entry that are preambular, correlates positively with that rule's ability to withstand judicial review.

To test these hypotheses, this study will collect and analyze the word count data for the preambular sections of Federal Register entries in isolation from the rule text itself. Statistical analysis will produce the requisite correlations between the length, and relative length, of the rules' preambles and success in the courts. The results of this statistical analysis will conclusively establish whether agencies, and their attorneys, have a sound empirical basis for devoting more time (and more words) to these sections of rulemakings.

When confronted with the empirical trend of increasing Federal Register pages per rule over time, administrative law scholars invariably pointed to an explanation external to the rulemaking agency—the number of public comments.[28] Legally, agencies must respond to significant comments in the preamble to the final rule,[29] so logic dictates that more comments would lead to more words in the Federal Register. Meanwhile, in the real world, use of personal

[28] *See* Cornelius M. Kerwin & Scott R. Furlong, *Time and Rulemaking: An Empirical Test of Theory*, 2 J. PUB. ADMIN. RSCH. & THEORY 113, 130–31 (1992) (finding a negative correlation, two political scientists, Cornelius Kerwin and Scott Furlong, previously tested for correlation between the number of public comments received and the time it took to write a rule).

[29] *See* Perez v. Mortg. Bankers Ass'n, 575 U.S. 92, 96 (2015) ("An agency must consider and respond to significant comments received during the period for public comment.").

computers,[30] access to the internet,[31] and awareness of regulations.gov[32] have all risen in parallel with rules getting longer and have all made commenting on rulemaking easier over time. The empirical picture would thus not be complete without examining the potential connection between the number of comments and the length of a rulemaking. The data necessary to examine that question was not part of the initial study in *Word Limited*. This study will collect and analyze that necessary data to test the following hypothesis:

> "Call-and-response" Hypothesis: the number of public comments on a rulemaking correlates positively with the length of the final rule's Federal Register entry, measured in number of words.

To test this hypothesis, this study will count the number of public comments submitted on each of the federal rules in the dataset. This hypothesis also depends on the word count data for Federal Register entries described above. Statistical analysis will produce the requisite correlations between the length of a final rule entry and the number of public comments. The results of this empirical analysis will establish whether the trend of increasing rule length can be explained by a corresponding rise in the number of public comments on those rules.

Two of the foremost empirical scholars of administrative law first took notice of both of the hypothesized sources of rule length tested here in the context of broader study on ossification.[33] Using a National Park Service rule as an

---

[30] *See* William Lehr & Frank R. Lichtenberg, *Computer Use and Productivity Growth in US Federal Government Agencies, 1987–92*, 46 J. INDUS. ECON. 257, 259, 267 (1998) ("us[ing] productivity data obtained from the Bureau of Labor Statistics' (BLS) Federal Productivity Measurement Program, which was established for the specific purpose of tracking the labor productivity (real output per employee-hour worked) of federal government organizations, linked to data on computer use obtained from Computer Intelligence Infocorp (CII), a private marketing research firm" and reporting data demonstrating that "computerization of public sector workplaces proceeded at a rapid pace [from 1987 to 1992]"); *see also* Lois Mayer Nichols, *Pencil and Paper Versus Word Processing: A Comparative Study of Creative Writing in the Elementary School*, 29 J. RSCH. ON COMPUTING EDUC. 159, 160 (1996) (finding, in a study of elementary school students, that those using computers wrote compositions with significantly more words and sentences than those writing with pencil and paper).

[31] *See* Max Roser et al., *Internet*, OUR WORLD IN DATA, https://ourworldindata.org/internet [perma.cc/PWM5-WEAK] ("[E]stimates for 1990 suggest that only half of a percent of the world population were online. . . . By the year 2000 almost half of the population in the US was accessing information through the internet. . . . [I]n 2016, three-quarters (76%) of people in the US were online and during these years countries from many parts of the world caught up: in Malaysia 79% used the internet; in Spain and Singapore 81%; in France 86%; in South Korea and Japan 93%; in Denmark and Norway 97%; and Iceland tops the ranking with 98% of the population online.").

[32] *See* Thomas A. Bryer, *Public Participation in Regulatory Decision-Making: Cases from Regulations.gov*, 37 PUB. PERFORMANCE & MGMT. REV. 263, 263–64 (2013) (describing regulations.gov, launched in 2003, as "an award-winning government Web site that has democratized the federal rulemaking process by making it easier for citizens to search, read, and comment on proposed rules").

[33] *See* Yackee & Yackee, *supra* note 1, at 1460.

example, they noted that the final Federal Register entry included approximately 6,000 words laying out historical context and justification (i.e., preambular explanation) and another 6,000 words responding to public comments.[34] In that same large study, they confirmed that the public demand for regulation has remained fairly steady over time; a fact that might influence participation in the way of comments.[35] Their work, despite these related observations, did not set out to empirically examine the length of rulemakings over time and test potential explanations thereof. This Work, and its predecessor, set out to document the four-decades-long trend towards longer and longer federal agency rulemakings and test important hypotheses to produce results that have the potential to guide policymakers towards greater efficacy and efficiency. The preambular hypotheses and the call-and-response hypothesis derive from bodies of administrative law scholarship and the intuition of policymakers and academics. The empirical analyses test those theories and intuitions to provide data on the effectiveness of the administrative state in relation to the length of the Federal Register.

II. PREAMBLES AND REGULATORY TEXT

A. Importance

The Administrative Procedure Act (APA) has long required an agency to "incorporate in the rules adopted a concise general statement of their basis and purpose."[36] Whatever the original conception of this procedural requirement, it is now common practice in Federal Register entries for agencies to explain in painstaking detail the components of the administrative record, and even the legal analysis, supporting a particular rulemaking.[37] Some commentators have analogized this function of the administrative record to judicial opinion writing, which "give[s] a 'reasoned elaboration' for . . . actions according to norms of consistent, neutral and candid decisional processes."[38] Judicial review of agency action is confined, except for a few narrow exceptions, to the administrative record.[39] Thus, agencies likely have determined that laying out explicitly the reasons

---

[34] *See id.* (citing Final Rule Concerning National Capital Parks Regulations, 51 Fed. Reg. 7566 (Mar. 5, 1986) (to be codified at 36 C.F.R. pt. 50)) ("NPS included in the final rule preamble a nearly 6000-word summary and response to the various comments received, in addition to another 6000-word summary of the regulation's history and a justification of the final rule adopted.").

[35] *Id.* at 1479.

[36] 5 U.S.C. § 553(c).

[37] *See, e.g.*, Repeal of the Clean Power Plan and Revisions to Emissions Guidelines, 84 Fed. Reg. 32520 (July 8, 2019) (to be codified at 40 C.F.R. pt. 60) (including two separate sections heavily laden with legal analysis, one laying out the legal basis for repeal of the Clean Power Plan and another laying out the authority for the regulation of electric generating units).

[38] Shapiro & Levy, *supra* note 27, at 412 (quoting White, *supra* note 27).

[39] Fla. Power & Light Co. v. Lorion, 470 U.S. 729, 744 (1985) ("If the record before the agency does not support the agency action, if the agency has not considered all relevant factors, or if the reviewing court simply cannot evaluate the challenged agency action on the basis of

for an action in the public record, rather than tying everything in the record together for the first time before a court, makes that action less likely to be deemed "arbitrary and capricious" under the APA.[40] This study empirically tests that assumption—an assumption fundamental to the allocation of agency rulemaking resources.

As a result of this agency practice, Federal Register entries for final rules can be divided into two main components—the preambular text and the regulatory text. Agencies, and the Federal Register itself, have somewhat inconsistently labeled the various parts of rulemakings over time, but have always maintained a clear demarcation between explanatory sections and the regulation itself. Thus, this study will employ that same general bifurcation of the Federal Register entry and focus specifically on the preambular text.

In parallel to individual rulemakings ballooning in length over time, observers have anecdotally griped about the length of preambular explanations. As one scholar put it,

> Anyone who has picked up the *Federal Register* and waded through a preambular explanation and a final rule will have encountered a familiar phenomenon: five or six pages of rule, preceded by fifty or more *Federal Register* pages setting forth detailed agency explanations and/or responses to the most technical and arcane comments.[41]

Although the APA indeed mandates some explanation for a rulemaking, it also commands that said explanation be "concise" and "general."[42] Consequently, not only might excessive preambular text be inefficient and unhelpful for judicial review, it might actually be counterproductive, creating another ground on which to argue a violation of the APA.[43] Admittedly, an APA claim based on an agency providing *too much* information or process seems unlikely to succeed. Regardless, agencies should strive for the ideals the statute prescribes. In this particular area, one might argue that increasingly lengthy and complex preambular text undermines the public participation, and ultimately democratic accountability, envisioned by the APA.[44]

---

the record before it, the proper course, except in rare circumstances, is to remand to the agency for additional investigation or explanation.").

[40] *See supra* notes 5–39, *infra* notes 41–74 and accompanying text; *see also* Richard W. Parker, *The Empirical Roots of the "Regulatory Reform" Movement: A Critical Appraisal*, 58 ADMIN. L. REV. 359, 397 (2006) ("The implicit assumption behind current practice seems to be that if the APA requires 'a concise general statement' of a rule's basis, then an extremely detailed, lengthy, and arcane explanation is so much the better.").

[41] Parker, *supra* note 40, at 395.

[42] 5 U.S.C. § 553(c).

[43] *See* Parker, supra note 40, at 396 ("Needless to say, an explanation of 'rationales' that occupies 29 *Federal Register* pages (or 52 such pages if one includes responses to comments) hardly qualifies as the 'concise and general' explanation for which the APA calls.").

[44] *See id*. at 397 ("It would appear that, goaded by judicial appeals and court decisions, agencies have developed the tradition of offering extremely long explanations densely packed with technical detail and responsive to a host of comments but targeted only at an insider

B. Previous Empirical Research

No prior empirical work has set out specifically to test the connection between the length of preambular sections, overall Federal Register entry length, and result on judicial review. Others have, however, studied related data points touching on some of the measures used here, as well as advanced arguments consistent with the preambular hypotheses.

