2023

Bureaucratic autonomy and the policymaking capacity of United States agencies, 1998–2021

Nicholas Bednar
University of Minnesota Law School, nbednar@umn.edu

Recommended Citation

Bureaucratic autonomy and the policymaking capacity of United States agencies, 1998–2021

Nicholas Ryan Bednar ⓘD

University of Minnesota Law School, Minneapolis, USA
Email: nbednar@umn.edu

Abstract

Despite a renewed interest in the health of the US administrative state, the absence of meaningful time-series measures of bureaucratic capacity hinders the testing of core theories of bureaucratic and executive politics. Using over 190 million personnel records, I estimate 5590 yearly policymaking-capacity scores for 261 unique agencies from 1998 to 2021. These measures provide an invaluable tool as either an independent or dependent variable in studies of administrative policymaking. To illustrate the value of these measures, I test longstanding theories about the relationship between bureaucratic autonomy and capacity. In contrast with emerging survey research, this study demonstrates that agencies with higher levels of structural independence have higher levels of policymaking capacity.

In 2019, the Bureau of Land Management (BLM) announced that it would move its headquarters from Washington, DC, to Grand Junction, Colorado—a city of fewer than 70,000 people and more than a four-hour drive to Denver. Rather than relocate, much of the workforce retired or found new jobs. BLM's headquarters lost 34 percent of its employees with over 25 years of experience. Employees told reporters that this loss of institutional knowledge made it difficult for the agency to weigh regulatory or legislative proposals, leading to delays in the policymaking process (Katz, 2021). According to one departed policy employee in the Department of the Interior, "There was an intentional effort to hobble the agency. Not just to reduce the workforce, but to diminish the capacity of the department" (Katz, 2023).

Scholars of bureaucratic institutions study the ways in which bureaucratic capacity affects policymaking. The expertise, experience, and teamwork within an agency's workforce are integral to this capacity. In the words of Justice Felix Frankfurter, "expertise is the lifeblood of the administrative process" (*Burlington Truck Lines, Inc v. United States*, 371 U.S. 156, 167 (1962)). Indeed, theories of delegation and agency design attribute the willingness of elected officials to relinquish control over policymaking to the agency's relative expertise and experience in a given policy domain (see, e.g., Gailmard, 2002; Huber and Sanford, 2004; Miller and Whitford, 2016). But scholars fail to test these theories with measures of capacity that tap into characteristics like expertise, experience, or teamwork. Instead, they rely on simple measures such as workforce size or self-reported perceptions of workforce skills.

The bureaucracy literature needs measures of capacity that better adhere to its conceptual foundations. In furtherance of this need, this study develops a new measure of policymaking capacity. I collect over 190 million personnel records from the Office of Personnel Management to better understand which bureaucrats perform tasks related to policymaking. I use the information

in these personnel records to create agency-level indicators that correlate with workforce expert-ise, experience, and retention. A Bayesian model aggregates these indicators to produce 5590 agency–year scores for 261 agencies from 1998 to 2021. The rank ordering of the scores comports with prior expectations about which agencies have the most-skilled and best-managed work-forces, and the scores predict the time it takes for an agency to finalize its rulemaking activities.

I use this measure to show that bureaucratic autonomy fosters the development of capacity. Autonomy attracts policy-motivated individuals to government service and encourages them to invest in their expertise and experience. In the United States, Congress and the president have sought to protect bureaucratic autonomy by designing agencies with certain insulating structures (Selin, 2015). Yet emerging research challenges the claim that insulating structures better protect capacity relative to structures that offer the president a wider degree of control (Devins and Lewis, 2023). If insulating structures adequately protect bureaucratic autonomy, then agencies with greater structural independence should exhibit higher levels of bureaucratic capacity. The results mirror this expectation and demonstrate how different methods of measuring capacity result in different empirical findings. I conclude by highlighting other theories that scholars may use these measures to test in the future.

1. Conceptualizing bureaucratic capacity

Bureaucratic capacity describes an agency's ability to complete the tasks delegated to it by Congress and the president. The most nuanced conceptualizations acknowledge that capacity has multiple dimensions, each of which must be assessed and measured separately (Geddes, 1994; Fukuyama, 2013).[1] Despite emerging discussions of "capacity" in the American literature, scholars have made few efforts to conceptualize or measure the concept as it pertains to the US bureaucracy. This study focuses on a key aspect of capacity—human capital—because scholars generally agree that the quality and management of an agency's workforce play a pivotal role in bureaucratic performance. All else equal, agencies whose workforces exhibit higher levels of human capital have a greater likelihood of enacting regulations, distributing benefits, and enforcing the law (Bolton *et al.*, 2015; Drolc and Keiser, 2020; Fisher and Shapiro, 2020).

Like the broader concept of capacity, human capital cannot be easily distilled to a single char-acteristic. A high-capacity workforce will have (1) substantive and procedural expertise, (2) an ability to recruit and retain skilled employees, and (3) an ability to organize itself for efficient team production. Many scholars describe these individual attributes in passing but few—if any—have sought to synthesize them as attributes of a single, latent concept.

For many scholars, bureaucratic capacity means expertise. Expertise describes the knowledge and skills possessed by the agency's workforce. But expertise also has various understandings (Fisher and Shapiro, 2020). Expertise may describe the scientific or technical knowledge brought to the agency by employees' educations and past careers (Miller and Whitford, 2016). It may also encompass the practical knowledge attained while working in public service. Experience hones the bureaucrat's professional judgment and teaches them how to use government procedures to advance the agency's interests (Gailmard and Patty, 2007; Potter, 2019; Fisher and Shapiro, 2020). As Mashaw (1983) states, experience teaches bureaucrats "what works and what doesn't when ferreting out information, what evidence is reliable and what is not—things that are an implicit part of the culture of the system but are not to be found in manuals or regulations" (67). These forms of expertise are not necessarily interchangeable and, therefore, a high-capacity agency will have both an educated and experienced workforce.

---

[1]Geddes's (1994) description of the bureaucracy as a machine requiring "human and material inputs to accomplish tasks" captures this multidimensionality: "The machine's ability to get things done can be undermined in three ways. Its material inputs can be inadequate, that is it can lack funds. Its human inputs can be inadequate, that is, it can lack sufficient expertise. Or, finally, its human inputs, having free will, can opt to pursue their own personal goals to the detriment of the agency's" (46).

