https://doi.org/10.51408/issi2025_089

An Empirical Study on the Distributional Characteristics of Policy Citation Behaviors in Climate Action Policies

Zheng Xinman[1], Liu Xiwen[2]

[1]*zhengxinman@mail.las.ac.cn*, [2] *liuxw@mail.las.ac.cn*
National Science Library, Chinese Academy of Sciences, Beijing (China)
Department of Information Resources Management, School of Economics and Management, University of Chinese Academy of Sciences, Beijing (China)

Abstract

This study empirically analyzes the distributional characteristics of policy citation behaviors in climate action policies. By examining policy documents from different institutional sources, including intergovernmental organizations (IGOs), governments, and think tanks, the study finds that citation annotations are prevalent across all three types of institutions, with an overall usage rate of 87%. IGOs exhibit the highest utilization of citation annotations at 97.1%, followed by think tanks at 92.7%, and governments at 80.4%. The chi-square test confirms a statistically significant difference in citation annotation usage among these institutions. The study identifies six common types of citation annotations: footnotes, endnotes, bibliographies, in-text citations, captions, and hyperlinks. Footnotes and bibliographies are the most frequently used types across all policy sources, accounting for over 60% of total citations. However, preferences vary among institutions; IGOs favor captions, think tanks prefer bibliographies and in-text citations, while governments predominantly use footnotes and hyperlinks. Think tank policies exhibit the highest citation frequency, while government policies have a relatively lower rate. These findings shed light on the differences in citation behaviors among various policymaking institutions and provide insights into the Science-Policy Interface in climate action policies.

Introduction

Policy document citations are citations of external information within policy texts, similar to citation data in academic papers, and have a wide range of research value and applications. By analyzing policy document citations, it is possible to analyze the Science-Policy Interface (SPI), which refers to the interaction and mutual influence between scientific research and policy-making, or to assess the social impact of scientific research publications (Bornmann, 2016; Haunschild and Bornmann, 2017; Bornmann, 2022). In addition, quantitative analyses of policy document citations can enrich the research scope of public policy analysis by providing statistical data on how policy draws on external information (Newson, 2018), thereby expanding the research paradigm of public policy. The quantitative analysis of policy citation data, which reflects policy citation behavior, refers to how external information is used in policy texts. Therefore, understanding the specific ways of the distribution of policy citation behavior helps to better understand the connotation of policy citation data and promote the further development of policy citation research.

Current research on the distribution of policy citation behavior faces a triple challenge: limited data availability, unclear institutional variations and methodological constraints. First, the policy document citations used in the current

study mainly come from policy citation data provided by Altmetric and Overton databases, which suffer from data coverage bias (Bornmann, 2016; Bornmann, 2022; Tattersall, 2018; Maleki, 2022), and some policy document citation behavior may not be supported and identified by the databases due to unstructured features (Overton, 2019). The resulting distribution of policy document citations based on citation behavior data from these databases appears to be similar in nature to the distribution of academic citations (Szomszor and Adie, 2022), but conclusions such as the generally low proportion of citations from academic papers in policy texts lead to difficulties in determining whether policy citation behaviors are systematic practices or accidental manipulations, which, in turn, undermines the ability to ability to verify the universal law of policy citation behavior. Second, different policymaking institutions (e.g., governments and think tanks) show differences in the use of academic citations in policymaking; for example, during the COVID-19 pandemic, the tendency to cite science in policy documents seems to have been concentrated mainly within intergovernmental organizations (IGOs), such as the World Health Organization (WHO), and to a much lesser extent in national governments, which mainly consume science through intergovernmental organizations indirectly consume science (Yin, 2021). Such differences may stem from differences in institutional resource endowments or knowledge translation mechanisms, but existing studies have not yet been able to further reveal and quantify the extent of differences in cross-agency policy citation behavior, nor the selection mechanisms and influencing factors behind them, due to a lack of comparative analysis of cross-agency policy citation behavior. More critically, the unstructured characteristics of policy texts and the existence of multiple citation styles make the automated extraction of citation data challenging, and it is difficult to extract citation data directly from policy texts and consumes a large amount of labor costs (Newson, 2018), which leads to the reliance of existing studies on the distribution of policy citation behaviors on small samples of manual annotations (Newson, 2018; Yu, 2023), making the comparison of policy citation behavior differences across sources lacking data support. These three obstacles together constitute the "black box" of the policy citation behavior distribution problem - do policies in all fields follow the same citation pattern? What structural factors drive the heterogeneity of citation patterns?