Following a wave of executive orders and general movement towards centralized oversight of the administrative state in the 1980s and 1990s, a strain of scholarship critically examining some of the oversight obligations emerged. The Office of Information and Regulatory Affairs (OIRA), a subagency within the Office of Management and Budget (OMB) founded by the Paperwork Reduction Act of 1980, stands as a popular target.[45] In Executive Order 12,291, President Reagan formalized the so-called "regulatory review" process while also requiring agencies to conduct "Regulatory Impact Analyses."[46] Directly impacting preambular text, the Order required agencies to "[m]ake a determination that the regulation is clearly within the authority delegated by law and consistent with congressional intent, and include in the Federal Register at the time of promulgation a memorandum of law supporting that determination."[47] These, and other requirements of the era, prompted critics to blame OIRA for the perceived "ossification" and politicization of agency rulemaking.[48] In support of reform, others looked for positive benefits of centralized oversight. One such scholar drew a line directly from OIRA to agencies better explaining, and documenting, their reasoning.[49] EPA officials themselves attested to centralized review helping them "clarify or gain support for their position."[50]

No study, as of yet, has set out to empirically tie the length of preambular text to these reform efforts, but some have at least considered the connection with ossification. In that context, one study surmised, without fully testing, that agencies in the period of theorized ossification "expended greater effort explaining and justifying their actions."[51] Comparing the number of words in Notices of Proposed Rulemaking from the National Park Service in pre-1975 and post-1975 time periods, the authors found almost a 200% increase in length during the later

audience. . . . Democratic accountability is defeated when agency explanations are so long, diffuse, and technical that no one but insiders can fathom them.").

[45] Paperwork Reduction Act of 1980, 44 U.S.C. §§ 3501–21.

[46] Exec. Order No. 12,291, 3 C.F.R pt. 127 (1981).

[47] *Id.*

[48] *See, e.g.*, Lisa Schultz Bressman & Michael P. Vandenbergh, *Inside the Administrative State: A Critical Look at the Practice of Presidential Control*, 105 MICH. L. REV. 47, 77 (2006).

[49] *See, e.g.*, Katzen, *supra* note 11, at 1507 ("In other words, the very existence of OIRA causes the agencies to do a better job in thinking through and documenting support for their proposals. That is, I would submit, a valuable contribution to decision-making.").

[50] *See* Bressman & Vandenbergh, *supra* note 48, at 77.

[51] Yackee & Yackee, *supra* note 1, at 1459.

period.[52] Foretelling a study focusing explicitly on the length of rulemakings, like this one and its predecessor, the authors also pointed out that final rules from the National Park Service showed similar increases in mean length between the two time periods,[53] suggesting that at least some of that increase came from the modern obligation to respond to comments individually.[54] The authors cited these figures as evidence of agencies improving their explanations, rather than evidence of potential inefficiency or wasted resources.[55] One scholar critical of these preambular sections unfortunately confined his arguments to anecdotal examples.[56]

Scholarly attention has thus focused for some time on what agencies are compelled to include within the preambular text of Federal Register entries. This study will be the first to empirically analyze the trend towards more preambular text and, more importantly, test hypotheses that might rationally explain that trend.

III. PUBLIC COMMENTS AND PUBLIC PARTICIPATION

A. Importance

The APA commands that agencies provide notice of and an opportunity for the public to comment on new rulemaking before a rule becomes final.[57] Courts interpreting the APA have made clear that the required process further demands agency response to any significant comments within the Federal Register entry

---

[52] *Id.* ("The mean for the early period is 1161 words, while the mean for the later period is 3206 words, an increase of 176%.").

[53] *Id.* ("There is a similar increase in the amount of words in NPS final rules, from a mean of just 849 to a mean of 4014.").

[54] *Id.* ("Some of this increase is due to the fact that agencies in the ossified period must respond in the final rule preamble to public comments submitted in response to the NPRM. Early final rules provided almost comically brief responses to public comments. For example, a typical final rule (here, from 1959) might recite that '[c]onsideration having been given to all relevant matters presented, it has been determined that the following proposed amendment shall become effective upon publication in the FEDERAL REGISTER.' " (footnote omitted) (quoting Final Regulation Concerning Zion and Bryce Canyon National Parks, 24 Fed. Reg. 6977 (Aug. 28, 1959))).

[55] *See* Yackee & Yackee, *supra* note 1, at 1459 ("A reasonable interpretation of the increase is that NPS, like other agencies, is now better at explaining its intentions.").

[56] Parker, *supra* note 40, at 396 (breaking down the components of a Federal Register entry for EPA, National Primary Drinking Water Regulations, Arsenic and Clarifications to Compliance and New Source Contaminants Monitoring, 66 Fed. Reg. 6976 (Jan. 22, 2001)).

[57] *See* 5 U.S.C. § 553(b)–(c) ("General notice of proposed rule making shall be published in the Federal Register, unless persons subject thereto are named and either personally served or otherwise have actual notice thereof in accordance with law. . . . After notice required by this section, the agency shall give interested persons an opportunity to participate in the rule making through submission of written data, views, or arguments with or without opportunity for oral presentation.").

for a final rule.[58] This prescribed procedure is widely understood as the method of ensuring public participation in the regulatory process. With the rise of the administrative state, the locus of substantive lawmaking (i.e., where the rules that govern daily life are written) shifted away from the democratically accountable Congress to the unelected federal bureaucracy. That power transfer, and the congressional acts delegating authority that facilitated it, did not come about by accident or happenstance. The APA represents a legislative effort to quell fears about the resulting unaccountability. Central to that effort was the concept of public participation, realized in the notice-and-comment process. However, the practice of rulemaking has not lived up to that ideal. As Donald Elliott colorfully put it, "Notice-and-comment rulemaking is to public participation as Japanese Kabuki theater is to human passions—a highly stylized process for displaying in a formal way the essence of something which in real life takes place in other venues."[59]

Scholars have written tomes on the shortcomings of administrative process in practice. Some of those commentaries speak with particular salience to the subject of this study—the amount of work, and words, devoted to the service of public comments. Though the APA prescribes the general structure of the rulemaking process, as laid out above, agencies make choices that can significantly affect the number of comments they receive and the amount of resources expended responding to them. For instance, the APA seems to contemplate that "an agency goes through the steps once; there is one [Notice of Proposed Rulemaking (NPRM)], one comment period, and one final rule."[60] As Anne Joseph O'Connell points out, agencies often elect to issue multiple NPRMs, sometimes labeled "advance" or "supplementary," resulting in multiple public comment periods for a single proposed rule.[61] On the other end of the spectrum, some agencies have attempted to minimize the opportunity for public comment by engaging in "direct final rulemaking."[62] Thus, the choices agencies make about the particulars of the rulemaking process can, and should, reflect strategic calculations about resource allocation, policy priorities, politics, and more.[63]

Because significant public comments must be responded to and agencies cannot act arbitrarily, every comment carries a risk of litigation on procedural and substantive grounds. Some scholars suggest the litigation risk translates directly into work within the agency to change a proposed rule, arguing that those changes come in proportion to the total number of comments (i.e., more

---

[58] *See* Perez v. Mortg. Bankers Ass'n, 575 U.S. 92, 96 (2015) ("An agency must consider and respond to significant comments received during the period for public comment.").

[59] E. Donald Elliott, *Re-Inventing Rulemaking*, 41 DUKE L.J. 1490, 1492 (1992).

[60] Anne Joseph O'Connell, *Agency Rulemaking and Political Transitions*, 105 NW. UNIV. L. REV. 471, 477 (2011).

[61] *See id.* (noting the use of "interim final" rules as well).

[62] Michael Kolber, *Rulemaking Without Rules: An Empirical Study of Direct Final Rulemaking*, 72 ALB. L. REV. 79, 80–81 (2009).

[63] *See* O'Connell, *supra* note 20, at 917 ("An agency must weigh the costs and benefits of various procedures, assuming that it has a choice among them.").

comments means more work).[64] Adherents to this line of reasoning point to its troubling logical consequence—"badly imbalanced stakeholder input into rules may lead to imbalanced outputs *because of*, not in spite of, judicial review."[65]

The EPA, in particular, historically chooses to offer relatively more opportunities for public participation than other agencies do or than the APA requires.[66] The EPA routinely provides individual responses to major comments in Federal Register entries for final rules, including summaries of those comments alongside the specific responses.[67] Unsurprisingly, this practice can add a good amount of textual content, not to mention employee labor hours, to a rulemaking. These effects are far from theoretical; rulemakings have been delayed based solely on an inundation of comments.[68] Delays and commitment of additional resources do not necessarily reflect failings in process. To the contrary, one might contend that they embody just the type of considered response to public input that the APA idealized.[69]

Until now, empirical study of public comments has largely fallen by the wayside. "While there are theories about the interaction between participation and deliberation and concerns about bureaucrats merely tallying comments as votes, few empirical studies exist using the volume of comments as an independent variable."[70] In making that observation, Stuart Shapiro explicitly pushed for

---

[64] *See* Wendy Wagner et al., *Rulemaking in the Shade: An Empirical Study of EPA's Air Toxic Emission Standards*, 63 ADMIN. L. REV. 99, 118 (2011); *see also* Patrick Schmidt, *Pursuing Regulatory Relief: Strategic Participation and Litigation in U.S. OSHA Rulemaking*, 4 BUS. & POL. 71 (2002) (finding that, in a study of Occupational Safety and Health Administration rules, comments were the most influential input to agency work on a rule because they posed immediate risks of litigation).

[65] Wendy Wagner, *Revisiting the Impact of Judicial Review on Agency Rulemakings: An Empirical Investigation*, 53 WM. & MARY L. REV. 1717, 1721–22 (2012).

