Development of a scale to measure perceived administrative burden, with broad applicability beyond direct policy clients

Kang, I.; Sievert, M.C.G.; Na, C.

Governance

WILEY

Governance

ORIGINAL ARTICLE OPEN ACCESS

Development of a Scale to Measure Perceived Administrative Burden, With Broad Applicability Beyond Direct Policy Clients

Inkyu Kang[1] | Martin Sievert[2] | Chongmin Na[3]

[1]Department of Public Administration and Policy, University of Georgia, Athens, Georgia, USA | [2]Institute of Public Administration, Leiden University, Den Haag, the Netherlands | [3]Graduate School of Public Administration, Seoul National University, Seoul, Korea

Correspondence: Chongmin Na (chongmin20@snu.ac.kr)

Received: 30 October 2024 | **Revised:** 22 April 2025 | **Accepted:** 29 April 2025

Funding: This work was supported by Seoul National University 10-10 Project.

Keywords: citizen-state interactions | perceived administrative burden | scale development and validation

ABSTRACT

Individuals form perceptions about the degree of administrative burdens associated with public agencies and programs, even without being (potential) policy clients themselves (e.g., a wealthy person's view of administrative burdens experienced by welfare program applicants) or without referring to a specific instance (e.g., perceptions of bureaucratic hurdles in the federal government overall). This study develops and validates a scale of perceived administrative burden that is applicable beyond direct policy clients to broad audiences and political stakeholders. The scale underwent psychometric evaluation using large-scale, representative citizen samples from the US and South Korea (n = 3000). Exploratory and confirmatory factor analyses on randomly split samples confirmed a unidimensional latent structure and the scale's construct validity. Cronbach's alpha and composite reliability indicated strong internal consistency, while a high average variance extracted, alongside additional correlational analyses, supported convergent and discriminant validity. We discuss the scale's potential applications in exploring how perceived administrative burden shapes policy processes, political behavior, and public administration.

1 | Introduction

The concept of administrative burden has garnered significant scholarly attention, serving as a valuable framework for understanding citizen-state interactions (Halling and Baekgaard 2024; Herd and Moynihan 2019; Moynihan et al. 2015). It also has important practical implications, as evidenced by recent high-profile policy changes in democratic nations, such as the Biden Administration's Executive Order 14058, which outlines initiatives to reduce administrative burdens.

Much of the growing body of research has focused on ways to reduce policy clients' experiences of administrative burdens (e.g., Compton et al. 2023; Herd et al. 2013; Kalucza and Sievert 2024). Yet any member of society can form perceptions about the extent of burdens imposed on policy clients, even if they are not clients themselves (e.g., an affluent individual's beliefs about the degree of burden on welfare program applicants) or have no direct reference to specific instances (e.g., beliefs about the extent of burden associated with the federal government in general). Furthermore, substantial variation in perceived burden can occur even among policy clients who undergo the same bureaucratic processes due to individual factors such as administrative literacy (e.g., Döring 2021). This suggests that perceived administrative burden should be studied and measured as a distinct concept—neither supplementary to nor less important than objective indicators of burdensome bureaucratic processes. Moreover, perceived burden among the

public and political stakeholders plays a pivotal role in various aspects of the policy process, political behavior, and public administration (Keiser and Miller 2020; Sievert and Bruder 2024; Stenderup and Pedersen 2024), with implications that extend beyond outcomes for policy clients at the implementation stage. For instance, the concept has the potential to advance scientific understanding of foundational issues in public administration research, such as bureaucratic reputation (e.g., Bustos 2021; Carpenter and Krause 2012) and citizen satisfaction (e.g., Jilke 2018; Nie and Wang 2023).

This study presents a scale for measuring perceived administrative burden, supported by a series of reliability and validity tests using a large-scale survey of representative citizen samples from the United States and South Korea (n = 3000). These tests included exploratory and confirmatory factor analyses on randomly split samples, average variance extracted, and correlational analyses using criterion variables, as well as Cronbach's alpha and composite reliability. Our results indicate that perceived administrative burden is a unidimensional construct, suggesting that individuals do not necessarily distinguish between specific cost categories (e.g., learning, compliance, and psychological cost) when perceiving burdens associated with public agencies and programs. We also observed notable cross-agency and cross-national variations. For instance, in both the US and South Korea, tax-related agencies—the Internal Revenue Service in the US and the National Tax Agency in South Korea—received the highest perceived burden ratings, while service-oriented agencies—the United States Postal Service in the US and the Community Service Center in South Korea—were rated the lowest among all randomly presented institutions in the survey. We believe this scale will facilitate research in relevant areas by providing a standardized and broadly applicable measurement tool that enables the accumulation of comparable, replicable data across diverse contexts and over time.

2 | Theoretical Relevance of Perceived Administrative Burden

A widely accepted definition of administrative burden is "an individual's experience of policy implementation as onerous" (Burden et al. 2012, 742). Perceived administrative burden, by contrast, refers to *an individual's assessment of the extent of administrative burdens imposed on clients of a public program or agency*. In our conceptualization, perceived administrative burden is not limited to policy clients who directly experience administrative demands (e.g., Porumbescu et al. 2024; Zhu et al. 2024). Instead, it applies to broad audiences and political stakeholders—primarily the mass public, which this study focuses on empirically, but also potentially policymakers and civil servants—whose perceptions of burden arise through mechanisms beyond firsthand experience. Perceptions of administrative burden are central to political debates, including those concerning the trade-offs of burdensome state actions and their justifications (Nisar and Masood 2023). For example, the Obama administration's Deferred Action for Childhood Arrivals (DACA) policy, introduced in 2012, allows young immigrants brought to the US as children to seek temporary protection from deportation. As of 2023, approximately 600,000 individuals are active DACA recipients, representing less than 0.2% of the population. Nevertheless, the burdensome state actions associated with DACA hold significant political salience for a broader audience beyond the program's direct beneficiaries (Ling et al. 2024), primarily due to its racialized policy feedback effects (Bell et al. 2024).

Even among direct policy clients who constitute a relatively small segment of the broader population, we argue that perceived administrative burden retains distinct conceptual value. One might suspect that for individuals with extensive firsthand experience navigating bureaucratic procedures (e.g., applying for welfare programs), there is little distinction between experienced and perceived burden. However, we maintain that these two constructs remain conceptually distinct, despite the potentially large correlation between the two due to learning effects—namely, experiences with burdensome state measures shape and inform perceptions of administrative burden. Moreover, we expect perceptions of burden to vary even among policy clients who experience the same bureaucratic process. For instance, individuals with high administrative literacy may view complex paperwork as relatively manageable. In contrast, those with low administrative literacy may find the same process overwhelming. In short, we argue that perceived administrative burden (e.g., perceptions of the number of paperwork required for program X) is a distinct concept and should not be treated as a supplementary or inferior measure of the objective aspects of burdensome bureaucratic processes (e.g., the number of paperwork required for program X).

If direct experience is not necessarily the primary source of perceptions of administrative burden, what other factors account for their formation? While a detailed examination of the specific determinants or mechanisms of perceived burden is beyond the scope of this study, we speculate that a range of environmental and individual-level factors—such as vicarious experiences through family and friends, media exposure, and preexisting beliefs about the state or a policy (Tomlinson et al. 2024)—can affect how individuals perceive administrative burdens. Perceptions of burden among the public have implications for attitudes toward government and state decision-making (Keiser and Miller 2020), ultimately influencing political behaviors such as engagement and activism. For example, the US Public Service Loan Forgiveness (PSLF) program, established in 2007, aims to alleviate student debt for individuals working in the public sector by forgiving remaining federal student loan balances after 120 qualifying monthly payments. However, the program has been widely criticized for its complex requirements and high rejection rates (Herd and Moynihan 2024). To put this into perspective, in the first three years of eligibility, approximately 227,000 borrowers applied for relief. Still, the Department of Education approved forgiveness for fewer than 4000 applicants—resulting in a denial rate of roughly 98% (Wu 2021). The program's complex paperwork and high rejection rates triggered widespread frustration, sparking activism that extended beyond those directly affected. Frustrated borrowers mobilized others through social media, petitions, and advocacy groups, pushing for policy reforms that ultimately prompted government efforts to simplify the process.

Perceptions of administrative burden may also play a significant role in shaping policy feedback effects. Existing studies have primarily focused on how administrative burdens of policy clients undermine their trust in government and influence civic dispositions. For example, Bell et al. (2024) argue that administrative burden leads to negative orientations toward the state, reflected in emotional responses (i.e., interpretive effects) and reduced access to benefits (i.e., resource effects). They tested this argument using survey data from applicants to Oklahoma's Promise, a means-tested scholarship program. While their findings offer valuable insights, policy feedback effects can extend beyond program participants and unfold on a much broader scale (Wang and Zhang 2024). For instance, individuals not directly involved with Oklahoma's Promise program may form perceptions about administrative burdens of its beneficiaries—through stories from acquaintances, media exposure, online information, or inferences based on prior beliefs about the agency, program, or target population. These perceptions—regardless of whether they are factually accurate—can shape public support for the program, attitudes toward the implementing agency, and broader trust in government.