The agency must also have the ability to recruit and retain skilled, public-service motivated employees. At baseline, an effective system of administrative governance must provide career civil servants with tenure (Geddes, 1994; Weber, 1978). The United States's robust system of civil-service laws offers that protection (Skowronek, 1982). However, agencies must also provide adequate financial remuneration to avoid rent-seeking behavior and corruption. Indeed, the Weberian ideal necessitates the employment of civil servants who need not rely on rent-seeking behaviors (Weber, 1978). Even absent rent-seeking behaviors, the lack of competitive salaries makes employment in the private sector a more attractive alternative to government service (Gailmard and Patty, 2007). The lack of competitive salaries for many civil servants poses a possible threat to bureaucratic capacity and, therefore, this threat should be minimized in agencies that offer higher salaries (Lewis, 2019).

A subtler attribute of human capital describes the managerial and organizational structures that allow the agency to make use of the diverse expertise within its workforce. Agencies must organize the workforce in ways conducive to team production (Williams, 2021). Agencies organize their workforces in ways that further their primary mission, neglecting tertiary tasks they see as less aligned with their core principles (Wilson, 1989; Barkow, 2013; Hickman, 2016). These structures are part cultural and part managerial. Culturally, agencies cultivate *esprit de corps* to unite the workforce behind the agency's mission. Managerially, the agency must train employees, organize them into teams, and find ways to maximize the exchange of information within the workforce. A poorly organized workforce may fail to see a key aspect of a problem confronting the agency or may move slowly in conveying information across teams.

A broader conceptual issue emerges when attempting to translate this definition into measures. We often think of capacity an amorphous concept that describes the agency's ability to do *something*. Too often, however, we fail to ask, "The capacity to do *what*?" While a workforce may have sufficient expertise and experience to complete one type of task (e.g., policymaking), it may lack the requisite capacity for a different type of task (e.g., law enforcement or adjudication). For example, the Federal Communications Commission (FCC)—an agency charged with regulating telecommunications—may operate comfortably with 1500 employees with law and telecommunications backgrounds. However, the same workforce would struggle to deliver 162 million pieces of mail to over 41,000 zip codes each day. Meanwhile, there is no guarantee that the Postal Service's 330,000 mail carriers would have sufficient expertise to manage the radio spectrum. Does the Postal Service have greater capacity than the FCC because it has more employees? Not necessarily. We can reduce errors in our estimates of capacity by tailoring measures to a single type of activity undertaken by a variety of agencies.

In addition, one should avoid confusing the concept of capacity with bureaucratic performance. Unlike performance, capacity concerns the agency's *prospective* ability to complete a task. The prospective focus of capacity means that an agency may have a latent ability to complete a particular task even if it chooses not to undertake the activity. Experienced careerists may use their knowledge of administrative procedure to delay policy proposals they dislike (Potter, 2019). Alternatively, Congress and the president may enact laws that limit the agency's authority to take certain actions. For example, subject to notable exceptions, only the Department of Justice has the authority to litigate on behalf of federal agencies (Selin and Lewis, 2018). Yet the nature of administrative work still requires many agencies to employ experienced litigators who would excel at defending their agency's actions in court. Capacity spills over from one task to another. Indeed, Congress and the president delegate new programs to existing agencies to harness this latent capacity rather than create new agencies from scratch (Hickman, 2016).

Finally, scholars ought to remain mindful of agency workloads when studying the relationship between bureaucratic capacity and performance. Poor performance may occur when an agency's workload exceeds its capacity. All else equal, two agencies with the same levels of capacity may experience different outcomes as a result of different workloads. Empirically, many studies of performance require scholars to control for both the capacity of the agency to perform a particular

task and the size of the workload attributable to that task. This study provides measures of human capital for the policymaking context, but scholars will also need measures of workload tailored to their particular task of interest.

2. Existing measures of human capital

Existing measures of human capital come in two varieties. The crudest measures rely on the number of full-time employees within an agency (see Bolton *et al.*, 2015; Potter and Shipan, 2019). Although workforce size is an important component of an agency's ability to complete many tasks, it does not capture the level of expertise or experience within the workforce. For many tasks, such as policymaking, expertise and experience play a greater role in task completion than workforce size. Moreover, measuring the entirety of the workforce introduces measurement error by failing to control for what proportion of the overall workforce is capable of performing the task of interest.

More sophisticated measures rely on surveys of civil servants. For example, Richardson *et al.* (2017) present federal executives with a list of federal agencies and ask, "In your view, how skilled are the workforces of the following agencies?" The authors then aggregate the responses to create measures of workforce skills. Surveys capture concepts, like teamwork, that researchers cannot easily discern from observational data. At the same time, however, they measure *perceptions* of capacity. Political leanings, recent events, and turf wars influence these perceptions. Moreover, survey-based measures are time bound and often suffer from low response rates within particular government agencies. Unless researchers regularly survey the federal government about its capacity, these measures provide limited utility outside of specific contexts.

The need for new measures of capacity is apparent in the literature. In recent years, theorists have formalized predictions about the effects of capacity on a variety of political phenomena (see Huber and Sanford, 2004; Ting, 2011; Turner, 2020). Moreover, the recent experience of the Trump Administration has raised questions about what strategies presidents use to build and deplete capacity within the civil-service workforce (see Benn, 2019; Bednar and Lewis, 2023). The measures developed here allow scholars to probe differences in human capital to better understand how capacity emerges within the federal bureaucracy and affects policy creation and outcomes.

3. Measuring policymaking capacity

The development of new measures is possible given the wealth of data on the US civil service. The scores developed here measure the ability of an agency's workforce to make policy. All agencies engage in some level of policymaking. Formally, agencies make policy by promulgating rules, drafting legislation or executive orders, and issuing enforcement guidance. They also make informal policy decisions by interpreting vague mandates in statutes to discern the extent of their authority and their obligations under the law. Agencies must adhere to similar procedures when engaged in certain forms of policymaking (see Potter, 2019). The high degree of standardization of policymaking activities within the executive branch makes it an ideal task for comparing capacity across agencies.