[66] *See* Bressman & Vandenbergh, *supra* note 48, at 50–51, 78 (reporting, based on surveys of federal agency employees, that the EPA scores well on accountability, which includes transparency and accountability).

[67] Wagner et al., *supra* note 64, at 155 ("In each rule there are often many—usually dozens of these individual comment-responses—to explain the changes made in the final rule.").

[68] *See, e.g.*, Delay in Issuing 2014 Standards for the Renewable Fuel Standard Program, 79 Fed. Reg. 73007, 73007–08 (proposed Dec. 9, 2014) (to be codified at 40 C.F.R. pt. 80) (delaying standards on the basis that "[t]he proposal has generated significant comment and controversy, particularly about how volumes should be set in light of lower gasoline consumption than had been forecast at the time that the Energy Independence and Security Act was enacted, and whether and on what basis the statutory volumes should be waived").

[69] *But see* Katzen, *supra* note 11, 1508 & n.72 (arguing that Bressman and Vandenbergh, and others who advocate for transparency and public participation, should more explicitly "acknowledge that confidentiality is often important to honest deliberations and that candor is an important ingredient in collegial decision-making"); Wagner et al., *supra* note 64, at 105 ("Yet while the opportunity to lodge comments is a vital step that ensures that the agency is adequately educated about the issues, the comment process, standing alone, does not ensure that the agency will take the comments seriously.").

[70] Stuart Shapiro, *Does the Amount of Participation Matter? Public Comments, Agency Responses and the Time to Finalize a Regulation*, 41 POL'Y SCIS. 33, 34 (2008) ("In particular, the impact of the number of comments on agency decision making has been virtually ignored in the literature."); *see, e.g.*, O'Connell, *supra* note 60, at 491 (noting explicitly that the

an answer to the question that the study herein asks—do higher numbers of comments result in longer rules that take more time and resources to draft?[71] The call-and-response hypothesis suggests an affirmative answer to that inquiry.

Over the four decades of the modern administrative state's existence, much has changed in society, particularly with respect to how we communicate, disseminate information, and engage with the government. Lagging slightly behind society at large, the federal government, through OMB, finally moved to digitize the rulemaking process in 2002. The OMB's electronic rulemaking platform, regulations.gov, launched that year, accompanied by estimates that it would increase public participation (i.e., comments) by 600%.[72] Scholars have since examined those claims in limited circumstances. Initially, some early literature parroted the government's claims, cautioning that while the number of comments would increase, their quality would not improve.[73] Others championed the internet's ability to engage previously left out segments of the population, but expressed skepticism that officials would truly consider these newly generated comments.[74] One later study of the Department of Transportation rules issued before and after the implementation of e-rulemaking observed increased participation only in the most prominent rulemakings.[75] The data analysis below adds to this relatively meager literature and provides new insight on e-rulemaking's effect on EPA processes.

Regardless of its theatrical merit,[76] the notice-and-comment process still carries tremendous weight to scholars and practitioners alike.[77] It remains the formal

database used in her study lacked "*all the information present in the Federal Register notices of rulemakings* (for instance, *the number of comments received in a notice of final rulemaking*)" (emphasis added)).

[71] Shapiro, *supra* note 70, at 34 ("Further, do additional comments mean that agencies take more time to finalize proposed rules, because they have to respond to the comments (even if only to reject them)? An affirmative answer to this last question would mean that a greater volume of comments could lead to further ossification of the rulemaking process.").

[72] MARK FORMAN, EXEC. OFF. OF THE PRESIDENT, OFF. OF MGMT. & BUDGET, E-GOVERNMENT STRATEGY: IMPLEMENTING THE PRESIDENT'S MANAGEMENT AGENDA FOR E-GOVERNMENT 27 (2002), http://xml.coverpages.org/OMB-egovstrategy200202.pdf [perma.cc/UCR3-VWQ3].

[73] Fred Emery & Andrew Emery, *A Modest Proposal: Improve E-Rulemaking by Improving Comments*, 31 ADMIN. & REGUL. L. NEWS 8, 8 (2005).

[74] *See* J. Woody Stanley & Christopher Weare, *The Effects of Internet Use on Political Participation: Evidence from an Agency Online Discussion Forum*, 36 ADMIN. & SOC'Y 503, 504, 522 (2004).

[75] *See* Steven J. Balla & Benjamin M. Daniels, *Information Technology and Public Commenting on Agency Regulations*, 1 REGUL. & GOVERNANCE 46, 61–62 (2007).

[76] *See* Elliott, *supra* note 59, at 1492.

[77] *See* Scott R. Furlong, *Interest Group Influence on Rule Making*, 29 ADMIN. & SOC'Y 325, 339–41 (1997) (documenting the importance industry places on communication with regulators); Richard J. Pierce, Jr., *Seven Ways to Deossify Agency Rulemaking*, 47 ADMIN. L. REV. 59, 84–86 (1995); Michael Asimow, *Nonlegislative Rulemaking and Regulatory Reform*, 1985 DUKE L.J. 381, 403 (1985) ("An invitation to submit comments stimulates outsiders to furnish data and other inputs, providing a source of low-cost information to agency decisionmakers. A rule is likely to be a better product if its drafters must consider seriously alternatives that

channel of communication between the regulated and their regulators. It also comprises a crucial piece of the administrative record, preserving (or foreclosing) arguments and evidence for judicial review purposes. It is thus not surprising that even in an age of informal, backchannel influence, comments continue to pour in from the full range of stakeholders—from ordinary citizens to interest groups to Fortune 500 corporations. The number of such comments also continues to vary from rulemaking to rulemaking. This study will empirically examine the impact of that fluctuation.

B. Previous Empirical Research

There has been some previous empirical work analyzing the number of comments and the identity of commenters, testing hypotheses about influence and regulatory capture.[78] At least one recent study combined public comment data with corporate disclosure data to expose unethical, and potentially illegal, double speak on environmental issues.[79] While the results produced reveal valuable truths about the regulatory state, they leave part of the picture to be uncovered. This study sets out to bring even more of the picture into focus.

Data compiled on the identity of commenters reveal the perhaps unsurprising truth that the majority of public comments come from well-funded corporate interests.[80] Furthermore, the powerful ability of those comments to change the

---

they might have overlooked or take account of practical problems that otherwise would crop up only after a rule goes into effect.").

[78] *See, e.g.*, Jason Webb Yackee & Susan Webb Yackee, *A Bias Towards Business? Assessing Interest Group Influence on the U.S. Bureaucracy*, 68 J. POL. 128, 131 (2006) (examining the effect of interest group comments on forty rulemakings across four federal agencies); Susan Webb Yackee, *Assessing Inter-Institutional Attention to and Influence on Government Regulations*, 36 BRIT. J. POL. SCI. 723, 725 (2006) (analyzing the effect of comments and political institutions on forty rulemakings); Wagner et al., *supra* note 64, at 103 (analyzing how the number of industry comments on proposed EPA emission standards translated into changes in final rules).

[79] *See* James W. Coleman, *How Cheap Is Corporate Talk? Comparing Companies' Comments on Regulations with Their Securities Disclosures*, 40 HARV. ENV'T L. REV. 47, 49, 54–55 (2016).

[80] *See* Yackee & Yackee, *supra* note 78, at 133 (finding, in a study of more than thirty separate rulemakings, that business interests submitted 57% of comments and nongovernmental organizations submitted 22%); Cary Coglianese, Challenging the Rules: Litigation and Bargaining in the Administrative Process 73 tbl.2-2 (1994) (Ph.D. dissertation, University of Michigan) (on file with author) (finding businesses participated in 96% of rulemakings studied, while national environmental groups participated in only 44%); Wagner et al., *supra* note 64, at 119, 128–29 (confirming the following hypothesis: "The formal comments lodged with the agency on a complex rule will come predominantly from regulated industry, and the changes made to the proposed rule in the final rule will track this imbalance and generally favor industry" with findings showing that industry comments comprised an average of more than 81% of those submitted on a dataset of Hazardous Air Pollutant regulations, as opposed to the 4% from public interest groups); *see also* Marissa Martino Golden, *Interest Groups in the Rule-Making Process: Who Participates? Whose Voices Get Heard?*, 8 J. PUB. ADMIN. RSCH. & THEORY 245, 247 (1998); William F. West, *Formal Procedures, Informal Processes, Accountability, and Responsiveness in Bureaucratic Policy Making: An Institutional Policy Analysis*, 64 PUB.

substance of proposed rules has also been proved empirically. Yackee and Yackee examined approximately 1,700 public comments across more than thirty rules and made two important findings: (1) "business commenters, but not non-business commenters, hold important influence over the content of final rules," and (2) "as the proportion of business commenters increases, so too does the influence of business interests."[81]

Compounding the problem, corporate disclosure data suggests that entities understand this influence disparity and manipulate it. Looking at a dataset including 3,955 unique public comments on the EPA's renewable fuel standards from 2010 to 2013,[82] James Coleman "empirically demonstrate[d] that oil companies facing adverse regulations tailor their messages to each audience–emphasizing the cost and economic danger of regulation to regulators while telling shareholders that regulation is merely a cost of doing business with few negative impacts."[83]

These previous studies undercut, with sound empirical evidence, the normative arguments championing the APA's public participation process. They suggest instead that notice-and-comment procedures, at least in current practice, tend to perpetuate inequities and do not resemble democratic accountability.[84] Reasons for this failing could be external to the process's design. Some adjacent work has revealed how resource disparities between public interest groups and industry groups contribute to the disparity in attention and influence. As Wendy Wagner has astutely pointed out, for every judicial review petition filed by an environmental nonprofit about an arbitrarily under-protective EPA rule, there may be ten or more similarly flawed rules that go unchallenged, half of those not even having been commented on.[85] Virtually no arbitrarily overprotective rules go similarly unnoticed.

