Academic Utilization and Issue Novelty *

Connor Dye

Abstract

Why do policymakers utilize academic expertise when alternative sources provide more accessible information? I argue that academia's evaluation system decouples research priorities from policymaker demand, enabling academics to accumulate expertise on emerging problems that utilization-oriented organizations systematically avoid until issues become politically salient. I test this theory examining expert witness selection at U.S. congressional hearings from 1969 to 2021. Using natural language processing methods to measure issue novelty through hearing descriptions, I find that academics constitute a significantly larger share of witnesses when issues are novel and a smaller share as issues mature. Think tanks, federal agencies, and analytical agencies show the opposite pattern, providing evidence that increased academic testimony on novel issues reflects compositional shifts rather than broader demand for expertise. These findings suggest that academic utilization stems from its unique evaluation system that creates incentives for academics to invest in problems that other experts avoid, and that efforts to make academic research more policy-relevant may undermine the distinctive value academics provide when policymakers confront unfamiliar issues.

Keywords: Academic Utilization, Policy Expertise, Information Processing, Organizational Design, Agenda Setting

*An appendix with supplementary material will be made available online.

Policymakers operate within an information environment made up of a diverse range of expert organizations that compete to supply policy-relevant expertise (Downs 1972; Kingdon 1984; Jones and Baumgartner 2005). Existing scholarship suggests that academics should rarely succeed in this competition. Academic information is communicated through abstract language and technical details, which create higher processing costs for policymakers relative to alternative sources of expertise (Caplan 1979; Dunn 1980; Ward, House, and Hamer 2009; Mead 2015). Moreover, the longer production timelines make academics less responsive to politically salient issues than alternative sources of expertise (Caplan 1979). Despite these disadvantages, policymakers consistently report utilizing information produced by academics (Landry, Lamari, and Amara 2003; Belkhodja et al. 2007; Jennings and Hall 2012; Cherney et al. 2015; Newman, Cherney, and Head 2016; Raymaekers et al. 2025; Xu et al. 2025). In a crowded information environment, why do policymakers rely on expertise produced by academia?

To this question we have no general answers. Existing research on academic utilization examines academia in isolation from the broader set of expert organizations competing for policymaker attention. Studies of research utilization describe how academic research enters the policy process (Weiss 1977; Weiss 1979; Head 2013), but do not explain why policymakers would choose academics over other expert organizations capable of supplying similar information. A separate line of work argues that academics are utilized because their research is credible and relatively insulated from political pressures (Boswell 2008; Montpetit 2011). Yet, credibility alone does not explain when academics will be favored in a crowded information environment where many expert organizations also provide credible analysis, often in formats more accessible and responsive to policymaker needs.

I argue that academic utilization is shaped by the independence of academia's evaluation system from policymaker demand for information. Policy-oriented expert organizations succeed by responding to political principals, clients, or donors, creating incentives to allocate analytical resources toward problems already attracting political attention where their work can yield immediate returns. Academia's peer review system decouples

academic research priorities from the political agenda by rewarding originality, methodological contribution, and advancing disciplinary knowledge rather than responsiveness to current policy debates. As a result, academia is more likely to accumulate expertise on problems that have yet to attract sustained political attention than policy oriented expert organizations.

The relative independence to policymaker demand created by academia's evaluation system generates predictable patterns of academic utilization across an issue's lifecycle. When problems first emerge on the political agenda, few expert organizations possess relevant information because utilization-oriented experts have had little reason to invest analytical effort in issues whose political salience was uncertain. Academics, whose priorities are less tethered to political demand, are more likely to have developed relevant expertise. Over time, as issues recur and become established on the political agenda, utilization-oriented organizations respond by producing their own analyses in formats more accessible to policymakers. As the information environment becomes denser with policy-ready expertise, academic utilization declines. Thus, the theory predicts that academics should constitute a larger share of expert participation when issues are novel to the political agenda and a smaller share as issues mature.

I focus in particular on explaining variation in academic utilization in the context of committee hearings in the U.S. Congress. Congressional hearings are an important case for three distinct reasons. First, hearings constitute the primary venue through which Congress, the principal policy-making body in the United States, gathers expert information to develop policy across the full range of national problems (Baumgartner and Jones 1993; Talbert, Jones, and Baumgartner 1995). Second, because experts view invitations to testify as consequential opportunities for influence (Leyden 1995; Rich 2004), hearings represent a setting for studying how policymakers select among academics, think tanks, agencies, and advocacy organizations. Finally, while several studies demonstrate that Congress utilizes a diversity of experts in the policy-making process (Esterling 2004; Workman, Shafran, and Bark 2017; Ban, Park, and You 2023; Ban, Park, and You 2025), the question of why academics are selected over alternative expert organizations remains

largely unexplored.

To measure issue novelty on the congressional agenda, I draw on an original linguistic approach that compares the language used in congressional hearing summaries to all prior hearings on the same policy topic. The Policy Agendas Project provides human-generated summaries that distill the main substantive content of all congressional hearings from the 80th to the 117th Congress. I group hearings by substantive topic and employ a retroactive comparison method in which each hearing summary is compared to all prior hearing summaries on the same topic. Next, I employ natural language processing techniques to generate three complementary measures of linguistic novelty at the hearing level. The first two measures use a bag-of-words approach to identify when hearings introduce entirely new vocabulary versus when they draw on established terminology. The third measure distinguishes substantive shifts in problem framing from lexical variation using semantic embeddings to capture conceptual novelty.

The results of fixed effects regressions broadly support my theory. When an issue first breaks out on the congressional agenda, academics make up a larger share of the witnesses that testify on the problem. Hearings that introduce new vocabulary, exhibit greater semantic distance from prior hearings on the same topic, and draw less on established terminology all feature significantly higher proportions of academic witnesses. However, as issues become established the proportion of academics testifying declines. Further, I leverage the fact that other expert organizations testify at congressional hearings to examine whether the patterns observed for academics reflect a general increase in expert utilization or a compositional shift in the types of experts consulted. I find that as an issue becomes more established, the proportion of think tanks, executive agencies, and legislative agencies increases. Thus, increased utilization on novel issues are unique to academia.

These results provide novel evidence on the role of academics as contributors to policy development, informing longstanding debates on the role of academics in the policy process (Mead 2015; Newman and Head 2015). Importantly, the evidence supports a general theory of academic utilization that is applicable across both substantive domains

and institutional contexts. I show that academic utilization is greater when Congress is addressing problems across the full range of issues from macroeconomics to environmental issues. Moreover, the results emerge in the highly competitive U.S. federal policy environment where well-funded think tanks, bureaucracies, and interest groups provide extensive alternative expertise. In contexts where expert infrastructure is less developed, academics should maintain influence across a broader portion of an issue's lifecycle.

While my focus is on academic utilization, the retroactive comparative approach to measuring issue novelty could be used by congressional scholars to study a wide range of policy dynamics, such as how legislative effectiveness, coalition structures, or committee jurisdictional conflict change as issues mature on the congressional agenda. More generally, my approach to creating longitudinal measures using summaries of government activity could be expanded to a wide range of other institutional contexts and forms of government activity. In particular, the pre-categorized longitudinal text data available through projects like the Comparative Agendas Project and Comparative Manifestos Project creates numerous opportunities to apply this method to understand how government processes around the world evolve across an issue's lifecycle.

Organizational Foundations of Information Availability

Existing scholarship emphasizes how academics are unique to other experts in how it communicates its information to policymakers. Academics embed their analysis in abstract language and technical details designed for peer review rather than policy consumption (Caplan 1979; Mead 2015). On the other hand, all other policy experts produce information that is packaged for policymaker consumption. Think tanks present information in accessible formats tailored to policymaker needs (Stone 2007; Craft and Howlett 2012; Fraussen and Halpin 2017). Interest groups frame their expertise to align with political opportunities (Austen-Smith 1993; Beyers 2004; Hall and Deardorff 2006). Bureaucrats translate technical knowledge into actionable recommendations (Workman 2015; Palus and Yackee 2016; Cross and Gluck 2020; Flatt and Do 2023).

The greater processing costs characterizing academic information reduces policymakers' utilization of academia. Policymakers face limits on the attention they can devote to gathering information and must allocate scarce analytical resources across competing demands (Simon 1955; Downs 1972; Jones and Baumgartner 2005). Under these constraints, policymakers prioritize nonacademic sources of expertise. The low cost expertise from nonacademics both enables policymakers to absorb a broader set of expertise with the same analytical resources, and frees attention that can be redirected toward tasks such as communicating with constituents, fundraising, or managing political principals. Thus, the higher processing burdens associated with academic information create a systematic disadvantage for academics relative to alternative expert organizations that convey substantively similar information in more accessible formats.

The effect of communication costs on utilization depends on the density of the information environment. Information density refers to the number of expert organizations that possess policy-relevant expertise on a particular problem at a given time. In dense information environments, many experts possess information on an issue (Jones and Baumgartner 2005). When confronted with several well-informed sources, policymakers prioritize expertise from organizations whose information costs fewer resources to process. By contrast, sparse information environments are characterized by few expert organizations possessing relevant expertise. Sparse environments arise when problems are difficult or costly to study, when administrative capacity around an issue is thinly developed, or when parties and interest groups have not mobilized around a problem (May 1991; Jones and Baumgartner 2005; Boswell 2012; May, Koski, and Stramp 2016). Because policymakers lack competing sources of expertise, the smaller set of available information reduces the role of communication costs in shaping utilization. Thus, the communication disadvantages of academia are less likely to affect utilization on issues where few alternative expert organizations possess relevant information.

Why would academics possess information on a problem when utilization-oriented organizations do not? Existing scholarship offers no answer to this question. Research examining how academics, think tanks, agencies, advocacy groups, and bureaucratic

offices selectively attend to issues in the political system typically analyzes each expert community in isolation (Halpin and Binderkrantz 2011; Workman 2015; Fraussen and Halpin 2017; Fagan, McGee, and Thomas 2021; Fagan, McGee, and Thomas 2021; Fagan and McGee 2022; Fagan 2024; Fagan, Furnas, et al. 2024). While scholars are beginning to document differences in attention across organizational types (Cantarelli, Belle, and Hall 2023; Walker et al. 2023), it has yet to explain why those differences arise. A separate body of research examines how institutional incentives shape expert agendas (Gray et al. 2005; Leech et al. 2005; Bertelli and Wenger 2009). Yet, this literature has not developed a unified theoretical framework that explains variation in expert agendas across different organizational structures.

Organizational Design and Expert Information Supply

I theorize how evaluation systems within expert organizations shape which problems receive analytical attention. The theory posits that the independence of academia's peer review evaluation system from the political agenda makes academics systematically more likely to possess information on problems with high utilization uncertainty. Evaluation systems affect the type of information supplied by different expert organizations by determining which problems they prioritize. Academic evaluation systems create incentives for academics to prioritize communicating their information to peer experts who systematically reward methodological innovation, identification of unexplored research questions, and theoretical contributions that advance disciplinary knowledge rather than immediate policy relevance. In contrast, alternative expert organizations such as think tanks, advocacy groups, and bureaucratic agencies are evaluated in part on whether their information is utilized by political principals. The comparative independence of academia from policymaker demand implies that academics are most likely to possess information on problems that alternative expert organizations avoid due to high uncertainty over future utilization.

This argument rests on four assumptions about how expert organizations allocate an-

alytical resources under uncertainty. First, developing expertise over a problem requires investing analytical resources. Understanding the causes, consequences, and potential solutions to policy problems requires expert organizations to invest time, staff, and analytical capacity in building specialized knowledge (Carpenter 2001; Gailmard and Patty 2007; Medvetz 2012). Second, expert organizations possess scarce analytical resources. Organizations are limited in time, information, and labor, which restricts how many problems they can address (Jones 2001; Simon 1997). While organizations vary in their total capacity, none can invest deeply in all problems within the policy environment. As with households and firms expert organizations must make trade-offs in the face of scarcity.