Furthermore, perceived administrative burden is deeply connected to foundational concepts in public administration research, such as bureaucratic reputation (Lee and Van Ryzin 2019; Zhang et al. 2024), perceived performance (van den Bekerom et al. 2021), trust in government (Grimmelikhuijsen and Meijer 2014; Peeters and Dussauge Laguna 2021), citizen satisfaction (Jilke 2018; Nie and Wang 2023), or blame attribution (Moynihan 2012; Wei et al. 2024). Measuring perceived administrative burden as a distinct concept allows scholars to enrich existing perspectives in novel ways. For instance, perceived administrative burden may be a critical determinant of bureaucratic reputation—a theoretical connection that prior studies have often referenced implicitly. One illustration comes from Lee (2022)'s vignette experiment on bureaucratic reputation, conducted with a representative sample of US adults. One hypothetical scenario featured a high-profile scandal at the Department of Veterans Affairs (VA), triggered by long wait times for veterans' medical appointments: "After Scandal, VA still Troubled by Long Waits for Medical Care ... Years after long waits for medical appointments for veterans exploded into a nationwide scandal, ..." (Lee 2022, 1136). The underlying assumption of this vignette design—which appears reasonable—is that, even though most US residents are not veterans and have no direct stake in the VA, their perceptions of the burdens experienced by veterans can significantly shape their judgments of the agency's reputation. Since administrative burden reflects everyday interactions between the public and the state (Martin et al. 2024), perceptions of administrative burden may heavily influence the reputation of government agencies—sometimes more so than the agencies' technical capacity or task performance.

3 | Scale Development

3.1 | Conceptual Comparison

It is worth outlining the conceptual differences between perceived administrative burden and other burden-related constructs, such as experienced burden (Jilke et al. 2024), anticipated burdens (Baekgaard and Madsen 2024), burden tolerance (Aarøe et al. 2021; Baekgaard et al. 2025; Halling et al. 2023), and burden legitimacy.

As previously explained, perceived administrative burden differs from experienced burden, which focuses exclusively on direct policy clients who navigate bureaucratic procedures (Halling and Baekgaard 2024). Jilke et al.'s (2024) recent scale contains items explicitly designed to measure prior personal experiences of burden. Perceived burden also differs from anticipated burden, which refers to a policy client's self-oriented projection of future burdens associated with a program or agency (Baekgaard and Madsen 2024). While there may be substantial overlaps between perceived and anticipated burdens for individuals with a realistic chance of future interaction with a program or agency, anticipated burden does not apply to the majority of individuals who have never been a client themselves and are unlikely to become one in the future. In such cases, perceived burden becomes a more relevant consideration. Furthermore, perceived administrative burden is a value-neutral assessment of the extent of burdens, in contrast to burden tolerance (referring to an individual's willingness to accept administrative burdens) or burden legitimacy (referring to the belief that burdens are justified for reasonable causes) (Aarøe et al. 2021; Halling et al. 2023) which are inherently value-laden.

3.2 | Item Generation

To generate an initial pool of survey items, we incorporated feedback from diverse experts who have published in this area. This step was essential for establishing content validity, defined as "the degree to which elements of an assessment instrument are relevant to and representative of the targeted construct for a particular assessment purpose" (Almanasreh et al. 2019, 214). We sought to include experts with varied primary research agendas, disciplinary backgrounds, and institutional affiliations. Specifically, we contacted experts in the following areas: administrative burden and citizen-state interactions, survey design and psychometrics, and public administration more broadly. In total, we contacted seven experts sequentially. We received unstructured, open-ended comments—exceeding the conventional recommendation of a minimum of three experts (Lynn 1986). We continuously went back and forth between existing research, expert feedback, and our survey items in order to refine the item wordings incrementally, which does not fit neatly into distinct, sequential stages.

One major revision prompted by expert feedback involved fundamentally reconsidering the target construct of our scale, shifting from "bureaucratic reputation for administrative burden" to "perceived administrative burden." A bureaucratic reputation scholar emphasized that reputation reflects a shared belief among multiple audiences (Carpenter and Krause 2012). Therefore, an individual-level survey alone cannot fully capture reputation unless follow-up analyses identify common patterns at a group (i.e., audience) level, which we agreed with. Other experts agreed and suggested that our scale essentially measures perceived administrative burden, although such perceptions

may have a bidirectional relationship with bureaucratic reputation. Multiple experts also noted that perceptions of administrative burden arise not only regarding public agencies but also regarding public programs. They mentioned that our scale should measure both, although we can focus on agency-related perceived burden for the purpose of psychometric validity testing in the empirical analysis. Additionally, several experts advised focusing on widely recognized instances of burden and using clear, universally understandable language while avoiding technical jargon—ensuring that the scale is broadly applicable, particularly to individuals with low literacy levels or no direct personal experience with bureaucratic processes.

Our scale encompasses the three widely recognized categories of administrative burden—learning costs, compliance costs, and psychological costs (Moynihan et al. 2015)—with a refined conceptualization of each cost category. We acknowledge that some scholars have noted the difficulty of disentangling these categories (e.g., Baekgaard and Tankink 2022; Tomlinson et al. 2024), while others have proposed complementary or alternative classifications (e.g., Daigneault 2024; Nisar and Masood 2023). Against this backdrop, we refine the original cost categories to address potential ambiguities and confusion while retaining their original terminology. Specifically, we argue that learning costs, compliance costs, and psychological costs can be more clearly understood as cognitive burdens, resource burdens, and emotional burdens, respectively—each representing a distinct yet interrelated domain of administrative burdens.

First, perceived learning cost refers to an individual's assessment of the *cognitive burdens* that policy clients face when trying to obtain necessary information—such as vague or confusing terms, obligations, and procedures in an application process (Barnes and Riel 2022). For the items representing perceived learning cost, we identified three key aspects of cognitive challenge: the amount of information, its complexity, and the lack of guidance. When individuals believe that clients of a public agency or program must process excessive or overly complex information without adequate guidance, they will rate the agency or program high on perceived learning costs.

Second, perceived compliance costs reflect an individual's assessment of the *resource burdens* associated with procedural tasks or requirements imposed on policy clients. For items representing perceived compliance costs, we identified time (e.g., long waits or travel to an office), money (e.g., application fees), and procedural labor (primarily associated with documentation and checklists) as the most commonly observed and studied resource burdens in bureaucratic encounters. While some agencies or policies may not charge fees or require paperwork at all, this can be reflected in the survey items through significantly low ratings. Furthermore, the disparity between perceived and objective resource burdens in bureaucratic processes does not diminish the importance of measuring perceived administrative burden; rather, it highlights its unique value as a distinct concept.

Third, perceived psychological cost refers to an individual's assessment of the *emotional burdens* that policy clients bear during their interaction with the state. To measure perceived psychological cost, we included indicators representing three aspects: feelings of discomfort, poor treatment during the process, and stress. We intentionally chose the expressions "feeling uncomfortable," "stress," and "being treated badly" over more specific terms such as trauma, stigma, or loss of autonomy for two key reasons. First, expert feedback suggested that terms such as trauma, stigma, or loss of autonomy might be ambiguous or overly technical, particularly for individuals without personal experience of these psychological consequences or those with lower literacy levels. Second, we deliberately avoided terms tied to specific emotional consequences to account for the possibility that important aspects of psychological costs have yet to be fully recognized in the literature. Baekgaard and Tankink (2022) argue that psychological costs can encompass various negative experiences not yet sufficiently studied, such as the uncertainty of not knowing whether one will receive benefits in the future. By using broader language, our scale remains adaptable and capable of capturing a wide range of emotional burdens that future research may further uncover.

Beyond the three types of costs, we introduce perceived generic burden as an additional domain. Perceived generic burden refers to an individual's assessment of the overall extent of administrative burden that does not map neatly or exclusively onto a specific cost category, which has been alluded to in prior research (e.g., Keiser and Miller 2020). We were concerned that survey items tailored exclusively to the three specific cost categories might not fully capture individuals' general disposition toward an agency or program with regard to administrative burdens—dispositions that tend to be shaped more by affective impressions than by deliberate cognitive reasoning (for instance, see Daigneault 2024). In practice, people often express perceptions of administrative burden using general language—for example, saying "the federal government is a headache"—without consciously classifying their judgment into one of the three cost categories, which are derived from scholarly reasoning. This tendency can be even more pronounced when individuals have never directly interacted with the program or agency and possess limited concrete knowledge about it. Comparable scale development efforts in public administration have acknowledged a similar issue. For instance, Lee and Van Ryzin (2019) included "general reputation" in their bureaucratic reputation scale—in addition to the four sub-dimensions of bureaucratic reputation identified in the existing literature.

Based on these considerations, we narrowed the item pool to 15 items representing the four content domains of perceived administrative burden: perceived learning cost (cognitive burden), perceived compliance cost (resource burden), perceived psychological cost (emotional burden), and perceived generic burden (see Table 1). Our initial item pool was intentionally kept concise and parsimonious to minimize survey fatigue, which can introduce measurement error and compromise data quality. The goal was to avoid including construct-irrelevant indicators while ensuring adequate content coverage—an important balance in the item generation stage (Almanasreh et al. 2019). In line with comparable scale development studies in the field (e.g., Jilke et al. 2024), we prioritized identifying items that are broadly applicable and adaptable across contexts. We also intentionally avoided using the term "client" in the survey items, despite its common usage in scholarly literature. Based on expert feedback,

TABLE 1 | Initial 15-item pool of perceived administrative burden scale.

Content domain	Item	*This agency (program)...*
Perceived generic burden	Item 1	puts many obstacles in people's way.
	Item 2	is burdensome for people to deal with.
	Item 3	is hassle-free for people to engage (R).
Perceived learning cost (perceived cognitive burden)	Item 4	lacks information for people to access services.
	Item 5	makes things difficult for people to understand.
	Item 6	provides people with too little guidance.
	Item 7	helps people learn about their services (R).
Perceived compliance cost (perceived resource burden)	Item 8	asks people for lots of paperwork.
	Item 9	takes up people's time.
	Item 10	charges people lots of fees.
	Item 11	reduces the work required for people (R).
Perceived psychological cost (perceived emotional burden)	Item 12	makes people feel uncomfortable.
	Item 13	treats people badly.
	Item 14	causes a lot of stress to people.
	Item 15	makes people feel respected (R).