There exists limited research on the effects of commenting dynamics on EPA resource allocation and workload. The one study that looked for a link between number of comments and the time it takes an agency to finalize a rule (a relatively good proxy for resource intensity) had a sample size too small to yield

---

ADMIN. REV. 66, 70 (2004); Susan Webb Yackee, *Sweet-Talking the Fourth Branch: The Influence of Interest Group Comments on Federal Agency Rulemaking*, 16 J. PUB. ADMIN. RSCH. & THEORY 103, 103–04 (2005).

[81] Yackee & Yackee, *supra* note 78, at 128. *But see* Bressman & Vandenbergh, *supra* note 48, at 87–88 (surveying EPA officials, a majority of whom acknowledged that White House and OIRA review and oversight pushed changes to proposed rules that served the national interest, as opposed to business or other special interests).

[82] *See* Coleman, *supra* note 79, at 61.

[83] *Id.* at 49.

[84] *See* Yackee & Yackee, *supra* note 78, at 128 (claiming that their findings "generally suggest that notice and comment procedures have not succeeded in 'democratizing' the agency policymaking process to the extent sometimes suggested in the normative rulemaking literature").

[85] *See* Wagner, *supra* note 65, at 1740.

statistically significant results.[86] Within that small sample, "the rules that took the longest were those that were either simple, with few comments, or complicated, with many."[87] The latter of those findings is consistent with the call-and-response hypothesis, while the former is not. Some related work has demonstrated that new forms of rulemaking—negotiated rulemaking in particular—increase satisfaction of all stakeholders, including within the agency.[88] That result is consistent with the increased attention paid to notice-and-comment alternatives in recent years. The study of the call-and-response hypothesis herein will thus provide crucial data for an ongoing debate about the most efficient and effective administrative processes.

IV. PREAMBULAR HYPOTHESES

A. Introduction and Preliminary Analyses

The twin hypotheses concerning the preambular text build on the theory, practice, anecdotal observations, and limited existing empirical literature described above.[89] Those sources suggest that agencies have devoted more resources (and words) to preambular sections in Federal Register entries in concert with those entries growing in overall length. Scholarship and practitioner behavior further suggest that a prominent strategic goal of the change in preambular text is insulation from judicial review. The following two hypotheses translate that theory and intuition into something capable of empirical examination:

> Preambular Hypothesis One: the length of a rule's preamble in the Federal Register, measured in number of words, correlates positively with that rule's ability to withstand judicial review.

> Preambular Hypothesis Two: the relative length of a rule's preamble sections in the Federal Register as opposed to the text of the rule itself, measured by the percentage of words in the Federal Register entry that are preambular, correlates positively with that rule's ability to withstand judicial review.

---

[86] *See* Shapiro, *supra* note 70, at 42 ("A sample size of nine cases is obviously not large enough to use statistical techniques to analyze the relationship between dependent and independent variables.").

[87] *Id.* at 46. "[A]gencies are likely to make changes to proposed rules only when there is a high volume of comments and the rule is complex. This can be explained as the situation when the agency is most likely to receive information from the public that will help it better understand the area in which it is regulating." *Id*. at 43.

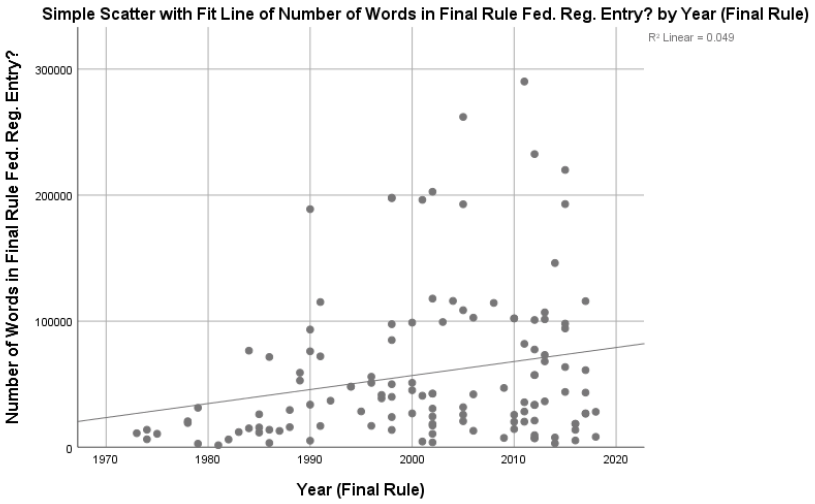
[88] *See* Laura I. Langbein & Cornelius M. Kerwin, *Regulatory Negotiation Versus Conventional Rule Making: Claims, Counterclaims, and Empirical Evidence*, 10 J. PUB. ADMIN. RSCH. & THEORY 599, 603–05 (2000) (reporting more satisfaction with negotiated rules and finding significantly more negative comments for conventional rules).

[89] *See supra* Part II.

Before testing those two hypotheses, some work needs to be done to establish the underlying assumption—that preambular word count grew in concert with overall word count. For that purpose, the dataset used for this study was subjected to two basic scatterplot and trendline analyses.

First, since the dataset represents only a subset of rules promulgated during the study period (1975–2020), initial analyses set out to confirm that these particular Federal Register entries conformed with the overall trend of increasing length over time. For that purpose, a scatterplot of the word count for each final rule Federal Register entry by the year of that rule's publication was produced.

FIGURE 2

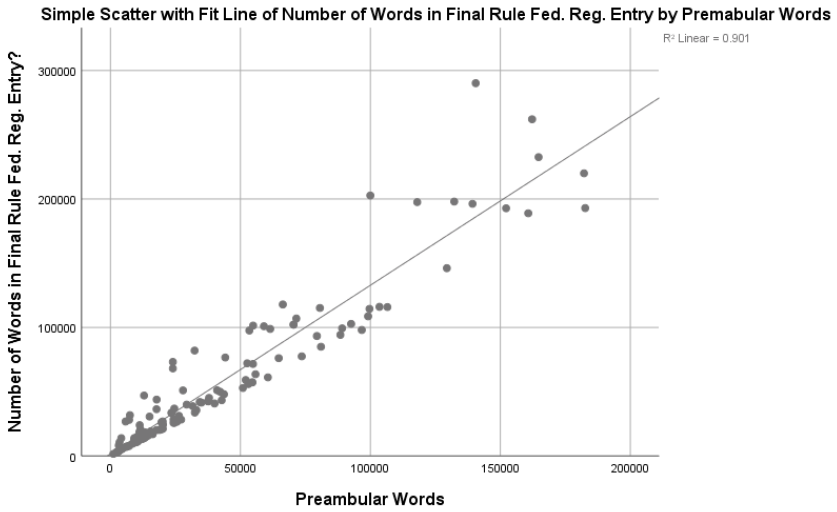


The sample shows a similar trend to the overall length-over-time data reproduced earlier in Figure 1. However, the sample is much noisier than the more inclusive dataset based on yearly average lengths of all rules published in the Federal Register. This is unsurprising; the study data represent individual Federal Register entry lengths in number of words, while the larger trend data utilized average lengths in pages. In other words, the data used for this study has a finer resolution on two dimensions and thus produces more noise. Importantly, however, the general trend is the same—Federal Register entries with more words came later in time.

The next, and more revelatory, preliminary analysis required for this study examined the connection between the amount of preambular text (measured in words) and the overall word count of Federal Register entries for final rules. In order to proceed to the analyses below trying to unpack empirically the reasons for increased preambular text, it was vitally important to connect that increase in preambular text back to the initial finding that prompted all of the analyses herein. This statistical analysis demonstrates empirically the theoretical intuition

that growth in preambular text has driven, at least in part, the growth in overall rule length over time.

FIGURE 3



MODEL SUMMARY

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.949[a]	.901	.900	18914.434

a. Predictors: (Constant), Number of words in statement of basis and purpose/preamble/supplementary materials?

The trend is clear and statistically quite powerful. As preambular word count grows, so grows the overall word count. Indeed, the regression model results indicate that the number of preambular words accounts for 90% of the variation in overall word count between Final Register entries. That result is significant and perhaps unsurprising. Consistent with the below analysis on "Preambular Hypothesis Two," preambular text comprises large percentages of many rule entries. Hence, one would expect, and the data confirm, that more preamble correlates positively and significantly with more rule text.

Having established that the sample exhibits the same trend of increasing rule length over time and demonstrates a strong positive correlation between preambular length and overall length, the table is set for the more complex analyses below. Before moving on, let us pause to contemplate the implications of these preliminary findings. The strong correlation between preamble word count and overall word count helps refine the foundational question of *Word Limited* and this study. No longer is the inquiry, "Why have *rules* been getting longer?" Instead, this study, and others that follow it, can, and should, ask why *preambles*

are getting longer. The analyses below test empirically some hypothetical explanations for that trend—first with respect to judicial review and then, in the next part, with respect to public participation.

B. Methodology

This analysis relied on a dataset comprised of 130 EPA rules from 1973 to 2018 and the opinions resulting from their judicial review in one of the United States Courts of Appeals or the United States Supreme Court. The data was extracted from the Federal Register entries for the final rules and the opinions reviewing them. That dataset was constructed by first identifying, via a search query in LexisNexis, United States Courts of Appeal and Supreme Court cases wherein EPA rules were mentioned alongside arbitrary and capricious review.[90] Those search results were then individually screened to create a dataset of 135 cases that actually involved arbitrary and capricious review of 130 EPA rules (as opposed to cases disposed of on other grounds, such as *Chevron* deference,[91] or cases involving rules from other administrative agencies). The 135 judicial opinions comprised the first half of the source data. From the list of cases, a list of the unique 130 rules subject to review in those cases was produced. The Federal Register entries for those final rules were analyzed to produce word counts for each rule entry in its entirety and the preamble portion of each. Those word counts made up the second half of the source data.