Citation annotations in policy documents are an important basis for analyzing policy citation behavior, which can provide key clues for deciphering the "black box" problem in current research. Existing research shows that citation annotations in policy documents are mainly divided into two categories: one is the use of specific wording or quotation marks to cite external sources of information in the body or table headings, footnotes and endnotes, such as "based on", "refer to" and other prompt words such as "see" (Huang, 2015; Overton, 2022), or quotation marks in the body text of policy documents (Ba, 2022); the other is referencing styles in common publications such as academic papers, such as footnotes, endnotes, hyperlinks, bibliographies, etc. (Newson, 2018; Yu, 2023). For example, Newson et al. found that approximately two-thirds of childhood obesity prevention policy documents issued by the New South Wales government in Australia between 2000 and 2015

contained references, and of these, more than one-third of the policies used footnotes, hyperlinks, or a combination of these forms (Newson, 2018).Yu et al. found that the standardized referencing style was the main form of reference when policies cite academic papers, as demonstrated by citing academic papers in the form of post-textual reference lists and including footnotes or endnotes in the body of the policy (Yu, 2023). It can be seen that policy citation annotations have the potential to analyze the distribution of policy citation behaviors.

The United Nations Sustainable Development Goals (SDGs) are complex global issues that involve a wide range of policymaking institutions and stakeholders, including governments (national and local), IGOs, the private sector, non-governmental organizations (NGOs) and academia. Among them, SDG 13 (Climate Action) has a wide range of impacts, covering a variety of fields such as environment, energy, economy, etc., and has produced a wealth of policy documents and scientific research results. As a common challenge faced by all mankind, the types of institutions that formulate climate action policies are many and numerous, and the degree of policy disclosure is high (Bornmann, 2022). As time progresses, global climate governance faces important challenges that call for more scientific policy development and greater citation of evidence, and therefore has the potential for generalized use of citation behaviors in its policies compared to policies on other topics, but current research based on policy citation databases suggests that climate action policies cite science at a low rate (Bornmann 2016) and it is not clear that policies from different sources have similar citation behavior. In addition, policymaking institutions have their own positions, and their processes and roles in policy development vary, which may lead to differences in whether and in what form references to sources of information are included in policy documents.

This study explores the prevalence and variability of the distribution of policy citation behaviors in the field by analyzing citation annotations in climate action policy documents. Three specific issues are analyzed: first, an analysis of policy availability, which explores the main policy sources and document styles of policies, as well as the availability of policies; second, an analysis of the prevalence of policy citation behaviors across different policy sources and formats; and finally, an examination of the differences in citation annotations choices and use across different policy sources. Based on these analyses, important support is provided for understanding the prevalence and differences in the distribution of policy citation behaviors.

Method

Policy Document Source Identification

We used the Overton policy document database to retrieve policy documents from different sources in the field of climate action for two reasons. The first is that the Overton database covers a wide range of institutional types of sources of policy documents. The other is that the Overton database maps policy documents to one or more of the SDGs. The Overton database defines a policy document very broadly as “documents written primarily by and for policy makers”. This idea is intended to

cover not only policy documents documenting the policy or legislation itself, but also documents intended to inform or influence decision-making (Szomszor and Adie, 2022). The policy scope of this paper is consistent with Overton's definition. This paper combines the SDG labels provided by the Overton database to select SDG 13 policy documents for climate action, totaling 22,352. These policy documents contain the types of Publication, Blog post, and Working paper. In this paper, we choose publication as the sample of policy citation annotations because it is a formally released document with a relatively standardized style, and we get 20,303 policy documents with three types of institutional sources, including 3,954 documents from IGOs, 10,330 documents from governments, and 6,019 documents from think tanks in various countries.

Policy Document Collection and Sampling

In order to understand the characteristics of policy citation annotations for each type of organization, this paper draws samples from the policies of each type of organization separately for fine-grained annotations by means of stratified sampling. Stratified sampling, also called type sampling, is a sampling method that divides the overall units into a number of types or strata according to their attribute characteristics, and then randomly selects sample units from the types or strata. Stratified sampling is characterized by the fact that the commonality between units in each type is increased through the delineation of types and strata, and it is easy to draw a representative survey sample. This method is suitable for the overall situation is complex, the difference between the units is large, more units, applicable to the application of this paper's scenario. The specific process of stratified sampling is to first calculate the sampling proportion of each institutional category, and for each category, multiply it by the total sample size to get the sample size that should be taken for that category. Random sampling is then performed in each category to ensure that the sample in each category is random. Rogers et al. consider the amount of literature data used for bibliometric analysis to be at least 200 . This study refers to this criterion and 200 policy documents were sampled to ensure that the sample size was sufficient for econometric analysis. The stratified sample yielded 38 IGO policies (proportion: 19%), 102 government policies (proportion: 51%), and 60 think tank policies (proportion: 30%).