The Office of Personnel Management (OPM)—the federal government's human-resources department—publishes annually its Enterprise Human Resources Integration (EHRI) database, which includes records of civilian employees working in federal agencies.[2] From 1998 to 2021, EHRI includes over 190 million personnel records.[3] Each entry in EHRI includes information about the employee's agency, salary, education, occupation, and length of service in a given

---

[2]OPM publishes these data annually in September for all years, which syncs the measure with the fiscal year of the federal government.

[3]Each personnel record is an employee-year, which explains why the number of records exceeds the number of unique federal employees over the study period.

year. The regular publication of EHRI permits the development of time-series measures of human capital by aggregating these data into agency–year units.

To allow for meaningful cross-agency comparisons, the measures must adequately control for differences in employees' responsibilities across agencies. Reliance on OPM's occupational classifications addresses this concern. OPM standardizes occupations for all federal agencies based on the knowledge and skills required for the job, its responsibilities, and the level of supervision received by the employee. Federal law requires most agencies to use these classifications when creating new positions and setting salaries. Certain occupations play a particular role in administrative policymaking (Walker, 2013). I use OPM's *Handbook of Occupational Groups and Families* to code occupations that mention responsibilities related to policymaking, such as "developing regulations" or "planning policies" (O.A.2).

Identifying agency policymakers based on standardized job descriptions has three main advantages. First, it ensures that individuals employed in the same occupation have similar responsibilities regardless of which agency employs them. Second, it sets a baseline standard for the qualifications and salary of a given occupation across all agencies. Third and finally, it allows us to identify individuals who perform tasks related to policymaking and remove employees in non-policymaking occupations from the sample. The chosen employees all have responsibilities related to policymaking, and OPM's standardization of their occupations creates a comparable level of necessary qualifications and compensation across agencies.[4]

Drawing on the broader conceptualization of human capital, I create five agency–year indicators correlated with the workforce's ability to make policy. Plausible alternatives exist for each indicator, but the measures are robust when estimated with these alternatives. I discuss other possible specifications in the Appendix (O.A.1).

To make policy, agencies need employees with scientific and technical expertise related to their policy domains. Bureaucrats generally acquire this scientific and technical expertise through post-secondary education (Miller and Whitford, 2016; Fisher and Shapiro, 2020). I measure the scientific and technical expertise within an agency's policymaking workforce as the average proportion of policymaking employees with a college education (mean: 0.74, SD: 0.17).[5] Employees with more experience in the federal government have a greater understanding of how to navigate complex procedures and reconcile competing political and policy considerations (Potter, 2019; Fisher and Shapiro, 2020). I measure the experience within the agency's policymaking workforce as the mean length of service of policymaking employees (mean: 15.56 years, SD: 3.31 years).[6]

Agencies must be able to recruit and retain these expert and experienced civil servants. To do this, they must offer sufficient salaries to prevent civil servants from engaging in rent-seeking behavior or pursuing other career paths (Weber, 1978; Gailmard and Patty, 2007). I measure the ability of the agency to recruit and retain civil servants using the average salary of the agency's policymaking employees (mean: \$115.85 thousand (2020 USD), SD: \$21.56 thousand (2020 USD)).[7] In addition to the employee's base salary, the salary incorporates any locality adjustment. The locality adjustment ensures that individuals in the same occupation receive comparable wages regardless of the employee's geographic location.

An agency must organize its policymaking workforce in ways conducive to team production. Rarely does a single bureaucrat provide all of the expertise and experience the agency needs to

---

[4]We cannot observe whether a given employee engages in policymaking. Nevertheless, their employment in a position that includes these responsibilities means they have the prospective potential to engage in policymaking.

[5]When the scores are estimated with the proportion of policymaking employees with a graduate education, the two measures correlate at $\rho = 0.93$.

[6]When estimated with the cumulative length of service of policymaking employees (logged), the two measures correlate at $\rho > 0.99$.

[7]All salaries are adjusted for inflation prior to estimation. Alternative constructions of this indicator incorporate the availability of outside options to agency employees. These alternatives, however, introduce additional measurement error and, in any event, are highly correlated with the preferred specification (O.A.1).

complete a particular task. Instead, bureaucrats work in teams, bringing their own unique perspectives and knowledge to the task (Williams, 2021). Measuring the management and organization of an agency's workforce is difficult with EHRI. The personnel records do not tell us anything about whether the agency effectively manages or organizes its workforce. We know, however, that agencies organize their workforces around tasks more central to their missions (Wilson, 1989; Barkow, 2013; Hickman, 2016). Consistent with the approach of other scholars (Bersch et al., 2017), I measure the centrality of policymaking to the agency's workforce as the proportion of total agency employees in policymaking occupations (mean: 50.40, SD: 0.24). This indicator assumes that an agency with a higher proportion of policymaking employees is more likely to have organized the workforce in ways that promote effective policymaking. This indicator positively correlates with other plausible measures of agency management and organization.[8] In addition, validation of the policymaking-capacity scores with time-series surveys of civil servants shows a positive correlation with the reported level of cooperation within the agency, suggesting the measures adequately capture these organizational concepts (O.A.3).

Finally, the number of policymakers within an agency informs how many policies the agency may pursue at one time. Empirical research shows that the number of employees within an agency is positively correlated with policymaking productivity (Bolton et al., 2015; Potter and Shipan, 2019). Accordingly, the measure includes the total number of policymaking employees (logged) in the agency (mean: 6.28, SD: 2.31).

A Bayesian model aggregates these five variables to produce 5590 yearly policymaking-capacity scores for 261 agencies from 1998 to 2021 (O.A.1). For the 15 cabinet departments, I create both aggregated and bureau-level measures. The resulting scores range from −2.38 (low capacity) to 3.59 (high capacity) (mean: 0, SD: 0.68). The results are robust to other specifications. Measuring these concepts with different indicators or incorporating other possible concepts into the measure produces highly correlated results (O.A.1).