The opinions included within the source data were then coded on a binary dimension indicating that the EPA rule subjected to arbitrary and capricious review was either: (1) upheld or (2) reversed or remanded to the agency.[92] That binary coding comprised one half of the data necessary to perform a statistical analysis. The dual word counts for each of the Federal Register entries—entire entry and preamble—for the identified rules comprised the other half of the data necessary to test the second iteration of the insulation hypothesis.

Statistical analysis was then performed using IBM SPSS Statistics software.[93] Specifically, partial correlation and regression analyses[94] were performed on the dataset and a subset of the dataset, as well as descriptive statistical

---

[90] The specific query searched the LexisAdvance database of published Federal Courts of Appeal and United States Supreme Court opinions for the following terms in combination: "EPA" and "rule" and "arbitrary and capricious." The search returned 1,703 opinions that satisfied the parameters (search results on file with author).

[91] *See infra* notes 93–94 and accompanying text.

[92] In the dataset, a code "2.0" indicated a rule being upheld, while a code "1.0" indicated a rule being reversed or remanded.

[93] IBM SPSS Statistics, Version 25.0 (2017).

[94] Correlation analysis identifies the association (magnitude and direction) between two variables, or the absence of a significant association. *See Introduction to Correlation and Regression Analysis*, BOSTON U. SCH. PUB. HEALTH, http://sphweb.bumc.bu.edu/otlt/MPH-Modules/BS/BS704_Multivariable/BS704_Multivariable5.html [perma.cc/ZAW8-ZW92] (Jan. 17, 2013). Regression analysis generates a model meant to predict the value of one variable based on the known value of the other.

analyses for both the complete dataset and the selected subset. The partial correlation analyses performed produced the zero-order correlations (i.e., raw correlations without any control variables) between the year of the final rule, the words in the preamble sections of that rule's Federal Register entry, and whether that rule was upheld on judicial review; the analyses also produced the partial correlation of the preambular word count with whether that rule was upheld on judicial review (controlling for the year the rule was promulgated).[95]

C. Results

1. Preambular Hypothesis One: The Length of a Rule's Preamble in the Federal Register, Measured in Number of Words, Correlates Positively with That Rule's Ability to Withstand Judicial Review.

Table 1 provides the results testing for partial correlations on the entire dataset of EPA rules between the preambular word count for a rule's Federal Register entry, whether that rule was upheld on judicial review, and the year of the final rule. The results show both zero-order correlations and correlations controlling for the year of the final rule.

[95] Controlling for the year was necessary to ensure that the strong, positive linear relationship between year and rule length, observed in the larger dataset and reported *supra*, did not influence the results in testing the insulation hypothesis.

TABLE 1

Control Variables			Reverse/ Remand/ Vacate?	Preambular Words	Year (Final Rule)
-none-[a]	Reverse/ Remand/ Vacate?	Correlation	1.000	-.127	.055
		Significance (2-tailed)	.	.155	.538
		df	0	124	124
	Preambular Words	Correlation	-.127	1.000	.227
		Significance (2-tailed)	.155	.	.010
		df	124	0	124
	Year (Final Rule)	Correlation	.055	.227	1.000
		Significance (2-tailed)	.538	.010	.
		df	124	124	0
Year (Final Rule)	Reverse/ Remand/ Vacate?	Correlation	1.000	-.144	
		Significance (2-tailed)	.	.109	
		df	0	123	
	Preambular Words	Correlation	-.127	1.000	.227
		Significance (2-tailed)	.109	.	
		df	123	0	

a. Cells contain zero-order (Pearson) correlations.

The zero-order correlation between the number of preambular words and its fate on judicial review is both fairly low (-0.127) and not statistically significant ($p = 0.155$). Furthermore, the sign of the correlation (negative) is the opposite of hypothesized; a significant negative correlation would indicate that final rules with fewer preambular words enjoyed more success when subjected to arbitrariness review. The partial correlation, controlling for the effect of the year of a rule's publication, is only slightly different (-0.144) and, though closer to significance, still falls just short ($p = 0.109$). The slight change from the zero-order correlation is due to the significant positive correlation between the year and the number of preambular words, as well as the insignificant positive correlation between the year and a rule's resilience to judicial review.

The next set of tables provide the results for regression analyses, which sought to determine whether the number of preambular words was a good predictor of whether a rule was upheld on judicial review. Two regression equations were constructed—one using the raw number of preambular words and a second

using a logarithmic transformation of the number of preambular words. Both equations sought to use their respective variable inputs (words and log(words)) to predict the result on judicial review. Neither of these equations proved useful.

Regression analyses rely on a number of assumptions about the underlying dataset and can be rendered invalid, or useless, if those assumptions are violated. The Hosmer and Lemeshow test provides a measure of the goodness of fit for a regression model. In other words, it measures how well the data fits the constructed model. The test relies on a comparison between observed event rates (here the number of rules reversed, remanded, or vacated in the dataset) and the predicted event rates (the number of reversals, remands, or vacations predicted by the model). A p-value of less than 0.05 signals a poor fit, and thus a model lacking statistical significance. Table 2 shows the Hosmer and Lemeshow results for the first two regression equations tested.

TABLE 2: HOSMER AND LEMESHOW TEST

Step	Chi-square	df	p-Sig.
1	16.683	8	.034

The low p-value (0.034) for these equations suggests a very poor fit—a fact that Table 3, comparing the observed to predicted results, confirms.

TABLE 3

	Observed		Predicted		
			Reverse/ Remand/Vacate?		Percentage Correct
			Y	N	
Step 1	Reverse/Remand/Vacate?	Y	10	46	17.9
		N	10	60	85.7
	Overall Percentage				55.6

While the regression models predicted success on judicial review for the rules within the dataset at a rate of over 85%, they failed to even reach 20% correct for rules that ultimately were overturned by the courts. These results confirm the poor fit indicated by the Hosmer and Lemeshow test and indicate that neither regression equation could function as a useful predictor of success on judicial review.

Nonetheless, Table 4 provides the specific parameters of the equations. Importantly, the "B" value indicates the coefficient in the equation for each of the variables.

TABLE 4: VARIABLES IN THE EQUATION

		B	S.E.	Wald	df	Sig.	Exp(B)	EXP(B) 95% C.I.	
								Lower	Upper
Step 1[a]	Preambular Words?	.000	.000	.656	1	.418	1.000	1.000	1.000
	preambular-words_log	.038	.322	.014	1	.905	1.039	.552	1.955
	Constant	.130	2.967	.002	1	.965	1.138		

a. Variable(s) entered on step 1: Number of words in statement of basis and purpose/preamble/supplementary materials?, preambularwords_log.

Most striking about these summary results was the "B" value of 0.000 for the variable that represents the raw number of preambular words. That result indicates an extremely small coefficient for that variable—less than 0.001. The probable reason for that tiny coefficient is the relatively large magnitude of the variable inputs—word counts were in the thousands. Thus, in order to get a complete picture, the regression analyses were run again, replacing the raw preambular word count variable with a variable based on thousands of preambular words. Unfortunately, dividing a variable by 1000 across the board will do nothing to change the fit of the model, producing the same poor results as shown in Tables 2 and 3. It did, however, reveal the B coefficient for the preambular word count variable as shown in Table 5.

TABLE 5: VARIABLES IN THE EQUATION

		B	S.E.	Wald	df	Sig.	Exp(B)	EXP(B) 95% C.I.	
								Lower	Upper
Step 1	Preambular-Words1000s	-.007	.008	.656	1	.418	.993	.977	1.010
	preambular-words_log	.038	.322	.014	1	.905	1.039	.552	1.955
	Constant	.130	2.967	.002	1	.965	1.138		

The "B" coefficient appears in this model, but it still registers quite low. And, furthermore, a coefficient of less than one suggests that the probability of surviving judicial review would decrease as the number of preamble words increased. The Exp(B) result helps to clarify this by showing the ratio-change in the odds of reversal for a one-unit change in the predictor. Hence, the Exp(B) for thousands of preambular words is equal to 0.993, which means that the odds of surviving judicial review (i.e., a "N" for reverse, remand, or vacate) for a rule

with 1,000 preambular words is 0.993 times the odds of survival for a rule with one preambular word, all other variables being equal.

The negative value of the "B" coefficient in the equation was thus inconsistent with Preambular Hypothesis One. The model predicated that as 1000s of preamble words went up, the chance of survival on judicial review went down. Fortunately, and unsurprisingly, the regression model remained insignificant. Thus, the length of a given rule's preamble sections is not an accurate or useful predictor of whether it will survive judicial review. This result is consistent with the regression results based on the final rule length variable that were previously reported.[96] Indeed, the combination of a finer resolution variable, preambular word count, and a more sophisticated regression model help to confirm that policymakers cannot confidently use devotion of more rule-writing resources alone as a strategy to insulate their decisions from judicial review.

2. Preambular Hypothesis Two: The Relative Length of a Rule's Preamble Sections in the Federal Register as Opposed to the Text of the Rule Itself, Measured by the Percentage of Words in the Federal Register Entry That Are Preambular, Correlates Positively with That Rule's Ability to Withstand Judicial Review.

The second preambular hypothesis focuses on the relative length of preambular text to the length of the regulatory text. Consequently, the source data for the analysis relies on the word counts for preamble sections and for the entire Federal Register entry to create percentage figures. These percentages indicate how much of a final rule's entry is comprised of preambular text. Table 6 provides the results testing for partial correlations on the entire dataset of EPA rules between the preambular word percentage for a rule's Federal Register entry and whether that rule was upheld on judicial review, controlling for the year of the final rule.

[96] *See* Moffa, *supra* note 12, at 755.