Note: (R) = Reverse-worded items.

we recognized that the business-like connotation of "client" could systematically bias responses, potentially inflating expectations about service accessibility and convenience. Instead, we opted for "people" as a more neutral term, with the provision that the survey introduction would clearly specify that "people" refers to policy clients interacting with the agency or program. Finally, to address potential acquiescence or agreement bias, we included four reverse-worded items—one for each content domain—in our initial pool of 15 items (DeVellis 2003).

As previously discussed, an important caveat is that the content domains represented by each item are conceptually distinct yet may be empirically interrelated, consistent with Baekgaard and Tankink (2022) observation that "the three costs may not be possible to separate empirically in many instances." This applies to the items listed in Table 1. For example, in the case of a complicated application process, perceived learning costs related to cognitive challenges (e.g., item 5: "this agency makes things difficult for people to understand") may also contribute to perceived compliance costs associated with resource burden (e.g., item 9: "this agency takes up people's time") or perceived psychological costs associated with emotional burden (e.g., item 14: "this agency causes a lot of stress to people").

3.3 | Survey Implementation

Our initial 15-item pool was included in a large-scale, cross-national survey conducted in the US and South Korea through Tillion and Dynata, two major global survey firms. The survey was administered online between November 14 and December 13, 2022. The theoretical population included adults between the ages of 20 and 69 residing in major metropolitan areas: Los Angeles and New York City in the US, and Seoul, Incheon, Daejeon, Daegu, Busan, and Gwangju in South Korea. We used stratified sampling based on age, gender, and region to enhance the representativeness of the samples. The survey collected 1500 responses from each country (n = 3000). Originally written in English, the survey was translated into Korean for the South Korean sample through an iterative refinement process. After the initial translation, members of the research team—including the authors and research assistants proficient in both languages—continually reviewed and revised the translation to ensure that the Korean version accurately captured the intended meaning of the English items.

It is important to clarify that the concept of perceived administrative burden applies to the general public and various political stakeholders, including policymakers and civil servants. However, in our empirical analysis, we specifically focused on the general public, given its direct relevance to the key aspects of citizen-state interactions outlined in the theory section—such as mass orientations toward government, policy feedback, citizen satisfaction, and trust in government. In addition, the scale applies to perceived administrative burden related to public agencies as well as specific programs. Program-related perceptions occur when individuals focus on specific policies (e.g., Item 8, "Policy X asks people for a lot of paperwork"). In contrast, agency-related perceptions reflect generalized beliefs about an entire organization (e.g., Item 8, "Agency X asks people for a lot of paperwork"). We suggest that these two aspects of perceptions are often deeply interconnected. For instance, when individuals hear about burdensome procedures in a specific program (e.g., long wait times for unemployment benefits), they may generalize this to the entire agency administering it (e.g., "The Department of Labor is slow and inefficient"). By contrast, a perception at the agency level (e.g., "DMV is a bureaucratic nightmare") may result in the assumption that all programs under its umbrella are burdensome. However, our empirical analysis focused on agency-related perceptions primarily due to their greater theoretical significance. Variations in perceptions of burden are expected to be more pronounced at

the agency level, as public agencies often carry out a wide range of disparate—even seemingly unrelated—tasks despite being misleadingly perceived by broad audiences as more homogeneous than they actually are (Carpenter and Krause 2012). We were also concerned that anchoring the survey to a narrow policy case might result in limited generalizability as well as respondent knowledge. While this is not a flaw of the scale itself, it highlights that the interpretation of the scale validation tests warrants caution.

In the survey, respondents were first presented with a brief introduction: "The [Name of agency] is a government agency that provides [description of service] to people. Please indicate the degree to which you agree or disagree with each of the following questions about [Name of agency]." This introduction not only provided respondents with basic information about the agency and its services but also clarified that the term "people" in the survey items specifically referred to policy clients who interact with the agency and its services. The 15 survey items were presented in a randomized order within a single block, without any indication of the content domain each item was intended to represent. In Table 2, the [name of agency] and [description of service] tabs were filled randomly. In the [name of agency] tab, one of four agencies was randomly assigned from the US and South Korea, respectively (Table 3). The US survey participants could receive one of the following agencies: Social Security Administration (SSA), Internal Revenue Service (IRS), United States Postal Service (USPS), or the Federal Government. South Korean participants could receive one of the following: the National Police (NP), National Tax Service (NTS), Community Service Center (CSC), or the Central Government. By diversifying the types of institutions through random assignment, we reduced the risk of our analysis being skewed toward any particular agency.

IRS in the US is directly comparable to the NTS in South Korea, which is highly relevant for the concept of administrative burden due to the cumbersome nature of tax services. The SSA and the USPS in the US and the NP and the CSC in South Korea are not directly comparable, but they were included in the survey given their high visibility and broad clientele. Lastly, the US federal government and the central government in South Korea were deliberately chosen to examine the applicability of our measurement scale to federal/national, state/provincial, and local governments in general. As previously discussed, perceptions about burdens often arise regarding a generic institution, such as the federal or central government, without direct reference to a particular occurrence of burdens. This is similar to how other perceptual or attitudinal variables in public administration research, such as citizen satisfaction, are measured with regard to federal/national, state/provincial, and local governments rather than a particular service being delivered.

Responses to each item were recorded on a five-point Likert scale, with one indicating "strongly disagree" and five indicating "strongly agree." Those who selected "I don't know" were treated as missing values and coded as 9. The percentage of missing values in each survey item was small, ranging from 7.53% to 9.73% in the US samples and from 2.93% to 5.33% in the South Korean samples (refer to Table 4 for more details). We noticed that missing values were dispersed across respondents rather than concentrated among a few. In this situation, listwise deletion could eliminate significant portions of valuable data with no evidence that the missing values occurred randomly. As such, we employed multiple imputations, generating three plausible replacements for the missing data based on the relationships among variables in the dataset (Horton and Lipsitz 2001; Rubin 1996). As a result, the following analytical steps were based on three imputed datasets, each containing 1500 observations for the US and South Korean cases, respectively.

TABLE 2 | Survey introduction and items (presented in random order).

Introduction		Questions
		This agency...
[Name of agency] is a government agency that provides [description of service] to people. Please indicate the degree to which you agree or disagree with each of the following questions about [name of agency].	Q1	puts many obstacles in people's way.
	Q2	is burdensome for people to deal with.
	Q3	is hassle-free for people to engage.
	Q4	lacks information for people to access services.
	Q5	makes things difficult for people to understand.
	Q6	provides people with too little guidance.
	Q7	helps people learn about their services.
	Q8	asks people for lots of paperwork.
	Q9	takes up people's time.
	Q10	charges people lots of fees.
	Q11	reduces the work required for people.
	Q12	makes people feel uncomfortable.
	Q13	treats people badly.
	Q14	causes a lot of stress to people.
	Q15	makes people feel respected.

TABLE 3 | Randomized items in the introduction.

Country	Name of Agency (Description of Service)
The US	Social Security Administration (social insurance such as retirement benefits and disability income)
	Internal Revenue Service (tax-related services such as tax filing and tax refunds)
	United States Postal Service (postal services such as delivery of mail and packages)
	US Federal Government (a wide range of public services)
South Korea	National Police (services related to public safety and law enforcement, such as crime accusations or traffic accident investigation)
	National Tax Services (tax-related services such as tax filing and tax refunds)
	Community Service Center (general administrative services such as government certificates)
	Central Government (a wide range of public services)

TABLE 4 | Missing value distribution.

		Missing values (%)	
	This agency...	US samples (n = 1500)	Korean samples (n = 1500)
Item 1	puts many obstacles in people's way.	114 (7.60%)	61 (4.07%)
Item 2	is burdensome for people to deal with.	110 (7.33%)	50 (3.33%)
Item 3	is hassle-free for people to engage (R).	138 (9.2%)	70 (4.67%)
Item 4	lacks information for people to access services.	113 (7.53%)	57 (3.80%)
Item 5	makes things difficult for people to understand.	118 (7.87%)	61 (4.07%)
Item 6	provides people with too little guidance.	113 (7.53%)	55 (3.67%)
Item 7	helps people learn about their services (R).	129 (8.60%)	60 (4.00%)
Item 8	asks people for lots of paperwork.	117 (7.80%)	57 (3.80%)
Item 9	takes up people's time.	114 (7.60%)	51 (3.40%)
Item 10	charges people lots of fees.	128 (8.53%)	80 (5.33%)
Item 11	reduces the work required for people (R).	116 (7.73%)	55 (3.67%)
Item 12	makes people feel uncomfortable.	112 (7.47%)	53 (3.53%)
Item 13	treats people badly.	116 (7.73%)	55 (3.67%)
Item 14	causes a lot of stress to people.	112 (7.47%)	44 (2.93%)
Item 15	makes people feel respected (R).	146 (9.73%)	59 (3.93%)

We verified that the bivariate correlation coefficients among the 15 items were relatively modest, and the standard deviation of each item was sufficiently large, ranging from 0.95 to 1.38 on a five-point scale (Table 5 and Table 6). These confirm decent variation in the responses to different items. Furthermore, the average response time was approximately 6.5 min, which is not short given our survey was relatively brief. This descriptive information helps mitigate concerns about speeding or inattentiveness.

