Scholarly Citations in Canadian Policy Documents

François Claveau, Jérémie Dion, Francis Lareau, Éric Montpetit, Mathieu Ouimet, Louis Renaud-Desjardins, Centre interuniversitaire de recherche sur la science et la technologie

Key messages:

- Overton is an emerging tool to understand the policy impact of academic research
- Federal government reliance on research has been stable over the last decade, with about one-third of policy documents citing scholarly work
- Provincial and territorial government reliance on research is low and falling
- While Canadian scholarly output is 2% of the global total, it represents 30% of the scholarship cited by Canadian policy documents

November 2025

This paper was commissioned by the CCA using funding provided by the Social Sciences and Humanities Research Council (SSHRC) under the Strategic Initiatives Fund.

This paper is one of eight commissioned to support the CCA's 2025 assessment on the State of Science, Technology, and Innovation in Canada by synthesizing evidence and providing new insights into the assessment charge. The expert panel considered each paper alongside many other sources of evidence. These papers are the work of the authors alone and were not approved or endorsed by the expert panel.

This paper is only available in the language of submission (English).
Ce document est uniquement disponible dans la langue dans laquelle il a été soumis (anglais).

Scholarly Citations in Canadian Policy Documents

A report for the Council of Canadian Academies

François Claveau[i,ii] Jérémie Dion[i,iii] Francis Lareau[i,ii],
Éric Montpetit[i,iv] Mathieu Ouimet[i,v] Louis Renaud-Desjardins[i,iii]

Main findings

- In the Canadian public sector, public agencies, such as health agencies and the Bank of Canada, demonstrate the highest attention to recent and international research, with this trend becoming clearer since 2005.
- The City of Toronto and provinces/territories in Canada show very low and declining attention to scholarly research in their policy documents.
- The federal government's attention to research has been relatively stable since 2014, with around 35% of policy documents citing scholarly work, but the age of cited research has increased, indicating less focus on up-to-date findings.
- Although documents on some policy subjects such as the environment tend to pay more attention to scholarly work, the variation in citation propensity is much stronger across sources (municipal, provincial, federal and agencies) than it is across subject categories.
- Around 30% of scholarly citations in Canadian policy documents are to Canadian research, even though Canada is responsible for roughly 2% of scholarly output worldwide.
- These results about attention to science must be interpreted with caution because bibliometric work on Overton (the database used) is still in its infancy and scholarly citations in policy documents are, at best, an imperfect proxy for the policy impact of science.

[i] Centre interuniversitaire de recherche sur la science et la technologie; [ii] Université de Sherbrooke; [iii] Université du Québec à Montréal (UQAM); [iv] Université de Montréal; [v] Université Laval

Corresponding author: François Claveau <francois.claveau@usherbrooke.ca>

This version: 2025-02-28. Do not circulate or cite without consent from CCA or the corresponding author.

1. Introduction

Science plays multiple roles in contemporary societies. One such role is to inform public policy. Calls for science-informed policies are frequent in Canada, but we currently lack a general assessment of the phenomenon. How much is scholarly research used in public policy at different levels of government in Canada? Is the use of research more widespread today than 20 years ago? Are governments using up-to-date research? This report offers tentative answers to these questions by extending standard bibliometric methods to the analysis of policy documents.

In a sample of 430,235 policy documents from Canadian governmental organizations, we find that attention to research is **very low at the City of Toronto**[1] (1% to 2% of citing documents in our data) and that it is **dwindling at the level of provinces & territories**, where an ever-smaller fraction of documents cite scholarly research. **Signals are mixed at the federal level**: in our data, the fraction of citing documents was relatively high in the late 2000s (around 30%) and has slightly increased after this point (to around 35%), but this fraction has been stable since 2014 and the age of cited research has increased significantly over the period. Hence, our data do not indicate increased attention to up-to-date research by the federal government. Finally, public agencies (such as health agencies and the Bank of Canada) are where explicit attention to recent and international research is by far the highest in our data, and this **exceptionalism of public agencies** has only become clearer since 2005.

These results are tentative for two main reasons. First, our large sample of policy documents from the Overton database is not statistically representative of the population of Canadian policy documents. Currently, there is no established way to determine the severity of bias generated by the unsystematic harvesting of policy documents by Overton. Second, even if we had access to all Canadian policy documents, explicit *attention* to research in these documents is, at best, an imperfect proxy for the policy *impact* of science in Canada.

2. Policy impact of research and scholarly citations

The policy impact of scholarly research is assessed with various methods (Boaz et al. 2009; Newson et al. 2018a). As for other types of impact, the most robust attributions of policy impact require labour-intensive data gathering techniques such as document analysis, surveys and interviews (Greenhalgh et al. 2016). To cut down on costs and allow large-scale assessments, many voices now propose a computational approach to policy impact using citation data in policy documents (Drew et al. 2016; Newson et al. 2018b; Szomszor and Adie 2022). This approach is an extension of a well-established method for assessing another type of impact: bibliometrics has, for decades, treated citations among scholarly documents as indicative of relative influence *inside* science (Gingras 2016).

[1] Toronto is the only Canadian city in Overton at the moment. Our results at the municipal level are thus retricted to the most populous Canadian city.

The extension of bibliometrics to scholarly citation patterns in policy documents is promising. Yet, it is crucial to understand its limitations and to recognize that the extent of its usefulness is still largely unknown. On the one hand, the fact that some policy documents cite a scholarly document does not imply that research has influenced public policy (Knott and Wildavsky 1980). Indeed, many explicit references to scholarly research in policy circles are for symbolic use – i.e., research is brought in strategically to support pre-existing positions or to project an image of being informed and proactive on a subject (Pelz 1978; Amara et al. 2004). On the other hand, the lack of academic references in policy documents does not necessarily mean that research did not influence policy. Researchers can impact policy by more direct means, such as briefing a minister directly, or by more indirect means, such as influencing an international organization or an organization such as the Council of Canadian Academies, which are themselves trusted sources for policymakers (Yin et al. 2021). Moreover, the absence of citations to scholarly work in policy documents may also stem from editorial guidelines applicable to certain types of texts, such as policy statements or program descriptions, which may instruct officials to avoid citing external or any sources.

