Kyle Higham, Bernardo Buarque
and Troy Baisden

# Evidence-Based Policymaking and Public Management

## emerging empirical approaches

**Abstract**

This article demonstrates how emerging data sources and analytical tools can be applied to better understand evidence-based policymaking and its relationship to public sector capabilities. By analysing policy documents and their citations, we show how these methods can explore uses of evidence in policy processes, highlight gaps in knowledge integration, and evaluate the balance between local and international research inputs. Using New Zealand environmental policy as a case study, we show how these tools may be applied to complex policy areas, with broader implications for public sector decision making.

**Keywords** evidence-based policymaking, research–policy interface, knowledge flows, environmental science, policy citations

**Kyle Higham**, **Bernardo Buarque** and **Troy Baisden** are all affiliated with Te Pūnaha Matatini Centre of Research Excellence, University of Auckland, and Motu Economic and Public Policy Research, Wellington.

Evidence-based policymaking is widely seen as a way to improve the quality of government decisions by grounding them in research and empirical analysis. The idea is straightforward: rather than relying on political intuition, ideology or tradition, policymakers should turn to systematic evidence to guide their choices. However, evidence must be interpreted, weighed against competing interests, and translated into actionable policy, and even when high-quality research is available, policymakers face institutional, political and practical constraints that shape whether and how it is used. As such, effective evidence-based policymaking depends crucially on the analytical capabilities and institutional capacities of the public sector, so understanding and measuring how evidence is actually used can directly inform efforts to support these capabilities.

In practice, evidence-based policymaking takes several forms. For example, it can be evaluative in nature: which policies worked and which did not? Other modes are more forward-looking, such as those that focus on compiling advice and policies in a more exploratory

fashion: what form should new policies take? For the purposes of this article, all forms of evidence-based policymaking are formulated based on information available to the policymaker – the 'evidence'.

Now, the phrase 'information available' is doing a lot of heavy lifting here, and its definition is crucial to understanding New Zealand's evidence-based policymaking landscape. First, this information should typically include retrospective analysis in the domestic context. Second, New Zealand policymakers are rarely the first to think about implementing a particular policy; we often have access to a wealth of academic and policy research in international contexts that speaks to potential impacts. Finally, policymakers can commission research from experts to fill any obvious gaps in their understanding. This latter source of information is vital to provide insight into how policy may interact with New Zealand-specific social, cultural, economic or environmental contexts, and even more so in policy domains of high complexity and uncertainty, and when conditions can rapidly change (Saul et al., 2013; Head and Alford, 2015; Manning et al., 2015).

Despite its appeal, evidence-based policymaking is often misused. Sometimes this is innocent: policymakers may lack the technical expertise to assess the reliability of competing studies (Newman, Cherney and Head, 2017), or may otherwise be pushed towards expediency rather than rigour by various structural and political incentives and constraints (Lindblom, 1959). However, a guise of evidence-based policymaking can also be used as a weapon to implement politically motivated policy change, whereby research is selectively used to justify decisions rather than inform them (Weiss, 1979; Strassheim and Kettunen, 2014; Parkhurst, 2017), sometimes referred to as 'policy-based evidence-making'. Such practices are often easy to spot, such as citations of articles in predatory journals (Brandts-Longtin et al., 2022; Albert, Lalu and Grudniewicz, 2025). More pernicious variations include a 'funding effect' of support from policy agencies (Krimsky, 2012), the prioritisation of easily measurable outcomes at the expense of either harder to quantify but equally important factors (Sanderson, 2002), such as social equity or ethical considerations, or those that are only measurable at long time horizons (Lindblom, 1959; Boston, 2017).

This article addresses what we perceive to be a deep irony: while evidence-based policymaking proponents call for policy to be based on evidence, there is surprisingly little research on how evidence is actually used in policymaking. The result of this fact is that much of what we assume about evidence-based policymaking remains theoretical or anecdotal, with limited real-world validation outside of a small number of qualitative works. This lack of research leaves open questions on both the policy and research ends of the policymaking process: do policies informed by scientific evidence actually perform better than those driven by other considerations? Which types of research are most influential? To what extent do different government agencies rely on internal expertise versus external scientific research? Do co-production and co-design initiatives lead to more research-driven policy decisions? This short article does not answer these questions; instead, we aim to provide an introduction to new analytical tools and data sources that may be used to do so in the future. After briefly discussing the current state of empirical research on evidence-based policymaking practice, we explore the use of policy documents and their citations as a source of data for future research in this area. Finally, we demonstrate how this data is being used in practice, via initial explorations of environmental science policy that form part of a larger Te Pūnaha Matatini-funded research project.