3.1 Content and discriminant validation

The value of the scores depends on their validity relative to other plausible measures of policymaking capacity. I validate the content of the policymaking-capacity scores against the number of policymaking employees in each agency. I use case-oriented content validation to check whether the measures comport with expectations about which agencies have higher levels of policymaking capacity (Adcock and Collier, 2001). To do so, I draw a set of priors about the levels of policymaking capacity across the administrative state:

1. Agencies whose core missions center around regulation should have higher scores than agencies whose core missions center around other activities, such as law enforcement or benefits distribution.
2. Agencies whose policy domains necessitate high levels of technical expertise should have higher scores.
3. Agencies with known patterns of mismanagement should have lower scores.

These heuristics are not infallible, but they provide a set of criteria by which to compare the two measures.

---

[8]OPM's Federal Employee Viewpoint Survey (FEVS) asks employees whether they agree or disagree with the statement: "Managers support collaboration across work units to accomplish work objectives." There is a positive correlation between the proportion of respondents within an agency agreeing with this question and the proportion of policymakers in the agency ($\rho = 0.33$). Another question asks respondents whether they agree or disagree with the statement: "The people I work with cooperate to get the job done." There is a positive correlation between the proportion of respondents within an agency agreeing with this question and the proportion of policymakers in the agency ($\rho = 0.30$). The FEVS questions themselves are not good candidates for inclusion within the measure because of the lack of available data for all agencies across the study period.
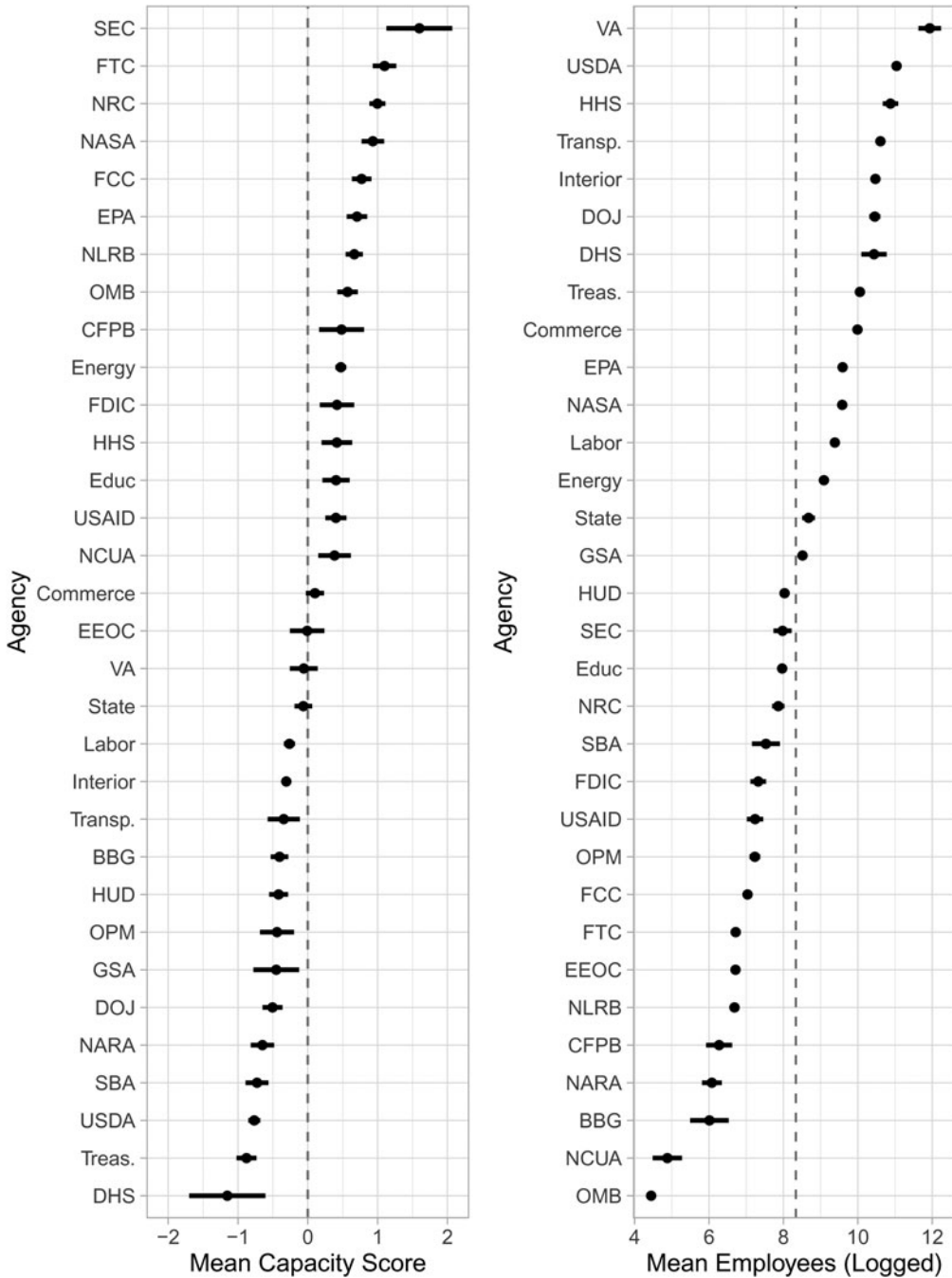
Figure 1. Policymaking capacity estimates and total policymaking employees for major agencies.

Figure 1 compares the average policymaking-capacity score for the 15 executive departments and the agencies classified as large independent agencies by OPM ($N = 34$) against the number of policymaking employees (logged) in the same subset of agencies. The policymaking-capacity scores better comport with our expectations than the measures of workforce size.

I begin by explaining the descriptive findings of the policymaking-capacity scores. First, eight of the top ten agencies have strong regulatory missions. These agencies enact regulations that govern nuclear energy and the environment (the Nuclear Regulatory Commission, the Environmental Protection Agency, the Department of Energy), finance and securities (Securities and Exchange Commission, Federal Trade Commission, Consumer Financial Protection Bureau), telecommunications (Federal Communications Commission), and labor (National Labor Relations Board). Although not a regulator, the Office of Management and Budget (OMB) acts as the nucleus of policymaking in the executive branch by reviewing agencies' proposed policies. Many of these agencies, such as NASA and the Department of Energy, are known for their employment of highly educated experts. For example, the Department of Energy's National Laboratories founded the Human Genome Project and identified the cause of the Cretaceous–Paleogene extinction event.