TABLE 6: CORRELATIONS

Control Variables			Reverse/ Remand/ Vacate?	Preamble Percentage?
Year (Final Rule)	Reverse/Remand/ Vacate?	Correlation	1.000	.115
		Significance (2-tailed)	.	.200
		df	0	123
	Preamble Percentage?	Correlation	.115	1.000
		Significance (2-tailed)	.200	.
		df	123	0

The initial results are consistent with the second preambular hypothesis—a small positive correlation (0.115) existed between the percentage of words in a final rule entry that were preambular and the rule's success on judicial review. Unfortunately, upon closer examination, the results reveal that correlation to lack statistical significance (p=0.200).[97] Without a statistically significant correlation, these initial findings cannot confirm a relationship exists between the amount of a relative preamble text for a final rule and its fate on judicial review. Nonetheless, at least the direction of the statistically insignificant correlation is consistent with, rather than contrary to, the hypothesis.

Hope for a useful regression model was slightly higher given the positive correlation. A regression equation was constructed using two input variables based on the preamble percentage—the raw preamble percentage and the log preamble percentage—to predict the result on judicial review. The preliminary analysis of the fit of the regression model is reported in Tables 7 and 8.

[97] A p-value of 0.05 or less would have indicated significance using a two-tailed measurement.

TABLE 7: HOSMER AND LEMESHOW TEST

Step	Chi-square	df	Sig.
1	10.208	8	.251

TABLE 8: CLASSIFICATION TABLE[a]

Observed		Predicted		
		Reverse/Remand/Vacate?		Percentage
		Y	N	Correct
Reverse/Remand/Vacate?	Y	9	47	16.1
	N	7	63	90.0
Overall Percentage				57.1

a. The cut value is .500.

The results, particularly the Hosmer and Lemeshow test, indicated a better fit than the regression model produced for the first preambular hypothesis. As described above, a test result with a p-value of less than 0.05 signals a poor fit, and thus a model lacking statistical significance. For this hypothesis, the p-value was 0.251—not a powerfully significant result, but certainly above the 0.05 threshold. On the other hand, the classification table shows that the model did a poor job of predicting the instances where a rule was reversed, remanded, or vacated, while it predicted success on judicial review much more accurately. Overall, the model performed only a few points better than the model produced for the first hypothesis.

The regression equation that the model produced is described by Table 9.

TABLE 9: VARIABLES IN THE EQUATION

		B	S.E.	Wald	df	Sig.	Exp(B)	EXP(B) 95% C.I.	
								Lower	Upper
Step 1[a]	Preamble Percentage?	-.017	.049	.121	1	.727	.983	.892	1.083
	preamblepercentage_log	1.778	2.946	.364	1	.546	5.917	.018	1905.997
	Constant	-6.090	8.976	.460	1	.497	.002		

a. Variable(s) entered on step 1: Preamble percentage?, preamblepercentage_log.

Unfortunately, even though the fitness tests suggest that the model might help confirm the second preambular hypothesis, or at least prove useful, the full results suggest otherwise. In terms of confirming the hypothesis, the "B" coefficient being less than one suggests exactly the opposite—the model decreases the probability of surviving judicial review as preamble percentage increases. The

Exp(B) result, the ratio-change in the odds, comes in very near, but less than, one, indicating that each percentage more of rule text taken up by preamble lowers the predicted success rate on judicial review at a ratio-change of 0.983. That equation is not only inconsistent with the hypothesis, it is inconsistent with the correlation results (showing an insignificant positive association between preamble percentage and success on judicial review). Among the other reported regression results, the Wald statistic helps to confirm that the parameter of preamble percentage actually does little to drive the model. The Wald statistic measures the ratio of the coefficient to its standard error, squared, and its significance level must be small (less than 0.05) for a parameter to register as useful to the model. Here, neither variable—preamble percentage nor the log of the figure—had significant Wald statistics. Accordingly, the regression model, which despite showing at least a better fit with the data conflicted with the correlation results and contradicted the hypothesis, did not prove useful when it came to the variable of interest.

The combined results of the statistical analyses for the preambular hypotheses failed to confirm either. There exists no empirical support for the idea that more preamble words—either absolutely or relatively—better insulates a final rule from judicial review. Those results complete the picture on the larger question of whether increased rule length can be logically explained by a strategic desire to avoid judicial reprimand on the part of federal agencies. The data say quite clearly that it cannot.

V. CALL-AND-RESPONSE HYPOTHESIS

A. Introduction

The "call-and-response" hypothesis originates from evolving judicial interpretation of the APA's requirements and the response from policymakers and scholars, especially after being confronted with the foundational data on rule length over time. The theory itself is quite intuitive.

> "Call-and-response" Hypothesis: the number of public comments on a rulemaking correlates positively with the length of the final rule's Federal Register entry, measured in number of words.

If proved true, this hypothesis provides a neat, simple, and logical explanation for the difference in the average length of rules from the dawn of the modern administrative state to today. Although the hypothesis does not directly speak to substantive judicial review, the now well-established requirement that agencies respond to significant comments ties together number of comments, overall length, and reversal or remand on procedural grounds. In other words, the connection between responding to comments and procedural challenge is clear and

direct such that more comments should lead to longer rules (provided agencies do not want rules remanded on any grounds—procedural or substantive).

The call-and-response hypothesis further coincides with the Preambular Hypotheses because the responses to comments appear in the preamble sections of a rule's Federal Register entry. Just as dividing the entries into preambular and regulatory text provided finer resolution data with which to analyze the growth in overall rule length, looking within the preamble at the responses to comments adds another level of detail. If, as the preliminary analysis above demonstrates,[98] increases in preambular words can explain some or all of the growth in Federal Register entries for final rules, then increases in comments could explain the growth within that preambular text. A result confirming the call-and-response hypothesis would thus help confirm an important underlying assumption and be consistent with the overarching conclusion of the Preambular Hypotheses.

B. Methodology

This analysis relied on the same foundational dataset as the insulation hypothesis—130 EPA rules from 1973 to 2018 and the opinions resulting from their judicial review.[99] However, in order to test the call-and-response hypothesis, the addition of data on the number of public comments for each of the rules was necessary. Unfortunately, the number of public comment submissions[100] could not be obtained for every rule on the initial list of 130. Two sources of information regarding public comments were utilized to compile the data necessary to complete the analysis. First, the Federal Register entry for each final rule was scrutinized for any mention of the number of public comments (or commenters) in the text. EPA did not uniformly report this information, so it was not consistently available.[101] Second, for final rules promulgated after 2002, public comment data was available on regulations.gov. Where both sources of data were available for a particular rulemaking, the data entry used for statistical analysis was the more precise figure or the lower magnitude figure if both sources reported the number with the same precision.

---

[98] *See supra* Figure 3.

[99] *See supra* Section IV.B (describing the data compilation methodology).

[100] Due to inconsistencies in recording and styles across multitudes of public commenters over a half century, this study analyzes the number of comment documents submitted to the EPA rather than the number of individual "comments" within those documents. In other words, if an interested party submitted a numbered list of criticisms of a particular rule in a letter to the agency, said letter would count as "one" comment for the purposes of this analysis. The limitations of this method are explored in Coleman, *supra* note 79 at 61 n.63 ("Many more comments were filed as part of letter-writing campaigns, but EPA does not include duplicative comments in its online docket. As of March 2015, the controversial 2014 rule had received 344,326 comments.").

[101] *Contra* O'Connell, *supra* note 60, at 491 (suggesting that Federal Register entries routinely "contain individual-level information about a rulemaking—when it commenced, when the public period opened, how many comments were received, when the final rule was issued, what the rule is about, whether the rule is significant, and so forth").

The number of comments on individual rulemakings varies widely, and the subset of rules providing the data for this study proved no exception. To present the most complete picture, and provide metrics on which to classify outliers, a descriptive statistical workup of the dataset was necessary. Table 10 provides those descriptive results.

TABLE 10

Statistics		
Number of Comments		
N	Valid	106
	Missing	875
Mean		37628.50
Median		147.00
Std. Deviation		122154.582
Range		756624
Minimum		1
Maximum		756625

The number of public comments covered a wide range—from a few rules that prompted only one comment to a small number that prompted half a million. Due to this wide range, two derivative datasets were constructed to allow for statistical analyses with outliers removed. Those datasets started from the median number of public comments (147) and the mean number of comments (37,629), in conjunction with the standard deviation (122,155). The first derivative dataset included only those rules generating comments within one standard deviation of that median number (i.e., rules with less than 122,302 comments).[102] The second derivative dataset included only those rules generating comments within two standard deviations of the mean (i.e., rules with less than 281,938 comments). Together with the raw data on the number of comments, that comprised one half of the data necessary to perform statistical analyses. The word count for each of the Federal Register entries for the identified rules comprised the other half of the data necessary to test the call-and-response hypothesis.

Statistical analysis was then performed using IBM SPSS Statistics software.[103] Specifically, partial correlation and regression analyses were performed on the dataset and a subset of the dataset,[104] as well as the descriptive statistical analyses reported above. The partial correlation analyses performed produced correlations between the number of comments on a rulemaking and the number of words in the final rule's Federal Register entry, controlling for the year, as well as separate correlations between the number of comments on a rulemaking and the year of that rulemaking, controlling for the rule's overall length. The

[102] The standard deviation only produced an upper bound because, obviously, no rules generated negative comments.

[103] IBM SPSS Statistics Version 25.0 (2017).

[104] *See supra* note 94.

descriptive statistics revealed, among other things, the means, medians, and standard deviations for the number of comments. The descriptive statistics allowed for the identification, and subsequent removal, of outliers. After the outliers were removed from both the complete dataset and the selected subset, the partial correlation and regression analyses were re-run, producing new sets of results for each.

C. Results

Table 11 provides the results testing for partial correlations on the entire dataset of EPA rules between the number of public comments on that rulemaking, the number of words in the Federal Register entry for the final rule, and the year of the final rule. The results show both zero-order correlations and correlations controlling for the year of the final rule.