… while evidence-based policymaking proponents call for policy to be based on evidence, there is surprisingly little research on how evidence is actually used in policymaking.

### The evidence for evidence-based policymaking

Here, we review the methods being used to study evidence-based policymaking in practice: that is, how can we quantify the input of research findings into policy decisions? Of course, the most direct way is to ask the decision makers. Surveys of this kind can provide useful insight into how policymakers perceive and use scientific evidence, including the barriers they face in accessing or applying it. While such findings often lack generalisability in isolation, a recurring finding is that policymakers express strong support for evidence-based policymaking in principle, but often struggle to integrate this evidence into policy, for a variety of reasons (Oliver et al., 2014; Cairney, 2016). However, a key limitation of survey-based approaches is that they usually measure attitudes rather than behaviour. That is, policymakers may claim to use or perceive evidence in a certain way, but whether they actually do is another question.

To move beyond self-reported data and directly observe policymaker behaviour, some researchers have turned to experimental methods (Kertzer and Renshon, 2022; Haaland, Roth and Wohlfart, 2023). Such an experiment may, for example, test whether policymakers who receive well-designed, policy-relevant research are more likely to incorporate evidence into their decisions (Callen et al., 2020; Hjort et al., 2021; Lee, 2022; Toma and Bell, 2024). However, as one might imagine, experiments like this are extremely sensitive to the context in which they are carried out, and thus the results are often complex and nuanced. In the above

example, political alignment (Furnas, LaPira and Wang, 2024), cognitive biases (Banuri, Dercon and Gauri, 2019) and the timing of research dissemination (Rose et al., 2020) all influence whether and how evidence is used. As such, while experiments can provide valuable (and potentially causal) insights, design limitations mean they usually focus on immediately observable outcomes in relatively narrow domains, rather than broader patterns of institutional behaviour.

> … the increasing availability of structured policy document datasets represents a major opportunity for assessing evidence-based policymaking from an empirical perspective.

A more recent approach, enabled by advances in computational tools and data collection, involves the systematic analysis of policy documents themselves. For example, resources like Overton (Szomszor and Adie, 2022) allow researchers to track how academic research is cited in government reports, white papers and legislative documents. Unlike surveys, which rely on subjective reporting, and experiments, which are often limited in scope, policy document analysis provides much more direct, if nuanced (Yu et al., 2023), evidence of research uptake. Most pertinently, policy documents are widely accessible and primarily composed of text, making them amenable to natural language processing: text provides data that is the basis of empirical analysis (Grimmer and Stewart, 2013; Gentzkow, Kelly and Taddy, 2019).

### Policy documents as data

Policy documents serve as both a record of government decision making and a reflection of the evidence that informs it, making them a valuable resource for studying the integration of research into policy. While policy documents contain non-textual elements such as graphs, important takeaways from these elements are usually elaborated on within the text. Further, while the automatic extraction of quantitative insights from text is an established area of study (Ash and Hansen, 2023), doing the same for non-textual elements at scale is much more challenging and is a field in its infancy as a general tool (Davila et al., 2020; Farahani et al., 2023; Huang et al., 2024). As such, the remainder of this article will focus on the systematic extraction of information from the text of policy documents for the purposes of developing a quantitative understanding of how research evidence is incorporated into policy.

The most explicit way to link policy documents to the evidence base on which they depend is to identify any citations they make to other documents. Such an exercise will result in a network of links between a given policy document and the broader literature in which it sits, which can include academic articles, commissioned reports and other policy documents (Bornmann, Haunschild and Marx, 2016; Szomszor and Adie, 2022). Of course, these cited documents may have links of their own to previous work, which link to others, and so on down the citation chain. The result of this iterative linking process is a large, complex web of connected ideas and the people who channel them (De Solla Price, 1965; Sorenson, Rivkin and Fleming, 2006; Phelps, Heidl and Wadhwa, 2012).

On top of this network structure sits metadata that describes the links and nodes (i.e., documents). In addition to the text of the documents themselves, this metadata includes authors, institutions and topics, from which we can derive a whole set of new citation networks between these entities. That is, a citation between documents is also a citation between the authors, institutions and topics of those documents. These derived networks can provide insights into, for example, how central they are to the flow of information in the network (Burt, 2004; Sandström and Carlsson, 2008; Yan and Ding, 2009) or how the networks evolve with time (Barabási et al., 2002; Powell et al., 2005). Analysing the text around the citation can also give us context (Nicholson et al., 2021; Yu et al., 2023), such as when it refers to supporting evidence or a counterargument.