Many of the bottom ten agencies have missions that revolve around law enforcement, benefits distribution, and other non-policymaking tasks. Although all of these agencies engage in some form of policymaking, policymaking is tertiary to their main missions, especially when compared to the missions of the top ten agencies. This does not mean that these agencies lack the capacity to perform their primary functions. However, it does mean that these agencies are less equipped to make policy relative to their high-capacity peers. The agency with the lowest average score—the Department of Homeland Security (DHS)—polices American borders, waterways, and airports while also adjudicating immigration applications and distributing benefits following disasters. The fact that DHS lacked a policy shop until 2005 highlights the relative inattention it gives to this function (McIntire, 2005). Many of the bureaus within the Department of the Treasury and the Department of Justice (DOJ) also have strong enforcement missions. The majority of Treasury's activities involve auditing tax returns, distributing refunds, and protecting the fisc (Hickman, 2016). DOJ's enforcement bureaus, such as the Bureau of Prisons and the Drug Enforcement Administration, employ over half of its workforce classified as policymaking employees (58.2 percent). DOJ's strong enforcement culture has caused it to neglect many other functions (Barkow, 2013).

Mismanagement has wracked many of these lower-ranked agencies. Both the Federal Emergency Management Agency and Secret Service within DHS have been the subjects of highly publicized scandals and performance failures (Lewis, 2008; Leonnig, 2021). In the last decade, the largest employer within Treasury—the Internal Revenue Service—has suffered from a significant loss of employees brought about by an aging workforce, government shutdowns, and budget cuts (Weiner, 2013; Kiel and Eisinger, 2018). Since 2010, IRS's appropriations have fallen by 20 percent, and the agency has lost 22 percent of its total staff (Congressional Budget Office, 2020). Likewise, commentators have raised concerns about the ability of certain bureaus within DOJ to make policy. Following President Biden's promise to use the Bureau of Alcohol, Tobacco, Firearms, and Explosives (ATF) to create gun-control policies, commentators expressed doubts about the capacity of the agency's workforce to accomplish the administration's objectives. Former ATF officials expressed "that the agency need[ed] to be restructured, modernized, given adequate resources and managed in a more proactive and aggressive way" if the agency hoped to implement the policies desired by the administration (Thrush et al., 2021).

By contrast, the size of the agency's policymaking workforce displays almost completely opposite trends. Enforcement agencies with known problems of mismanagement—DHS, DOJ, and Treasury—appear toward the top of the list. In addition, the agency with the greatest number of policymaking employees, the Department of Veterans Affairs, experienced one of the largest scandals related to workforce attrition and recruitment in the last two decades (Katz, 2022). The regulatory agencies that appear on the top of the left-hand plot appear on the bottom of the right-hand plot. The Office of Management and Budget—the overseer of all policymaking in the executive branch—appears at the very bottom of the list.

Comparisons of the two measures suggest that workforce size does not adequately capture the concept of policymaking capacity as scholars describe it. The agencies with the highest number of

policymakers are not those that scholars describe as more expert or carefully managed. By contrast, the policymaking-capacity scores seem to comport with prior expectations of which agencies have the workforce and management needed to engage in policymaking.

3.2 Construct validation

An alternative means of validating the policymaking-capacity scores is to examine a well-supported hypothesis in the literature (Adcock and Collier, 2001). Scholars of agency rulemaking have identified a positive relationship between workforce size and successful completion of the rulemaking process (see Potter and Shipan, 2019). Specifically, Bolton *et al.* (2015) demonstrate that the Office of Information and Regulatory Affairs (OIRA) takes longer to review rules when it has fewer employees. This result should extend to all rulemaking agencies. Rulemaking provides an appropriate test of the policymaking-capacity scores because rulemaking is a notoriously arduous process that requires significant technical expertise and experience (see Potter, 2019). If the scores properly measure policymaking capacity, then agencies with higher scores should take less time to finalize their rules.

The dataset consists of substantive rulemakings begun during the first terms of the Bush, Obama, and Trump Administrations. I measure *Duration of Rulemaking* as the number of days between the issuance of the agency's Notice of Proposed Rulemaking and its final submission of the final rule to the OIRA (mean: 488.34 days, SD: 662.84 days).[9] I use the date of the agency's final submission to OIRA because, at that point, the agency has completed its work on the rulemaking. Like most scholars studying rulemaking (Potter, 2019; Potter and Shipan, 2019), I examine the subset of rules deemed "significant" by the agency. These rules have greater policy implications compared to other rules and represent the greatest test of the agency's policymaking capabilities. I discuss the research design in greater detail in the Appendix (O.A.4).

The explanatory variable of interest is the agency's *Capacity*. I measure the rulemaking agency's level of *Capacity* in the fall before the start of the new presidential administration. I control for other common predictors of rulemaking success. Agencies may choose to slow down the rulemaking process when they disagree with the proposals of the current president (Potter, 2019). I use measures of agency ideology to construct a binary indicator of whether the agency aligns with the current president (*Aligned President*: 41.6 percent) (Richardson *et al.*, 2017). I also control for *Agency Ideology* separately because left-leaning agencies may have a stronger regulatory bent than right-leaning agencies (Potter and Shipan, 2019). A negative value indicates a more liberal agency and a positive value indicates a more conservative agency. Agencies with greater structural independence often face fewer procedural requirements in the rulemaking process because they are not required to submit their rules to the White House for review. I measure *Independence* using Selin's measure of decision-maker independence (Selin, 2015). A negative value indicates an agency with less independence and a positive value indicates an agency with more independence. Consistent with the need to consider capacity in the context of agency workloads, I control for the agency's workload using the logged number of entries in the *Unified Agenda* in the fall before the inauguration. I stratify the results by presidential administration to control for time-varying policies within administrations that affect rulemaking procedures.[10] To allow for ease of comparison, I normalize all continuous variables prior to estimation.