Third, the development of expertise precedes policymaker demand. Developing expertise over the causes and consequences of a problem as well as potential solutions takes time (Sabatier 1988). Policymakers, however, require information immediately when issues reach the agenda (Kingdon 1984). Expert organizations must therefore make investment decisions about developing expertise before policymakers express demand for that information. Finally, expert organizations face uncertainty over whether an issue will reach the political agenda. The policy environment contains far more potential problems than the limited agenda space political institutions can accommodate (Baumgartner and Jones 1993; Jones and Baumgartner 2005). The issues that attract sustained political attention depend on unpredictable factors such as focusing events, shifts in political coalitions, or changes in the broader policy environment (Birkland 1998; Kingdon 1984). Thus, organizations make decisions over how to invest their limited analytical resources in the face of uncertainty over whether the information will be demanded by policymakers.

Organizational evaluation systems shape which problems receive sustained organizational attention. In the context of expert organizations, evaluation systems dictate how experts respond to the uncertainty of future policymaker demand. Evaluation systems that reward members based on their capacity to achieve utilization create incentives to account for utilization probability when allocating analytical effort. Members in utilization oriented organizations benefit from prioritizing issues likely to generate policymaker demand and face costs for investing in issues unlikely to attract political attention. By

contrast, evaluation systems that reward members based on criteria independent of policymaker utilization create weaker incentives for their members to account for utilization probability.

Policy-oriented expert organizations operate under evaluation systems that reward aligning analytical resources with predictable political demand. The survival of policy oriented expert organizations is dependent on demonstrating immediate utility to external stakeholders. Think tanks rely on donors who value visible policy engagement as a return on investment (McGann 2016; Fraussen and Halpin 2017; Fagan 2021; Fagan and McGee 2022); federal agencies rely on political principals to protect their budgets and bureaucratic autonomy (Carpenter 2001); and advocacy groups rely on members and donors who expect evidence of policy impact (Walker 1983; Hojnacki 1997). Dependence on demonstrable policy relevance for organizational survival creates incentives to minimize uncertainty over whether investments in expertise will generate downstream utilization.

Academia's peer review system, by contrast, decouples research priorities from policymaker demand for information. Academic evaluation rewards novelty, methodological contribution, and disciplinary advancement rather than whether their information is utilized by decision-makers. Tenure decisions rest on publication in peer-reviewed journals that evaluate theoretical contribution and methodological rigor rather than policy impact (Merton 1957; Whitley 1987; Stephan 1996). Moreover, research funding agencies reward innovation and scientific merit rather than political responsiveness (Azoulay, Graff Zivin, and Manso 2011). The structural independence of academic advancement from their capacity to achieve policy impact removes the professional penalty scholars face for investing in emerging problems that alternative expert organizations avoid. As a result, academia maintains a broad distribution of expertise across the issue space, allocating attention to problems regardless of whether they present high or low uncertainty of future policy demand.

The structural differences between academic and policy-oriented evaluation systems create divergent incentives for expertise allocation under uncertainty. When organizations decide whether to invest in developing knowledge about a policy problem, policy-oriented

experts face professional pressure to account for the probability of future utilization, while academics face no such constraint. Thus, policy-oriented organizations will prioritize issues with low utilization uncertainty while academia will distribute analytical effort more evenly across the policy environment regardless of uncertainty levels.

Expert Information Supply and Issue Dynamics

Whether the structural differences between academic and policy-oriented evaluation systems produce systematically different research agendas depends on the extent to which expert organizations face varying levels of utilization uncertainty across policy problems. An issue's history on the political agenda shapes how much utilization uncertainty expert organizations face when deciding whether to invest analytical resources.

The policy environment contains a large number of potential problems, but the limited agenda space of political institutions means most issues are not considered on the political agenda (Baumgartner and Jones 1993). For novel issues to break through requires exogenous shocks, focusing events, or changing political alignments (Birkland 1998; Kingdon 1984), which are difficult to predict. By contrast, issues that have already achieved sustained political attention are more likely to receive future attention due to reduced agenda friction (Baumgartner and Jones 1993). After an issue breaks out onto the political agenda, stable policy subsystems form as legislative committees claim jurisdictional authority, executive agencies develop specialized capacity, and organized interests mobilize around shared policy images (Redford 1960; Baumgartner and Jones 1993; Jones and Baumgartner 2005). Subsystem actors depend on continued issue salience for organizational survival, which creates an ongoing demand for expertise to assist in crafting legislation, designing programs, and advocating for policy change. Thus, issues that have established themselves on the political agenda present low utilization uncertainty, while novel issues that have never attracted sustained attention present high uncertainty.

The interaction between evaluation systems and agenda history creates predictable patterns in the distribution of expertise across an issue's lifecycle. For problems that

have never attracted sustained political attention, utilization uncertainty is at its highest. Policy-oriented organizations operating under utilization-dependent evaluation systems avoid investing in developing expertise on problems with high uncertainty about whether investment will generate utilization. Academics operating under evaluation systems decoupled from the political agenda are more likely to develop relevant expertise despite high uncertainty because they do not face the same penalty for investing in problems that fail to generate utilization. The reduced penalties that academics face for investing in issues with uncertain political futures produce an information environment where expertise on problems that have yet to reach the political agenda is sparse and dominated by academics.

As issues emerge onto and recur on the political agenda, uncertainty about future utilization diminishes. When novel issues first break out, uncertainty about future recurrence remains high and policy-oriented organizations face risks in committing resources to problems that lack established patterns of political attention. Academic informational advantages persist at the point of initial emergence as policy-oriented organizations await evidence that political attention will be sustained. As issues establish track records of sustained political attention, reduced uncertainty changes the investment calculus for policy-oriented organizations. Issues that once presented high risk now offer predictable returns to analytical effort. Policy-oriented organizations respond by committing analytical resources, and the information environment shifts from sparse to dense as multiple expert sources accumulate expertise. Early informational advantages held by academia erode as policy-oriented organizations commit resources to problems previously avoided. Established issues become informationally dense environments where academics constitute one source among many expert organizations.

Thus, the distribution of available expertise systematically shifts as issues mature on the political agenda. When issues first break out, academic investment in problems that policy-oriented expert organizations avoided creates a comparative informational advantage. As issues recur and become established, the academic monopoly on relevant information erodes as policy-oriented organizations accumulate expertise and the

informational environment becomes more dense.

Issue Dynamics and Academic Utilization

The changing structure of the supply of expert information interacts with how policymakers select information. Policymakers select experts through a two-stage process. In the screening stage, decision-makers rapidly filter potential sources based on topical relevance (Tversky 1972; Payne 1976; Pirolli and Card 1999). In the selection stage, policymakers choose among screened experts by evaluating accessibility, political alignment, and ease of processing (Simon 1955; Simon 1956; Tversky and Kahneman 1973; Tversky and Kahneman 1974). The screening stage is less resource-intensive because it involves simpler assessments that allow decision-makers to quickly eliminate irrelevant options, whereas selection requires more costly evaluation of attributes among viable alternatives (Pirolli and Card 1999). From the perspective of policymakers, assessing topical fit from an expert's research portfolio or publication titles is more straightforward than evaluating the political implications of their findings or determining how accessible their expertise will be for consumption.

The density of the information environment determines which stage is selective. In sparse environments where few organizations possess relevant expertise, screening is highly selective. Policymakers must conduct comprehensive searches to identify the limited set of qualified experts, and most potential sources are filtered out for lacking topical relevance. In dense environments, screening becomes non-selective as many experts possess relevant expertise. Policymakers halt screening early and shift resources to the selection stage, where they evaluate accessibility, political alignment, and processing costs among the pool of qualified experts.

The interaction between expert supply and policymaker information processing creates predictable comparative advantages for different expert communities across an issue's lifecycle. When issues first emerge and information environments are sparse, screening limits the pool of viable experts, elevating the small set of organizations that possess

relevant knowledge. Academics are more likely to have invested in issues that have not previously occupied space on the political agenda, which allows them to pass the screening stage at higher rates and be more frequently selected despite the higher processing costs associated with their work. By contrast, utilization-oriented organizations struggle to survive the initial screening stage due to a lack of expertise on the problem. As issues recur and information environments become denser, screening becomes less restrictive as many organizations possess relevant expertise on the problem. When faced with many expert organizations that possess information on a problem, policymakers prioritize experts that package information in more accessible formats. Thus, as other expert communities respond to the political agenda by producing issue-specific analyses, the comparative informational advantage of academics diminishes, and policymakers increasingly rely on experts whose information is easier to process.

These dynamics imply that academic utilization depends on the agenda history of an issue. Academics are most involved early in an issue's agenda history and participate less as the issue becomes established and recurs more often.

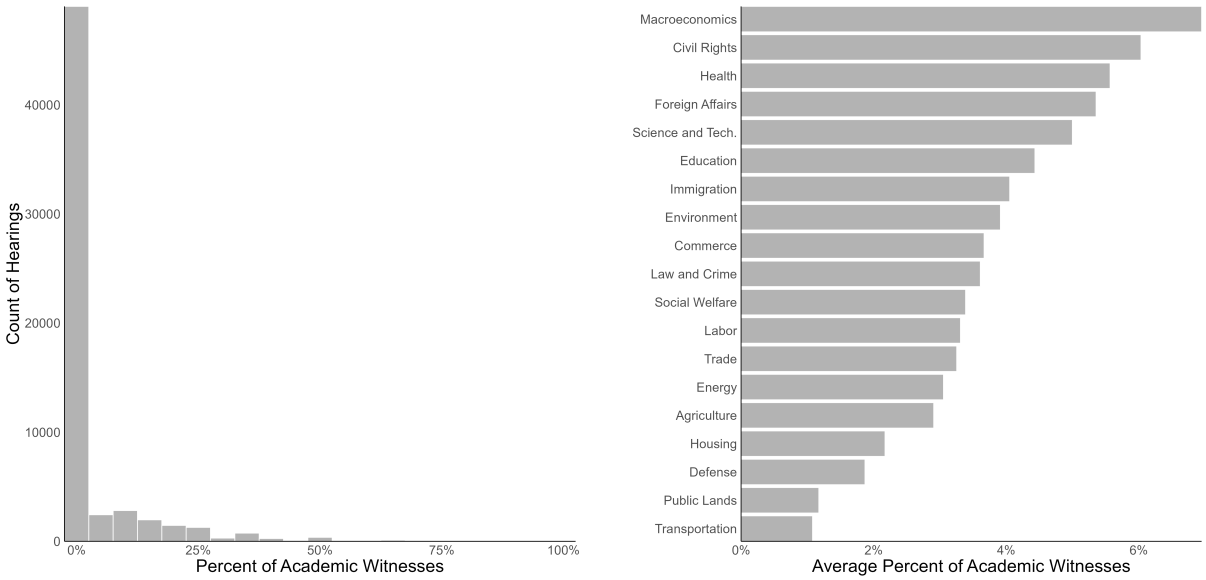
Setting: Congressional Committee Hearings

I focus on academic utilization in congressional committee hearings. The U.S. Congress is the central policy-making institution that confronts the full range of policy problems in the United States. Policymakers divide the labor of collecting information on problems across different committees, which are specialized bodies within Congress that are provided staff and resources to collect information on problems within their respective jurisdictions (Weingast and Marshall 1988; Krehbiel 1991). Congressional committees collect information on novel and established problems in the political system through hearings (Baumgartner 2015). Limitations in staff and member resources mean that committees can only hold hearings on a subset of issues. Thus, the set of problems that are the subject of the hearing represent the subset of problems in the systemic agenda that receive attention by the political system (Baumgartner and Jones 1993). Commit-

tees routinely hold hearings on both emerging issues and established problems, generating variation in the novelty of policy content discussed in each hearing (Baumgartner, Jones, and MacLeod 2000).