In sum, citation networks are very useful for understanding the flow of information in a system, and are amenable to the tools of network science (Coscia, 2021). As such, once this data structure is in place, we can start to ask targeted, domain-specific questions: which voices shape policy debates? Are certain perspectives or research communities systematically over- or under-represented in decision-making processes? Longitudinal analysis of citations in regular government reports offers another intriguing avenue for studying science–policy interactions, and such an analysis is demonstrated later in this article.

From a computational point of view, citation linkages are generally easy to identify. However, it is also possible to use other properties of documents to infer relationships, such as text similarity (Vijaymeena and Kavitha, 2016; Arts, Cassiman and Gomez, 2018). These methods are very useful for comparing the content of a small, predefined set of documents, but such comparisons can quickly become very computationally expensive. A common way around this problem is to use machine-learning methods to 'embed' or 'project' documents into a low-dimensional space wherein similar documents are closer in this space (Le and Mikolov, 2014; Devlin et al., 2018; Morris et al., 2023). Readers may be familiar with such projections in other contexts, such as the distillation of political views onto a two-dimensional 'political compass' on which distance reflects similarity in those views.[1] Notably, these techniques avoid many of the limitations of citations, such as the assumption that all citation links are meaningful.

Once documents are embedded, one can analyse them using spatial metrics, such as distance and density (Kusner et al., 2015; Shibayama, Yin and Matsumoto, 2021; Ganguli et al., 2024). Dynamic patterns in the output of authors or organisations can be studied within the space describing this evolving body of literature. For example, we may see the set of policy documents produced by two ministries start to become more or less similar over time as government priorities change.

In summary, the increasing availability of structured policy document datasets represents a major opportunity for assessing evidence-based policymaking from an empirical perspective. By applying new analytical techniques to these sources, we can start to move beyond theoretical discussions and begin systematically measuring how evidence is used in governance, while also uncovering the limitations of policy documents as data.

### Exemplar: environmental science policy and greenhouse gas inventories

In this section, we will present examples from our research on environmental policy in New Zealand to illustrate the use of policy documents as a data source and its implications for public sector capabilities. We note that this policy area was chosen as this article is based on ongoing work within a larger Te Pūnaha Matatini-funded project with this focus. The base data for the following analyses is Overton, a large database of citations made by policy documents to both science and other policy documents. We collected from this database 1,406 policy documents (hereafter 'base documents') published between 1998 and 2023 by three New Zealand governmental institutions working on topics related to environmental policy – the Ministry for Primary Industries, the Ministry for the Environment and the parliamentary commissioner for the environment (hereafter 'base institutions'). This triad captures the enhancement of outputs of the primary industries, policies monitoring and regulating environmental impacts, and the independent 'watchdog' agency charged with reporting on the management of the environment. In addition to citations made by these documents, we also obtain the publication dates, titles and sources of those cited documents, resulting in a rich dataset with which to start to measure the policy and science diets of these three institutions that are used to develop their public policy. We shall look at the characteristics of policy and science citations in turn.

### Citations to policy

In total, our base documents cite 1,701 other policy documents, which we can dissect to obtain insights into the sources policymakers turn to for information. These insights are presented below as a series of stylised facts, the implications of which we will discuss thereafter.

First, Table 1 shows that each institution makes many citations to other documents that they produced, and to documents produced by either of the other two environmental policy institutions: 40% of all citations fall into one of these categories, and primarily the former. This is not surprising: one would expect New Zealand's new environmental policy to build on (or at least reference) existing New Zealand environmental policy.

Interesting patterns exist among those citations made between institutions, however, with the Ministry for the Environment citing many parliamentary commissioner for the environment and Ministry for Primary Industries documents, but receiving very few in return. These patterns are displayed in Table 2. Indeed, the Ministry for the Environment is not even among the top 20 most cited sources of the other two: knowledge flows into the ministry but does not appear to flow out, at least according to the citation record.