As *Duration of Rulemaking* exhibits right censoring for uncompleted or withdrawn rules, I estimate the model using a Cox proportional-hazard model. Because presidents care about finalizing rules so that they can claim credit for the resulting policy, I also estimate the model by

---

[9]If the final rule was exempted from OIRA review, then I use the day the final rule appeared in the *Federal Register*.

[10]I do not include any measures of divided government because (1) all three presidents had some level of unified government during the first Congress of their first term and (2) the alignment of the second Congress may be a function of presidential policymaking success and, therefore, post-treatment. Stratifying by presidential administration controls for at least some of the dynamics between Congress and the president.

Table 1. Estimated days to final rule: first term of the Bush, Obama, and Trump administrations

	Hazard rate of rulemaking		
	(1)	(2)	(3)
Capacity	0.09*	0.11*	0.17***
	(0.04)	(0.04)	(0.04)
Ideologically aligned		0.15**	0.21**
		(0.06)	(0.08)
Agency ideology		−0.09*	−0.07
		(0.04)	(0.05)
Independence		0.08	0.14*
		(0.06)	(0.06)
Workload (logged)		0.07	0.03
		(0.07)	(0.08)
Presidency strata	No	Yes	Yes
First-term censor	No	No	Yes
Estimator	Cox	Cox	Cox
Observations	1705	1705	1705

Note: Standard errors clustered at the agency level. All continuous variables normalized prior to estimation.
* $p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

censoring all rules not completed by the end of the president's first term in office. A positive coefficient indicates that the agency takes less time to finalize its rules.

Table 1 reports the results. In all three models, *Capacity* is a positive and significant predictor of rulemaking finalization. For model (2), an agency with *Capacity* one standard deviation above the mean is 12 percent (hazard ratio: 1.12) more likely to complete an ongoing rulemaking at a fixed point in time. By contrast, an agency with *Capacity* one standard deviation below the mean is 10 percent (hazard ratio: 0.90) less likely to complete a rulemaking at a fixed point in time. The effect size is larger when examining the agency's ability to complete the rule within a president's first term. The results mirror the expectations set by Bolton *et al.* (2015), extending their findings to show that capacity has a broad effect on rulemaking duration in all agencies—not just OIRA. Overall, these results suggest that the policymaking-capacity scores measure the concept of interest: an agency's ability to produce policy.

4. Bureaucratic autonomy and policymaking capacity

How do bureaucratic institutions acquire and retain this policymaking capacity? A common theory attributes the acquisition and erosion of human capital within an agency to bureaucratic autonomy (see Gailmard and Patty, 2007; Richardson, 2019). Autonomy describes the degree of freedom that agencies enjoy from Congress and the president. Civil servants in policymaking positions select into government service to influence policy in the direction of their sincerely held preferences. Scholars identify two possible mechanisms by which autonomy increases bureaucratic capacity. First, autonomy encourages civil servants to invest in their own expertise to attain their preferred policy outcomes while working for government (Gailmard and Patty, 2007). Second, autonomy induces civil servants to remain in public service, foregoing more lucrative private-sector jobs in favor of the opportunity to influence policy without political interference (Gailmard and Patty, 2007; Richardson, 2019). The policymaking-capacity scores afford scholars the opportunity to test the relationship between bureaucratic autonomy and capacity.

Agencies derive their autonomy, in part, from the structures that govern their interactions with Congress and the president. Structural independence insulates agencies from congressional and presidential control. Most independent agencies exist outside of the executive departments, removed from the president's direct control (Lewis, 2003). The most insulated agencies are independent regulatory commissions. The commissioners for these agencies serve for fixed terms and

have for-cause removal protections (Selin and Lewis, 2018). In designing these agencies, Congress expressed a belief that independence would give commissioners "an opportunity to acquire the expertness in dealing with these special questions concerning industry that comes from experience" (S. Rep. No. 63–597, 11 (1914)).

Recent work by Devins and Lewis (2023) calls into question whether these insulating structures promote capacity building. Devins and Lewis (2023) use the Richardson *et al.* (2017) survey measures to compare the levels of workforce skills between independent commissions and other agencies. Devins and Lewis find that there is no statistically distinguishable difference in workforce skills between these agencies. The authors conclude that "independent agencies are not particularly expert, influential or independent" (p. 4). The scores allow us to test the conventional wisdom against the claims of Devins and Lewis (2023) using measures derived from observational sources of data.

The dependent variable is an agency's *Capacity* in a given year. Using measures developed by Selin (2015), I estimate the effect of two types of structural independence: decision-maker independence and political-review independence. *Decision-Maker Independence* measures how much influence the president has over the selection and removal of key agency officials (mean: -0.12, SD: 0.87). *Political-Review Independence* measures the ability of the White House to review the agency's policy proposals (mean: 0.10, SD: 0.95).

I control for several other structural variables to reduce the possibility that agency independence correlates with some other cause of capacity building. I control for *Agency Ideology* on the grounds that liberal politicians may be more likely to invest in bureaucratic capacity and have stronger incentives to do so in agencies that align with their policy preferences (Benn, 2019; Bednar and Lewis, 2023) (mean: -0.07, SD: 1.04). I use Richardson *et al.* (2017) measure of agency ideology, which measures the agency's stable ideological leaning in both Democratic and Republican administrations. I also control for the agency's policy domain. I code each agency's main policy area by reading its mission statement and assigning it a topic area according to the Comparative Agendas Project's major-topic codes. The inclusion of topic fixed effects controls for the possibility that Congress and the president invest heavily in agencies concerned with certain issues and simultaneously give these same agencies greater autonomy (Binder and Spindel, 2017; Miller and Whitford, 2016). I control for the *Age* of the agency (in hundreds of years) because newer agencies have had less time to build capacity than older agencies (mean: 0.57, SD: 0.50). Finally, I include year fixed effects to account for time trends that influence policymaking capacity in all agencies. I estimate the model using an OLS regression with standard errors clustered at the agency level.

Table 2 reports the results. Figure 2 plots the predicted effect, highlighting the 2021 scores of notable agencies. A one standard deviation increase in either decision-maker or political-review independence increases bureaucratic capacity by more than a quarter of a standard deviation. The effect is even stronger when Selin's measures are substituted for an indicator of whether the agency is an independent commission (O.A.5). Consistent with conventional wisdom, agencies with higher levels of structural independence exhibit higher levels of policymaking capacity.