Policy Citation Annotation Coding

The original text of the sampled policy documents was downloaded according to the URL provided in the Overton database. The coding yielded information about each policy document, including two categories of policy document basic information and policy citation annotations information. The basic information of the policy document includes the title of the policy document, the source country, the name of the source organization, the type of the source organization, the link to the original policy text, the date, the availability of the policy document, and the type of the policy document format. Most of the information comes from data items exported from the database. Policy document availability, policy document format type, and total number of pages in the policy document are manually coded, and are judged and

counted when the original policy text is downloaded. Policy citation annotation information includes information on whether it contains citation annotations, the type of citation annotations, and the number of times the citation annotations were used. Policy citation annotations were obtained manually by scanning the full text of each policy through a combination of manual identification and content analysis to find and record the types of citation annotations that appeared in the main text and appendices and the frequency of use of that type of citation annotation in the policy documents.

Determines whether the policy contains citation annotations

Considering the universality and consistency of citation formats in academic papers, this paper identifies and records citation annotation types based on the reference citation formats commonly used in academic papers. If a policy document contains at least one type of reference citation format, it is registered as "containing citation annotations." By summarizing the commonly used citation annotation types in various publications, along with their positions and forms, citation annotations can be classified into six types: Footnotes, Endnotes, Captions (below tables or figures), In-text citations, Bibliographies, and Hypertext Links. Among these, In-text citations differ from Footnotes, Endnotes, and Hypertext Links in the formatting of the markers used when directly quoting content within the text. In-text citations typically adopt a parenthetical format indicating the author-date next to the quoted content, formatted as (Author, Year), for example, (Smith, 2019). Different academic writing style guides (such as APA, MLA, Chicago, etc.) may exhibit some variations in the formatting of in-text citations. However, regardless of the citation style, it is advocated to provide basic information about the cited content within the text (such as the author's name, publication year, title of the article or book, etc.), enabling readers to accurately understand the source and context of the quoted content. Footnotes and Endnotes commonly use numerical or symbolic markers. A Hypertext Link, also known as a Hyperlink or simply a Link, is used in web pages or electronic documents to direct users to other pages, resources, or locations when clicked. Hypertext Links are usually presented in text form and are often highlighted by changing the color of the link text or by underlining it.

Determining the type of citation annotations

Check whether the policy documents contain six common types of annotations, such as “reference lists”, “in-text markup”, “footnotes”, “endnotes”, “notes below charts”, “hypertext links”, etc., and determine whether these types of annotations play the role of citation. “Six common types of annotations, including footnotes, endnotes, and hypertext links, were examined to determine whether or not they played a role in citation annotations. We take into account the cases where footnotes, endnotes, and notes underneath charts and tables may play a non-citation annotations role, such as terminology explanations only, and so on. Since the intent of this paper is to observe citation behaviors in policy documents, only annotation types that play a citation role are registered in this paper. When a certain annotation type provides external sources in the policy documents, it can be regarded as playing the role of citation, registering

this annotation type as the citation annotations of the policy, and taking the frequency of this annotation in the text as the frequency of the annotation use. When a certain type of annotations only explain the role, there is no citation annotations, do not register the type of annotations. If there is a citation of external sources of information in the text, but the citation annotations do not belong to the six common types of annotations, the type will be registered as other types.

Counting the frequency of use of citation annotations

In order to facilitate counting and reduce labor costs, the total number of times a certain type of citation annotations appear in a single text as the citation annotation frequency, without the need to distinguish one by one which content is a citation and which is an explanation, in order to ensure that the identification of which types of annotations play the role of citation, and greatly improve the efficiency of manual labeling.

Results

Policy documents accessibility

The integration of URLs obtained through sampling and the subsequent download of original policy documents yielded a comprehensive dataset comprising 55 think tank policies, 34 intergovernmental organization (IGO) policies, and 92 government policies, resulting in a total sample size of 181 policies. The sample encompasses contributions from 21 countries, 15 IGOs, and 39 think tanks, indicating a diverse and extensive range of sources contributing to policies related to Sustainable Development Goal (SDG) 13 on climate action.

Table 1. Distribution of Policy Sample Sources.

Source Type	*#countries/regions*	*#institutions*	*Top 3 institutions/countries by frequency*
IGO	——	15	UNEP, World Bank, FAO
Government	21	53	USA, EU, UK
Think Tank	13	39	USA, UK, Belgium, Germany

The vast majority of policy documents in the sample (90.5%) were publicly accessible through existing or archived websites, with 87% available in PDF format and 3.5% in HTML format. However, 19 documents were unavailable due to inaccessible web pages (e.g., "page not found" or "404 - file or directory not found"), lack of access rights, or misclassification as policy documents (e.g., conference proceedings unrelated to institutional policies). These 19 documents, which could not be retrieved or were deemed irrelevant, were categorized as "other," accounting for 9.5% of the sample, as illustrated in the table below. The findings indicate that the availability of policies across different sources exceeds 80%, reflecting a relatively high level of accessibility. This availability rate is notably higher than that

of the Overton database, which itself surpasses the percentage of valid policy data in Altmetric.com (71%) (Yu H, 2023). These results underscore the robustness of the dataset and the comparative advantage of the Overton database in terms of policy data accessibility.

Table 2. Accessibility of policies from different sources.