Second, 60% of the cited policy documents do not originate with our base institutions. We will group these other institutions according to country and type. The bulk of the cited institutions in this set are intergovernmental organisations (IGOs), such as the Organisation for Economic Co-operation and Development (OECD) or the Intergovernmental Panel on Climate Change (IPCC), producing

**Table 1: Unique citations to other policy documents according to country and type**

| Country | Type | No. of docs | No. of DOIs cited | DOIs per doc. |
|---|---|---|---|---|
| IGO | IGO | 373 (43.4%) | 95072 (88.2%) | 254.9 |
| New Zealand | Government | 137 (15.9%) | 2061 (1.9%) | 15 |
| USA | Government | 68 (7.9%) | 2135 (2.0%) | 31.4 |
| UK | Government | 49 (5.7%) | 807 (0.7%) | 16.5 |
| USA | Think tank | 47 (5.5%) | 1359 (1.3%) | 28.9 |
| New Zealand | Think tank | 46 (5.4%) | 474 (0.4%) | 10.3 |
| EU | Government | 37 (4.3%) | 977 (0.9%) | 26.4 |
| UK | Think tank | 19 (2.2%) | 199 (0.2%) | 10.5 |
| France | Think tank | 17 (2.0%) | 2538 (2.4%) | 149.3 |
| Australia | Government | 17 (2.0%) | 526 (0.5%) | 30.9 |

Note: digital object identifiers (DOIs) are unique codes that link to specific digital objects; for our purposes, these refer to published academic research (including preprints).

**Table 2: Top ten most cited policy sources, by institution**

| Ministry for the Environment | | Ministry for Primary Industries | | Parliamentary commissioner for the environment | |
|---|---|---|---|---|---|
| MfE | 486 (39.7%) | MPI | 144 (31.2%) | OECD | 25 (20.0%) |
| OECD | 69 (5.6%) | WHO | 52 (11.3%) | PCE | 19 (15.2%) |
| WHO | 52 (4.2%) | FAO | 28 (6.1%) | IPCC | 11 (8.8%) |
| PCE | 51 (4.2%) | DOC | 18 (3.9%) | Treasury | 7 (5.6%) |
| EPA | 42 (3.4%) | IPCC | 17 (3.7%) | CCC[1] | 6 (4.8%) |
| DOC | 38 (3.1%) | Motu | 15 (3.3%) | OP EU[2] | 6 (4.8%) |
| Motu | 37 (3.0%) | MOH | 12 (2.6%) | UNEP | 5 (4.0%) |
| Treasury | 34 (2.8%) | FSANZ[3] | 11 (2.4%) | WHO | 5 (4.0%) |
| IPCC | 33 (2.7%) | OECD | 10 (2.2%) | IUCN[4] | 4 (3.2%) |
| MPI | 32 (2.6%) | IUCN | 9 (2.0%) | IEA | 3 (2.4%) |

[1] Climate Change Commission [2] Publications Office of the European Union [3] Food Standards Australia New Zealand
[4] International Union for Conservation of Nature

43% of all cited documents not originating with our base institutions. A further 16% cite documents produced by other New Zealand government institutions, such as the Department of Conservation or Treasury. Policy produced by foreign government organisations make up 22%, while independent organisations such as Motu produce the remaining 18%.

We can also compare the policy citations across the three institutions to highlight obvious differences and similarities. For example, the OECD is the most cited non-base source of policy information for the parliamentary commissioner for the environment (20%) and the Ministry for the Environment (5%), but is only the ninth most cited non-base source for the Ministry for Primary Industries (2%). Perhaps unsurprisingly, the Food and Agriculture Organization of the United Nations (FAO) is the second most cited non-base source for the Ministry for Primary Industries (6%), but very rarely cited by the other two base institutions. Some sources, like the IPCC and the World Health Organization (WHO), are frequently cited by all three institutions.

Altogether, these results provide an overview of knowledge influences and potential sharing across governmental agencies in New Zealand and their relationship to other international authorities, highlighting the capabilities and specialisations across these institutions. They also provide a good sanity check for the use of citation data in the first place: the relative ranking of different policy knowledge sources appears to make intuitive sense.

The simple explorations above reveal the potential for a network-analytic approach to studying public management. Understanding how knowledge flows around a policy system and how to measure these flows in a useful way can, for example, highlight unexpected gaps in knowledge transfer between government agencies or identify central actors in this knowledge exchange network.

We also note that knowledge flow in policy networks is a chronically understudied topic, at least from an empirical standpoint; even with policy citation data in its infancy, most research that uses this information is focused on ties between policies and published science. While these links are important, we suspect that they are less likely to have influenced the direction of policy than other policy documents or commissioned research: this is a topic of ongoing research within the larger project of which this article is a part. We now describe basic patterns in the published science cited by our base institutions.