Given these two limitations, we do not know the extent to which policy documents cite all or only the scholarly research that significantly contributed to their production, and whether different types of policy documents conform better to this norm. There are, however, some hints that citations in policy documents and the policy impact of research are related. Szomszor and Adie (2022, pp. 638-642) study the relation between Overton citation counts and policy influence outcomes reported by researchers to UK funding agencies in the context of impact assessments. They find a generally positive correlation ($r = 0.42$) between citation counts in Overton and reported policy influence. They also report wide variations across subject categories – from close to perfect in “Mechanical Engineering” to almost nonexistent in “Food science & nutrition.” These results are restricted to the UK and to a specific method of impact assessment. Much more validation work is required before we can confidently draw general conclusions about the extent to which citation counts in policy documents track real policy impact.

Given the novelty of the approach, we advise caution in the interpretation of our citation analysis below. To avoid misinterpretation, we will refer to scholarly citations in policy documents as being a measure of *attention* in policy circles toward some research and not a measure of impact.

3. Data and methods

This report utilizes data from Overton, a database of policy documents launched in 2019. Overton was founded by Eugene Adie, who also created the Altmetric database. The term "Altmetric" has come to denote new data sources that aim to capture the broader influence of research beyond academia. Aiming to uncover connections between science and other actors, such as the media, policymakers, and firms, these sources are meant to help with research evaluation (Tattersall and Carroll 2018; Orduna-Malea et al. 2017; Thelwall et al. 2013; Kousha and Thelwall 2017; Ravenscroft et al. 2017).

The Overton.io team scrapes the websites of 2400 sources associated with 44,000 organizations identified as producers of policy documents, defined as “materials written by a policy maker or primarily for a policy maker” (Szomszor and Adie 2022, 627). These organizations are classified into the following categories: government, intergovernmental organizations, think tanks, and others. Overton grew by an average of 300 sources annually between 2021 and 2024. The content within Overton is diverse, including legal documents, clinical or technological guidance, parliamentary transcripts, policy reports and more. However, as of early 2025, items in the database are not classified by document type. Since Overton is actively developing document summarization and categorization tools, this limitation might be lifted in the future.

A crucial feature of Overton is that it extracts scholarly citations from policy documents, which is the main reason why it has generated a lot of research since its launch.[2] As a nascent research evaluation tool that has not undergone thorough critical evaluation by the scientometric community, detailed information about Overton's curation and extraction procedures is limited. Overton employs a set of heuristics or "features" to identify potential references and scores the similarity of these character strings against CrossRef entries. The results are calibrated to achieve a minimum accuracy of 98% and a minimum recall of 80% (Overton.io, 2022).

Although Overton covers most countries, we restrict our corpus to Canadian governmental documents with a publication date between 2005 and September 2024. In total, 430,235 documents from 19 sources correspond to these criteria.

Table 1: *Number of documents from Canadian governmental sources (2005-2024)*

Source type	Source	Documents
City of Toronto	City of Toronto	51 672
Provinces & territories	Alberta	1882
Provinces & territories	British Columbia	42 559
Provinces & territories	Manitoba	34 930
Provinces & territories	New Brunswick	7786
Provinces & territories	Newfoundland and Labrador	7414
Provinces & territories	Northwest Territories	17 883
Provinces & territories	Nova Scotia	17 117
Provinces & territories	Nunavut	6019
Provinces & territories	Ontario	15 251
Provinces & territories	Prince Edward Island	4744
Provinces & territories	Saskatchewan	1156
Provinces & territories	Quebec	160 944
Provinces & territories	Yukon	3271
Federal government	Government of Canada	48 177
Public agencies	Bank of Canada	693
Public agencies	Canadian Agency for Drugs and Technologies in Health	5478

2 Research using Overton includes Pinheiro et al. 2021; Yin et al. 2021; Bornmann et al. 2022; Szomszor and Adie 2022; Yin et al. 2022; Malkov et al. 2023; Cabral and Salles-Filho 2024; Dorta-González et al. 2024; Fang et al. 2024.

Public agencies	Cancer Care Ontario	223
Public agencies	Institut national de santé publique du Québec	3036

We group the sources in four types: City of Toronto (n = 51 672), Provinces & territories (n = 320 956), Federal government (n = 48 177) and Public agencies (n = 9 430). Table 1 presents the number of documents by sources. Many signs suggest that this corpus is not a representative sample of Canadian governmental documents in the last two decades. First, it includes documents from only one city, the City of Toronto. Second, although it contains documents from all 13 provinces & territories, the distribution indicates some disproportion in the number of documents from one jurisdiction to the next. A striking case is the contrast between Ontario and Quebec: although the government of the former has a budget roughly a third larger than Quebec, its number of policy documents in Overton is not even a tenth of the latter (15 251 vs. 160 944 documents). Finally, the type 'Public agencies' contains only four sources, three of them being health agencies. This skewed representation is another reason – after the caveat in the previous section regarding the distinction between attention and policy impact – why our results must be interpreted with care.

Each document in Overton is labelled with **subject categories** based on the hierarchical classification provided by the International Press Telecommunications Council (IPTC). We retain the highest-level categories, grouping some together, excluding the least frequent ones and using shorthand names for some categories below.[3] The same document can be labelled with more than one subject category. Subject categories allow us to investigate whether scholarly citations are significantly more frequent on some topics than others, irrespective of the source. The result of this investigation can be used to test a plausible hypothesis: if one policy source cites research significantly moe than another (e.g., public agencies vs. the City of Toronto), it might be because it specializes in a subject that relies heavily on science (e.g., health). Note that subject categories are characteristics of policy documents and not of the cited scholarly documents. While we can expect that policy documents on some topics (e.g., the economy) cite some disciplines more often (e.g., economics), we are not measuring the intensity by which disciplines are *cited*, but rather the percentage of *citing* documents.

Our main variable of interest is thus the **fraction of citing documents**, defined as the ratio of policy documents containing at least one scholarly citation to all policy documents. This fraction can be computed for various subsets: by source type, by subject category, by year of publication, and as an intersection of more than one of these properties. By choosing this variable, we leave out the intensity by which a given policy document cites research, which could be approximated by the number of scholarly sources in a document. Since most policy documents do not cite research at all, it seems more pressing to study when some citations happen, and not the intensity of citations when they occur.

for wide sampling disparities among provinces & territories, we weigh each jurisdiction by its share of governmental expenses[4] when computing values for this source type. Hence, Ontario's value influences

3 Please contact the authors for more information on data preprocessing.
4 Statistics Canada. Table 36-10-0450-01, "General governments expenditure."

the aggregate value of 'Provinces & territories' roughly a third more than Quebec's value, even though it accounts for a lot less documents in Overton than Quebec.