Source Type	*PDF Format*	*Html Format*	*Unable to Obtain*
IGO	89.5%（n=34）	0.0%（n=0）	10.5%（n=4）
Government	89.2%（n=91）	1.0%（n=1）	9.8%（n=10）
Think Tank	81.7%（n=49）	10.0%（n=6）	8.3%（n=5）
Total	87.0%（n=174）	3.5%（n=7）	9.5%（n=19）

Distribution of climate action policy citation behaviors

The utilization rate of citation annotations in climate action policies: The analysis of citation annotation usage rates across policies from different institutional types revealed that, overall, 87% (n=181) of climate action policies included citation annotations. This high percentage underscores the prevalence of citation practices within climate action policies, suggesting that referencing and acknowledging sources is a common and integral aspect of policy development in this domain.

Usage rates of citation annotations for different institution types: In terms of the type of institution, the proportions of citation annotation policies originating from governments, think tanks, and intergovernmental organizations (IGOs) within the sample set of policies from their respective sources are 80.4%, 92.7%, and 97.1%, respectively. These figures indicate a higher prevalence of citation annotations in policies issued by these three types of institutions, as illustrated in the table below. Yin et al. posited that IGOs exhibit a more pronounced tendency to cite scientific research in their policy documents compared to national governments (Yin, 2021). The findings of this study corroborate this assertion, revealing that IGOs have the highest utilization of citation annotations, while governments have the lowest. This suggests that the policy citation behaviors, whether or not it pertains to scientific research, is more prevalent among IGOs than among governments.

Table 3. Citation annotation usage rate of policies from different sources.

Source Type	*With Citations*	*Without Citations*
IGO	80.4%	19.6%
Government	92.7%	7.3%
Think Tank	97.1%	2.9%
Total	87%	13%

After conducting the chi-square test, the obtained test values for Pearson's chi-square and the likelihood ratio were 0.016 and 0.010, respectively. Both of these values are less than the conventional significance level of 0.05 (p-value). This indicates that there is a statistically significant relationship between the type of institution and the rate of policy use of citation annotations. Specifically, there is a significant difference in the rate of policy citation annotations usage among governments, think tanks, and IGOs. As detailed in the table below, IGOs and think tanks exhibit a greater inclination to utilize citation annotations in their policy documents compared to governments.

Table 4. Chi-Square Tests for the Relationship Between "Source Type" and "Citation Annotation Usage Rate".

Test Type	*Value*	*Degrees of Freedom*	*Asymptotic Significance (2-sided)*
Pearson Chi-Square	8.289	2	0.016
Likelihood Ratio	9.196	2	0.010
Linear-by-Linear Association	7.740	1	0.005
Valid Cases	181	——	

Utilization of citation annotations across various file format types: The statistics presented in Figure 1 provide insights into the utilization of policy citation annotations across different document format types. Upon analyzing Figure 1, it becomes evident that the proportion of policy documents employing citation annotations is notably higher in both PDF and HTML formats. This suggests that policy citation behaviours are prevalent in documents where citations are explicitly manifested.

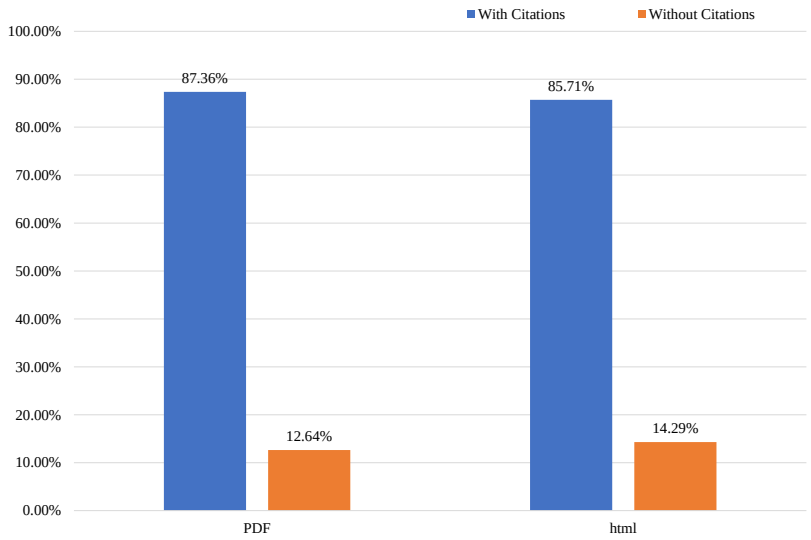

Figure 1. Proportion of Citation Annotations in Sample Policies.