I examine patterns of academic utilization by analyzing the composition of witness panels across hearings. Committees interact with the supply of available information on a problem through their selection of witnesses that provide information during hearings. Information is transferred from witnesses to committee members through a prepared statement by the witness and a time allotment afforded to each committee member to ask witnesses questions related to their testimony. A wide range of experts and non-experts testify at hearings, but committee resources constrain the amount witnesses that can testify in each hearing. Expert organizations actively seek opportunities to testify because selection signals political influence. Interest groups view witness selection as an indicator of their standing in policy debates (Leyden 1995), think tanks pursue testimony to demonstrate policy impact and maintain funding (Rich 2004), and government agencies compete to shape policy discussions within their jurisdictions (Workman 2015).

Because expert organizations compete for limited slots, the composition of witness panels represents the expert information sources committees prioritize when confronting different problems. Figure 1 summarizes academic participation in congressional hearings from 1969 to 2021. Panel (a) shows that academics appear at only 21 percent of hearings and make up just 18 percent of witness panels when present. This pattern reflects the institutional purposes of many hearings focused on oversight, appropriations, and statutory compliance, where committees need information about administrative performance and political accountability rather than scholarly expertise. Panel (b) groups academic testimony based on the topic of the hearing. While academics testify across all major issue areas, the extent of academic testimony varies by issue. Academics are most prevalent in hearings concerning macroeconomics and civil rights, while they are least prevalent in hearings addressing public lands and transportation issues. While the extent of academic testimony varies across issue domains, I examine whether the novelty of issues within each major topic systematically shapes academic utilization.



(a) Distribution of Academic Witness Share Across Hearings

(b) Average Academic Witness Share by Issue Area

Figure 1: Academic Testimony in Congressional Hearings, 1969–2021.

One concern with relying on committee hearings as venues of information gathering is that little meaningful information is actually transmitted to committee members. Indeed, hearings are organized by committee chairs who often have incentives beyond the collection of information about policy problems and committee members often spend their allocated speaking time during hearings grandstanding instead of engaging in substantive policy engagement with the witnesses (Park 2021). However, I argue that the information provided by witnesses informs how problems are understood for two reasons. First, both committee members and staff process information throughout the preparation and execution of hearings. Staff prepare briefing materials, develop potential lines of questioning, and synthesize testimony, ensuring that relevant information enters the committee's deliberative process even if not all members actively engage with the information during the hearing. Second, testimony delivered during hearings is entered into the public record. All hearings are transcribed, and many contemporary hearings are televised, making witness information publicly accessible to journalists, policymakers, and the broader public.

Hearings serve as venues through which committees gather information on both emerging and established issues by engaging with expert and non-expert organizations that

provide testimony and respond to member questions. Academics constitute a small but meaningful share of these witnesses, contributing to the information environment through their research and expertise. My theory predicts that the presence of academic witnesses will be greater in hearings addressing novel dimensions of policy problems and will decline as those problems become more established on the political agenda.

A Linguistic Approach to Measuring Issue Novelty

Despite the large body of scholarship examining congressional information processing in dynamic policy environments, empirically assessing the novelty of a problem under examination is difficult. One challenge is measuring the degree to which an issue has established itself on the political agenda. While policy change often occurs in short bursts, the process through which an issue gains and maintains its presence on the agenda is typically incremental, as actors, institutions, and coalitions slowly develop around the issue (Cobb and Elder 1971; Baumgartner and Jones 1993). Thus, the degree to which each issue is established on the political agenda changes over time. For instance, issues concerning climate change gradually emerged on the congressional agenda, receiving only limited attention in the 1970s and 1980s before becoming an established policy concern as scientific understanding deepened, political coalitions shifted, and the issue's policy dimensions multiplied. A second challenge is identifying the degree to which an issue on the political agenda aligns with the set of issues examined in the past. While new technologies and exogenous shocks can cause the emergence of new issue areas, more often new dimensions emerge within existing problems. The inability to identify novel dimensions of issues has led scholars to emphasize the emergence of entirely new issue areas (Jones, Theriault, and Whyman 2019), while overlooking the majority of the evolutionary dynamics characterizing the policy process.

I address these challenges by examining how the language used to summarize congressional hearings evolves over time within the same topic. Human generated hearing summaries reduce the dimensions of lengthy hearing transcripts to their most important

elements, filtering out the perfunctory questions, procedural exchanges, and tangential discussions that may appear in full transcripts but do not reflect the core substantive content of the hearing. At the same time, hearing summaries are written to provide detail that capture the novel elements of the substance, including new problem framings, emerging policy dimensions, and shifting interpretations within established issue areas. Indeed, scholars widely use hearing summaries to categorize both the substantive content of an issue (Baumgartner and Jones 1991; Jones, Sulkin, and Larsen 2003; Jones and Baumgartner 2005) as well as the type of congressional activity taking place within a topic (Talbert, Jones, and Baumgartner 1995; McGrath 2013; Ban, Park, and You 2023; Ban, Park, and You 2025).

To assess the novelty of language used in hearings, I employ a retroactive comparative method that compares the language in each hearing to past hearings held on the same policy topic.[1] I rely on the Policy Agendas Project coding system, which assigns each hearing to one of 20 different policy categories covering a broad spectrum of issues including macroeconomics, defense, energy, and public lands (Jones, Peterson, et al. 2016). After grouping hearings into their policy categories, I compare the language used in each hearing to the language used in all prior hearings within the same policy category. Measuring novelty relative to the historical record within each policy area accounts for when hearings introduce novel terminology, framings, or problem definitions that have not previously appeared in congressional hearings of that topic. Moreover, the retroactive comparative method captures the gradual establishment of issues, as the language associated with a problem becomes less novel over time when subsequent hearings adopt similar terminology and framings. To ensure that early hearings are compared against a stable reference corpus, I draw on hearing summaries from the 80th to the 117th Congresses, but begin calculating novelty with the 91st Congress to establish an initial baseline of language usage within each policy topic.

Using the retroactive comparison approach, I develop three measures that capture

[1]While text-as-data approaches have been used to measure legislative text reuse (Wilkerson, Smith, and Stramp 2015; Casas, Denny, and Wilkerson 2020), such measures detect repetition rather than innovation and cannot identify whether hearings introduce new concepts or framings.

different dimensions of issue novelty. The first two measures use a bag-of-words approach to identify differences in language used to describe a hearing. The bag-of-words approach treats each document as an unordered collection of words, disregarding grammar and word order while preserving word frequency.[2] The first measure captures the introduction of entirely new vocabulary by counting the number of tokens in a hearing summary that have never appeared in prior hearings within the same policy category. New problems typically require new terminology to describe. For instance, when cryptocurrency first emerged as a policy concern, hearings would necessarily introduce terms like "blockchain," and "Bitcoin" that had no precedent in prior financial services hearings. As a result, I identify all unique words that do not appear in any previous hearing summary on that topic where a higher count of new words indicates that the hearing is addressing aspects of the policy problem that have not been previously discussed in congressional hearings.

The second measure identifies the degree that an issue has established itself on the congressional agenda by examining the extent to which hearings draw on established vocabulary. As an issue matures and becomes a regular feature of congressional deliberation, the language used to discuss it stabilizes and becomes more frequent. Indeed, words like “emissions” and “renewable” have been repeatedly used over the past two decades to describe environmental and climate policy, reflecting how what was once a novel concern has become a routine part of the legislative agenda. To measure the degree to which hearings draw on established vocabulary, I calculate the average frequency with which tokens in a hearing summary have appeared in past hearings. For each token in a hearing summary, I count its total occurrences across all prior hearings within the same policy category, then average these counts across all tokens in the summary. Higher average word frequencies indicate that the hearing is discussing the policy problem using well-established terminology that has been repeatedly invoked in previous congressional discussions. Finally, I take the natural log of the average token count to account for its right-skewed distribution (mean = 153.5, SD = 145.3), which reduces the influence of

[2]To ensure that the measures focus on substantive content rather than common linguistic structures, I clean the data by removing all stopwords, punctuation, numbers, lowercasing all text, and stemming each word in the corpus.

hearings with extremely high token overlap. Taken together, the first and second measures capture the two ends of the same underlying process, where issues evolve from novel to established as the introduction of new vocabulary declines and reliance on established language increases.

While the first two measures identify lexical innovation and stabilization, they cannot distinguish between linguistic and conceptual change. For example, hearings referring to "global warming" and "climate change" would appear novel despite addressing the same issue. As a result, I develop a measure of novelty using an embedding approach that represents each sentence as real-valued vectors derived from the contextual patterns in which they appear in natural language, whereby sentences with similar vectors convey similar meanings. One challenge with embedding-based similarity measures is that distances concentrate in high-dimensional spaces as the comparison set grows, reducing semantic contrast so that similarity scores reflect global geometry rather than meaning (Beyer et al. 1999). To mitigate this concern, I restrict comparisons to a rolling 5-congress window of prior hearings within the same major topic. I use the SentenceTransformer model all-mpnet-base-v2, which encodes the semantic content for each hearing description by mapping each text into a 768-dimensional vector (Reimers and Gurevych 2019).

To measure semantic novelty, I calculate the cosine similarity between each hearing's embedding and the mean embedding of hearings on the same major topic from the preceding five congresses. Cosine similarity measures the angle between two vectors in the embedding space, with values closer to 1 indicating greater semantic similarity and values closer to -1 indicating opposite semantic orientation.[3] Although cosine similarity is defined on $[-1, 1]$, in practice sentence embeddings for related texts tend to cluster within the positive orthant of the embedding space, where most observed values fall between 0 and 1.[4] I define semantic novelty as one minus the cosine similarity, such that higher values indicate greater semantic deviation from recent topic-specific discourse. Figure 2 illustrates the distributions of the three linguistic novelty measures across congressional

[3]Formally, cosine similarity S_{ij} between two vectors $\mathbf{v}_i$ and $\mathbf{v}_j$ is defined as $S_{ij} = \frac{\mathbf{v}_i \cdot \mathbf{v}_j}{\|\mathbf{v}_i\| \, \|\mathbf{v}_j\|}$, where $\mathbf{v}_i \cdot \mathbf{v}_j$ denotes the dot product and $\|\mathbf{v}_i\|$ and $\|\mathbf{v}_j\|$ denote the Euclidean norms of the vectors.

[4]Only one observation in the sample exhibits a negative cosine similarity value.

hearings.