4 | Scale Validation

4.1 | Exploratory Factor Analysis

Given the lack of a solid theoretical foundation for the latent structure of perceived administrative burden, we began with an exploratory factor analysis (EFA), using principal component factoring with orthogonal rotation (DeVellis 2003). For each of the three multiply imputed datasets from the US and South Korea, we randomly split the samples, using the first half for EFA and reserving the second half for confirmatory factor analysis (CFA)—a common practice in scale development research (e.g., Lee and Van Ryzin 2019). Following the Kaiser-Guttman rule (Kaiser 1960), we retained the first two factors with eigenvalues greater than 1.

According to the EFA results, all reverse-worded items (i.e., items 3, 7, 11, and 15) loaded onto a separate factor (Factor 2) in both the US and South Korean samples. In contrast, the positively worded items had factor loadings on Factor 2 that were mostly close to zero. First, this pattern may be explained by the possibility that respondents processed the reverse-worded items differently from the positively worded ones. Research in

TABLE 5 | Bivariate correlation across items in the initial pool (US samples).

		Correlation													
Item #	**SD**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
1	1.26	.													
2	1.25	0.72	.												
3	1.20	0.07	0.02	.											
4	1.26	0.67	0.64	0.17	.										
5	1.27	0.71	0.71	0.06	0.67	.									
6	1.24	0.65	0.63	0.15	0.66	0.66	.								
7	1.11	0.09	0.10	0.61	0.13	0.09	0.14	.							
8	1.19	0.47	0.50	0.11	0.44	0.49	0.45	0.17	.						
9	1.20	0.56	0.57	0.05	0.52	0.52	0.49	0.10	0.62	.					
10	1.24	0.51	0.50	0.15	0.49	0.48	0.49	0.10	0.50	0.54	.				
11	1.20	0.01	−0.02	0.58	0.08	0.03	0.07	0.53	0.14	0.04	0.16	.			
12	1.26	0.60	0.59	0.03	0.55	0.58	0.51	0.04	0.59	0.63	0.58	0.05	.		
13	1.28	0.62	0.59	0.06	0.58	0.59	0.55	0.05	0.53	0.61	0.56	0.09	0.75	.	
14	1.25	0.60	0.61	0.03	0.53	0.57	0.51	0.04	0.62	0.67	0.57	−0.02	0.71	0.66	.
15	1.22	0.15	0.10	0.46	0.19	0.09	0.19	0.47	0.19	0.19	0.21	0.59	0.16	0.19	0.10

Note: This table presents the figures averaging across the three imputed datasets.

TABLE 6 | Bivariate correlation across items in the initial pool (South Korean samples).

		Correlation													
Item #	**SD**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
1	1.13	.													
2	1.09	0.59	.												
3	1.02	0.19	0.02	.											
4	1.06	0.60	0.60	0.12	.										
5	1.09	0.66	0.67	0.10	0.62	.									
6	1.11	0.67	0.63	0.14	0.66	0.71	.								
7	0.95	0.00	−0.12	0.40	−0.10	−0.14	−0.15	.							
8	1.00	0.49	0.52	0.02	0.48	0.55	0.51	−0.11	.						
9	1.04	0.57	0.60	0.04	0.55	0.61	0.58	−0.10	0.63	.					
10	1.06	0.55	0.44	0.19	0.45	0.50	0.50	0.00	0.53	0.56	.				
11	0.99	0.01	−0.12	0.38	−0.06	−0.10	−0.10	0.51	−0.08	−0.11	0.06	.			
12	1.11	0.62	0.63	0.06	0.58	0.63	0.65	−0.14	0.60	0.70	0.58	−0.11	.		
13	1.09	0.63	0.61	0.11	0.60	0.62	0.67	−0.14	0.57	0.66	0.57	−0.13	0.72	.	
14	1.11	0.62	0.66	0.04	0.59	0.66	0.66	−0.14	0.61	0.71	0.54	−0.13	0.74	0.72	.
15	1.02	0.13	−0.07	0.39	0.00	−0.03	−0.04	0.54	−0.01	−0.02	0.17	0.59	−0.03	−0.04	−0.06

Note: This table presents the figures averaging across the three imputed datasets.

psychology on cognitive and perceptual asymmetries supports this interpretation (e.g., Baumeister et al. 2001). It is also possible that perceiving the *absence* of burdens reflects a conceptually distinct process from perceiving their presence rather than simply representing opposite ends of a single continuum. Another possible explanation is respondent inattentiveness or speeding; however, we believe this is unlikely, given the sufficient variation in responses across items—as indicated by modest correlation coefficients and large standard deviations (shown in Table 5 and Table 6)—and the reasonable average response time in our survey, as previously denoted. Based on these considerations, we excluded the reverse-worded items from the scale in subsequent analyses.

In both the US and South Korean samples, the positively worded items predominantly loaded onto Factor 1, which accounted for approximately 46%–48% of the total variance. These results suggest that perceived administrative burden is a

unidimensional construct, with the four content domains contributing directly to the overall concept without intermediary layers of latent structure. To reiterate, all survey items were presented in a fully randomized order within a single block, with no indication of their associated content domains—strengthening the validity of the results by minimizing the potential for response bias or domain priming.

Two central insights from the EFA results merit emphasis. First, individuals do not necessarily distinguish between the cost categories—learning, compliance, and psychological—when forming perceptions of administrative burden. While these categories serve important analytical purposes, they do not appear to represent a clear conceptual framework in people's minds. This suggests that research on administrative burden should avoid conflating these analytical distinctions with the ways individuals perceive, interpret, and assess administrative burdens. Second, all 12 positively worded items in the initial pool exhibited strong factor loadings and were retained in the final scale. This strong performance may reflect the conceptual clarity and parsimony emphasized during the item development process. Indeed, content validity grounded in literature reviews and expert input should not be seen as less rigorous than subsequent validity assessments based on numerical factor loadings (Almanasreh et al. 2019).

Tables 7 and 8 present the EFA results, showing each item's loadings on the first two dominant factors for both samples.

4.2 | Confirmatory Factor Analysis

The model fit indices from CFA allow us to confirm the validity of the scale's latent structure (Worthington and Whittaker 2006). For the CFA, we returned to the reserved split-half samples. Prior to analysis, we assessed multivariate normality using the Doornik–Hansen test. The results indicated that both the US and South Korean datasets violated the assumption of multivariate normality. The p-values below 0.05 led us to reject the null hypothesis, suggesting that at least one item in each dataset did not follow a normal distribution. In response, we employed Robust Maximum Likelihood (MLR) as the estimation method to relax the normality assumption. The pooled fit statistics across the three imputed datasets were summarized using Rubin's rules (Rubin 1987), with the resulting estimates and standard errors presented in Table 9.

Our analysis obtained four model fit indices that are widely used in the literature: Chi-square (χ^2), root mean square error of approximation (RMSEA), comparative fit index (CFI), and Tucker-Lewis Index (TLI). Smaller χ^2 indicates a superior fit, and significant χ^2 indicates that the null hypothesis of perfect fit is rejected. However, aspiring to reject the null hypothesis of perfect fit tends to be highly unrealistic and hinges on the multivariate normality assumption, a criterion that should be cautiously considered as it is not the sole indicator of poor model fit (Hu and Bentler 1998; MacCallum and Austin 2000): "the null hypothesis of exact overall fit tested by the χ^2 test is often not of general interest" (West et al. 2012, 211). As such, research has suggested referring to additional indices such as RMSEA, CFI, or TLI to evaluate the model fit (Hu and Bentler 1999). Although the threshold for an acceptable fit may be context-dependent, we followed the cut-offs that are generally used in the literature, which are as follows: 0.10 or lower of RMSEA, 0.90 or greater of CFI, and 0.90 or greater of TLI.

The survey items within each cost category may share variance that is not fully accounted for by the latent factor. For example, items representing the same cost category are similarly phrased, which may introduce method effects (Quilty et al. 2006). Accordingly, we allowed the error terms within each cost category to covary. The results indicated that the 11-item, unidimensional scale of perceived administrative burden performed well, as evidenced by acceptable model fit in both the US and South Korean samples. In the US sample, the model fit indices were as follows: RMSEA = 0.10, CFI = 0.96, and TLI = 0.93. In the South Korean sample, the model showed an even stronger fit: RMSEA = 0.07, CFI = 0.98, and TLI = 0.97.

As a robustness check, we estimated a two-level structure where items loaded onto their respective subdimensions, which in turn loaded onto a higher-order construct of perceived administrative burden. The coefficients linking the four subdimensions to the higher-order construct were approximately 0.50 or lower, suggesting a weak predictive relationship. This finding further supports our conclusion that perceived administrative burden is best modeled as a unidimensional construct. Figure 1 displays the standardized factor loadings for both the US and South Korean samples.

To further verify construct validity, we conducted measurement invariance testing across agencies and countries. Specifically, we examined the results of the score test (also known as the Lagrange Multiplier test) to assess whether relaxing certain constraints would significantly improve model fit. Our findings support partial metric invariance. For cross-agency comparisons, only one of the 11 scale items failed the test in the US (item 12, $p = 0.02$) and South Korea (item 4, $p = 0.00$), respectively. The follow-up likelihood ratio (LR) test comparing the configural model to the metric invariance model suggested a p-value of 0.29. This suggests that respondents in each country understood the relationships among constructs across the four public institutions randomly presented in the survey in a similar way. In contrast, we found less support for metric invariance in cross-national comparisons between the US and South Korea. Three of the 11 scale items failed the score test (item 2: $p = 0.00$, item 4: $p = 0.00$, item 6: $p = 0.00$), and the follow-up likelihood ratio (LR) test comparing the configural model to the metric invariance model suggested a p-value of 0.00. This confirms substantive differences in how respondents associate the latent construct with its indicators across the two national contexts. This could be due to broader societal and cultural differences between the two countries in terms of public perceptions of administrative burdens.