**Table 3: Top ten most cited journals and institutions (all base institutions, unique citations)**

| Journal | | Institution | |
|---|---|---|---|
| *New Zealand Journal of Marine and Freshwater Research* | 910 (27.6%) | NIWA | 957 (0.28%) |
| *Journal of Food Protection* | 486 (14.7%) | University of Auckland | 470 (0.14%) |
| *Marine Ecology Progress Series* | 281 (8.5%) | University of Otago | 449 (0.13%) |
| *International Journal of Food Microbiology* | 274 (8.3%) | Manaaki Whenua | 259 (0.08%) |
| *Science* | 262 (7.9%) | University of Canterbury | 258 (0.08%) |
| *Applied Environmental Microbiology* | 241 (7.3%) | AgResearch | 257 (0.07%) |
| *Chemosphere* | 223 (6.8%) | Massey University | 214 (0.06%) |
| *Science of the Total Environment* | 208 (6.3%) | Victoria University of Wellington | 207 (0.06%) |
| *Marine Pollution Bulletin* | 208 (6.3%) | Ministry of Health | 201 (0.06%) |
| *PLOS One* | 204 (6.2%) | Lincoln University | 166 (0.05%) |

**Table 4: Top ten most cited countries**

| Country | Unique institutions | Unique DOIs | Unique citations per institution |
|---|---|---|---|
| United States | 1042 | 4780 | 4.6 |
| New Zealand | 76 | 4118 | 54.2 |
| Great Britain | 301 | 1586 | 5.3 |
| Australia | 161 | 1421 | 8.8 |
| Canada | 126 | 825 | 6.5 |
| Germany | 179 | 539 | 3 |
| France | 233 | 484 | 2.1 |
| Netherlands | 66 | 464 | 7 |
| Spain | 136 | 463 | 3.4 |
| Japan | 151 | 430 | 2.8 |

### Citations to science

The pathways between research and policies are complex, and citation data offers several perspectives on these pathways. We can look at the published research that is directly referenced by the base institutions, and Overton identifies 23,252 unique digital object identifiers (DOIs) (which act as identifiers for published research) cited in this way. However, research is often embedded in other policy documents. This is particularly true for documents produced by IGOs: if we only look at those IGO policy documents cited by the base institutions, we find over 92,000 unique DOIs cited. Indeed, IGO policy documents cited about 16 times more published science than the average base document: most of the science that is being incorporated into New Zealand policy actually comes indirectly via external policy references.

Despite the volume citations made by the IGO documents, most of the science cited by the base institutions is not duplicated from those documents. This implies that there is significant complementarity between the knowledge they source directly and that coming indirectly from the IGOs. Indeed, compiling a comprehensive evidence base for decision making is costly, so it pays to source it from other policy relevant institutions. However, there are unique social and cultural features

of New Zealand that are likely to be outside the IGO's field of view, so local agencies may need to source additional evidence to ensure policy relevance for a local context. Such an arrangement represents a sensible division of labour, given the limited resources for policymaking in the New Zealand context.

So, where do New Zealand environmental policymakers source their knowledge from? We link Overton's extracted DOIs to an open bibliometric database, OpenAlex (Priem, Piwowar and Orr, 2022), to obtain metadata for these articles to map these sources. Aggregated statistics for all citations from base documents at the journal, institution and country level are displayed in Tables 3 and 4.

Contrary to what one might expect based on the local 'context' argument above, the plurality of citations made by the base institutions are to US-based researchers, though New Zealand researchers are a close second. This is perhaps unsurprising, given the sheer volume of research produced by the United States, and, all else equal, any given research article from New Zealand is still far more likely to be cited than articles produced elsewhere. Further, this finding is partially an artefact of the research/policy document dichotomy in the Overton database; four of the five most cited US institutions are actually government agencies publishing in peer-reviewed journals. However, none of these institutions crack the top ten most cited institutions by the base institutions – these are all New Zealand-based, highlighting the importance of local sources of scientific knowledge for environmental policy.

Beyond aggregated statistics across all base documents, we can also compare information sources between the three base institutions. One way to do this is to simply look at the overlap of the academic journals, researchers and institutions cited by each of these institutions. Because each of these entities is often relatively specialised, such an analysis may provide insight into the academic domains from which research evidence is being sourced.