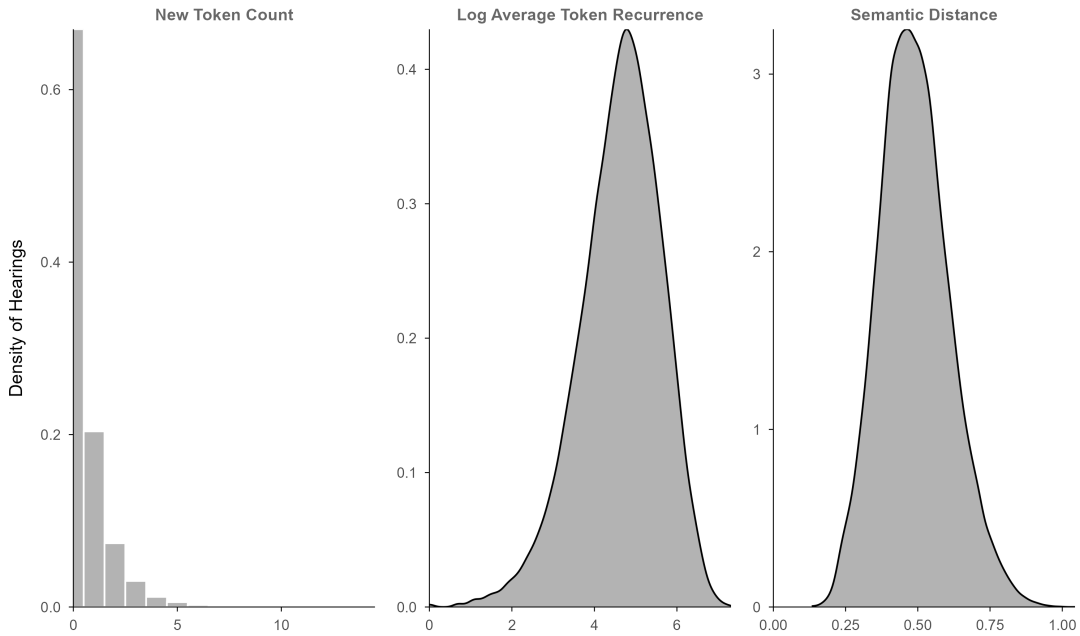


Figure 2: Distributions of three measures of linguistic novelty across congressional hearings, 1969 to 2021. *Note:* Each panel displays the distribution of a different measure of linguistic novelty across hearings. The histogram in the left panel and the kernel density estimates in the middle and right panels are scaled as probability densities, so the area under each curve equals one. Panel scales differ because each measure spans a distinct numerical range.

I take several steps to validate that these measures capture the extent to which an issue is novel or established on the congressional agenda. First, I examine the hearing descriptions with the ten highest and ten lowest scores for each measure of novelty. The hearings with higher novelty and semantic novelty scores clearly address problems that were new for their time.[5] Second, I assess the pairwise correlations among the three measures, which behave in theoretically consistent ways. Hearings that introduce new vocabulary are moderately more semantically distinct from prior hearings ($r = 0.27$), while those that rely more heavily on established vocabulary are less semantically novel ($r = -0.36$). Hearings introducing new words also tend to draw less on established terminology ($r = -0.15$), though the weak relationship suggests that these measures capture distinct aspects of novelty and stabilization. Finally, I demonstrate that the measures of linguistic novelty behave consistently with the evolution of how Congress has

[5]Examples provided in the online appendix.

discussed issues related to the internet. Congressional attention to computer networking and digital communication first appeared sporadically in the early 1980s, largely in the context of research and defense. The Internet, however, began to emerge as a distinct policy domain in the early 1990s, as the public introduction of the World Wide Web in 1991 transformed it into a mass communication and economic platform. As shown in Figure A2, the trajectory of the three measures of issue novelty are consistent with institutionalization of Internet issues. As Congress's attention to the Internet evolved from an emerging to an established policy domain, the number of new terms used in hearings declined, reliance on established vocabulary increased, and the semantic distance between hearings narrowed.[6]

Research Design

To evaluate how hearing-level novelty shapes academic utilization, I measure the proportion of witnesses from Research 1 (R1) universities who testify at each congressional hearing. The Carnegie Classification of Institutions of Higher Education designates universities as R1 if they have "Very High Research Spending and Doctorate Production." (American Council on Education 2025). As a result, universities designated as R1 are most likely to evaluate faculty on their research productivity rather than their teaching or service. Thus, focusing on R1-affiliated witnesses captures scholars from institutions where the evaluative emphasis on research is most pronounced.

I construct the measure of academic utilization by first compiling raw witness data from ProQuest Congressional and creating a dictionary of R1 university names and acronyms to identify university-affiliated witnesses. University affiliation alone does not reliably indicate academic expertise. Senior administrators such as presidents, provosts, and deans typically testify in representative capacities rather than as subject-matter experts. To identify witnesses appearing as subject-matter experts rather than institutional representatives, I develop two dictionaries classifying job titles as either administrative

[6] A full explanation of how this example was constructed, along with a detailed description of the case, is provided in the appendix.

positions or research positions.[7] For witnesses whose titles were not in either dictionary, research assistants manually verify their role by reviewing the hearing transcripts and testimony to determine whether they testify primarily as researchers or administrators. The approach identifies 22,747 research oriented academics that testified at committee hearings. Finally, to capture the extent to which committees rely on researchers in high-research-activity positions, I calculate the proportion of research-oriented academics among all witnesses at each hearing.

I model the proportion of academics testifying at each hearing using a quasi-binomial regression with a logit link and a fixed effects design that exploits variation in novelty across hearings within committees and issue areas over time. The formal specification is:

$$\text{Academic Proportion}c, i, t = \alpha + \gamma c + \omega_t + \delta_i + \theta_{ch} + \beta_1 \text{Hearing Novelty}_{c,i,t} + \mathbf{X}_{c,i,t} + \epsilon_{c,i,t} \quad (1)$$

where γ_c represents committee fixed effects, accounting for baseline differences in each committee's expertise requirements and witness selection patterns; ω_t represents Congress fixed effects, controlling for temporal changes in both the information environment and the extent to which hearings are used for information gathering versus symbolic or political purposes; δ_i represents major topic fixed effects, accounting for systematic differences in academic involvement across policy domains; and θ_{ch} represents chamber fixed effects, controlling for differences between House and Senate hearing practices. The error term $\epsilon_{c,i,t}$ is clustered by committee to account for within-committee correlation in witness selection patterns.

The primary independent variables are the issue-based measures of novelty developed in the previous section. Lexical innovation captures the introduction of entirely new vocabulary by counting tokens in a hearing summary that have never appeared in prior hearings within the same policy category, with higher counts indicating discussion of previously unaddressed aspects of the policy problem. Lexical recurrence measures the extent to which hearings draw on established vocabulary by calculating the log average

[7] If a witness held both a faculty position and an administrative role (such as a dean who also maintains a faculty appointment), I categorized them as holding an administrative position.

frequency with which tokens in a hearing summary have appeared in past hearings within the same policy category, where higher values indicate reliance on well-established terminology that has been repeatedly invoked in previous congressional discussions. Semantic novelty captures how conceptually distant a hearing is from previous hearings on the same major topic using sentence embeddings, testing whether academic utilization increases when examining unexplored dimensions of established policy areas.

The specification includes several control variables to isolate the effect of issue novelty from other factors that shape academic testimony. First, I control for the possibility that the relationship between academic utilization and issue novelty is driven by the functions committee perform during hearings. In addition to collecting information, committee use hearings to perform routine functions that include authorizing and funding agency programs, reviewing pending legislation, and conducting bureaucratic oversight. Each type of hearing systematically features different witness compositions than information-gathering hearings (Ban, Park, and You 2023). Authorization and appropriations hearings prioritize agency officials and congressional budget staff who address program implementation and resource needs. Oversight hearings center on agency officials and inspectors general who can speak to bureaucratic performance. Legislative review hearings draw heavily on stakeholders and advocacy organizations affected by pending bills. I account for whether the relationship between novelty and utilization is driven by the functions being performed by committees by controlling for whether hearings serve appropriation, legislative referral, or oversight functions.

In addition to hearing level controls, I also account for who is organizing the hearing. The Democratic party has historically pursued policy that expands the role of government in public life, implying that Democrats are more likely to introduce new problems onto the political agenda. In addition, academia has historically been tied to liberal normative values that align more with the Democratic party (Buckley 1951). To account for the possibility that greater academic utilization is driven by partisan dynamics rather than the novelty of the problem under examination, I include a binary variable that takes a value of one if the committee chair was a Democrat.

Finally, I control for the state of the broader policy environment. Major policy changes often emerge when previously unexplored dimensions of problems become salient, prompting policymakers to search more broadly across expert communities (Baumgartner and Jones 1993; Jones and Baumgartner 2005; Baumgartner 2015). If academic utilization primarily reflects these moments of issue salience rather than informational sparsity on novel problems, the theory would not explain variation within stable periods of attention. To address this, I control for the proportion of hearings held on each major topic in each Congress.

Academic Testimony and Issue Novelty

Table 1 presents results from six quasi-binomial models estimating the proportion of academic witnesses invited to testify at congressional hearings from the 91st to the 117th Congress. Because coefficients from quasi-binomial models are expressed in log-odds and are difficult to interpret directly, figure 3 displays predicted probabilities from the fully specified model for each measure of novelty.

Model 1 estimates the relationship between novel vocabulary and academic participation while controlling only for committee, Congress, chamber, and majortopic fixed effects. The count of novel tokens used in hearing descriptions is positively associated with the proportion of academic witnesses, and this relationship holds in Model 2 after adding the full set of controls. Substantively, holding other variables at their means, increasing the novel token count by one is associated with roughly a one percentage-point increase in the predicted academic witness share. Together, these results indicate that committees systematically turn to academic expertise when hearings are examining topics that are described with novel words.

Models 3 and 4 examine whether conceptually unfamiliar hearings rely more heavily on academic expertise. Cosine Distance ranges from 0 representing hearing language nearly identical to the prior way in which the topic was discussed to 1 representing completely novel conceptual content. The results show a large and consistent relationship.

Table 1: Quasi-Binomial Model Predicting Proportion of Academic Testimony at Every hearing (91st to 117th Congress)

	(1)	(2)	(3)	(4)	(5)	(6)
Novel Token Count	0.056***	0.043***				
	(0.011)	(0.008)				
Semantic Distance			1.309***	1.113***		
			(0.184)	(0.191)		
Lexical Recurrence					−0.147***	−0.110***
					(0.022)	(0.026)
Appropriations Hearing		−0.739***		−0.667***		−0.702***
		(0.082)		(0.082)		(0.084)
Legislative Hearing		−0.372***		−0.355***		−0.365***
		(0.061)		(0.063)		(0.062)
Oversight Hearing		−0.164***		−0.164***		−0.148***
		(0.040)		(0.039)		(0.042)
Democratic Control		0.146		0.146		0.142
		(0.081)		(0.083)		(0.082)
Topic Salience		−1.272		−1.360		−0.976
		(1.796)		(1.820)		(1.775)
N	60 992	56 878	60 992	56 878	60 992	56 878
Congress FE	X	X	X	X	X	X
Chamber FE	X	X	X	X	X	X
Issue Area FE	X	X	X	X	X	X

* p <0.05, ** p <0.01, *** p <0.005
Standard errors clustered by committee.

Substantively, moving from a hearing one standard deviation below the mean of conceptual familiarity of 0.3 to one standard deviation above the mean of 0.62, the predicted share of academic witnesses increases by about three percentage points. When controls are added in Model 4, the estimated effect remains large and statistically significant. A shift from one standard deviation below the mean of 0.3 to one standard deviation above the mean of 0.62 produces an increase of roughly six percentage points in the proportion of academic witnesses testifying at the hearing.