4.3 | Convergent and Discriminant Validity

Next, we assessed the scale's construct validity by examining both convergent and discriminant validity (DeVellis 2003).

TABLE 7 | Exploratory factor analysis (US samples).

		Factor 1			Factor 2		
Item	**Mean (SD)**	**Data1**	**Data2**	**Data3**	**Data1**	**Data2**	**Data3**
Item 1	3.29 (0.05)	**0.84**	**0.83**	**0.83**	0.01	0.00	−0.01
Item 2	3.40 (0.05)	**0.83**	**0.83**	**0.83**	−0.01	−0.03	−0.01
Item 3	3.50 (0.04)	0.04	0.03	0.04	**0.82**	**0.83**	**0.82**
Item 4	3.24 (0.05)	**0.78**	**0.78**	**0.78**	0.15	0.12	0.14
Item 5	3.27 (0.05)	**0.81**	**0.82**	**0.81**	0.02	0.00	0.02
Item 6	3.34 (0.05)	**0.75**	**0.75**	**0.74**	0.15	0.12	0.16
Item 7	3.60 (0.04)	0.06	0.08	0.07	**0.82**	**0.81**	**0.82**
Item 8	3.54 (0.05)	**0.72**	**0.72**	**0.72**	0.20	0.21	0.16
Item 9	3.43 (0.05)	**0.78**	**0.78**	**0.78**	0.13	0.12	0.11
Item 10	3.37 (0.05)	**0.71**	**0.72**	**0.72**	0.18	0.19	0.17
Item 11	3.43 (0.04)	0.03	0.04	0.02	**0.86**	**0.86**	**0.85**
Item 12	3.26 (0.05)	**0.83**	**0.82**	**0.82**	0.06	0.08	0.05
Item 13	3.09 (0.05)	**0.81**	**0.81**	**0.82**	0.09	0.13	0.10
Item 14	3.48 (0.05)	**0.81**	**0.81**	**0.82**	0.01	0.01	0.00
Item 15	3.31 (0.04)	0.17	0.14	0.16	**0.76**	**0.76**	**0.75**
Percentage of variance		46%	46%	46%	19%	19%	18%

Note: $N = 750$ (first half of random-split sample). Data = Imputed dataset. Minimum residual (minres) in R with varimax rotation was performed. Bolded items are selected for the perceived administrative burden (PAB) scale.

TABLE 8 | Exploratory factor analysis (South Korean samples).

		Factor 1			Factor 2		
Item	**Mean (SD)**	**Data1**	**Data2**	**Data3**	**Data1**	**Data2**	**Data3**
Item 1	2.97 (0.04)	**0.81**	**0.80**	**0.81**	0.15	0.15	0.15
Item 2	3.42 (0.04)	**0.79**	**0.79**	**0.80**	−0.10	−0.10	−0.09
Item 3	3.26 (0.04)	0.18	0.16	0.18	**0.68**	**0.68**	**0.68**
Item 4	3.32 (0.04)	**0.78**	**0.78**	**0.78**	−0.03	−0.04	−0.03
Item 5	3.29 (0.04)	**0.82**	**0.82**	**0.82**	−0.06	−0.07	−0.05
Item 6	3.20 (0.04)	**0.81**	**0.82**	**0.81**	−0.08	−0.08	−0.09
Item 7	3.48 (0.04)	−0.15	−0.15	−0.16	**0.79**	**0.79**	**0.78**
Item 8	3.46 (0.04)	**0.72**	**0.73**	**0.72**	−0.07	−0.07	−0.06
Item 9	3.36 (0.04)	**0.80**	**0.81**	**0.80**	−0.06	−0.06	−0.05
Item 10	3.09 (0.04)	**0.71**	**0.72**	**0.72**	0.15	0.16	0.15
Item 11	3.36 (0.04)	−0.11	−0.11	−0.11	**0.80**	**0.80**	**0.81**
Item 12	3.22 (0.04)	**0.85**	**0.85**	**0.85**	−0.09	−0.09	−0.09
Item 13	3.15 (0.04)	**0.83**	**0.84**	**0.84**	−0.08	−0.07	−0.08
Item 14	3.29 (0.04)	**0.84**	**0.85**	**0.84**	−0.11	−0.11	−0.10
Item 15	3.17 (0.04)	−0.04	−0.03	−0.03	**0.83**	**0.83**	**0.84**
Percentage of variance		47%	48%	48%	17%	17%	17%

Note: $N = 750$ (first half of random-split sample). Data = Imputed dataset. Minimum residual (minres) in R with varimax rotation was performed. Bolded items are selected for the perceived administrative burden (PAB) scale.

The high factor loadings from the CFA suggest adequate convergent validity. To further verify this, we considered the Average Variance Extracted (AVE) and the Pearson correlation between the scale and theoretically related criterion variables.

First, AVE values of 0.50 or higher are generally considered acceptable. In our CFA based on the third imputed dataset, the AVE scores were 0.60 for the US sample and 0.57 for the South Korean sample—both indicating acceptable convergent validity. Second, for criterion-related convergent validity, we used

TABLE 9 | Fit statistics from CFA.

	χ^2 (p value)	RMSEA	CFI	TLI
Perceived administrative burden				
US Samples	342.52 (0.000)	0.10	0.96	0.93
South Korean samples	312.54 (0.000)	0.07	0.98	0.97
Acceptable fit	p-value greater than 0.05	0.10 or lower	0.90 or higher	0.90 or higher

Note: The fit statistics presented in the table are point estimates pooled across three imputed datasets using the Rubin's rule (Rubin 1987).

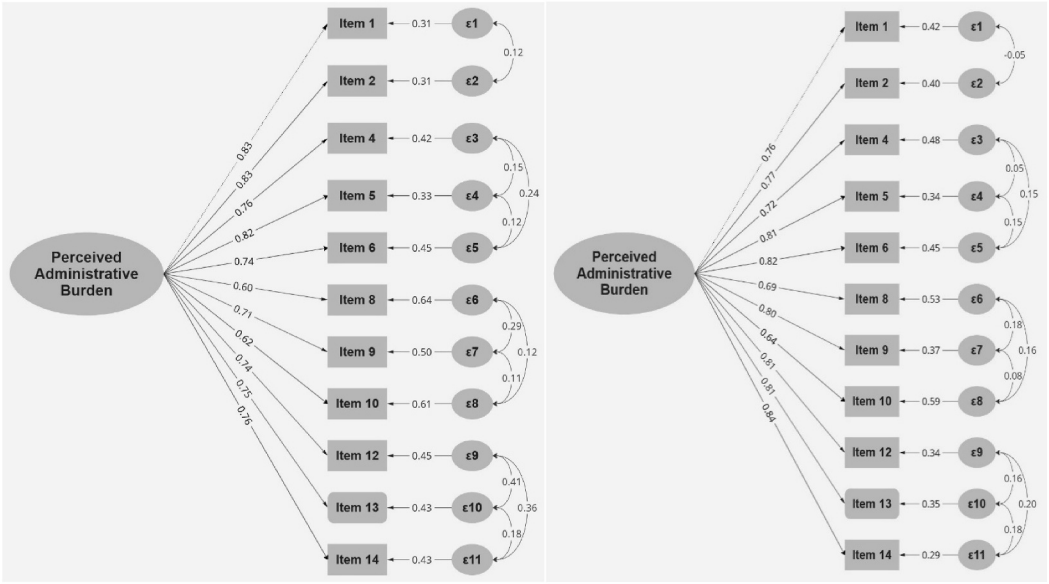

FIGURE 1 | Confirmatory factor analysis (standardized coefficients, US and South Korean samples).

agency favorability, agency reputation, satisfaction with the agency, and trust in the agency as criterion variables. Perceptions of administrative burden may be positively or negatively associated with these outcomes, depending on whether the burdens are perceived as legitimate and necessary or excessive and harmful. For this reason, we focused on the strength rather than the direction of the correlations, avoiding assumptions that could be arbitrary and theoretically tenuous. The results showed modest negative correlations between perceived administrative burden and the criterion variables in both countries. In the US sample, Pearson's r was −0.21 (favorability), −0.16 (reputation), −0.17 (satisfaction), and −0.24 (trust). In the South Korean sample, the correlations were stronger: −0.45, −0.42, −0.44, and −0.40, respectively.

In addition to assessing convergent validity, we examined criterion-discriminant validity by testing the correlation between the scale and a theoretically unrelated criterion variable: support for the government's international collaboration in combating the COVID-19 pandemic. Because this variable reflects attitudes toward international affairs, it should bear little relation to perceptions of administrative burden. The results aligned with this expectation. Pearson's r was 0.02 in the US sample and 0.10 in the South Korean sample, indicating extremely low correlations. Tables 10 and 11 present the pairwise Pearson correlations between perceived administrative burden and the criterion variables for the US and South Korean samples, respectively.

Although not part of the criterion validity tests, we conducted an exploratory analysis to examine whether the frequency of interactions with public agencies predicts perceived administrative burden. Our survey included a measure of interaction frequency, asking: "Including both online and in-person contact, how often do you interact with the presented agency?" (5-point Likert scale; 1 = Never, 5 = Always). We ran an OLS regression with interaction frequency as a categorical predictor (using 'Never' as the reference category), perceived administrative burden as the continuous dependent variable, and several basic demographic variables—age, gender, education, family income, and political orientation—as covariates. In the US sample, respondents who reported 'Always' interacting with the agency perceived significantly higher burden (b = 0.238, p = 0.046). In the South Korean sample, those in the 'Often' category also reported significantly higher perceived burden (b = 0.557, p = 0.000). One possible explanation is that frequent users of public agencies may translate repeated exposure to burdensome interactions into heightened perceptions of administrative burden. Alternatively, individuals with no experience may underestimate the extent of burden due to a lack of firsthand knowledge. This is an interesting exploratory finding that warrants further investigation in future research.