Subject categories are not uniformly distributed across source types. For instance, public agencies are overrepresented in the Health category. To isolate variation in citation propensity attributable specifically to subject categories, we estimated a mixed effect logit model. See the Appendix for details. With this model, we can retrieve estimated probabilities of being a cited document for each subject category while controlling for sources and the interaction between source and subject.

We study two other variables. First, the **age of cited research**, defined as the difference between the publication year of the policy document and the publication year of the scholarly work it cites. To retrieve publication years of scholarly work, we use data from Crossref.[5] Since publication lags have skewed distributions, we use the median lag through time as a measure of central tendency for this variable. Second and finally, we measure the **relative attention to Canadian research**, defined as the proportion of cited scholarly research that has at least one Canadian author. For this task, we use data from Scopus because, in contrast to Crossref, it has reliable information on author affiliations (Visser et al. 2021). However, Scopus has limited coverage, especially of non-English and more local journals (Mongeon and Paul-Hus 2016). By querying Scopus API, we were able to retrieve author information for 64% of the scholarly documents cited by our Canadian policy documents.[6]

4. Results

We present five main results: two results concern how the fraction of citing documents varies across source types and subject categories, two others are about temporal tendencies and the last one is about the attention to Canadian science. We provide supplementary results in the Appendix.

We queried the Crossref API with the Digital object identifier (DOI) provided by Overton. Cases where the lag was negative were interpreted as data errors and excluded from the results (as in Szomszor and Adie 2022). They correspond to less than 1% of cited scholarly documents.

6 Statistics Canada. Table 36-10-0450-01, "General governments expenditure."

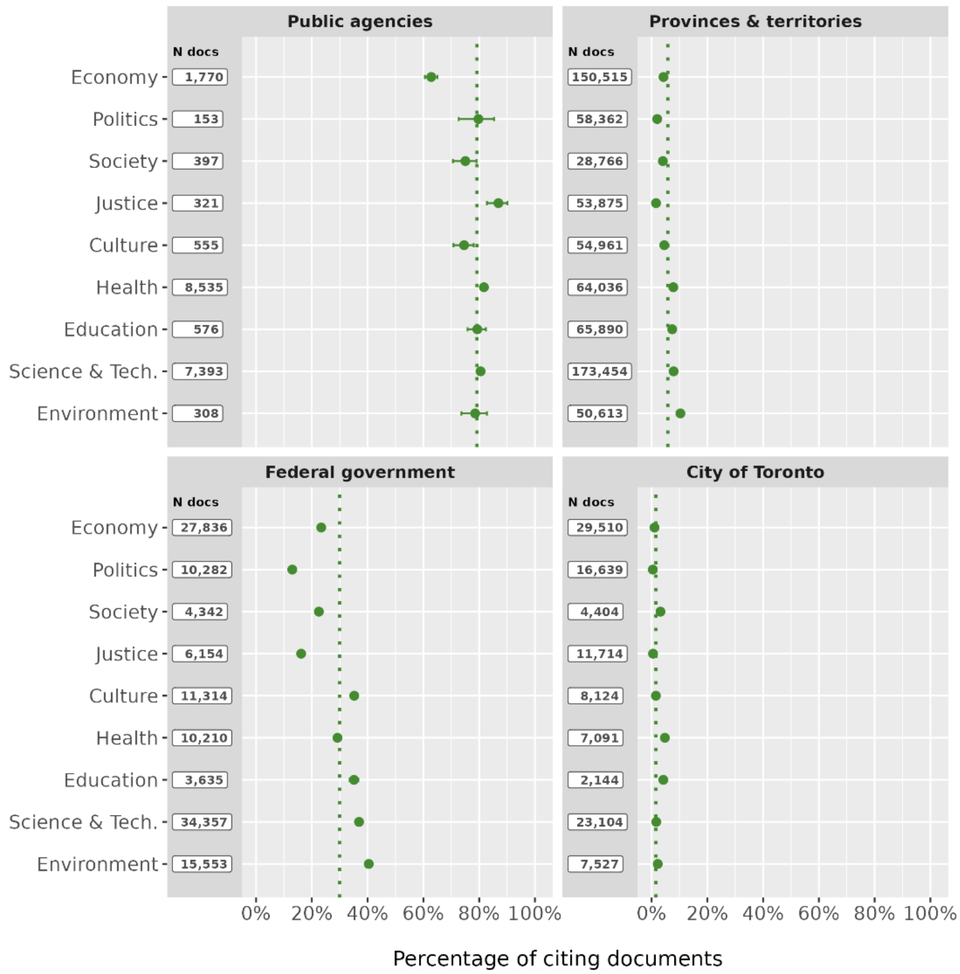

Figure 1: *Percentage of policy documents with at least one scholarly citation. Vertical dotted lines are overall percentages for each source type. Source types are aligned to facilitate the relevant comparisons.*

Citation propensity by source type. Overall, only 8.55% of documents in our Canadian sample cite at least one scholarly source, but this propensity to cite varies widely across the four source types (see Figure 1). Documents from public agencies cite most often (79%), followed by the federal government (30%), the provinces & territories (6%) and, finally, the City of Toronto (2%).[7] This ordering in citation propensity holds across the nine subject categories. In other words, for any topic (e.g., education, the economy, health), the propensity of policy documents to include scholarly citations increases with the level of government (from the municipal to the federal level) and reaches its maximum in public agencies.

7 This ordering is mostly preserved if we take each source separately, see Figure A1 in the Appendix. The four public agencies have citation propensities between 53% (Bank of Canada) and 89% (Cancer Care Ontario), while the province with the highest propensity is Ontario, with only 7%. The only exception to the pattern is the Province of Saskatchewan, which has only 0.7% citing documents, slightly below the City of Toronto at 1.1%. Note that similar results are produced with our logit model, see Figure A2 in the Appendix.