Variations in citation behaviors across policy sources

Citation Frequency in Policy Documents

An analysis of the cumulative distribution of citation counts for policies originating from diverse sources unveils a pronounced imbalance in the frequency of policy citations among these sources. Specifically, it is observed that 27% of government policies account for 80% of the total citation counts, 30% of policies from IGOs contribute to 80% of the citations, and 37% of think tank policies are responsible for 80% of the citations. This distribution pattern indicates that think tank policies generally exhibit a higher citation frequency, whereas government policies demonstrate a relatively lower citation rate, highlighting disparities in policy-making practices across different types of organizations. This phenomenon, wherein a minority of policies garner the majority of citations, aligns with the Pareto Principle, which posits that in numerous instances, roughly 20% of the factors (policies) generate approximately 80% of the outcomes or impacts (citations).

Upon analysing the citation frequency (total count of citation annotations per policy) and citation density (number of citation annotations per page of policy) for individual policies across different source institution types, it is evident that think tank policies exhibit the highest average number of citation annotations. This is followed by policies from IGOs, with government policies trailing behind. Notably, there is a higher prevalence of outliers in the citation data for government policy documents, indicating a greater degree of variation in citation behaviours among government policies compared to those from think tanks and IGOs.

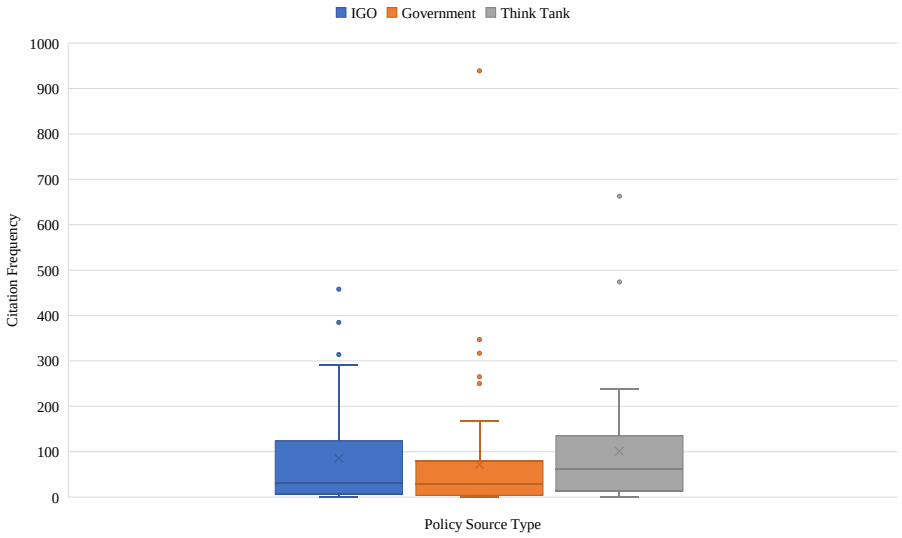

Figure 2. Distribution of Citation Frequency for Policies from Different Policy Source Type.

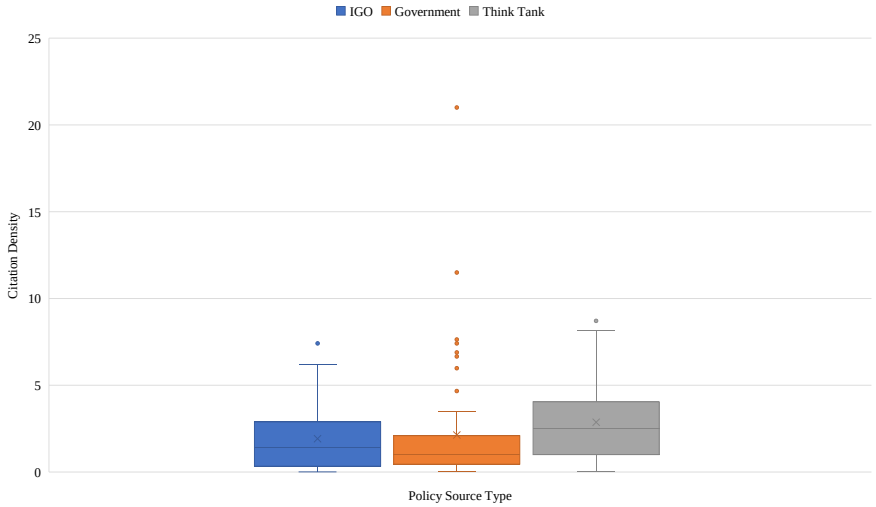

Figure 3. Distribution of Citation Density for Policies from Different Policy Source Type.

Usage of policy citation annotation types

Types of policy citation annotations: The annotation results reveal that there are a limited number of text box (Box) annotations that contain citation annotations within the policies. Upon examination, it becomes apparent that both the chart annotations and text box annotations share some similarities in their citation behaviour. Specifically, citations in these two cases typically occur in prominent and distinct locations within the text, and sometimes exhibit more independent citation patterns. In the text, citations are usually introduced using source or note prompt words. For the purpose of simplifying the analysis, this study groups these two types of

annotations into a single category, referred to as caption. By combining the common citation annotations found in policy documents with the types of annotations within the text, we ultimately identify six types of policy annotations: footnotes, endnotes, bibliographies, in-text citations, captions and hyperlinks.