Finally, although not central to the main analysis, the scale showed a near-zero correlation with political orientation (r = 0.05 in the US; r = 0.00 in South Korea), suggesting that perceived administrative burden reflects a value-neutral

TABLE 10 | Bivariate correlation analysis (US samples).

	1	2	3	4	5	6
1 Perceived administrative burden	.					
2 Agency favorability	−0.21	.				
3 Trust in the agency	−0.24	0.73	.			
4 Satisfaction with agency	−0.17	0.73	0.62	.		
5 Agency reputation	−0.16	0.74	0.74	0.62	.	
6 Support for the government's international collaboration in response to the pandemic	0.02	−0.23	−0.28	−0.15	−0.24	.

Note: This table presents Pearson's pairwise correlation coefficients pooled across three imputed datasets following the Rubin's rule (1987).

TABLE 11 | Bivariate correlation analysis (South Korean samples).

	1	2	3	4	5	6
1 Perceived administrative burden	.					
2 Agency favorability	−0.45	.				
3 Trust in the agency	−0.40	0.68	.			
4 Satisfaction with agency	−0.44	0.75	0.64	.		
5 Agency reputation	−0.42	0.77	0.75	0.69	.	
6 Support for the government's international collaboration in response to the pandemic	0.11	−0.04	−0.04	−0.05	−0.05	.

Note: This table presents Pearson's pairwise correlation coefficients pooled across three imputed datasets following the Rubin's rule (1987).

assessment of burden itself—distinct from value-laden concepts such as burden tolerance or burden legitimacy.

Lastly, we assessed Cronbach's alpha for the 11-item perceived administrative burden scale, as well as for each of the four cost categories, to provide reference for researchers who may wish to use a subset of indicators based on their research context. The Cronbach's alpha for the full 11-item scale was 0.94, indicating exceptionally high internal consistency. The alpha scores for the perceived generic burden (items 1 and 2), perceived learning cost (items 4, 5, and 6), perceived compliance cost (items 8, 9, and 10), and perceived psychological cost (items 12, 13, and 14) were 0.82, 0.85, 0.78, and 0.88, respectively.

Based on the CFA results and a range of additional tests, we conclude that our perceived administrative burden scale is validated across both national contexts, confirming its utility for researchers in the field. The wording of the items in the final scale is presented in Table 12. The scale is designed to serve as a broadly applicable instrument for measuring perceived administrative burden. However, researchers interested in focusing on a specific cost category may choose to use a relevant subset of items. For example, a study examining perceptions of compliance costs may use items 8, 9, and 10, while items 1 and 2 may be most relevant for research focused on general perceptions of administrative burden.

4.4 | Agency-Wide Comparison

We compared perceived administrative burden across the four agencies within each country (full results are presented in Table 13). In the US survey, the IRS had the highest perceived administrative burden, with an average rating of 3.56 on a five-point Likert scale. The federal government was rated second at 3.44, followed by the SSA at 3.20. The USPS received the lowest rating, with an average of 3.04. Notably, the ranking of the three federal agencies—IRS, SSA, and USPS—excluding the Federal Government mirrors the favorability ratings reported in the 2023 Pew Research survey (Oliphant and Cerda 2023). This parallel suggests that perceived administrative burden may be meaningfully related to the broader reputational standing of public agencies (Figure 2).

There was also a fair amount of agency-wide variation in perceived administrative burden in the South Korean data. The NTS had the highest perceived administrative burden, with an average score of 3.37. This was followed by the NPA, which had an average perceived administrative burden score of 3.36. The central government had the third-highest perceived administrative burden, with an average score of 3.31. Finally, the CSC was rated the lowest on perceived administrative burden, with an average score of 2.95. The CSC, which is the smallest government unit located in close vicinity of residents to serve their basic administrative needs, was seen as the easiest agency to interact with.

Interestingly, in both countries, tax-related agencies (i.e., IRS in the US and NTA in South Korea) were rated the highest on perceived administrative burden, while service-oriented agencies (i.e., USPS in the US and CSC in South Korea) were rated the lowest. While these patterns may reflect factual differences in the level of burdens, they could also stem from stereotypical beliefs about public sector organizations and their performance (e.g., Marvel 2016), which in turn may shape perceptions of how burdensome those organizations are. For example, tax agencies

TABLE 12 | Perceived administrative burden (PAB) scale.

Domain	Question	Strongly disagree (1)—Strongly agree (5)
Perceived generic burden	This agency (program) puts many obstacles in people's way.	1 2 3 4 5
	This agency (program) is burdensome for people to deal with.	1 2 3 4 5
Perceived Learning cost (perceived cognitive burden)	This agency (program) lacks information for people to access services.	1 2 3 4 5
	This agency (program) makes things difficult for people to understand.	1 2 3 4 5
	This agency (program) provides people with too little guidance.	1 2 3 4 5
Perceived compliance cost (perceived resource burden)	This agency (program) asks people for lots of paperwork.	1 2 3 4 5
	This agency (program) takes up people's time.	1 2 3 4 5
	This agency (program) charges people lots of fees.	1 2 3 4 5
Perceived psychological cost (perceived emotional burden)	This agency (program) makes people feel uncomfortable.	1 2 3 4 5
	This agency (program) treats people badly.	1 2 3 4 5
	This agency (program) causes a lot of stress to people.	1 2 3 4 5

Note: As detailed in the text, the term "people" was used in place of "client", whose business-like connotation could bias responses—such as by inflating expectations about service accessibility and convenience. Therefore, we strongly recommend including an introductory statement—similar to that used in this study—which clarifies that "people" in the items refers to policy clients of the agency and its services.

TABLE 13 | Cross-agency comparison of perceived administrative burden.

	Mean (SD)			
US samples ($n = 1500$)	SSA	IRS	USPS	Federal Gov.
PAB scale	3.20 (0.06)	3.56 (0.04)	3.04 (0.05)	3.44 (0.05)
Perceived generic burden	3.18 (0.07)	3.58 (0.06)	3.01 (0.06)	3.47 (0.06)
Perceived learning cost	3.11 (0.06)	3.45 (0.05)	3.02 (0.06)	3.30 (0.06)
Perceived compliance cost	3.32 (0.06)	3.67 (0.05)	3.21 (0.05)	3.62 (0.05)
Perceived psychological cost	3.20 (0.07)	3.57 (0.05)	2.92 (0.06)	3.37 (0.06)
South Korean samples ($n = 1500$)	NPA	NTS	CSC	Central Gov.
PAB scale	3.36 (0.05)	3.37 (0.04)	2.95 (0.05)	3.31 (0.04)
Perceived generic burden	3.37 (0.05)	3.32 (0.05)	2.86 (0.05)	3.24 (0.05)
Perceived learning cost	3.39 (0.05)	3.41 (0.05)	2.96 (0.05)	3.33 (0.05)
Perceived compliance cost	3.34 (0.05)	3.42 (0.04)	3.05 (0.05)	3.37 (0.04)
Perceived psychological cost	3.36 (0.05)	3.31 (0.05)	2.89 (0.05)	3.29 (0.05)

Note: This table presents the mean and the standard deviation pooled across three imputed datasets. Cross-national and cross-agency comparisons of latent means should be interpreted with caution, as the score test did not support scalar invariance.
Abbreviations: CSC = Community Service Center, IRS = Internal Revenue Service, NPA = National Police Agency, NTS = National Tax Service, PAB = Perceived Administrative Burden, SSA = Social Security Administration, USPS = United States Postal Service.

are often perceived as more bureaucratic and inflexible, with limited personalized customer service. In contrast, agencies such as the USPS in the US or the Community Service Center (CSC) in South Korea may be viewed as more user-friendly. These remain untested speculations, and future research may further explore how perceived administrative burdens vary across agencies in distinct policy domains and examine the implications for citizen-state interactions using our scale.

The variation across subcategories of perceived administrative burden within each agency was less pronounced than the differences observed across agencies. This pattern appears consistent with the EFA and CFA results, which support a unidimensional structure of the concept—suggesting that individuals do not strongly differentiate between subcategories of administrative burden in their perceptions.

5 | Discussion

This study developed and validated a scale for perceived administrative burden, applicable beyond direct policy clients to

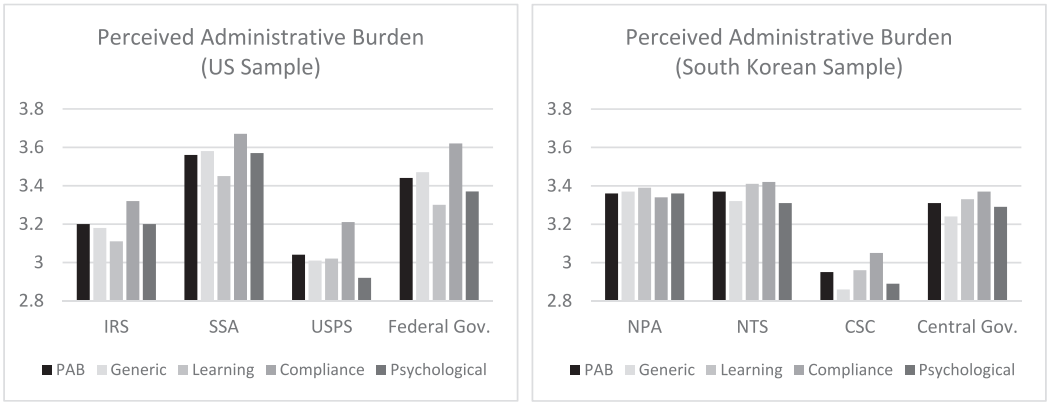

FIGURE 2 | Cross-agency comparison of perceived administrative burden. Cross-national and cross-agency comparisons of latent means should be interpreted with caution, as the score test did not support scalar invariance.

the general public and political stakeholders. We tested the scale's psychometric properties using large-scale, representative citizen samples from the US and South Korea ($n = 3000$), two countries with distinct cultural and historical backgrounds. The samples were obtained through stratified sampling across key demographic variables. The survey included four government institutions in each country (randomly presented to the respondents) that are relevant to administrative burdens, introducing additional variation that strengthens the scale's broad applicability.