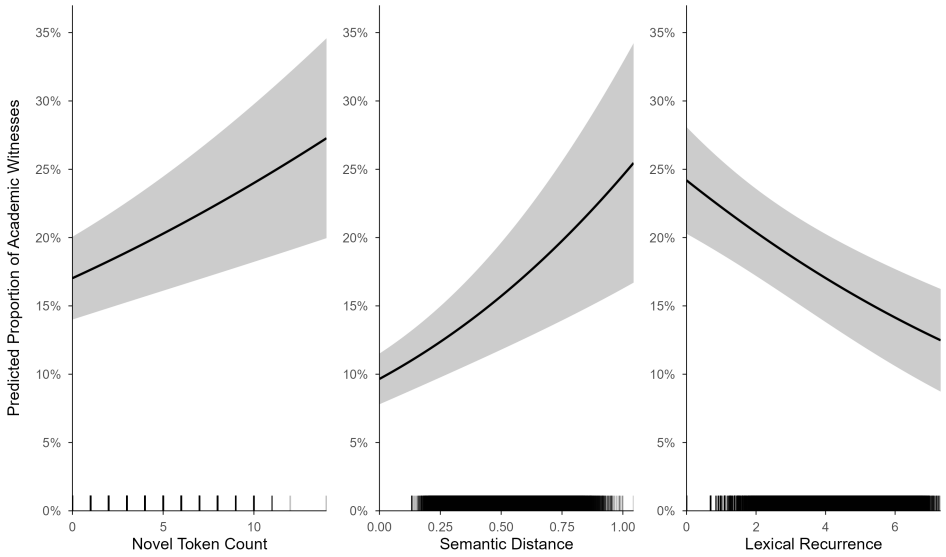


Figure 3: Predicted Probability of Academic Witnesses. *Note:* Each panel plots the predicted proportion of academic witnesses across the full observed range of a given novelty measure (Novel Token Count, Lexical Recurrence, and Semantic Distance). Shaded regions represent 95% confidence intervals computed from quasi-binomial models with Congress, major topic, committee, and chamber fixed effects, and standard errors clustered by committee. Predictions condition on the House Science Committee, Science and Technology issues, and the 110th Congress, with remaining covariates held at their means.

Finally, models 5 and 6 examine how the establishment of a policy problem shapes academic utilization. Model 5 shows that lexical recurrence is negatively associated with academic participation, and the relationship remains robust when controls are added in Model 6. As issues become more entrenched on the congressional agenda, the language used to describe them becomes more repetitive. When problems are routinely discussed, academics make up a smaller share of the witnesses that testify on the issue. Substantively, a one-unit increase in the logged average number of prior uses of the terms

appearing in a hearing description is associated with an approximate one percentage-point decline in the share of academic witnesses. As the vocabulary used in a hearing more closely resembles language that committees have relied on in the past, academic testimony becomes less prevalent.

One limitation of the quasi-binomial approach is that it treats the academic witness share as a continuous proportion, potentially conflating changes in the composition of testimony with variation in the total number of witnesses across hearings. This raises the concern that estimated effects could reflect differences in hearing size rather than shifts in the likelihood of academic participation. To address this issue, Table A1 in the appendix presents negative binomial models that treat academic testimony as a count outcome and include the total number of witnesses as an offset, explicitly accounting for variation in exposure. The results are substantively identical, suggesting that the results in table 1 do not depend on the proportional specification.

The results indicate that academic utilization varies systematically with the novelty of issues on the congressional agenda. When issues first break out onto the congressional agenda academics make up a larger share of the witness panel. However, as committees repeatedly engage with a problem, academics make up a smaller portion of the witness panel.

Distinguishing Academic and Expert Utilization

The positive association between issue novelty and academic testimony documented in the previous section may reflect two distinct causal mechanisms. First, committees may systematically increase their reliance on academic expertise when confronting novel policy problems due to academics' comparative advantage in analyzing unfamiliar phenomena. Alternatively, the relationship may capture a general tendency for committees to prioritize expert information when examining unfamiliar issues (Baumgartner 2015).[8] Under this interpretation, the observed relationship would not reflect an increase in academic

[8]Some research suggests the opposite dynamic where expert participation may decrease when Congress engages issues lacking established policy frameworks Esterling (2004).

utilization relative to other expert sources but rather a general congressional demand for expertise.

To assess whether the positive relationship between issue novelty and academic testimony reflects a broader increase in expert engagement, I examine the utilization of three alternative expert organizations: think tanks, federal agencies, and analytical agencies. Despite serving different constituencies, all three organizations are more reliant on policymakers utilizing their information than academics. Think tanks operate as external sources of expertise and must demonstrate policy impact to retain donors who expect their contributions to influence debate among ideological allies (Fagan 2021). Federal agencies function as internal resources answering primarily to the executive branch and require adoption of their recommendations to maintain bureaucratic autonomy and expand policy authority (Carpenter 2001). Analytical agencies like the Congressional Research Service and Government Accountability Office serve Congress directly and must establish their value through consistent utilization to secure continued appropriations (Fagan and McGee 2022).

The stronger dependence on demonstrated utility among these three organizations provides analytical leverage for distinguishing between the two causal mechanisms. Organizations that rely directly on policymaker utilization for their funding have strong incentives to concentrate on established problems where demand for expertise is predictable and contributions are readily measurable. If the positive association between issue novelty and academic testimony reflects a general congressional tendency to seek expert information on unfamiliar issues, we should observe similar increases in utilization of think tanks, federal agencies, and analytical agencies. Conversely, if academics possess a comparative advantage in analyzing novel policy problems, these organizations should not exhibit the same positive relationship with issue novelty. I identify witnesses from these three organization types using the same hearing data and coding procedures described previously, categorizing witnesses based on their institutional affiliation at the time of testimony.[9]

[9]Detailed coding procedures for each organization type are provided in the appendix.

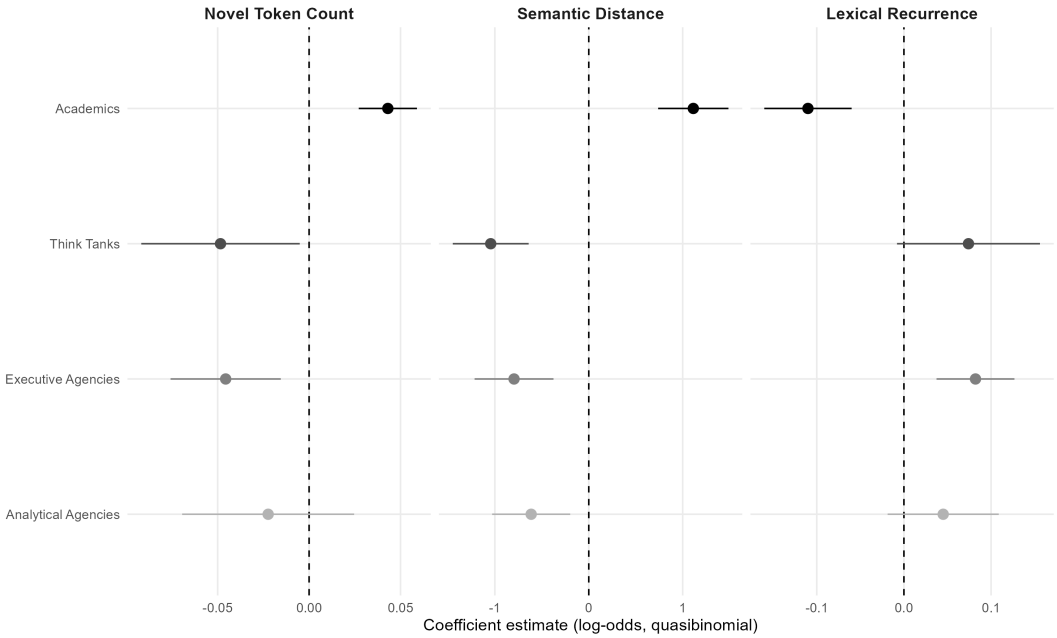


Figure 4: Effect of Issue Novelty on Expert Participation Across Organizations. *Note:* Each panel displays the estimated coefficient and 95% confidence interval for one measure of issue novelty (Novel Token Count, Semantic Distance, and Lexical Recurrence). All coefficients are drawn from fully specified models that include the complete set of controls, topic, committee, chamber, and Congress fixed effects, with standard errors clustered by committee. Full model specifications for every organization and novelty measure are provided in the appendix.

Figure 4 present the relationship between linguistic novelty and expert testimony across the four organization types.[10] Across the three measures of issue novelty, think tank, executive agency, and analytical agency testimony patterns differ from academics. For novel tokens, both think tanks and federal bureaucrats demonstrate significant negative associations with novel token count, declining by 0.001 and 0.009 percentage points respectively for each additional novel token. The share of bureaucrats from analytical agencies is also negatively associated with the count of novel words used to describe hearings, though the estimate is imprecisely estimated. Panel B demonstrates a similar pattern for semantic distance, with all three expert organizations making up a smaller share of the witness panels on hearings addressing issues that are more semantically dissimilar from past hearings on the same topic. All three organizations surpass conventional levels of statistical significance. Finally, panel C displays the relationship between token recurrence and expert testimony. Federal bureaucrats and think tanks exhibit a positive and significant relationship with issue recurrence indicating that both experts make up a larger share of the witness panel as an issues recurs more frequently in the historical record. The share of bureaucrats from analytical agencies are also positively associated with the degree of an issues recurrence, but the coefficient is imprecisely estimated.

Together, the negative novelty coefficients and positive recurrence coefficients across the different types of nonacademic experts demonstrate that increased academic utilization does not stem from Congress broadly expanding expert engagement when problems are unfamiliar. Instead, the composition of expert witnesses changes as the informational environment surrounding an issue becomes more established. When a problem first appears on the congressional agenda, the information environment is sparse because organizations that depend on policymaker utilization do not yet have relevant analysis or established positions to offer. As a result, academics face little competition from utilization-oriented organizations. As issues become established, utilization-oriented organizations generate policy-relevant information that is tailored for policymakers. Academic utilization declines as Congress prioritizes experts with information that is easier

[10]Full model specifications for each organization are provided in the appendix.

for policymakers to process and is tailored to fit the ideological orientation of each member.

Conclusion

In this article, I identify an unappreciated determinant of academic utilization in the policy process. Through academia's system of peer review, academics are more likely to invest analytical resources into developing expertise over problems that are not established on the political agenda than utilization-oriented organizations. As a result, academic utilization is greater when an issue first emerges on the political agenda, then declines as the issue recurs more frequently on the agenda. I test this core proposition using original witness data consisting of all academic researcher testimony at congressional hearings over the last 50 years. To identity the novelty of the issue examined in each hearing, I develop an original measure of issue novelty that captures the relative novelty of the issue under examination relative to how the issue was discussed in the past. I demonstrate that academics are more likely to testify on hearings addressing more novel issues and are less likely to testify on established issues. Subsequent results demonstrate that this effect is driven by committees prioritizing academics over experts on novel issues.

While this study focuses on the U.S. committee system, most governments at the national, state, and local levels draw on multiple expert organizations, and there is little reason to think that similar dynamics would not apply in these institutions as well. Congress is one of the most competitive expert environments in the world, with federal agencies, analytical agencies, think tanks, advocacy groups, and industry experts competing to shape national policy. In such dense environments, academia's informational advantage is replaced quickly once issues gain sustained political attention. By contrast, many policy-making venues operate in the context of less developed expert infrastructures. State legislatures, city councils, and small executive offices often face emerging problems with far fewer utilization-oriented experts capable of producing timely analysis. Future research should test whether academia's informational advantage persists longer

in institutional contexts with less developed expert infrastructures.

In addition to explaining variation in academic utilization, the framework of information supply dynamics can be applied to explain utilization patterns across a wide variety of expert organizations. For example, external and internal sources of expertise are the subject of oversight and evaluation by different actors in the political system. Most existing studies provide demand-side theories of utilization, highlighting the strategic incentives to utilize different experts (Esterling 2004; Hall and Deardorff 2006); however, expert segmentation dynamics provides a robust framework for theorizing the impact of information availability on variation in the types of experts providing information. More broadly, this framework highlights how evaluation systems shape where expertise accumulates; the extent to which these mechanisms structure information availability in other organizational settings beyond public policy is an important open question.

Finally, this argument has implications for ongoing efforts to make academic research more policy relevant. Many reform proposals urge scholars to align their agendas more closely with the immediate needs of policymakers or to emulate think tanks by packaging research around current debates (Mead 2015; Cairney and Oliver 2020). Yet, my findings provide evidence that at least part of academia's value to policymakers is in its capacity to generate expertise over problems in advance of political demand. Thus, efforts to make academic research more policy relevant risk weakening the very qualities that make academics uniquely useful when policymakers need them most.