The results suggest that insulating structures intended to protect bureaucratic autonomy promote capacity building. These results stand in stark contrast to those found by Devins and Lewis (2023). While agencies like the Securities and Exchange Commission (SEC) and the Federal Energy Regulatory Commission (FERC) have moderate skills according to survey-based measures, these agencies rise above executive agencies in the policymaking-capacity scores. Reassuringly, there is some overlap between the two measures. For example, both the scores and the survey measures show that the Federal Trade Commission has comparatively high levels of capacity and the Transportation Security Administration has comparatively low levels of capacity.

The difference in the results of this study and Devins and Lewis (2023) underscores the need to better understand the concept of bureaucratic capacity and how to measure it. The results of Devins and Lewis (2023) suggest that federal executives perceive no benefit to these insulating
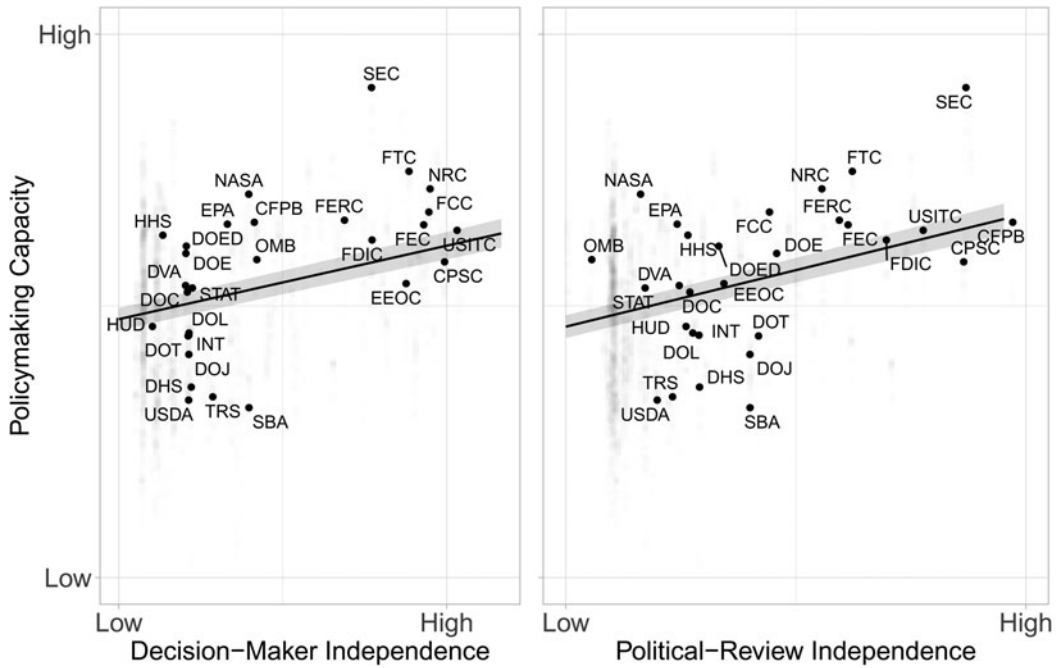
Figure 2. Predicted effects of structural independence on policymaking capacity.

Table 2. Model estimates of effect of independence on capacity

	Policymaking capacity			
	(1)	(2)	(3)	(4)
Decision-maker independence	0.24***	0.23***		
	(0.04)	(0.05)		
Political-review independence			0.19***	0.20***
			(0.04)	(0.05)
Agency ideology		−0.14*		−0.09
		(0.06)		(0.07)
Age (hundreds of years)		−0.19**		−0.22***
		(0.07)		(0.07)
Constant	0.03	−0.19	−0.02	−0.27
	(0.04)	(0.18)	(0.04)	(0.18)
Agency fixed effects	No	Yes	No	Yes
Topic fixed effects	No	Yes	No	Yes
Estimator	OLS	OLS	OLS	OLS
Observations	4713	4322	4713	4322
R^2	0.09	0.30	0.07	0.30

Note: Standard errors clustered at agency level.
*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

structures. Yet the policymaking-capacity scores suggest that these agencies still do a better job recruiting and retaining higher educated and more experienced civil servants—even when controlling for the agency's policy domain. In addition, the policymaking-capacity scores tell us specifically about the agency's ability to engage in *policymaking* whereas the survey measures do not ask respondents to consider a particular task performed by these agencies. However, the survey measures may capture hard-to-measure components of human capital, such as teamwork

and management, better than the policymaking-capacity scores. In the end, deciding whether these structures adequately promote capacity building requires us to further engage with what we mean by capacity and by what standard we evaluate an agency's capacity.

5. Conclusion

Bureaucratic capacity describes more than the size of an agency's workforce. When scholars theorize about delegation or administrative policymaking, they rely on a concept that incorporates expertise, experience, and teamwork. Current measures fail to capture these essential elements of bureaucratic capacity and human capital. Progress is possible, however, thanks to the availability of rich data from the federal government. This study presents a new time-series measure of policymaking capacity. Scholars should be careful, however, to use this measure appropriately. Measuring and testing capacity requires scholars to think through the task of interest and what skills the agency's workforce needs to complete the task. The measures of policymaking capacity will not measure, for example, an agency's ability to engage in law enforcement. Thankfully, this study provides a helpful blueprint for scholars to easily extend these measures to any other task such as agency enforcement, adjudication, or benefits distribution.

   The measures developed here provide a way to test a wealth of existing theories in the bureaucratic-politics canon. For studies of policymaking, the measures are a suitable independent variable. For example, Turner (2020) predicts that presidents have greater incentives to pursue legislation over unilateral action as capacity within the implementing agency increases. For studies of capacity building and investment, the measures are a suitable dependent variable. For example, Benn (2019) theorizes that conservative presidents intentionally deplete policymaking capacity in regulatory agencies to undermine the policy initiatives of their predecessors. These measures provide a simple but essential contribution to advancing the empirical study of bureaucratic politics.