The scale demonstrated strong psychometric performance across multiple tests. Item development combined a literature review with expert feedback (Hinkin 1995; Worthington and Whittaker 2006), emphasizing face validity and parsimony. Exploratory and Confirmatory Factor Analyses (EFA and CFA), conducted on randomly split samples across three imputed datasets, confirmed a unidimensional latent structure. Cronbach's alpha and Composite Reliability (CR) indicated strong internal consistency, while a high Average Variance Extracted (AVE) supported convergent validity. Finally, the scale's criterion validity was supported by additional correlational analyses using criterion variables that, from a theoretical standpoint, should be either strongly related (i.e., agency favorability, agency reputation, trust in the agency, and satisfaction with the agency) or unrelated (i.e., support for the government's international collaboration during the COVID-19 pandemic) to perceived administrative burden.

Our study contributes to the consistent use of a standardized measure of perceived administrative burden—an essential step toward improving the comparability and replicability of findings in public administration and policy research (Jilke et al. 2015; Jilke et al. 2017; Overman et al. 2024). The presented scale offers both methodological and practical advantages. For example, experimental studies often manipulate perceptions of burden, such as within survey vignette designs, rather than direct experiences due to practical and ethical constraints (e.g., Keiser and Miller 2020; Sievert and Bruder 2024). Our scale strengthens the rigor of such studies by serving as a reliable manipulation check, helping ensure that interventions shape perceived burdens as intended. In addition, because the scale captures a value-neutral assessment of burden—as partly evidenced by its weak correlation with political orientation—it may help mitigate social desirability bias in survey responses (Brenner and DeLamater 2016). However, further safeguards may still be necessary depending on the research context. The scale's brevity and intuitive wording also help reduce respondent fatigue and cognitive burden, contributing to improved data quality. Its design facilitates cross-national and cross-regional comparisons in administrative burden research. From a practical standpoint, the scale offers policymakers and government leaders a straightforward tool for monitoring how key stakeholders perceive the accessibility and complexity of public programs and agencies. By systematically tracking perceived burdens over time and across policy areas, practitioners can identify pain points, refine administrative processes, and improve service delivery in ways that enhance citizen experiences.

Through what mechanisms do perceptions of administrative burden emerge beyond firsthand experience? How does direct policy experience—or its absence—shape these perceptions? In what ways does perceived administrative burden influence public attitudes toward government and engagement with public programs? What are the broader political and policy-making implications of perceived burden, including its effects on agenda setting, program design, and policy feedback—extending beyond outcomes for policy clients? By offering a standardized measurement tool, we hope our scale will facilitate innovative research on these theoretically significant yet underexplored questions.

Ethics Statement

This work was approved by the Institutional Review Board of Seoul National University (IRB NO. 2211/002-021)

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

Aarøe, L., M. Baekgaard, J. Christensen, and D. P. Moynihan. 2021. "Personality and Public Administration: Policymaker Tolerance of Administrative Burdens in Welfare Services." *Public Administration Review* 81, no. 4: 652–663. https://doi.org/10.1111/puar.13381.

Almanasreh, E., R. Moles, and T. F. Chen. 2019. "Evaluation of Methods Used for Estimating Content Validity." *Research in Social and Administrative Pharmacy* 15, no. 2: 214–221. https://doi.org/10.1016/j.sapharm.2018.03.066.

Baekgaard, M., A. Halling, and D. Moynihan. 2025. "Burden Tolerance: Developing a Validated Measurement Instrument Across Seven Countries." *Public Administration Review* 85, no. 2: 519–546. https://doi.org/10.1111/puar.13835.

Baekgaard, M., and J. K. Madsen. 2024. "Anticipated Administrative Burdens: How Proximity to Upcoming Compulsory Meetings Affect Welfare Recipients' Experiences of Administrative Burden." *Public Administration* 102, no. 2: 425–443. https://doi.org/10.1111/padm.12928.

Baekgaard, M., and T. Tankink. 2022. "Administrative Burden: Untangling a Bowl of Conceptual Spaghetti." *Perspectives on Public Management and Governance* 5, no. 1: 16–21. https://doi.org/10.1093/ppmgov/gvab027.

Barnes, C. Y., and V. Riel. 2022. "'I Don't Know Nothing About That': How 'Learning Costs' Undermine COVID-Related Efforts to Make SNAP and WIC More Accessible." *Administration & Society* 54, no. 10: 1902–1930. https://doi.org/10.1177/00953997211073948.

Baumeister, R. F., E. Bratslavsky, C. Finkenauer, and K. D. Vohs. 2001. "Bad Is Stronger Than Good." *Review of General Psychology* 5, no. 4: 323–370. https://doi.org/10.1037/1089-2680.5.4.323.

Bell, E., J. E. Wright, and J. Oh. 2024. "Does Administrative Burden Create Racialized Policy Feedback? How Losing Access to Public Benefits Impacts Beliefs About Government." *Journal of Public Administration Research and Theory* 34, no. 3: 432–447. https://doi.org/10.1093/jopart/muae004.

Brenner, P. S., and J. DeLamater. 2016. "Lies, Damned Lies, and Survey Self-Reports? Identity as a Cause of Measurement Bias." *Social Psychology Quarterly* 79, no. 4: 333–354. https://doi.org/10.1177/0190272516628298.

Burden, B. C., D. T. Canon, K. R. Mayer, and D. P. Moynihan. 2012. "The Effect of Administrative Burden on Bureaucratic Perception of Policies: Evidence From Election Administration." *Public Administration Review* 72, no. 5: 741–751. https://doi.org/10.1111/j.1540-6210.2012.02600.x.

Bustos, E. O. 2021. "Organizational Reputation in the Public Administration: A Systematic Literature Review." *Public Administration Review* 81, no. 4: 731–751. https://doi.org/10.1111/puar.13363.

Carpenter, D. P., and G. A. Krause. 2012. "Reputation and Public Administration." *Public Administration Review* 72, no. 1: 26–32. https://doi.org/10.1111/j.1540-6210.2011.02506.x.

Compton, M. E., M. M. Young, J. B. Bullock, and R. Greer. 2023. "Administrative Errors and Race: Can Technology Mitigate Inequitable Administrative Outcomes?" *Journal of Public Administration Research and Theory* 33, no. 3: 512–528. https://doi.org/10.1093/jopart/muac036.

Daigneault, P.M. 2024. "Reconceptualizing Administrative Burden Around Onerous Experiences." *Perspectives on Public Management and Governance*: Advance online publication. https://doi.org/10.1093/ppmgov/gvae010.

DeVellis, R. F. 2003. *Scale Development: Theory and Applications*. Sage Publications.

Döring, M. 2021. "How-to Bureaucracy: A Concept of Citizens' Administrative Literacy." *Administration & Society* 53, no. 8: 1155–1177. https://doi.org/10.1177/0095399721995460.

Grimmelikhuijsen, S. G., and A. J. Meijer. 2014. "Effects of Transparency on the Perceived Trustworthiness of a Government Organization: Evidence From an Online Experiment." *Journal of Public Administration Research and Theory* 24, no. 1: 137–157. https://doi.org/10.1093/jopart/mus048.

Halling, A., and M. Baekgaard. 2024. "Administrative Burden in Citizen–State Interactions: A Systematic Literature Review." *Journal of Public Administration Research and Theory* 34, no. 2: 180–195. https://doi.org/10.1093/jopart/muad023.

Halling, A., P. Herd, and D. Moynihan. 2023. "How Difficult Should it Be? Evidence of Burden Tolerance From a Nationally Representative Sample." *Public Management Review* 25, no. 11: 2053–2072. https://doi.org/10.1080/14719037.2022.2056910.

Herd, P., T. DeLeire, H. Harvey, and D. P. Moynihan. 2013. "Shifting Administrative Burden to the State: The Case of Medicaid Take-Up." *Public Administration Review* 73, no. s1: S69–S81. https://doi.org/10.1111/puar.12114.

Herd, P., and D. Moynihan. 2024. "Reducing Administrative Burdens: The Case of the US Student Loan Forgiveness Website." In *Pathways to Positive Public Administration*, 83–87. Edward Elgar Publishing. https://doi.org/10.4337/9781803929170.00014.

Herd, P., and D. P. Moynihan. 2019. *Administrative Burden: Policymaking by Other Means*. Russell Sage Foundation.

Hinkin, T. R. 1995. "A Review of Scale Development Practices in the Study of Organizations." *Journal of Management* 21, no. 5: 967–988. https://doi.org/10.1177/014920639502100509.