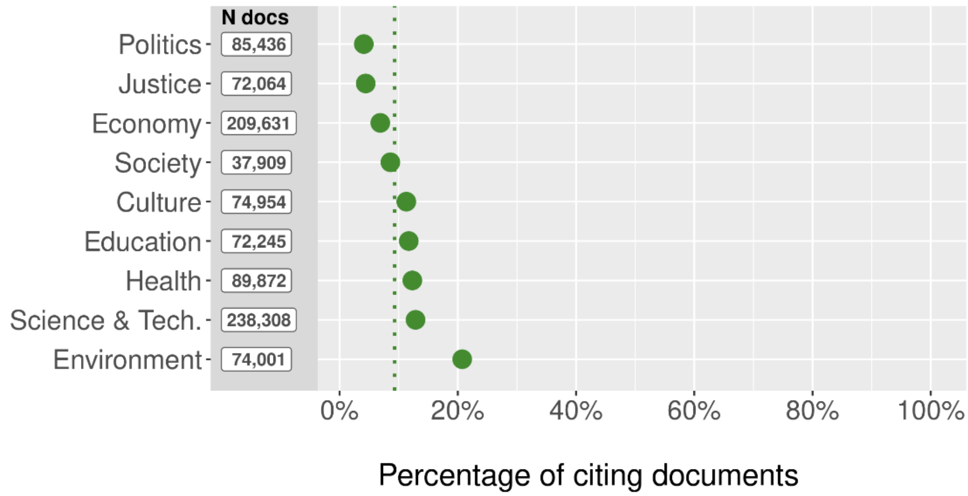

Figure 2: *Percentage of policy documents with at least one scholarly citation. Values for each subject category are estimated with a mixed model logit regression to correct for variations attributed to sources and to the interaction of a source and a subject.*

Citation propensity by subject category. Figure 1 already hints at the fact that the fraction of citing documents varies less across subject categories than it does across source types. Figure 2 makes this feature more evident by reporting the estimated baseline percentages from our mixed model logit regression: the variation from the least citing subject category (Politics) to the most citing category (Environment) is 16.5 percentage points, while it is 77 percentage points between public agencies and the City of Toronto. Among subject categories, only the Environment stands out as being substantially more likely to include scholarly citations.

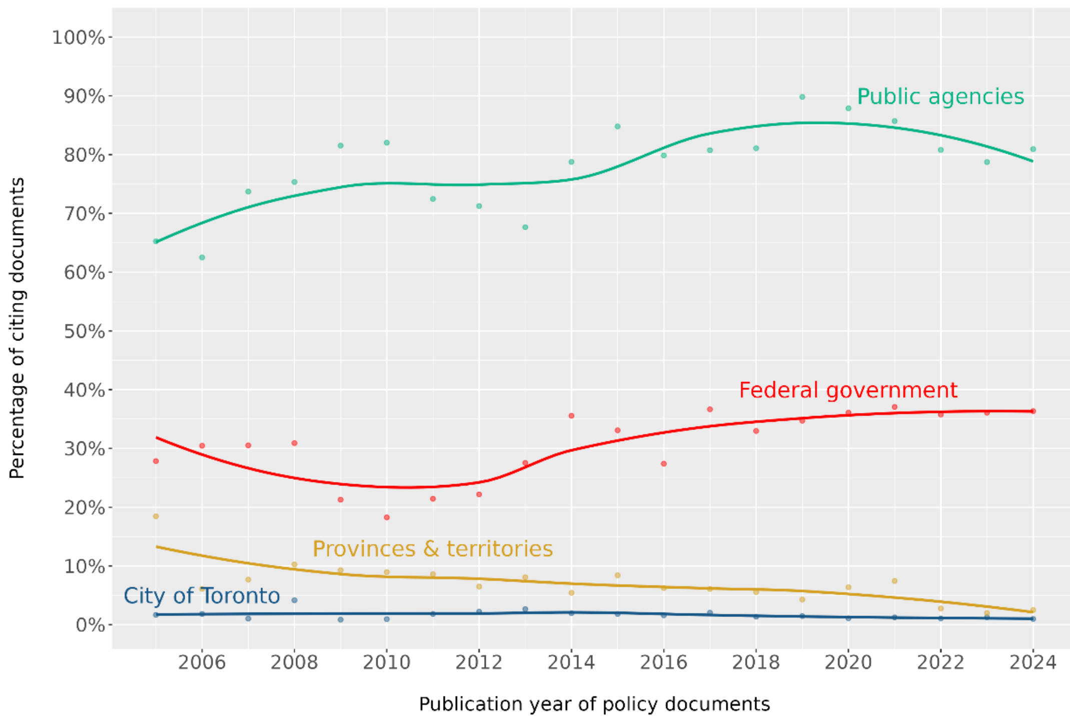

Figure 3: *Percentage through time of policy documents with at least one scholarly citation. Curves are smoothed using local polynomial regression.*

Temporal change in citation propensity. The fraction of citing documents evolves differently from 2005 to 2024 across source types. As Figure 3 demonstrates, public agencies follow a general upward trend from 65% in 2005 to a maximum of 90% in 2019, and a regression to around 80% since then. Documents from the federal government follow a shallow u-shape: from a roughly stable 30% at the beginning of the period, the fraction of citing documents dropped to around 20% around 2010 and then rose again starting in 2013, to stabilize around 35% since 2017. The fraction of citing documents for provinces & territories has trended down over the entire two decades. At the end of the period, the value of this source type is barely higher than the City of Toronto, which is stable around 1% to 2%.

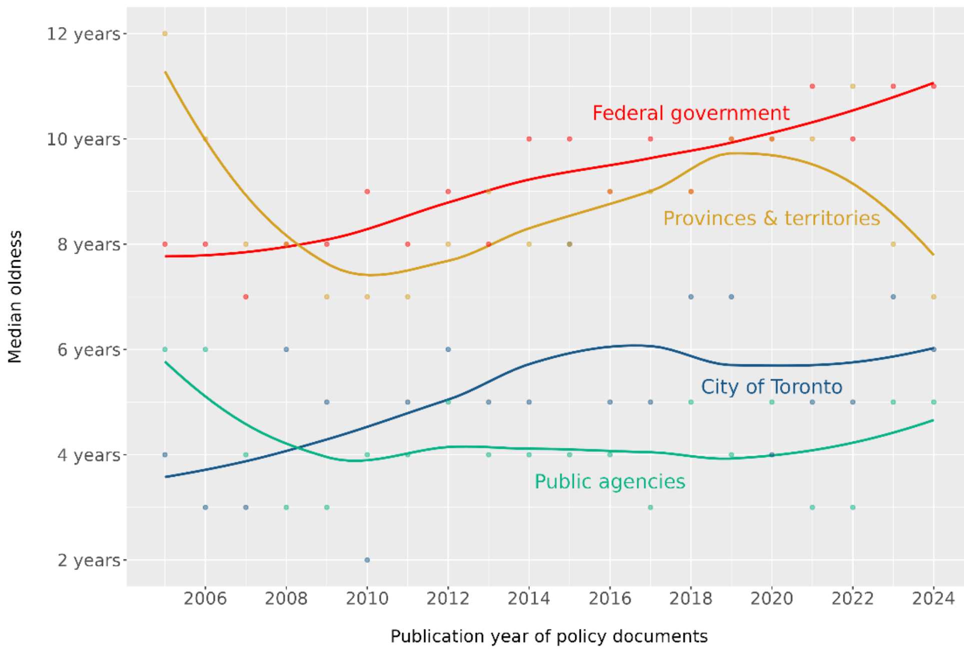

Figure 4: *Median age of cited scholarly sources. Curves are smoothed using local polynomial regression.*