In terms of the choice of citation annotation types, an analysis of the frequency and distribution of each annotation type within policy documents provides insights into the characteristics of policies with regard to their citation practices. The following table presents an overview of these findings:

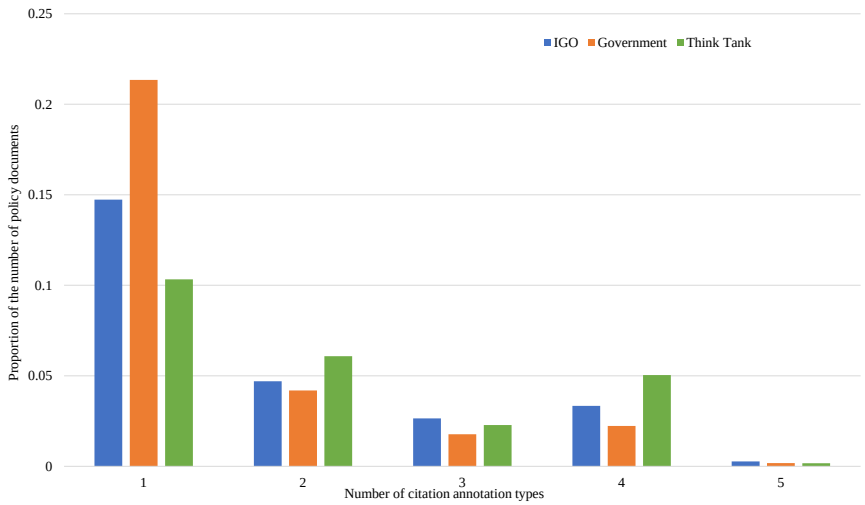

Figure 4. Number of citation annotation types used per policy (Normalized).

Analysing the provided figure, it is evident that the majority of policies utilize only 1 type of citation annotation. Additionally, there are policies that employ up to 5 types of citation annotations. Notably, policies containing fewer than 2 types of citation annotations constitute 60.5% of the policy document citations. This finding indicates that policies from different institutional types exhibit a common trend in their selection of citation annotations; specifically, they tend to use fewer than 2 types of citation annotations.

Analyse whether there is a common use of a certain citation annotations type across institutions. An examination of the types of citation markers contained in climate action policies was conducted, with the percentage of policies from each institution type utilizing each citation marker type calculated and presented in Table 2 below. As is shown in the table, over 40% of policies from all three types of institutions employed hyperlinks and footnotes, indicating that these are the most frequently used citation marker types in climate action policies. A chi-square test was employed for analysis, and the results revealed no significant differences ($P > 0.05$) in the usage rates of footnotes and hyperlinks among the three institution types. This result aligns with existing research on policy citation annotations by type of government agency. Newson (2018) noted that "more than one-third of policy documents do not list references in individual lists or appendices, but instead use footnotes, hyperlinks, or

a combination of these methods." This consistency in findings reinforces the notion that footnotes and hypertext links are widely accepted and utilized as effective means of citing sources in policy documents, particularly in the context of climate action policies.

Further compare whether there are significant differences in the types of policy citation annotations preferred by different institution types. Based on the observations, it appears that different types of institutions have distinct preferences when it comes to the types of citation annotations used in their climate action policies. Bibliographies and captions emerge as the most common citation annotation types for think tanks ($P<0.05$). IGOs favours using captions as citation annotations, which involve direct source citations at pictures, tables, and separate text boxes, ($P<0.05$). In contrast, government policies predominantly use footnotes and hypertext links as their most common citation annotations. Notably, governments are less likely to use endnotes compared to think tanks and IGOs ($p < 0.05$). Considering the varying primary modes of dissemination for policies among different institution types, there are also differences in the preferred types of citation markers. Different citation marker types serve distinct purposes and effects. For instance, the combination of bibliographies at the end of the document and in-text citations is the most common type in academic publishing, while hyperlinks are a convenient citation marker type for online publishing and dissemination.

Table 5. Comparison of the Proportion of Policies with Specific Citation Annotations by Source Type.

Citation Annotation Type	*Government*	*Think Tank*	*IGO*	*P Value*
Hyperlink	42.4	41.8	55.9	0.350
Footnotes	54.3	41.8	44.1	0.286
Endnotes	3.3	25.5	11.8	0.000
Captions	22.8	50.9	55.9	0.000
In-text citations	15.2	47.3	32.4	0.000
Bibliography	19.6	50.9	26.5	0.000

Common types of citation formats from different sources

Common citation annotation types in various policy sources. Considering that the citation content contained in in-text citations overlaps with bibliographies or endnotes, in-text citations were excluded from the analysis. The usage frequencies of the remaining citation annotation types were counted, resulting in a stacked percentage bar chart of citation annotation usage, as shown in the figure below. The results indicate that both footnotes and bibliographies account for 60% or more of the total citation annotation usages in policies from the three types of institutions.