TABLE 11: CORRELATIONS

Control Variables			Number of Words in Final Rule Fed. Reg. Entry?	Number of Comments	Year (Final Rule)
-none-[a]	Number of Words in Final Rule Fed. Reg. Entry?	Correlation	1.000	.303	.246
		Significance (2-tailed)	.	.002	.012
		df	0	102	102
	Number of Comments	Correlation	.303	1.000	.303
		Significance (2-tailed)	.002	.	.002
		df	102	0	102
	Year (Final Rule)	Correlation	.246	.303	1.000
		Significance (2-tailed)	.012	.002	.
		df	102	102	0
Year (Final Rule)	Number of Words in Final Rule Fed. Reg. Entry?	Correlation	1.000	.248	
		Significance (2-tailed)	.	.012	
		df	0	101	
	Number of Comments	Correlation	.248	1.000	
		Significance (2-tailed)	.012	.	
		df	101	0	

a. Cells contain zero-order (Pearson) correlations.

The zero-order correlation between the number of comments and the number of words in the Federal Register entry reports as modestly positive (0.303) and significant ($p = 0.002$). Because, within the dataset of EPA rules (consistent with the broader general trend), the correlation between a rule's length and the year of its promulgation was positive, the partial correlations controlling for year provide the more useful picture. Very little change is observed when year is controlled for. The partial correlation between the number of comments and the number of words in the Federal Register entry remains modestly positive (0.248) and significant ($p = 0.012$).

That significant, independent positive correlation is consistent with the call-and-response hypothesis. Within the sample, as the number of comments on a

rulemaking increased, the overall length of the final rule increased in proportion. The correlation being less than one suggests, perhaps unsurprisingly, that the number of public comments did not function as the lone driver of rule length. However, the strong statistical significance (a p-value of far less than 0.05) indicates that public comments have consistently contributed to the trend of agencies devoting more Federal Register words to final rule entries over time.

In addition to just controlling for the year, because of the rise in e-rulemaking, the correlation between the number of comments and the year of the rulemaking (independent of rule length) deserved some attention as well. This data begins to paint the next layer of the picture—if rules are getting longer at least in part because of increasing numbers of public comments (as the above analysis confirms), the next logical question begs an explanation for the growth in public commenting. As discussed in Part III above, there exists some practical and scholarly literature claiming relevant positive effects of e-rulemaking on public participation. The data reported here do not suffice to empirically settle the many questions about the effects of e-rulemaking and regulations.gov on the commenting habits of the public. Nonetheless, they provide some important context for the call-and-response hypothesis and a foundation for future research. Table 12 provides the results of correlation analyses looking at the interaction between the number of comments and the year of the final rule, here controlling for final rule length in the partial correlation.

TABLE 12: CORRELATIONS

Control Variables			Number of Comments	Year (Final Rule)	Number of Words in Final Rule Fed. Reg. Entry?
-none-	Number of Comments	Correlation	1.000	.303	.303
		Significance (2-tailed)	.	.002	.002
		df	0	102	102
	Year (Final Rule)	Correlation	.303	1.000	.246
		Significance (2-tailed)	.002	.	.012
		df	102	0	102
	Number of Words in Final Rule Fed. Reg. Entry?	Correlation	.303	.246	1.000
		Significance (2-tailed)	.002	.012	.
		df	102	102	0
Number of Words in Final Rule Fed. Reg. Entry?	Number of Comments	Correlation	1.000	.247	
		Significance (2-tailed)	.	.012	
		df	0	101	
	Year (Final Rule)	Correlation	.247	1.000	
		Significance (2-tailed)	.012	.	
		df	101	0	

The results examining commenting over time are remarkably similar in magnitude and significance to those reported above demonstrating positive correlation between the number of comments and rule length. Looking to the partial correlation, the results show a modest positive correlation (0.247) at a high degree of statistical significance (p=0.012) between the year of a final rule and the number of public comments on the underlying rulemaking. The direction of that result—positive—generally comports with the limited prior empirical study of e-rulemaking discussed in Part III.[105] However, the magnitude of the correlation (0.247) does not suggest the type of 600% increase touted by OMB.[106] Thus, while this data provides empirical support for the notion that e-rulemaking has contributed to an increase in public comments over time, more detailed analysis is required. The data utilized for this study, for instance, lacks the precision to

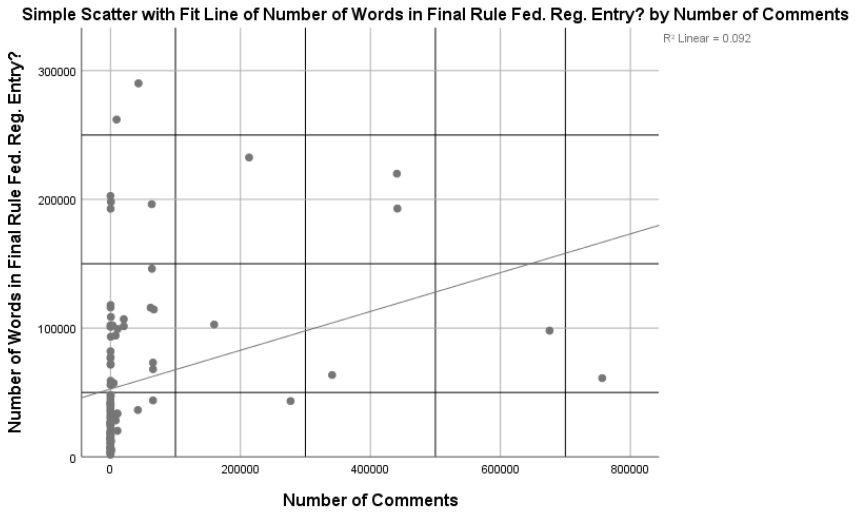
[105] *See supra* notes 73–76 and accompanying text.

[106] FORMAN, *supra* note 72, at 27.

examine the potential effects of individual developments within the technology and e-rulemaking space (e.g., the availability of high-speed internet, the utilization of regulations.gov, and the online outreach activities of federal agencies).

Returning to the call-and-response hypothesis itself, Figure 5 provides a scatterplot and summary results for a regression analysis, which sought to determine whether the number of public comments in a given rulemaking was a good predictor of the word count of the Federal Register entry for the final rule.

FIGURE 5



MODEL SUMMARY

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.303[a]	.092	.083	58655.564

a. Predictors: (Constant), Number of Comments

The regression results indicate that the variation in the number of words in Federal Register entries cannot be explained by just the number of comments on those rules. More precisely, the R-square value (0.092) suggests that the regression model can explain about a tenth of the variation (9.2%) in the data. This suggests that the number of comments on a proposed rule can accurately predict the length of the Federal Register entry for the final rule about one out of ten times. This is consistent with the correlation results reported above, which indicated a statistically significant positive effect but not at an overwhelming magnitude—again, one must conclude that other factors beyond the number of public comments contribute to rule length. The relative inaccuracy of the regression model likely also precludes any practical application of it. In other words, contrary to the apparent strategy behind some public commenting campaigns,

stakeholders cannot be confident that commenting more will result in the agency spending more time and resources on the final rule. Although the regression analysis precludes some forward-looking utility, the positive correlations between words and comments reported above confirm the call-and-response hypothesis with respect to rules in the sample from the last four decades of EPA activity.

In addition to displaying the graphical representation of the regression equation, the scatterplot, perhaps more importantly, reveals the presence of some prominent outliers. A cluster comprised of the majority of rules in the dataset emerges at the far-left end of the plot. In other words, some rulemakings in the dataset generated a massive amount of comments, while the rest generated much smaller relative numbers. To present a fuller account of the story that the data tells, and to ensure that these attention-hoarding rules were not skewing the sample, the correlation and regression analyses were performed on the dataset with only those rules generating less than 122,302 (one standard deviation from the median) comments.

TABLE 13: CORRELATIONS

Control Variables			Number of Comments	Number of Words in Final Rule Fed. Reg. Entry?	Year (Final Rule)
-none-[a]	Number of Comments	Correlation	1.000	.404	.313
		Significance (2-tailed)	.	.000	.002
		df	0	94	94
	Number of Words in Final Rule Fed. Reg. Entry?	Correlation	.404	1.000	.192
		Significance (2-tailed)	.000	.	.061
		df	94	0	94
	Year (Final Rule)	Correlation	.313	.192	1.000
		Significance (2-tailed)	.002	.061	.
		df	94	94	0
Year (Final Rule)	Number of Comments	Correlation	1.000	.369	
		Significance (2-tailed)	.	.000	
		df	0	93	
	Number of Words in Final Rule Fed. Reg. Entry?	Correlation	.369	1.000	
		Significance (2-tailed)	.000	.	
		df	93	0	

a. Cells contain zero-order (Pearson) correlations.

As the results indicate (df), there were ninety-three rules that generated comments within the defined range. Comparing that to the above results for all rules that data was available, eight outliers have thus been excluded. With those high-comment rules excluded, the data show stronger (0.369) and more significant (p=0.000) positive correlations between the number of comments and the number of words in a final rule's Federal Register entry. Importantly, the sample size remained quite large, and therefore powerful, even after cutting it at one standard deviation from the median. The change in the correlation results was in the direction of confirming the call-and-response hypothesis—within this subset of the data, an increase in the number of comments *more* strongly and directly linked to an increase in overall final rule word count. Such a result bolsters the above finding confirming the call-and-response hypothesis.

Table 14 provides the summary results for a regression analysis with the one-standard-deviation outliers excluded. Again, the regression attempts to model whether the number of comments on a proposed rule can accurately predict the length of the Federal Register entry for the final rule.

TABLE 14: MODEL SUMMARY

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.404[a]	.163	.154	52107.026

a. Predictors: (Constant), Number of Comments

The regression analysis for this subset of the data suggests a better fitting predictive model, though one that probably still lacks practical utility. The R-square value is now higher (0.163), suggesting that this regression model can explain about a sixth of the variation (16.3%) in the data. Other factors beyond the number comments still clearly contribute, and without identifying them, measuring them, and including them as variables in the equation, the regression model is too imprecise to be of use.