Temporal change in the age of cited research. The median lag between the publication year of a policy document and the publication year of its cited scholarly sources has a distinctive temporal shape for each source type (see Figure 4). Two source types – the City of Toronto and public agencies – have relatively low citation oldness throughout the two decades, and two source types – provinces & territories and, again, public agencies – have decreasing oldness over the period. It thus leaves the federal government as the source type with a relatively high median age of citations at the beginning of the period (8 years) and the strongest increase over the two decades, with 3 more years added to its median age. At the other extreme, public agencies stand out over most of the period as being the source type most attentive to recent research.

Table 2: *Percentage of scholarly citations with at least one Canadian author for different source types. Only cited documents found on Scopus are considered.*

Source type	Number of scholarly citations	With Canadian authorship
Provinces & territories	113,221	29%
Federal government	168,923	34%
City of Toronto	3621	32%
Public agencies	96,323	16%

Table 3: *Percentage of scholarly citations with at least one Canadian author for different subject categories. Only cited documents found on Scopus are considered.*

Subject category	Number of scholarly citations	With Canadian authorship

Environment	147,099	41%
Culture	73,982	39%
Science & Tech.	328,984	30%
Politics	11,313	30%
Society	20,875	28%
Economy	119,363	27%
Education	44,286	25%
Justice	15,390	24%
Health	193,342	22%

Attention to Canadian research. Among the scholarly citations in our corpus that are present in Scopus, 30% feature at least one Canadian author. Tables 2 and 3 breakdown the percentages of scholarly citations with Canadian authorship by source type and subject category. Public agencies are roughly half as likely as the three levels of government to cite domestic research (16% vs. values between 29% and 34%). Domestic research is most prevalent on the subject of the Environment (41%) and least on Health (22%). All of these fractions are far higher than the estimated share of scientific publications of Canada, which is around 2% of the world (National Science Board 2023).

5. Discussion

The results in the previous section are suggestive of key characteristics of the attention given to scholarly research by Canadian policy circles. In this section, we discuss these characteristics while highlighting areas of uncertainty and suggesting further research.

Public agencies give far more attention to research than the three levels of government in Canada: federal, provincial and municipal. Our results support this claim, which is rather intuitive to begin with. Indeed, one main mission of public agencies, such as public health bodies or central banks, is to establish the state of knowledge in their domain of expertise – either to advise government or to enact policies – which nowadays necessitate an explicit reliance on science (Carpenter 2010; Claveau and Dion 2018).

Public agencies possibly play the role of conduits of science toward the three levels of government, especially regarding international research. Yin et al. (2021) found that intergovernmental organizations (IGOs) such as the World Health Organization (WHO) play such a role: they cite scholarly work directly, but national and subnational governments tend to rely on IGOs instead. By making the policy impact of research indirect, IGOs and, perhaps, public agencies exemplify the point made earlier that the absence of scholarly citations at one level does not imply a lack of influence of science on policy. Part of this indirect impact could be captured in future work with Overton data, by investigating the policy documents that are cited by a source with an otherwise low propensity to cite research.

In our data, the federal government cites research 5 times more than the provinces & territories and 15 times more than the city of Toronto. Although the ordering could be expected by extrapolating the result of Yin et al. (2021) to a trickle down from national to more local governments, the extent of the difference is striking. Part of this difference may be attributed to a composition problem in the documents attributed to the Government of Canada: some units, such as the Pest Management Regulatory Agency, which operates under the authority of Health Canada, are subsumed under the 'federal government' although they have advisory and regulatory mandates, which make them akin to 'public agencies,' and not to federal departments. If an important fraction of documents attributed to the federal government are, in fact, the product of public agencies, the measured citation propensity of the federal government would be artificially raised.

Another possible interpretation of the difference across levels of government is that the culture of science advice was perhaps promoted earlier and more strongly at the federal level. A typical reference point is the 1999 report "Science Advice for Government Effectiveness (SAGE)" that provided "guidance on how to ensure that government decisions are informed by sound science advice." (CSTA 1999, 1) If such federal efforts to increase attention to science were effective, it would explain part of the important gap in citation propensity between the federal and more local levels when our period started in 2005.

Another hypothesis for the significant variation in citation propensities across different source types is that research is perceived as more relevant for certain policy subjects. Sources specialized in these subject areas would there have a higher proportion of citing documents. While we observe some variation across subject categories (Fig. 2), these differences are considerably smaller compared to those between source types. Consequently, our findings do not support the hypothesis that subject specialization accounts for most of the variation in citation propensities.[8]

Temporal changes in our data are also noteworthy. A spontaneous interpretation of the u-shaped evolution of the federal government (Fig. 3) is to invoke ruling parties: with the Conservative Party ruling between 2006 and 2015, and the Liberal Party being in power for the rest of the period. Indeed, the Conservative rule led to an open confrontation between the Harper government and members of the scientific community, culminating in the 2012 "Death of evidence" march by 2000 scientists in Ottawa (Nature 2012; O'Hara and Dufour 2014; Quirion et al. 2016).

However, our results cannot be interpreted as an indication that attention to scholarly research is generally stronger under a Liberal government. Indeed, the fraction of citing documents started increasing again in 2013, well before the change in the ruling party, and it has been relatively stable since 2014. This stability during the last decade occurred despite some very visible Liberal actions to promote science for policy, such as the reinstatement of a Chief Science Advisor in 2017 (Henderson 2017). Similarly, the Province of Quebec follows the general downward trend in citation propensity among provinces & territories despite the existence since 2011 of a Chief Scientist (Courchesne 2011, art. 67). The lack of correlation between these efforts to structure science advice and our indicator of attention to scholarly research requires further investigation.

8 It is also possible that there is a strong interaction between sources and subject categories, such that the nature of, say, 'Health' documents at the provincial level is fundamentally different from health documents at the federal level, but both levels would produce more similar documents (and thus more similar citation propensities) for other subjects, say, 'Justice.' However, our logit regression finds a weak variance for the interaction terms between sources and subjects. Consequently, our results do not support the view that much of the variation in citation propensities comes from the source-specific variations of some subjects.