References

American Council on Education. 2025. *The Carnegie Classification of Institutions of Higher Education.*

Austen-Smith, David. 1993. "Information and Influence: Lobbying for Agendas and Votes". In: *American Journal of Political Science* 37: 799.

Azoulay, Pierre, Joshua S. Graff Zivin, and Gustavo Manso. 2011. "Incentives and Creativity: Evidence from the Academic Life Sciences". In: *The RAND Journal of Economics* 42: 527–554.

Ban, Pamela, Ju Yeon Park, and Hye Young You. 2023. "How Are Politicians Informed? Witnesses and Information Provision in Congress". In: *American Political Science Review* 117: 122–139.

Ban, Pamela, Ju Yeon Park, and Hye Young You. 2025. "Bureaucrats in Congress: The Politics of Inter-branch Information Sharing". In: *The Journal of Politics*: 735440.

Baumgartner, Frank R. 2015. *The Politics of Information: Problem Definition and the Course of Public Policy in America.* Chicago: University of Chicago Press.

Baumgartner, Frank R. and Bryan D. Jones. 1991. "Agenda Dynamics and Policy Subsystems". In: *The Journal of Politics* 53: 1044–1074.

Baumgartner, Frank R. and Bryan D. Jones. 1993. *Agendas and Instability in American Politics.* Chicago: University of Chicago Press.

Baumgartner, Frank R., Bryan D. Jones, and Michael C. MacLeod. 2000. "The Evolution of Legislative Jurisdictions". In: *The Journal of Politics* 62: 321–349.

Belkhodja, Omar, Nabil Amara, Réjean Landry, and Mathieu Ouimet. 2007. "The Extent and Organizational Determinants of Research Utilization in Canadian Health Services Organizations". In: *Science Communication* 28: 377–417.

Bertelli, Anthony M. and Jeffrey B. Wenger. 2009. "Demanding Information: Think Tanks and the US Congress". In: *British Journal of Political Science* 39: 225–242.

Beyer, Kevin, Jonathan Goldstein, Raghu Ramakrishnan, and Uri Shaft. 1999. "When Is "Nearest Neighbor" Meaningful?" In: *Database Theory — ICDT'99*. Ed. by Catriel Beeri and Peter Buneman. Berlin: Springer: pp. 217–235.

Beyers, Jan. 2004. "Voice and Access: Political Practices of European Interest Associations". In: *European Union Politics* 5: 211–240.

Birkland, Thomas A. 1998. "Focusing Events, Mobilization, and Agenda Setting". In: *Journal of Public Policy* 18: 53–74.

Boswell, Christina. 2008. "The Political Functions of Expert Knowledge: Knowledge and Legitimation in European Union Immigration Policy". In: *Journal of European Public Policy* 15: 471–488.

Boswell, Christina. 2012. "How Information Scarcity Influences the Policy Agenda: Evidence from U.K. Immigration Policy". In: *Governance* 25: 367–389.

Buckley, William Frank. 1951. *God and Man at Yale: The Superstitions of "Academic Freedom"*. Washington, DC: Regnery Publishing.

Cairney, Paul and Kathryn Oliver. 2020. "How Should Academics Engage in Policymaking to Achieve Impact?" In: *Political Studies Review* 18: 228–244.

Cantarelli, Paola, Nicola Belle, and Jeremy L. Hall. 2023. "Information Use in Public Administration and Policy Decision-making: A Research Synthesis". In: *Public Administration Review* 83: 1667–1686.

Caplan, Nathan. 1979. "The Two-Communities Theory and Knowledge Utilization". In: *American Behavioral Scientist* 22: 459–470.

Carpenter, Daniel P. 2001. *The Forging of Bureaucratic Autonomy: Reputations, Networks, and Policy Innovation in Executive Agencies, 1862-1928*. Princeton, N.J: Princeton University Press.

Casas, Andreu, Matthew J. Denny, and John Wilkerson. 2020. "More Effective Than We Thought: Accounting for Legislative Hitchhikers Reveals a More Inclusive and Productive Lawmaking Process". In: *American Journal of Political Science* 64: 5–18.

Cherney, Adrian, Brian Head, Jenny Povey, Michele Ferguson, and Paul Boreham. 2015. "Use of Academic Social Research by Public Officials: Exploring Preferences and Constraints That Impact on Research Use". In: *Evidence & Policy* 11: 169–188.

Cobb, Roger W. and Charles D. Elder. 1971. "The Politics of Agenda-Building: An Alternative Perspective for Modern Democratic Theory". In: *The Journal of Politics* 33: 892–915.

Craft, Jonathan and Michael Howlett. 2012. "Policy Formulation, Governance Shifts and Policy Influence: Location and Content in Policy Advisory Systems". In: *Journal of Public Policy* 32: 79–98.

Cross, Jesse M. and Abbe R. Gluck. 2020. "The Congressional Bureaucracy". In: *University of Pennsylvania Law Review* 168: 1541–1683.

Downs, Anthony. 1972. "Up and Down with Ecology-the Issue-Attention Cycle". In: *Public Interest* 28.

Dunn, William N. 1980. "The Two-Communities Metaphor and Models of Knowledge Use: An Exploratory Case Survey". In: *Knowledge* 1: 515–536.

Esterling, Kevin M. 2004. *The Political Economy of Expertise: Information and Efficiency in American National Politics*. Ann Arbor: University of Michigan Press.

Fagan, EJ. 2021. "Issue Ownership and the Priorities of Party Elites in the United States, 2004-2016". In: *Party Politics* 27: 149–160.

Fagan, EJ. 2024. *The Thinkers: The Rise of Partisan Think Tanks and the Polarization of American Politics*. New York, NY: Oxford University Press.

Fagan, EJ, Alexander Furnas, Chris Koski, Herschel Thomas, Samuel Workman, and Corinne Connor. 2024. "The Dynamics of Issue Attention in Policy Process Scholarship". In: *Policy Studies Journal* 52: 481–492.

Fagan, EJ and Zachary A. McGee. 2022. "Problem Solving and the Demand for Expert Information in Congress". In: *Legislative Studies Quarterly* 47: 53–77.

Fagan, EJ, Zachary A. McGee, and Herschel F. Thomas. 2021. "The Power of the Party: Conflict Expansion and the Agenda Diversity of Interest Groups". In: *Political Research Quarterly* 74: 90–102.

Flatt, Henry and Nhat-Dang Do. 2023. "Analytic Bureaucracy and the Policy Process: Evidence from California". In: *The American Review of Public Administration* 53: 243–257.

Fraussen, Bert and Darren Halpin. 2017. "Think Tanks and Strategic Policy-Making: The Contribution of Think Tanks to Policy Advisory Systems". In: *Policy Sciences* 50: 105–124.

Gailmard, Sean and John W. Patty. 2007. "Slackers and Zealots: Civil Service, Policy Discretion, and Bureaucratic Expertise". In: *American Journal of Political Science* 51: 873–889.

Gray, Virginia, David Lowery, Matthew Fellowes, and Jennifer L. Anderson. 2005. "Legislative Agendas and Interest Advocacy: Understanding the Demand Side of Lobbying". In: *American Politics Research* 33: 404–434.

Hall, Richard L. and Alan V. Deardorff. 2006. "Lobbying as Legislative Subsidy". In: *American Political Science Review* 100: 69–84.

Halpin, Darren R. and Anne S. Binderkrantz. 2011. "Explaining Breadth of Policy Engagement: Patterns of Interest Group Mobilization in Public Policy". In: *Journal of European Public Policy* 18: 201–219.

Head, Brian W. 2013. "Evidence-Based Policymaking – Speaking Truth to Power?" In: *Australian Journal of Public Administration* 72: 397–403.

Hojnacki, Marie. 1997. "Interest Groups' Decisions to Join Alliances or Work Alone". In: *American Journal of Political Science* 41: 61.

Jennings, E. T. and J. L. Hall. 2012. "Evidence-Based Practice and the Use of Information in State Agency Decision Making". In: *Journal of Public Administration Research and Theory* 22: 245–266.

Jones, Bryan D. 2001. *Politics and the Architecture of Choice: Bounded Rationality and Governance*. Chicago London: University of Chicago Press.

Jones, Bryan D. and Frank R. Baumgartner. 2005. *The Politics of Attention How Government Prioritizes Problems*. Chicago: University of Chicago Press.

Jones, Bryan D., Tracy Sulkin, and Heather A. Larsen. 2003. "Policy Punctuations in American Political Institutions". In: *American Political Science Review* 97: 151–169.

Jones, Bryan D., Sean M. Theriault, and Michelle Whyman. 2019. *The Great Broadening: How the Vast Expansion of the Policymaking Agenda Transformed American Politics*. University of Chicago Press.

Jones, Michael D., Holly L. Peterson, Jonathan J. Pierce, Nicole Herweg, Amiel Bernal, Holly Lamberta Raney, and Nikolaos Zahariadis. 2016. "A River Runs Through It: A Multiple Streams Meta-Review". In: *Policy Studies Journal* 44: 13–36.

Kingdon, John W. 1984. *Agendas, Alternatives, and Public Policies*. Boston: Little, Brown.

Krehbiel, Keith. 1991. *Information and Legislative Organization*. Ann Arbor: University of Michigan Press.

Landry, Réjean, Moktar Lamari, and Nabil Amara. 2003. "The Extent and Determinants of the Utilization of University Research in Government Agencies". In: *Public Administration Review* 63: 192–205.

Leech, Beth L., Frank R. Baumgartner, Timothy M. La Pira, and Nicholas A. Semanko. 2005. "Drawing Lobbyists to Washington: Government Activity and the Demand for Advocacy". In: *Political Research Quarterly* 58: 19–30.

Leyden, Kevin M. 1995. "Interest Group Resources and Testimony at Congressional Hearings". In: *Legislative Studies Quarterly* 20: 431.

May, Peter J. 1991. "Reconsidering Policy Design: Policies and Publics". In: *Journal of Public Policy* 11: 187–206.

May, Peter J., Chris Koski, and Nicholas Stramp. 2016. "Issue Expertise in Policymaking". In: *Journal of Public Policy* 36: 195–218.

McGann, James G. 2016. *The Fifth Estate: Think Tanks, Public Policy, and Governance*. Washington, D.C: Brookings Institution Press.

McGann, James G. 2020. *Global Go to Think Tank Index Report*.

McGrath, Robert J. 2013. "Congressional Oversight Hearings and Policy Control". In: *Legislative Studies Quarterly* 38: 349–376.

Mead, Lawrence M. 2015. "Only Connect: Why Government Often Ignores Research". In: *Policy Sciences* 48: 257–272.

Medvetz, Thomas. 2012. *Think Tanks in America*. Paperback ed. Chicago, Ill.: Univ. of Chicago Press.

Merton, Robert K. 1957. “Priorities in Scientific Discovery: A Chapter in the Sociology of Science”. In: *American Sociological Review* 22: 635.

Montpetit, Éric. 2011. “Scientific Credibility, Disagreement, and Error Costs in 17 Biotechnology Policy Subsystems”. In: *Policy Studies Journal* 39: 513–533.

Newman, Joshua, Adrian Cherney, and Brian W. Head. 2016. “Do Policy Makers Use Academic Research? Reexamining the “Two Communities” Theory of Research Utilization”. In: *Public Administration Review* 76: 24–32.

Newman, Joshua and Brian Head. 2015. “Beyond the Two Communities: A Reply to Mead’s “Why Government Often Ignores Research”". In: *Policy Sciences* 48: 383–393.