Cutting the dataset at one standard deviation from the median represents a hybrid method of defining outliers that acknowledges the imbalance of the data in one direction (in this case, overwhelmingly large numbers of comments on a few rulemakings). Another, more traditional method would measure two standard deviations from the mean and deem anything outside of that range an outlier worthy of exclusion—in doing so the included data would capture roughly 95% of the values in the dataset.[107] Consistent with that standard approach in statistical analyses, one final subset of the data was constructed including only those rules generating comments within two standard deviations of the mean (i.e., rules with less than 281,938 comments and more than zero comments). Table 15 provides the results of correlation analyses performed on that subset.

[107] *See* RAND R. WILCOX, FUNDAMENTALS OF MODERN STATISTICAL METHODS: SUBSTANTIALLY IMPROVING POWER AND ACCURACY 32 (2d ed. 2010) (describing the traditional method of "[d]eclar[ing] a value to be an outlier if it is more than two standard deviations from the mean. In symbols, declare the value X to be an outlier if $|X - \mu| > 2\sigma$, the idea being that there is a low probability that the distance of an observation from the mean will be greater than two standard deviations. For a normal probability curve, this probability is 0.046.").

TABLE 15: CORRELATIONS

Control Variables			Number of Comments	Number of Words in Final Rule Fed. Reg. Entry?	Year (Final Rule)
-none-[a]	Number of Comments	Correlation	1.000	.335	.269
		Significance (2-tailed)		.001	.007
		df	0	97	97
	Number of Words in Final Rule Fed. Reg. Entry?	Correlation	.335	1.000	.206
		Significance (2-tailed)	.001		.041
		df	97	0	97
	Year (Final Rule)	Correlation	.269	.206	1.000
		Significance (2-tailed)	.007	.041	
		df	97	97	0
Year (Final Rule)	Number of Comments	Correlation	1.000	.297	
		Significance (2-tailed)		.003	
		df	0	96	
	Number of Words in Final Rule Fed. Reg. Entry?	Correlation	.297	1.000	
		Significance (2-tailed)	.003		
		df	96	0	

a. Cells contain zero-order (Pearson) correlations.

As the results indicate, this cut captured a few more rules than did the one-standard-deviation from the median. In this subset, ninety-six rules generated comments within the defined range, and thus only five outliers were excluded. The resulting correlation (0.297) between the number of comments and the word count of the final rule, controlling for year, came in between the reported correlation with no outliers excluded (0.247) and the reported correlation with a broader definition of outliers (0.369). Unsurprisingly, the results remained quite statistically significant (p=0.003).

Interestingly, the regression model for this subset, reported in Table 16, performed the worst of the three.

TABLE 16: MODEL SUMMARY

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.335[a]	.112	.103	55735.227

a. Predictors: (Constant), Number of Comments

Overall, the results of these analyses confirm the call-and-response hypothesis, which posited that the number of public comments on a rulemaking correlates positively with the length of the final rule's Federal Register entry, measured in number of words. For the approximately one hundred EPA rules in the dataset for which the relevant data was available, a statistically significant, modest positive correlation emerged. At least at the EPA over the last five decades, the agency devoted more written explanation, and thus resources, to rulemakings that generated more public comments. The data further confirm that the relationship between comments and rule length cannot be explained by a general trend towards more public comments as the agency matured. In other words, the agency has been consistent over time in its devotion of more rule-writing resources to rules with more public comments. The significant positive correlation between quantity of comments and final rule words adds an important layer to the developing picture of Federal Register entries growing longer over time. Taken together with the prior finding confirming that as rules have gotten longer those rules have each also tended to provide more net benefits to society, the empirical evidence now provides two important, rational explanations for increasing rule length—more public engagement and more public benefits.

VI. IMPLICATIONS OF FINDINGS

The empirical line of inquiry that this study contributes to seeks to unpack the underlying factors contributing to the increasing length of new regulations. The data demonstrate a clear historical trend exemplified by lower numbers of longer rules being added to the Federal Register each year. Regulatory reform advocates and champions of the administrative state alike must acknowledge the significance of that initial finding. Unlike empty political rhetoric about overzealous regulators or stacks of paper, the hard numbers show something has been changing in how agencies use the words of the Federal Register over time. Determining whether that change has marked an improvement in administrative process, good governance, and society writ large begins with the empirical analyses of this and similar studies.

Prior study hypothesized two common explanations for the increasing length of Federal Register rule entries. This study added to those explanations by drilling deeper on the question of insulation from judicial review and introducing the previously untested variable of public comment activity. The combined results of these studies confirmed two hypotheses outright—the socially beneficial hypothesis and the call-and-response hypothesis. Thus, over the past five decades,

rules with higher quantified societal benefits have tended to be longer, as have rules that generate more public comments. Furthermore, the positive correlations with rule length for both of these variables—net benefits and number of comments—reported as similar in magnitude (moderately positive, less than 0.6) and statistical significance (very high, p-values far less than 0.05). That information is revealing in a few important ways. First, it indicates that neither the amount of projected net benefits nor the number of comments is perfectly correlated with a final rule's length; in other words, neither variable alone could be responsible for the observed trend. Indeed, the insignificant regression analyses for both variables individually confirmed that finding. Second, the strong significance results for both variables indicate that, at least for the rules in the dataset, the projected benefits of a rulemaking *and* the number of public comments on it had separate non-random effects on the length of the final rule. And both of those effects tended to increase the rule's overall length.

On the other hand, the data failed to confirm the insulation hypothesis or either of the preambular hypotheses, all of which concerned the link between rule length and success on judicial review. Starting with the more general hypothesis, prior study failed to demonstrate any significant relationship between a final rule's length and its ability to withstand judicial review. Looking specifically at the section of a rulemaking designed to protect it from judicial review—the preamble—this study confirmed the lack of a significant relationship. Neither the raw nor relative length of a rule's preamble demonstrated a significant positive correlation with a rule's ability to withstand judicial review. Thus, one can state, based on empirical evidence, that adding content to a rule's preamble, and thereby lengthening its Federal Register entry, has no conclusive effect on the rule's resilience to challenge in court. That finding cuts against a strain of logic popular among policymakers and critics of allegedly searching review of administrative actions.

The study undertaken here set out to determine whether trend towards longer rules had any rational explanation(s). My prior study identified one such partial explanation—rules that confer more benefits are longer. Over time, the net societal benefits per word of final rules' Federal Register entries have remained fairly constant. Hence, we may have fewer individual rules, each comprised of more words, but agencies produce similar annual benefits through rulemaking activity. That result provided some comfort, while leaving open the possibility that more benefits could be generated through a more efficient allocation of rulemaking resources. The result confirming the call-and-response hypothesis added another rational explanation for the steady increase in final rule length over time—increasing numbers of public comments. Increasing a final rule's length in proportion to the number of public comments received on the proposal makes sense because current administrative law precedent demands agencies respond to significant comments. If a final rule entry ignored such comments, and was shorter as a result, it would be vulnerable to a relatively straightforward procedural challenge. Put another way, the law demands that agencies devote rule-writing

resources—both words and attention—to responding to individual public comments.

That legal reality invites a more fundamental query—does encouraging, facilitating, and ultimately responding to more public comments improve governance? Does the administrative state function more effectively and/or efficiently as a result of the legal necessity that increasing amounts of final rule text speak directly to input from public participation? The prevailing scholarly and practical thinking for many years stressed the importance of public participation in the administrative process, endorsing the tools that made it easier and celebrating the rules that garnered hundreds of thousands of comments. Some more recent scholarship takes a more nuanced view, unpacking the actual *influence* of different types of participation and participators. The empirical findings reported here help to quantify some of the costs of increased public participation. We now know definitively how more comments on a given rulemaking has translated to more words in the final rule. Assuming that agencies have limited rule-writing resources with a maximum number of rule words they can put in the Federal Register in a given year, one must consider whether better uses might be found for those words were responding to so many comments not necessary. A definitive answer to that question is not knowable, empirically or otherwise. Those convinced of the inefficiency of the current system, however, given the findings of this study, might suggest interesting reforms. Such reforms could either target the judicial precedent requiring responses to significant comments (joining with existing voices attacking this now old foe) or, more intriguingly, rethink the form of final rule entries to address comments more adeptly and succinctly. The results reported here should inform the debate over such potential reforms and ground it in empirics.

On the other hand, after this study, which scrutinized the words agencies use to explain their actions, there still exists no empirical support for any insulating effect of detailed explanations for rules. This study drilled down and looked specifically at the sections of the rule meant to provide that explanation, and therefore insulation. The data showed that these sections have grown in length consistent with the overall trend. However, the increased length of these sections has not correlated with survival on judicial review. That result should give policymakers pause, whose current approach across federal agencies reflects continued subscription to a theory of insulation. If those in power take the findings reported herein seriously, they would reallocate agency resources, particularly legal resources, away from the crafting of detailed analyses for the preamble sections of final rules. The data suggest that the amount of preambular content has no measurable effect on the outcome of a substantive challenge to the rule. Thus, the resources currently utilized to draft extremely detailed explanatory preambular content are likely not being utilized optimally. The results of this study, and its predecessor, demand a recalibration.

CONCLUSION

Much has been made of regulatory reform efforts over the past three to four decades. Too often, advocates use imprecise or irrelevant figures to make the case that the administrative state is inefficient or, worse, completely failing. The line of empirical work that this study continues set out to provide data-driven explanation for an observable trend—rules have been getting longer. The hope remains that uncovering statistically significant factors behind that trend will lead to more productive discussions about how to improve the administrative process. Having confirmed one of the three hypotheses tested here—the call-and-response hypothesis—a new dimension has been added to analysis of the public comment process as it relates to agency resource allocation. As the data reveal more about the words agencies use, so too will other dimensions of existing reform debates emerge. And if just one of those dimensions leads to improvement in process, the data has done its job.

[THIS PAGE INTENTIONALLY LEFT BLANK]