The age of cited scholarly work also indicates that attention to scholarly research in the federal government is not steadily improving. In 2005, the median age of citations at this level was close to the stable median lag around 7 years that Szomszor and Adie (2022, Fig. 8) found for the full Overton corpus. It is now well above this median (at 11 years). Among scientists, citation aging is associated with later career stages, when authors allegedly read less of the recent literature while still citing what they perceive as classics (Gingras et al. 2008, Fig. 3). Similarly, the significant citation aging of the federal government means that it has shifted its relative attention away from recent research. The opposite is true for public agencies, with a consistently low median age over the period. These agencies are clearly where contemporary research is most scrutinized.

In sum, Overton's data suggest that public agencies pay more attention to recent and international scholarly research than ministries, which is not surprising given differences in their mission. The federal government pays more attention to research than the provinces and municipalities, and this difference increases over time. That said, these results must be interpreted with great caution, as Overton does not provide us with representative samples of governmental documents. Moreover, the database does not include information on the very different types of documents that the same governmental organization produces. Some types of documents call for scholarly citations (e.g., Health Technology Assessments), while it might be explicitly forbidden to cite research in other documents (e.g., policy statements). Further research on how to disentangle the document types is necessary. Finally, if citations are a valid indicator of attention, governmental attention to scholarly research does not necessarily mean that research impacts policy.

References

Amara, Nabil, Mathieu Ouimet, and Réjean Landry. 2004. "New Evidence on Instrumental, Conceptual, and Symbolic Utilization of University Research in Government Agencies." *Science Communication* 26 (1): 75–106. https://doi.org/10.1177/1075547004267491.

Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. "Fitting Linear Mixed-Effects Models Using lme4." *Journal of Statistical Software* 67 (1): 1–48. https://doi.org/10.18637/jss.v067.i01.

Boaz, Annette, Siobhan Fitzpatrick, and Ben Shaw. 2009. "Assessing the Impact of Research on Policy: A Literature Review." *Science and Public Policy* 36 (4): 255–70. https://doi.org/10.3152/030234209X436545.

Bornmann, Lutz, Robin Haunschild, Kevin Boyack, Werner Marx, and Jan C. Minx. 2022. "How Relevant Is Climate Change Research for Climate Change Policy? An Empirical Analysis Based on Overton Data." *PLOS ONE* 17 (9): e0274693. https://doi.org/10.1371/journal.pone.0274693.

Cabral, Bernardo, and Sergio Salles-Filho. 2024. "Mapping Science in Artificial Intelligence Policy Development: Formulation, Trends, and Influences." *Science and Public Policy*, September, scae052. https://doi.org/10.1093/scipol/scae052.

Carpenter, Daniel. 2010. *Reputation and Power: Organizational Image and Pharmaceutical Regulation at the FDA*. Princeton, NJ: Princeton University Press.

Claveau, François, and Jérémie Dion. 2018. "Quantifying Central Banks' Scientization: Why and How to Do a Quantified Organizational History of Economics." *Journal of Economic Methodology* 25 (4): 349–66. https://doi.org/10.1080/1350178X.2018.1529216.

Courchesne, Michelle. 2011. *Projet de Loi N° 130, Loi Abolissant Le Ministère Des Services Gouvernementaux et Mettant En Œuvre Le Plan d'action 2010-2014 Du Gouvernement Pour La Réduction et Le Contrôle Des Dépenses En Abolissant et En Restructurant Certains Organismes et Certains Fonds*. https://www.publicationsduquebec.gouv.qc.ca/fileadmin/Fichiers_client/lois_et_reglements/LoisAnnuelles/fr/2011/2011C16F.PDF.

CSTA. 1999. "Science Advice for Government Effectiveness (SAGE)." Ottawa: Council of Science and Technology Advisors Secretariat, Industry Canada.

Dorta-González, Pablo, Alejandro Rodríguez-Caro, and María Isabel Dorta-González. 2024. "Societal and Scientific Impact of Policy Research: A Large-Scale Empirical Study of Some Explanatory Factors Using Altmetric and Overton." *Journal of Informetrics* 18 (3): 101530. https://doi.org/10.1016/j.joi.2024.101530.

Drew, Christina H., Kristianna G. Pettibone, Fallis Owen Finch, Douglas Giles, and Paul Jordan. 2016. "Automated Research Impact Assessment: A New Bibliometrics Approach." *Scientometrics* 106 (3): 987–1005. https://doi.org/10.1007/s11192-015-1828-7.

Fang, Zhichao, Jonathan Dudek, Ed Noyons, and Rodrigo Costas. 2024. "Science Cited in Policy Documents: Evidence from the Overton Database." arXiv.Org. July 13, 2024. https://arxiv.org/abs/2407.09854v1.

Gingras, Yves. 2016. *Bibliometrics and Research Evaluation: Uses and Abuses*. Cambridge, MA: The MIT Press.

Gingras, Yves, Vincent Larivière, Benoît Macaluso, and Jean-Pierre Robitaille. 2008. "The Effects of Aging on Researchers' Publication and Citation Patterns." *PLOS ONE* 3 (12): e4048. https://doi.org/10.1371/journal.pone.0004048.

Greenhalgh, Trisha, James Raftery, Steve Hanney, and Matthew Glover. 2016. "Research Impact: A Narrative Review." *BMC Medicine* 14 (1): 78. https://doi.org/10.1186/s12916-016-0620-8.

Henderson, Mark. 2017. "CSA Appointment Lauded but Debate Swirls over Scope and Breadth of Advisory Role." Research Money Inc. 2017. https://researchmoneyinc.com/article/csa-appointment-lauded-but-debate-swirls-over-scope-and-breadth-of-advisory-role.

Knott, Jack, and Aaron Wildavsky. 1980. "If Dissemination Is the Solution, What Is the Problem ?" *Knowledge* 1 (4): 537–78. https://doi.org/10.1177/107554708000100404.

Kousha, Kayvan, and Mike Thelwall. 2017. "Are Wikipedia Citations Important Evidence of the Impact of Scholarly Articles and Books?" *Journal of the Association for Information Science and Technology* 68 (3): 762–79. https://doi.org/10.1002/asi.23694.

Malkov, Dmitry, Ohid Yaqub, and Josh Siepel. 2023. "The Spread of Retracted Research into Policy Literature." *Quantitative Science Studies* 4 (1): 68–90. https://doi.org/10.1162/qss_a_00243.