Palus, Christine Kelleher and Susan Webb Yackee. 2016. “Clerks or Kings? Partisan Alignment and Delegation to the US Bureaucracy”. In: *Journal of Public Administration Research and Theory* 26: 693–708.

Park, Ju Yeon. 2021. “When Do Politicians Grandstand? Measuring Message Politics in Committee Hearings”. In: *The Journal of Politics* 83: 214–228.

Payne, John W. 1976. “Task Complexity and Contingent Processing in Decision Making: An Information Search and Protocol Analysis”. In: *Organizational Behavior and Human Performance* 16: 366–387.

Pirolli, Peter and Stuart Card. 1999. “Information Foraging”. In: *Psychological Review* 106: 643–675.

Raymaekers, Pieter, Koen Migchelbrink, Valérie Pattyn, and Peter De Smedt. 2025. “Organizational and Individual Factors of Evidence Informed Policy Making in Public Organizations”. In: *Public Administration Review*: puar.13946.

Redford, Emmette S. 1960. “A Case Analysis of Congressional Activity: Civil Aviation, 1957-58”. In: *The Journal of Politics* 22: 228–258.

Reimers, Nils and Iryna Gurevych. 2019. *Sentence-BERT: Sentence Embeddings Using Siamese BERT-Networks.*

Rich, Andrew. 2004. *Think Tanks, Public Policy, and the Politics of Expertise.* Cambridge University Press.

Sabatier, Paul A. 1988. "An Advocacy Coalition Framework of Policy Change and the Role of Policy-Oriented Learning Therein". In: *Policy Sciences* 21: 129–168.

Simon, H. A. 1956. "Rational Choice and the Structure of the Environment." In: *Psychological Review* 63: 129–138.

Simon, Herbert A. 1955. "A Behavioral Model of Rational Choice". In: *The Quarterly Journal of Economics* 69: 99.

Simon, Herbert A. 1997. *Administrative Behavior, 4th Edition.* 4th ed. New York: Free Press.

Stephan, Paula E. 1996. "The Economics of Science". In: *Journal of Economic Literature* 34: 1199–1235.

Stone, Diane. 2007. "Recycling Bins, Garbage Cans or Think Tanks? Three Myths Regarding Policy Analysis Institutes". In: *Public Administration* 85: 259–278.

Talbert, Jeffery C., Bryan D. Jones, and Frank R. Baumgartner. 1995. "Nonlegislative Hearings and Policy Change in Congress". In: *American Journal of Political Science* 39: 383.

Tversky, Amos. 1972. "Elimination by Aspects: A Theory of Choice". In: *Psychological Review* 79.

Tversky, Amos and Daniel Kahneman. 1973. "Availability: A Heuristic for Judging Frequency and Probability". In: *Cognitive Psychology* 5: 207–232.

Tversky, Amos and Daniel Kahneman. 1974. "Judgment under Uncertainty: Heuristics and Biases: Biases in Judgments Reveal Some Heuristics of Thinking under Uncertainty." In: *Science* 185: 1124–1131.

Walker, Jack L. 1983. "The Origins and Maintenance of Interest Groups in America". In: *American Political Science Review* 77: 390–406.

Walker, Richard M., Jiasheng Zhang, Yanto Chandra, Binzizi Dong, and Yao Wang. 2023. "Revisiting the Academic–Practitioner Divide: Evidence from Computational Social Science and Corpus Linguistics". In: *Public Administration Review* 83: 1599–1617.

Ward, Vicky, Allan House, and Susan Hamer. 2009. "Knowledge Brokering: The Missing Link in the Evidence to Action Chain?" In: *Evidence & Policy* 5: 267–279.

Weingast, Barry R. and William J. Marshall. 1988. "The Industrial Organization of Congress; or, Why Legislatures, Like Firms, Are Not Organized as Markets". In: *Journal of Political Economy* 96: 132–163.

Weiss, Carol H. 1979. "The Many Meanings of Research Utilization". In: *Public Administration Review* 39: 426–431.

Weiss, Carol H. 1977. "Research for Policy's Sake: The Enlightenment Function of Social Research". In: *Policy Analysis* 3: 531–545.

Whitley, Richard D. 1987. *The Intellectual and Social Organization of the Sciences*. Reprint. Oxford: Clarendon Pr.

Wilkerson, John, David Smith, and Nicholas Stramp. 2015. "Tracing the Flow of Policy Ideas in Legislatures: A Text Reuse Approach". In: *American Journal of Political Science* 59: 943–956.

Workman, Samuel. 2015. *The Dynamics of Bureaucracy in the U.S. Government: How Congress and Federal Agencies Process Information and Solve Problems*. Cambridge: Cambridge University Press.

Workman, Samuel, JoBeth Shafran, and Tracey Bark. 2017. “Problem Definition and Information Provision by Federal Bureaucrats”. In: *Cognitive Systems Research* 43: 140–152.

Xu, Chengxin, Yuan (Daniel) Cheng, Shuping Wang, Weston Merrick, and Patrick Carter. 2025. “Evaluating Use of Evidence in U.S. State Governments: A Conjoint Analysis”. In: *Public Administration Review* 85: 1217–1235.

Appendix

Validating Measures of Issue Novelty

Illustrative Hearings by Linguistic Novelty Measures

Congress	Hearing Description	Novel Token Count	Group
105	Examine FDA advanced notice of proposed rule-making regarding transition from metered dose inhalers propelled by chlorofluorocarbons to CFC-free MDI's under the Montreal Protocol on Substances That Deplete the Ozone Layer.	14	Top 10
109	Considers the following nominations: Gale A. Buchanan (Dean and Dir Emeritus, Coll of Agric and Environmental Sciences, Univ of Ga) to be Under Secretary, Research, Education, and Economics, USDA. Marc L. Kesselman (Dep Gen Counsel, OMB) to be General Cou.	14	Top 10
93	Prevent certain disruptions of ocean commerce between Hawaii, Guam, Samoa, or TTPI and the West coast by permitting petition for an injunction or restraining order against strike or lockouts.	12	Top 10
104	Examine availability to congress of accurate data on legal immigrations, in light of report by INS of recent decline while failing to disclose projections indicating an expected rise.	11	Top 10
105	Consider H.R. 2265 the No Electronic Theft (NET) Act to amend the Copyright Act to establish criminal fines and penalties for reproducing and distributing copyrighted works byelectronic means including the Internet even if the perpetrator does not bene.	11	Top 10
105	Review the merits of filling existing judgeship vacancies and creating new judgeships on the U.S. Court of Appeals for the Fourth, Fifth and Eleventh Court.	11	Top 10
106	Hearing in Springfield, Ill to examine proposals to revise EPA reformulated gasoline (RFG) program to replace methyl tertiary-butyl ether (MTBE) with grain-based or biomass-derived ethanol, in light of concerns that MTBE may be contaminating the water sup.	11	Top 10
105	Consider following bills- act to authorize naturalization of among refugees who served with U.S. supported guerilla units in Laos, and establish pilot project to let U.S. citizens use passports instead of ins boat landing permits to reenter U.S. after tri.	10	Top 10
106	Hearing before the Subcom on the Constitution to consider H.R. 3590 (text, p. 2-3), the ADA Notification Act, to amend the Americans with Disabilities Act (ADA) of 1990 to prohibit disabled individuals from bringing lawsuits against businesses to compel c.	10	Top 10
106	Hearing before the Subcom on Crime to consider H.R. 469 (text, p. 3), Jeremy and Julia's Law, to establish a new Federal crime for day care providers to knowingly make a false representation regarding the provider or the care given by the provider to pare.	10	Top 10
91	New Jersey-New York airport commission compact.	0	Bottom 10
91	Public health service hospital closings.	0	Bottom 10
91	Watershed project and soil conservation.	0	Bottom 10
91	Amend the CIA retirement act.	0	Bottom 10
91	Operation of the selective service system under Executive order.	0	Bottom 10
91	Benefits information service for servicemen and their family.	0	Bottom 10
91	Merchant marine shipbuilding and status of shipyards.	0	Bottom 10
91	Operations of military exchanges and commissaries within US and overseas.	0	Bottom 10
91	Government roles in combating youth crime, juvenile delinquency and youth gang criminal activity in Philadelphia.	0	Bottom 10
91	Need for legislation to control drug abuse.	0	Bottom 10

Congress	Hearing Description	Score	Group
98	Exam US military interventions in Grenada.	1495.00	Top 10
113	To examine FY2014 authorization requests for DOD programs.	1424.50	Top 10
114	To review FY2015 authorization requests for DOD programs.	1415.83	Top 10
112	To examine FY2013 authorizaion request for DOD programs.	1405.80	Top 10
114	To examine military defense posture and FY2015 authorization request for the DOD.	1369.50	Top 10
112	To review the FY2013 request for programs in the DOD.	1337.80	Top 10
112	To review the FY2013 requests for DOD programs.	1337.80	Top 10
112	To examine FY2012 authorization requests for DOD programs.	1329.00	Top 10
114	To examine DOD missile defense programs.	1315.40	Top 10
114	To consider FY2015 authorization request for DOD programs.	1313.67	Top 10
92	Death with dignity.	0.00	Bottom 10
93	Recycling of petrodollars.	0.00	Bottom 10
93	Facilitate the incorporation of the townsite of page, Arizona as a municipality.	0.00	Bottom 10
93	Daylight saving time.	0.00	Bottom 10
93	Inmate furloughs.	0.00	Bottom 10
93	Ten-year term for FBI director.	0.00	Bottom 10
93	DOT enforcement regarding schoolbus safety; safety design standards for schoolbus.	0.00	Bottom 10
94	US in the UN.	0.00	Bottom 10
94	Stratosphere ozone depletion.	0.00	Bottom 10
94	UN assessment.	0.00	Bottom 10

Congress	Hearing Description	Semantic Distance	Group
107	Decontamination of Anthrax and others Biological agents.	1.04	Top 10
93	Research on short-term weather phenomena.	1.00	Top 10
92	Status of current technology to identify seismic events as natural or man made.	1.00	Top 10
107	Examine use of anabolic steroids to improve performance in sports.	0.99	Top 10
95	Developments in cell biology and genetics.	0.99	Top 10
112	To examine the science behind the testing lab for human growth hormorne.	0.98	Top 10
110	To examine safety of phthalates and bisphenol-A.	0.97	Top 10
103	violent and demeaning imagery in popular music.	0.97	Top 10
96	Health effects of exposure to Agent Orange, other toxic pesticides.	0.96	Top 10
103	Review the status of experimental research into gene therapy.	0.96	Top 10
95	HUD housing and community development programs.	0.13	Bottom 10
100	US trade policies.	0.14	Bottom 10
92	Federal housing and urban development programs.	0.15	Bottom 10
98	Federal education programs.	0.15	Bottom 10
101	US exports and foreign trade.	0.16	Bottom 10
105	Examine issues relating to US trade policy.	0.16	Bottom 10
96	US trade policy.	0.16	Bottom 10
110	To examine U.S. trade policy issues and developments.	0.16	Bottom 10
92	Federal housing and community development programs.	0.17	Bottom 10
104	Examine various issues relating to U.S. trade policy.	0.17	Bottom 10

Validation Using Internet Policy Hearings

I assesses whether the measures of novelty behave in theoretically consistent ways within a policy domain whose emergence and institutionalization are well documented, I trace the measures of novelty across the evolution of congressional attention to Internet policy. Congressional references to computer networking and digital communication first appeared sporadically in the early 1980s, primarily in hearings concerning research infrastructure and defense communications. The Internet began to emerge as a distinct policy domain in the early 1990s, as the public introduction of the World Wide Web in 1991 transformed digital connectivity into a mass communication and commercial platform. As the internet became a more central feature of American life, the Internet transitioned from a technical innovation to a routine and enduring feature of the congressional agenda with discussions increasingly centered on issues of access, content regulation, encryption, and privacy.