From the perspective of usage frequency, footnotes and bibliographies are the most frequently used citation annotation types across policy texts from various sources.

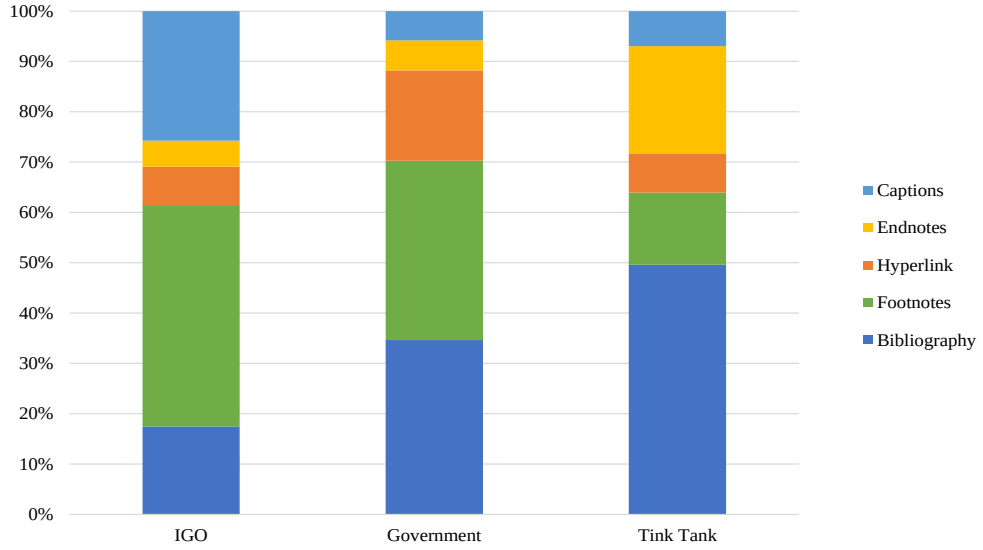

Figure 5. Percentage Stacked Bar Chart of Citation Annotation Usage.

Preferences in the usage frequencies of various citation annotation types. Upon analysing the figure below, it can be observed that the citation annotation types with the highest average usage frequencies in IGO policies are footnotes and bibliographies. In think tank policies, the citation annotation types with the highest average usage frequencies are in-text citations, bibliographies, and endnotes. Similarly, in government policies, the citation annotation types with the highest average usage frequencies are also in-text citations, bibliographies, and endnotes. Evidently, bibliographies emerges as the most frequently used citation annotation type across policies from all three sources.

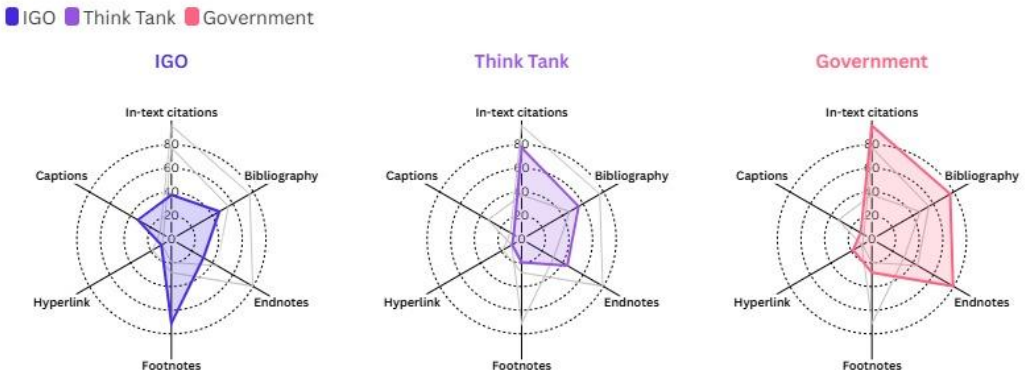

Figure 6. Average Usage Frequency of Each Annotation Type in Different Policy Sources.

Discussion

The sources of climate action policies are diverse, encompassing IGOs, governments, and think tanks across countries as the primary origins, and these policy documents are generally highly accessible, facilitating the direct extraction of policy citation data from the policy texts. Given the prevalence of policy citations in climate action policies, the direct extraction of such data holds significant importance for analyzing the sources of scientific evidence in the policy-making process, assessing policy impacts, and promoting the interaction between science and policy. Furthermore, with advancements in text mining and natural language processing technologies, it has become feasible to directly extract policy citation data from policy texts.