Mongeon, Philippe, and Adèle Paul-Hus. 2016. "The Journal Coverage of Web of Science and Scopus: A Comparative Analysis." *Scientometrics* 106 (1): 213–28. https://doi.org/10.1007/s11192-015-1765-5.

National Science Board. 2023. "Publications Output: U.S. Trends and International Comparisons." Science and Engineering Indicators 2024. Alexandria, VA: National Science Foundation. https://ncses.nsf.gov/pubs/nsb202333/publication-output-by-region-country-or-economy-and-by-scientific-field.

Nature. 2012. "Death of Evidence." *Nature* 487 (7407): 271–72. https://doi.org/10.1038/487271b.

Newson, Robyn, Lesley King, Lucie Rychetnik, Andrew Milat, and Adrian Bauman. 2018. "Looking Both Ways: A Review of Methods for Assessing Research Impacts on Policy and the Policy

Utilisation of Research." *Health Research Policy and Systems* 16 (1): 54. https://doi.org/10.1186/s12961-018-0310-4.
Newson, Robyn, Lucie Rychetnik, Lesley King, Andrew Milat, and Adrian Bauman. 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): 55. https://doi.org/10.1186/s12961-018-0326-9.
O'Hara, Kathryn, and Paul Dufour. 2014. "How Accurate Is the Harper Government's Misinformation? Scientific Evidence and Scientists in Federal Policy Making." In *How Ottawa Spends, 2014-2015: The Harper Government - Good to Go?*, edited by G. Bruce Doern and Christopher Stoney, 178–91. McGill-Queen's University Press. https://doi.org/10.1515/9780773584983-014.
Orduna-Malea, Enrique, Mike Thelwall, and Kayvan Kousha. 2017. "Web Citations in Patents: Evidence of Technological Impact?" *Journal of the Association for Information Science and Technology* 68 (8): 1967–74. https://doi.org/10.1002/asi.23821.
Overton.io. 2022. "Characteristics of Overton's Data." April 8, 2022. https://www.overton.io/blog/characteristics-of-overtons-data.
Pelz, Donald C. 1978. "Some Expanded Perspectives on Use of Social Science in Public Policy." In *Major Social Issues: A Multidisciplinary View*, edited by J. M. Yinger and S. J. Cutler, 346-357. New York: Free Press. https://cir.nii.ac.jp/crid/1571980075532987136.
Pinheiro, Henrique, Etienne Vignola-Gagné, and David Campbell. 2021. "A Large-Scale Validation of the Relationship between Cross-Disciplinary Research and Its Uptake in Policy-Related Documents, Using the Novel Overton Altmetrics Database." *Quantitative Science Studies* 2 (2): 616–42. https://doi.org/10.1162/qss_a_00137.
Quirion, Remi, Arthur Carty, Paul Dufour, and Ramia Jabr. 2016. "Reflections on Science Advisory Systems in Canada." *Palgrave Communications* 2 (1): 1–10. https://doi.org/10.1057/palcomms.2016.48.
Ravenscroft, James, Maria Liakata, Amanda Clare, and Daniel Duma. 2017. "Measuring Scientific Impact beyond Academia: An Assessment of Existing Impact Metrics and Proposed Improvements." *PLOS ONE* 12 (3): e0173152. https://doi.org/10.1371/journal.pone.0173152.
Szomszor, Martin, and Euan Adie. 2022. "Overton: A Bibliometric Database of Policy Document Citations." *Quantitative Science Studies* 3 (3): 624–50. https://doi.org/10.1162/qss_a_00204.
Tattersall, Andy, and Christopher Carroll. 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 (January). https://doi.org/10.3389/frma.2017.00009.
Thelwall, Mike, Stefanie Haustein, Vincent Larivière, and Cassidy R. Sugimoto. 2013. "Do Altmetrics Work? Twitter and Ten Other Social Web Services." *PLOS ONE* 8 (5): e64841. https://doi.org/10.1371/journal.pone.0064841.
Visser, Martijn, Nees Jan van Eck, and Ludo Waltman. 2021. "Large-Scale Comparison of Bibliographic Data Sources: Scopus, Web of Science, Dimensions, Crossref, and Microsoft Academic." *Quantitative Science Studies* 2 (1): 20–41. https://doi.org/10.1162/qss_a_00112.
Yin, Yian, Yuxiao Dong, Kuansan Wang, Dashun Wang, and Benjamin F. Jones. 2022. "Public Use and Public Funding of Science." *Nature Human Behaviour* 6 (10): 1344–50. https://doi.org/10.1038/s41562-022-01397-5.
Yin, Yian, Jian Gao, Benjamin F. Jones, and Dashun Wang. 2021. "Coevolution of Policy and Science during the Pandemic." *Science* 371 (6525): 128–30. https://doi.org/10.1126/science.abe3084.