References

Adcock R and Collier D (2001) Measurement validity: a shared standard for qualitative and quantitative research. *American Political Science Review* **95**, 529–546.

Barkow RE (2013) Prosecutorial administration: prosecutor bias and the department of justice. *Virginia Law Review* **99**, 271–342.

Bednar NR and Lewis DE (2023) Presidential investments in administrative capacity. *American Political Science Review*, First View, 1–16.

Benn A (2019) Tying Hands: Strategic Incentives to Undermine Bureaucratic.

Bersch K, Praça S and Taylor MM (2017) State capacity, bureaucratic politicization, and corruption in the Brazilian State. *Governance* **30**, 105–124.

Binder S and Spindel M (2017) *The Myth of Independence: How Congress Governs the Federal Reserve*. Princeton, NJ: Princeton University Press.

Bolton A, Potter RA and Thrower S (2015) Organizational capacity, regulatory review, and the limits of political control. *Journal of Law, Economics, and Organization* **32**, 242–271.

Congressional Budget Office (2020). Trends in Internal Revenue Service's Funding and Enforcement. Technical report.

Devins N and Lewis DE (2023) The Independent Agency Myth. *Cornell Law Review*, 108.

Drolc CA and Keiser LR (2020) The importance of oversight and agency capacity in enhancing performance in public service delivery. *Journal of Public Administration Research and Theory* **31**, 1–17.

Fisher E and Shapiro SA (2020) *Administrative Competence: Reimaging Administrative Law*. Cambridge, UK: Cambridge University Press.

Fukuyama F (2013) What is governance?. *Governance* **26**, 347–368.

Gailmard S (2002) Expertise, subversion, and bureaucratic discretion. *Journal of Law, Economics, and Organization* **18**, 536–555.

Gailmard S and Patty JW (2007) Slackers and zealots: civil policy and bureaucratic expertise. *American Journal of Political Science* **51**, 873–889.

Geddes B (1994) *Politician's Dilemma: Building State Capacity in Latin America*. Berkeley, CA: University of California Press.

Hickman KE (2016) Pursuing a single mission (or something closer to it) for the IRS. *Columbia Journal of Tax Law* **17**, 169–193.

Huber GA and Sanford CG (2004) Accountability and coercion: is justice blind. *American Journal of Political Science* **48**, 247–263.

Katz E (2021) Land management agency still facing vacancies from headquarters move after workers fled. *Government Executive*, November 18, https://www.govexec.com/workforce/2021/11/land-management-agency-still-facing-vacancies-headquarters-move-after-workers-fled/186945.

Katz E (2022) Staffing shortages are growing worse at VA after years of improvement. *Government Executive*, July 7, https://www.govexec.com/workforce/2022/07/staffing-shortages-are-growing-worse-va-after-years-improvement/373951.

Katz E (2023) It's glacial: one agency is still struggling to overcome the assault on its workforce. *Government Executive*, March 7, https://www.govexec.com/workforce/2023/03/one-agency-struggling-overcome-assault-workforce/38350.

Kiel P and Eisinger J (2018) How the IRS was gutted. *ProPublica*, December 11, https://www.propublica.org/article/how-the-irs-was-gutted.

Leonnig C (2021) *Zero Fail: The Rise and Fall of the Secret Service*. New York City, NY: Random House.

Lewis DE (2003) *Presidents and the Politics of Agency Design: Political Insulation in the United States Government Bureaucracy, 1946-1997*. Palo Alto, CA: Stanford University Press.

Lewis DE (2008) *The Politics of Presidential Appointments*. Princeton, NJ: Princeton University Press.

Lewis DE (2019) Deconstructing the administrative state. *Journal of Politics* **81**, 767–789.

Mashaw JL (1983) *Bureaucratic Justice: Managing Social Security Disability Claims*. New Haven, CT: Yale University Press.

McIntire K (2005) DHS policy shop may reshape agency priorities. *Government Executive*.

Miller GJ and Whitford AB (2016) *Above Politics: Bureaucratic Discretion and Credible Commitment*. New York City, NY: Cambridge University Press.

Potter RA (2019) *Bending the Rules: Procedural Politicking in the Bureaucracy*. Chicago, IL: University of Chicago Press.

Potter RA and Shipan CR (2019) Agency rulemaking in a separation of powers system. *Journal of Public Policy* **39**, 89–113.

Richardson MD (2019) Politicization and expertise: exit, effort, and investment. *Journal of Politics* **81**, 878–891.

Richardson MD, Clinton JD and Lewis DE (2017) Elite perceptions of agency ideology and workforce skill. *Journal of Politics* **80**, 303–308.

Selin JL (2015) What makes an agency independent?. *American Journal of Political Science* **59**, 971–987.

Selin JL and Lewis DE (2018) *Sourcebook of united states executive agencies*. Administrative Conference of the United States, Washington, DC.

Skowronek S (1982) *Building a New American State: The Expansion of National Administrative Capacities, 1877–1820*. Cambridge, UK: Cambridge University Press.

Thrush G, Hakim D and McIntire M (2021) How the A.T.F., Key to Biden's Gun Plan, Became an N.R.A. Whipping Boy. *New York Times*, May 2, https://www.nytimes.com/2021/05/02/us/politics/atf-nra-guns.html.

Ting MM (2011) Organizational capacity. *Journal of Law, Economics, and Organization* **27**, 245–271.

Turner IR (2020) Policy durability, agency capacity, and executive unilateralism. *Presidential Studies Quarterly* **50**, 40–62.

Walker CJ (2013) Inside agency statutory interpretation. *Stanford Law Review* **67**, 999–1079.

Weber M (1978) *Economy and Society: An Outline of Interpretive Sociology*. Berkeley, CA: University of California Press.

Weiner R (2013) What the IRS scandal means for health-care reform. *The Washington Post*, May 15, https://www.washingtonpost.com/news/the-fix/wp/2013/05/15/what-the-irs-scandal-means-for-health-care-reform.

Williams MJ (2021) Beyond state capacity: bureaucratic performance, policy implementation and reform. *Journal of Institutional Economics* **17**, 339–357.

Wilson JQ (1989) *Bureaucracy: What Government Agencies Do and Why They Do It*. New York, NY: Basic Books.