To identify hearings relevant to Internet policy, I construct a dictionary of terms commonly used to describe Internet technologies and related issue ("internet", "world wide web", "online", "digital", "e-commerce", "electronic commerce", "internet privacy", "online privacy", "encryption", "digital signature", "internet service provider", "ISP", "net neutrality", "fiber optic", "firewall", "domain name", "broadband", "cyberspace"). All hearing descriptions containing one or more terms were classified as Internet-related. Next, I manually reviewed the identified hearing descriptions and excluded cases that did not align with substantive Internet issues. For example, hearings describing "digital audio recording standards" or "digital sound broadcasting technologies" were excluded because they pertained to copyright and media hardware rather than to Internet policy. In total, the process resulted in 625 hearings related to the internet. Figure A1 illustrates the expansion of congressional attention to the Internet, with the number of hearings devoted to Internet-related issues rising steadily as the topic became a regular component of the congressional agenda.

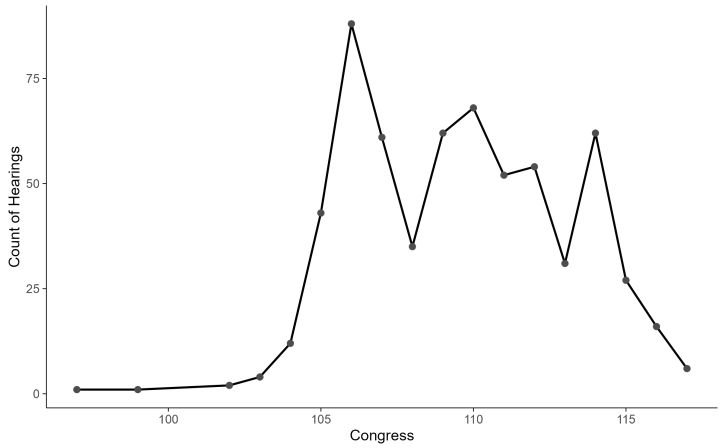


Figure A1: Congressional Attention to Internet Policy Across Time

Figure A2 plots the evolution of these three measures for Internet-related hearings over time. The first panel shows that the average number of new terms was initially high in the early 1990s, when congressional discourse surrounding the Internet was still forming. As hearings became more frequent and the policy domain stabilized, the introduction of new vocabulary declined sharply. The second panel presents the opposite pattern for lexical recurrence, as congressional hearings increasingly reused established terms,

reflecting the development of a shared policy lexicon. The third panel shows the trajectory of semantic distance. Although the average semantic distance between hearings declined, indicating greater conceptual coherence over time, the decline was modest compared to the lexical indicators. This pattern suggests that Internet policy remains conceptually distinct within the broader technology agenda, differing from other topics such as space exploration or weather monitoring even after it became an entrenched component of the congressional agenda.

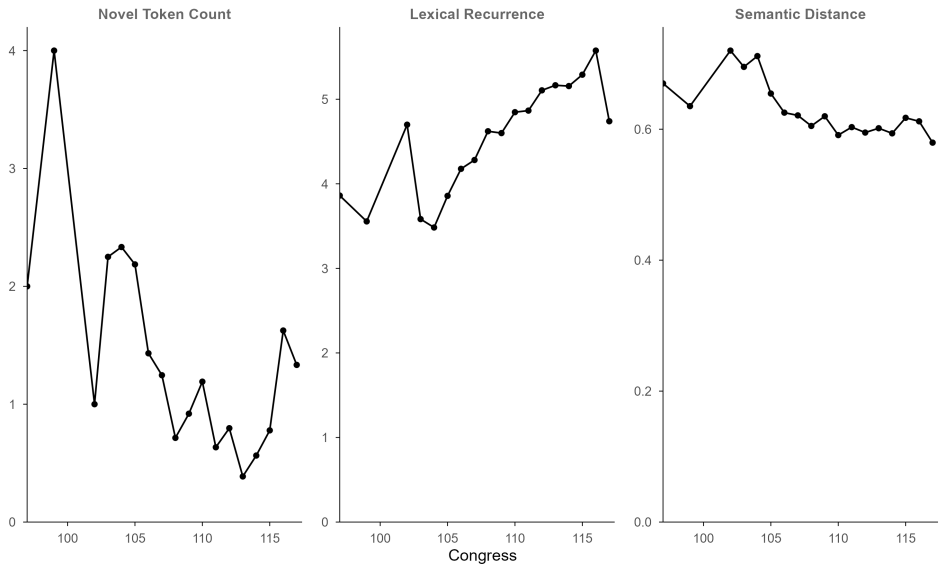


Figure A2: Evolution of Linguistic Novelty in Internet-Related Congressional Hearings

Together, these trends demonstrate that the linguistic novelty measures behave consistently with theoretical expectations about issue institutionalization. The Internet began as a highly novel policy concern characterized by new terminology and conceptual distinctiveness but gradually evolved into a stable and recurring domain of congressional activity.

Gathering Testimony Data of Alternative Expert Information Sources

Witness data for non-academic expert organizations were collected using the same procedure used to identify academic witnesses. I extracted witness affiliation and title information from ProQuest Congressional hearing records, which report each witness's institutional affiliation at the time of testimony. For each organizational category, I con-

structed a dictionary of organization names and commonly used acronyms and matched these dictionaries to the affiliation strings listed in the witness records.

Think tanks were identified using the Global Go-To Think Tanks Reports, which provides a comprehensive list of leading think tanks worldwide (McGann 2020). I included organizations that were headquartered in the United States and that appeared on at least one edition of the report. Using this list, I matched think tank names and acronyms to witness affiliations and then manually reviewed all matches to remove false positives, such as witnesses affiliated with similarly named private firms or advocacy organizations. This process identified 6,366 instances of think tank testimony.

Federal bureaucratic witnesses were identified by compiling a list of active and defunct federal executive agencies using information from usa.gov, the Federal Register, and Wikipedia. Agency names and acronyms were matched to witness affiliation fields, and all matches were manually reviewed to exclude non-governmental organizations. This procedure yielded 127,182 instances of testimony by bureaucrats from executive-branch and independent agencies.

Legislative support agencies were treated as a distinct category. I identified the Government Accountability Office, Congressional Research Service, Congressional Budget Office, and Office of Technology Assessment as the primary legislative agencies tasked with producing analytical expertise for Congress. Dictionary matching with manual validation identified 7,145 instances of legislative branch testimony.

Additional Models

Table A1: Negative Binomial Model Predicting Count of Academic Testimony at Every hearing (91st to 117th Congress)

	(1)	(2)	(3)	(4)	(5)	(6)
Novel Token Count	0.045***	0.035***				
	(0.009)	(0.010)				
Semantic Distance			1.287***	1.137***		
			(0.078)	(0.082)		
Lexical Recurrence					−0.149***	−0.125***
					(0.012)	(0.013)
Appropriations Hearing		−0.458***		−0.378***		−0.409***
		(0.046)		(0.046)		(0.046)
Legislative Hearing		−0.321***		−0.297***		−0.310***
		(0.025)		(0.025)		(0.025)
Oversight Hearing		−0.102***		−0.106***		−0.090***
		(0.027)		(0.027)		(0.027)
Democratic Control		0.166***		0.169***		0.169***
		(0.045)		(0.044)		(0.045)
Topic Salience		−1.798**		−2.062***		−1.567*
		(0.669)		(0.667)		(0.668)
Constant	−2.900***	−2.712***	−3.618***	−3.363***	−2.457***	−2.346***
	(0.104)	(0.108)	(0.126)	(0.131)	(0.129)	(0.132)
N	61 050	56 944	61 050	56 944	61 050	56 944

* p <0.05, ** p <0.01, *** p <0.005

Outcome is the count of academic witnesses; models include log(total witnesses) as an offset, so coefficients are interpretable as effects on the expected rate (or proportion) of academic witnesses among all witnesses.

Table A2: Quasi-Binomial Models Predicting the Proportion of Think Tank Testimony at Every hearing (91st to 117th Congress)

	(1)	(2)	(3)	(4)	(5)	(6)
Novel Token Count	−0.040	−0.048*				
	(0.025)	(0.022)				
Semantic Distance			−0.762**	−1.044***		
			(0.279)	(0.206)		
Lexical Recurrence					0.032	0.074
					(0.055)	(0.042)
Appropriations Hearing		−0.794***		−0.863***		−0.815***
		(0.157)		(0.161)		(0.160)
Legislative Hearing		−0.425***		−0.437***		−0.433***
		(0.128)		(0.128)		(0.130)
Oversight Hearing		−0.021		−0.027		−0.038
		(0.047)		(0.046)		(0.043)
Democratic Control		0.028		0.025		0.030
		(0.119)		(0.117)		(0.119)
Topic Salience		1.904		1.942		1.776
		(1.800)		(1.772)		(1.775)
N	60 787	56 689	60 787	56 689	60 787	56 689
Congress FE	X	X	X	X	X	X
Chamber FE	X	X	X	X	X	X
Issue Area FE	X	X	X	X	X	X

* p <0.05, ** p <0.01, *** p <0.005
Standard errors clustered by committee.

Table A3: Quasi-Binomial Models Predicting the Proportion of Bureaucratic Testimony at Each Hearing (91st–117th Congress)

	(1)	(2)	(3)	(4)	(5)	(6)
Novel Token Count	−0.073***	−0.046***				
	(0.016)	(0.015)				
Semantic Distance			−1.095***	−0.796***		
			(0.235)	(0.214)		
Lexical Recurrence					0.139***	0.082***
					(0.028)	(0.023)
Appropriations Hearing		0.993***		0.948***		0.970***
		(0.064)		(0.061)		(0.062)
Legislative Hearing		−0.439***		−0.453***		−0.447***
		(0.035)		(0.037)		(0.036)
Oversight Hearing		0.249***		0.253***		0.242***
		(0.033)		(0.033)		(0.032)
Democratic Control		−0.031		−0.033		−0.030
		(0.077)		(0.076)		(0.077)
Topic Salience		−1.817		−1.677		−1.938
		(0.964)		(0.978)		(1.002)
N	61 037	56 929	61 037	56 929	61 037	56 929
Congress FE	X	X	X	X	X	X
Chamber FE	X	X	X	X	X	X
Issue Area FE	X	X	X	X	X	X

* p <0.05, ** p <0.01, *** p <0.005
Standard errors clustered by committee.

Table A4: Quasi-Binomial Models Predicting the Proportion of Legislative Bureaucrat Testimony at Every hearing (91st to 117th Congress)

	(1)	(2)	(3)	(4)	(5)	(6)
Novel Token Count	−0.021	−0.022				
	(0.025)	(0.024)				
Semantic Distance			−0.334	−0.613**		
			(0.223)	(0.212)		
Lexical Recurrence					0.021	0.045
					(0.036)	(0.033)
Appropriations Hearing		−0.377***		−0.422***		−0.393***
		(0.121)		(0.121)		(0.119)
Legislative Hearing		−0.731***		−0.742***		−0.735***
		(0.064)		(0.063)		(0.064)
Oversight Hearing		0.239***		0.242***		0.235***
		(0.055)		(0.055)		(0.055)
Democratic Control		−0.010		−0.008		−0.008
		(0.135)		(0.133)		(0.135)
Topic Salience		−1.936		−1.859		−1.965
		(1.481)		(1.436)		(1.495)
N	60 770	56 754	60 770	56 754	60 770	56 754
Congress FE	X	X	X	X	X	X
Chamber FE	X	X	X	X	X	X
Issue Area FE	X	X	X	X	X	X

* p <0.05, ** p <0.01, *** p <0.005
Standard errors clustered by committee.