Horton, N. J., and S. R. Lipsitz. 2001. "Multiple Imputation in Practice." *American Statistician* 55, no. 3: 244–254. https://doi.org/10.1198/000313001317098266.

Hu, L., and P. M. Bentler. 1998. "Fit Indices in Covariance Structure Modeling: Sensitivity to Underparameterized Model Misspecification." *Psychological Methods* 3, no. 4: 424–453. https://doi.org/10.1037/1082-989X.3.4.424.

Hu, L., and P. M. Bentler. 1999. "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives." *Structural Equation Modeling: A Multidisciplinary Journal* 6, no. 1: 1–55. https://doi.org/10.1080/10705519909540118.

Jilke, S. 2018. "Citizen Satisfaction under Changing Political Leadership: The Role of Partisan Motivated Reasoning." *Governance* 31, no. 3: 515–533. https://doi.org/10.1111/gove.12317.

Jilke, S., M. Bækgaard, P. Herd, and D. Moynihan. 2024. "Short and Sweet: Measuring Experiences of Administrative Burden." *Journal of Behavioral Public Administration* 7. https://doi.org/10.30636/jbpa.71.340.

Jilke, S., B. Meuleman, and S. Van de Walle. 2015. "We Need to Compare, but How? Measurement Equivalence in Comparative Public Administration." *Public Administration Review* 75, no. 1: 36–48. https://doi.org/10.1111/puar.12318.

Jilke, S., N. Petrovsky, B. Meuleman, and O. James. 2017. "Measurement Equivalence in Replications of Experiments: When and Why it Matters and Guidance on How to Determine Equivalence." *Public Management Review* 19, no. 9: 1293–1310. https://doi.org/10.1080/14719037.2016.1210906.

Kaiser, H. F. 1960. "The Application of Electronic Computers to Factor Analysis." *Educational and Psychological Measurement* 20, no. 1: 141–151. https://doi.org/10.1177/001316446002000116.

Kalucza, J., and M. Sievert. 2024. "Alleviating Administrative Burden through Digitalization and Redesign: A Laboratory Experiment on Student Financial Aid in Germany." *International Journal of Social Welfare* 33, no. 4: 786–805. https://doi.org/10.1111/ijsw.12649.

Keiser, L. R., and S. M. Miller. 2020. "Does Administrative Burden Influence Public Support for Government Programs? Evidence From a

Survey Experiment." *Public Administration Review* 80, no. 1: 137–150. https://doi.org/10.1111/puar.13133.

Lee, D. 2022. "Managing Bureaucratic Reputation in the Face of Crises: An Experimental Examination of the Effect of Strategic Communication." *Public Administration Review* 82, no. 6: 1124–1137. https://doi.org/10.1111/puar.13520.

Lee, D., and G. G. Van Ryzin. 2019. "Measuring Bureaucratic Reputation: Scale Development and Validation." *Governance* 32, no. 1: 177–192. https://doi.org/10.1111/gove.12371.

Ling, I. C., H. Y. Choi, and M. Sudhinaraset. 2024. "Navigating a Climate of Administrative Burden: The Perspectives of Young Adult Undocumented Immigrants in Applying for COVID-19 Disaster Relief Assistance for Immigrants in California." *Frontiers in Public Health* 12. https://doi.org/10.3389/fpubh.2024.1304704.

Lynn, M. R. 1986. "Determination and Quantification of Content Validity." *Nursing Research* 35, no. 6: 382–385. https://doi.org/10.1097/00006199-198611000-00017.

MacCallum, R. C., and J. T. Austin. 2000. "Applications of Structural Equation Modeling in Psychological Research." *Annual Review of Psychology* 51, no. 1: 201–226. https://doi.org/10.1146/annurev.psych.51.1.201.

Martin, L., L. Delaney, and O. Doyle. 2024. "Everyday Administrative Burdens and Inequality." *Public Administration Review* 84, no. 4: 660–673. https://doi.org/10.1111/puar.13709.

Marvel, J. D. 2016. "Unconscious Bias in Citizens' Evaluations of Public Sector Performance." *Journal of Public Administration Research and Theory* 26, no. 1: 143–158. https://doi.org/10.1093/jopart/muu053.

Moynihan, D. P. 2012. "Extra-Network Organizational Reputation and Blame Avoidance in Networks: The Hurricane Katrina Example." *Governance* 25, no. 4: 567–588. https://doi.org/10.1111/j.1468-0491.2012.01593.x.

Moynihan, D. P., P. Herd, and H. Harvey. 2015. "Administrative Burden: Learning, Psychological, and Compliance Costs in Citizen-State Interactions." *Journal of Public Administration Research and Theory* 25, no. 1: 43–69. https://doi.org/10.1093/jopart/muu009.

Nie, L., and H. Wang. 2023. "Government Responsiveness and Citizen Satisfaction: Evidence From Environmental Governance." *Governance* 36, no. 4: 1125–1146. https://doi.org/10.1111/gove.12723.

Nisar, M. A., and A. Masood. 2023. "Are All Burdens Bad? Disentangling Illegitimate Administrative Burdens through Public Value Accounting." *Asia Pacific Journal of Public Administration* 45, no. 4: 385–403. https://doi.org/10.1080/23276665.2022.2088581.

Oliphant, J. B., and A. Cerda. 2023. "Americans Feel Favorably About Many Federal Agencies." *Especially the Park Service, Postal Service and NASA*. https://policycommons.net/artifacts/3531909/americans-feel-favorably-about-many-federal-agencies-especially-the-park-service-postal-service-and-nasa/4332994/.

Overman, S., E. Ropes, and W. Vandenabeele. 2024. "Measuring Microfoundations of Governance: A Behavioral Perspective." In *Handbook on Measuring Governance*, 187–203. Edward Elgar Publishing. https://doi.org/10.4337/9781802200645.00022.

Peeters, R., and M. I. Dussauge Laguna. 2021. "Acting Out or Playing along: A Typology of Citizens' Low Trust Responses to Public Organizations." *Governance* 34, no. 4: 965–981. https://doi.org/10.1111/gove.12631.

Porumbescu, G., S. Walsh, and A. Hetling. 2024. "Can Reducing Learning Costs Improve Public Support for Means-tested Benefit Programs?" *Policy Studies Journal* 53, no. 1: 135–157. https://doi.org/10.1111/psj.12578.

Quilty, L. C., J. M. Oakman, and E. Risko. 2006. "Correlates of the Rosenberg Self-Esteem Scale Method Effects." *Structural Equation Modeling: A Multidisciplinary Journal* 13, no. 1: 99–117. https://doi.org/10.1207/s15328007sem1301_5.

Rubin, D. B. 1987. *Multiple Imputation for Nonresponse in Surveys*. John Wiley & Sons.

Rubin, D. B. 1996. "Multiple Imputation after 18+ Years." *Journal of the American Statistical Association* 91, no. 434: 473–489. https://doi.org/10.1080/01621459.1996.10476908.

Sievert, M., and J. Bruder. 2024. "Unpacking the Effects of Burdensome State Actions on Citizens' Policy Perceptions." *Public Administration* 102, no. 3: 915–935. https://doi.org/10.1111/padm.12957.

Stenderup, R., and M. J. Pedersen. 2024. "Compliance Burden versus Program Integrity Protection: A Survey Experiment on Citizen Attitudes toward Administrative Documentation Requirements in Public Service Policy." *Public Administration Review*: Advance online publication. https://doi.org/10.1111/puar.13919.

Tomlinson, J., E. Kasoulide, J. Meers, and S. Halliday. 2024. "Direct and Vicarious Administrative Burden: Experiences of UK Public Services as Homes for Ukraine Host." *Journal of Refugee Studies* 37, no. 3: 768–784. https://doi.org/10.1093/jrs/feae036.

van den Bekerom, P., J. van der Voet, and J. Christensen. 2021. "Are Citizens More Negative About Failing Service Delivery by Public Than Private Organizations? Evidence From a Large-Scale Survey Experiment." *Journal of Public Administration Research and Theory* 31, no. 1: 128–149. https://doi.org/10.1093/jopart/muaa027.

Wang, P., and Y. Zhang. 2024. "Cross-domain Policy Feedback Effects on Mass Publics." *Governance* 38, no. 2. https://doi.org/10.1111/gove.12905.

Wei, W., N. Petrovsky, and X. Ni. 2024. "Citizen Blame Attributions for Government Fiscal Crises: Experimental Evidence From China." *Governance* 37, no. 4: 1299–1320. https://doi.org/10.1111/gove.12849.

West, S. G., A. B. Tylor, and W. Wu. 2012. "Model Fit and Model Selection in Structural Equation Modeling." In *Handbook of Structural Equation Modeling*, edited by R. R. Hoyle, 209–231. Guilford Press.

Worthington, R. L., and T. A. Whittaker. 2006. "Scale Development Research." *Counseling Psychologist* 34, no. 6: 806–838. https://doi.org/10.1177/0011000006288127.

Wu, R. 2021. "America's Unforgiving Forgiveness Program: Problems and Solutions for Public Service Loan Forgiveness." *Hastings Law Journal* 72, no. 3: 959.

Zhang, J., K. Yang, and H. Li. 2024. "Legislative Capacity, Bureaucratic Reputation, and Delegation From a Trust Perspective: A Survey Experiment." *Governance* 37, no. 1: 281–301. https://doi.org/10.1111/gove.12763.

Zhu, J., Y. Han, and Z. Chen. 2024. "The Effect of Administrative Burden Perception on Citizen Participation: A Barrier or a Spur?" *Public Performance and Management Review* 47, no. 5: 1138–1165. https://doi.org/10.1080/15309576.2024.2354304.