Policies from different sources exhibit significant variations in citation behavior. Think tanks and international governmental organizations (IGOs) tend to cite a wide range of literature to support their policy proposals, reflecting a strong commitment to evidence-based decision-making. In contrast, government agencies may cite relatively fewer references, as they often prioritize practical implementation and immediate effects. This disparity reflects differing approaches to evidence-based decision-making among various decision-making bodies: think tanks and IGOs emphasize the foundational role of scientific research in policy formulation, while government agencies focus more on the timeliness and operationalizability of policies. When promoting the interaction between science and policy, it is essential to fully consider the characteristics and needs of different decision-making bodies.

Limitation and Future Work

Limitation. This paper collects data from Overton, a process that inevitably involves a certain degree of selection bias regarding policy data sources, resulting in a higher likelihood of capturing only those policy data sources that are readily accessible. This limitation in data sources may imply that some non-public or hard-to-access policy documents are omitted, thereby affecting a comprehensive and in-depth analysis of policy citation behavior.

We employ explicit citation markings to explore policy citation behavior; however, the behavior itself is exceedingly complex. Extracting policy citation data solely from policies containing citation markings clearly cannot fully capture the entirety of policy citations. As policy documents often do not disclose whether specific evidence evaluation criteria have been applied, there exists the possibility that some studies, although utilized, are not explicitly cited. This constitutes a limitation of our research method.

Future Work. This study has conducted an in-depth exploration of policy accessibility and citation marking styles and preferences, laying a foundational groundwork that provides valuable experience and insights for the design of subsequent automated citation extraction methods. By understanding the avenues for obtaining policy documents and the diversity and preferences in citation markings, we can design algorithms and models more targetedly to enhance the accuracy and efficiency of automated citation extraction. This will significantly promote the large-scale acquisition and utilization of policy citation data, offering new tools and methods for policy research.

The extraction of policy citation data not only provides abundant material for research on the interaction between science and policy, but also enables the assessment of the impact of different types of publications on policy. By analyzing the sources, types, and frequencies of citations in policy documents, we can reveal which publications have exerted significant influence on policy formulation and how this influence occurs. This will facilitate a deeper understanding of the interaction between science and policy, providing scientific evidence for policymakers and policy-oriented references for publication editors and authors.

Comparing citation behaviors across policies in different domains is an important and intriguing issue. For instance, there may be significant differences in policy citations between clinical research and public health research. By conducting a comparative analysis of the citation characteristics of policies in these two domains, we can uncover the sources of scientific evidence and decision-making logic underlying policy formulation in different fields, as well as their demands and preferences for scientific research. This will contribute to a more comprehensive understanding of the nature and patterns of policy citation behavior, providing targeted suggestions and guidance for policymakers and researchers in various domains.

Acknowledgments

The policy document data were shared with us by Overton on September 11, 2023.

References

Ba, Z., Zhao, Y. C., Liu, X., & Li, G. (2022). Spatio-temporal dynamics and determinants of new energy policy diffusion in China: A policy citation approach. Journal of Cleaner Production, 376, 134270.

Bornmann, L., Haunschild, R., & Marx, W. (2016). Policy documents as sources for measuring societal impact: How often is climate change research mentioned in policy-related documents?. *Scientometrics*, *109*, 1477-1495.

Haunschild, R., & Bornmann, L. (2017). How many scientific papers are mentioned in policy-related documents? An empirical investigation using Web of Science and Altmetric data. *Scientometrics*, *110*, 1209-1216.

Huang, C., Su, J., Xie, X., Ye, X., Li, Z., Porter, A., & Li, J. (2015). A bibliometric study of China's science and technology policies: 1949–2010. Scientometrics, 102, 1521-1539.

Maleki, A., & Holmberg, K. (2022). Comparing coverage of policy citations to scientific publications in Overton and Altmetric. com: Case study of Finnish research organizations in Social Science. Informaatiotutkimus, 41(2–3), 92-96.

Newson, R., Rychetnik, L., King, L., Milat, A., & Bauman, A. (2018). Does citation matter? Research citation in policy documents as an indicator of research impact–an Australian obesity policy case-study. *Health Research Policy and Systems*, *16*, 1-12.

Overton. How does Overton find citation contexts? Overton. Retrieved December 7, 2024 from: https://help.overton.io/article/how-does-overton-find-citation-contexts/#supported-referencing-styles

Szomszor, M., & Adie, E. (2022). Overton: A bibliometric database of policy document citations. *Quantitative science studies*, *3*(3), 624-650.

Tattersall, A., & Carroll, C. (2018). What can Altmetric. com tell us about policy citations of research? An analysis of Altmetric. com data for research articles from the University of Sheffield. *Frontiers in research metrics and analytics*, *2*, 9.

Yin, Y., Gao, J., Jones, B. F., & Wang, D. (2021). Coevolution of policy and science during the pandemic. *Science*, *371*(6525), 128-130.

Yu, H., Murat, B., Li, J., & Li, L. (2023). How can policy document mentions to scholarly papers be interpreted? An analysis of the underlying mentioning process. *Scientometrics*, *128*(11), 6247-6266.