Appendix

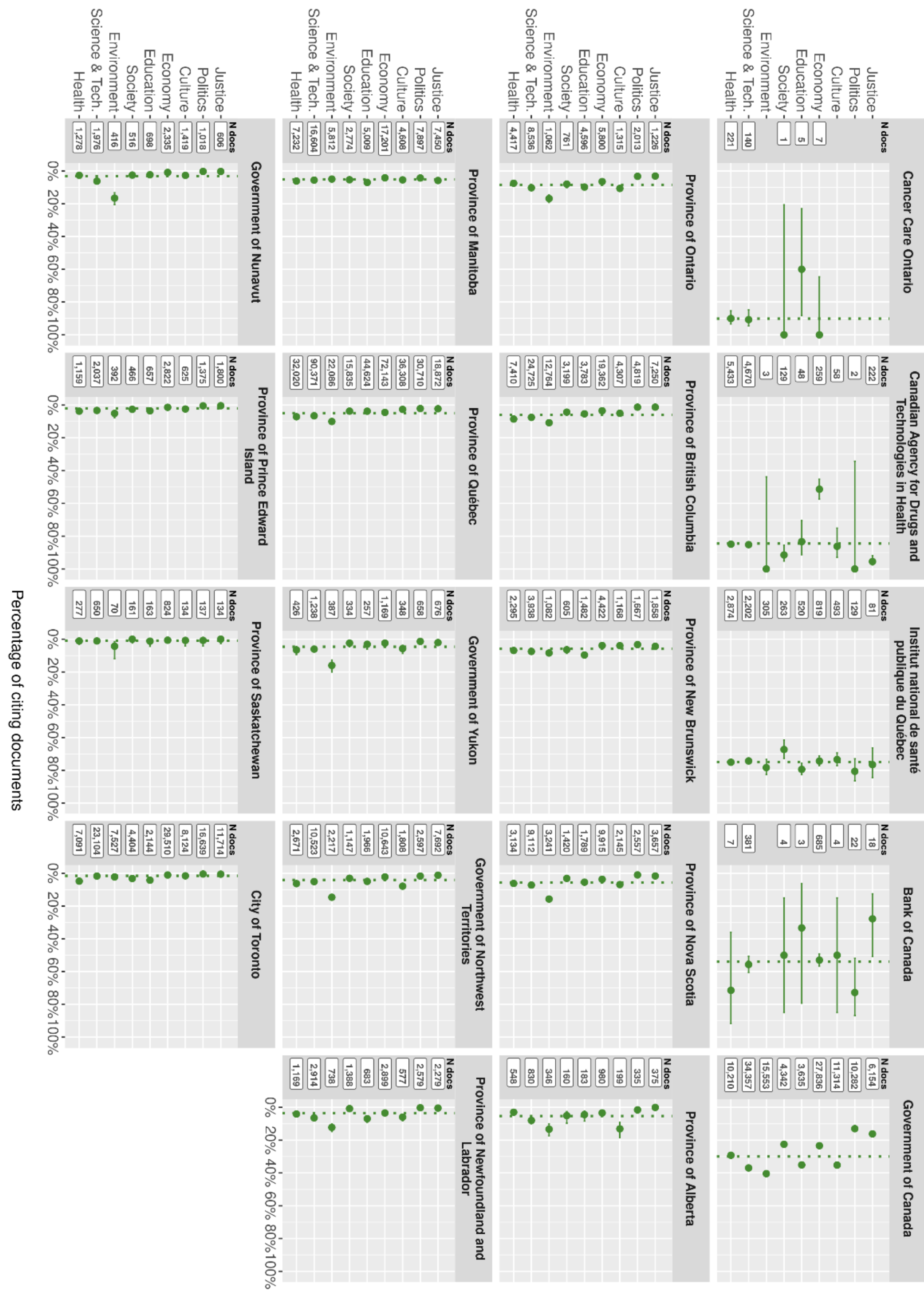

Figure A1: *Percentage of policy documents with at least one scholarly citation. Vertical dotted lines are overall percentages for each source. See Figure 1 for a version aggregated by source type.*

Figure 2 is based on estimates of citation propensity from a generalized linear mixed model with a binomial distribution and logit link function. The model includes random intercepts for the 19 sources (u_i), 9 subject categories (v_j), and 166 interactions between sources and subjects (w_{ij}). Formally, the model is:

$$\text{logit}(\ P(\text{citing}_{ijk} = \text{True})) = \beta_0 + + u_i + v_j + w_{ij}$$

where citing$_{ijk}$ represents whether document k (produced by source i and categorized as subject j) cites a scholarly document or not. The random effects are assumed to follow normal distributions:

$$u_i \sim N(0, \sigma^2_{\text{source}}),\ \ \text{vj} \sim N(0, \sigma^2_{\text{subject}}), \text{ and } w_{ij} \sim N(0, \sigma^2_{\text{source:subject}}).$$

The model was fitted using maximum likelihood estimation (Laplace approximation) with the "`glmer`" function in the lme4 R package (Bates et al. 2015). We have 954,420 observations since the same document can have more than one theme. The model summary is in Table A2.

Table A2: *Summary of the Generalized Linear Mixed Model*

Effect	Estimate	Std. Error	Variance
Fixed Effect			
Intercept	-2.280	0.046	—
Random Effects			
Source	—	—	3.935
Subject	—	—	0.356
Source × Subject	—	—	0.252

The fixed effect intercept corresponds to a baseline citation probability of approximately 9.3%. The random effect variances indicate substantially more variation across sources ($\sigma^2 = 3.935$) than across subjects ($\sigma^2 = 0.356$). The interaction (source × subject; $\sigma^2 = 0.252$) indicates some additional source-specific differences across themes.

With this model, we then estimate the mean effects for each subject category (using function "ranef" in lme4). Each effect is on the log-odds scale. We combine it to the estimate for the fixed effect (i.e., the baseline of -2.280, also on a log-odds scale) before converting the value to a probability of being a citing document for each subject category.

The same estimation and conversion can be made for the sources, which gives us a probability of being a citing document for each source, see Figure A2. These estimates corroborate our result in Figure 1 to the effect that there is a wide difference between citation propensities, with Toronto and Provinces & territories very low, and public agencies very high.

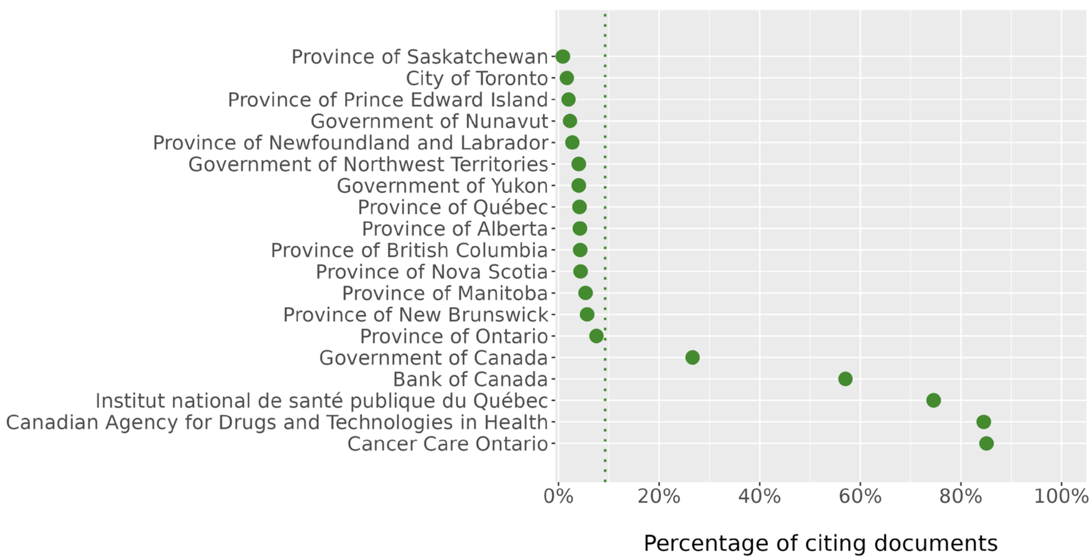

Figure A2: *Percentage of policy documents with at least one scholarly citation. The value for each source is estimated with a mixed model logit regression to correct for variations attributed to subject categories and to the interaction of a source and a subject.*