Figure 1 shows the results of this analysis. We find that there are strong correlations between the journals and institutions cited by the Ministry for the Environment and the parliamentary commissioner for the

**Figure 1: The shares of (A) journals, (B) institutions and (C) authors cited by the institution shown in the row labels (left), that are also cited by the institution shown in the column labels (bottom)**
For example, 48% of journals cited by MfE were also cited by MPI, and 36% of those cited by MPI were also cited by MfE

environment, but the Ministry for Primary Industries' sources are more distinct. This finding makes complete sense, as it reflects the overlap in policy domains under the remit of each of these institutions. The same pattern of overlaps (or lack thereof) is found at the cited researcher level, noting that researchers cannot produce papers at the same rate as an entire journal or institution, and thus citations are spread more evenly across researchers.

Combined with the differences in policy sources referenced by each institution, we can start to build a picture not only of the knowledge sourcing practices of individual government agencies, but also of complementarities in these knowledge sources and the potential for more efficient sharing of knowledge and inter-agency collaboration. Similar ideas have been explored extensively in other contexts, such as technological change (Hidalgo et al., 2007), industrial policy (Foray, 2014) and scientific research (Larivière et al., 2016). The wealth of information becoming available in the policy domain makes it possible to lean on this prior work to map knowledge flows both within policy circles and across the science–policy interface. We are optimistic that such a map could provide insight into improved knowledge management practice between and within government agencies to enable efficient evidence-based policymaking, and, ultimately, enable these agencies to make the best policy decisions possible with the resources available.

### Longitudinal opportunities: greenhouse gas inventories

One possible strategy to cut through the noise brought by the different publication and citation practices among the policy documents in the sample is to zoom in and focus more narrowly. We focussed on New Zealand's Greenhouse Gas Inventory (GGI), a well-referenced, annual publication parallel to those of other nations meeting requirements of the United Nations Framework Convention on Climate Change (UNFCCC). The national GGI reports serve a clear purpose, documenting emissions in ways that can support emissions-reduction policies.

Thus, the GGI offers a recurrent publication with consistent periodicity and citation standards. It offers a well-structured reference list from which most off-the-shelf reference extraction algorithms can extract citations to a high degree of accuracy.

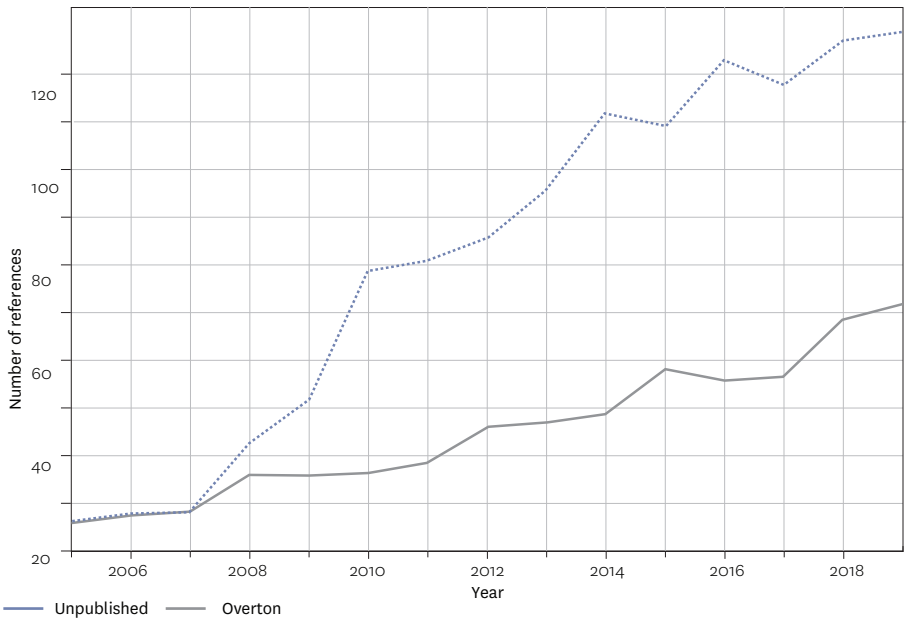
For the purpose of this demonstration, we identified 16 GGI published by the New Zealand Ministry for the Environment from 2005 to 2020, from which we can extract citations and quantify dynamics of knowledge inputs over this period. Noting that most citations are carried over between subsequent reports, dynamics that are easily studied include the number and diversity of scientific sources, the prevalence of new sources, and the longevity of cited evidence. A convenience of the New Zealand GGIs is that they are in an analytical 'sweet spot': they provide enough data to run various informative quantitative analyses, yet are small enough for their accuracy to be entirely manually validated in a timely manner.

While our focus lies in the New Zealand corpus, another advantage of using the GGIs is that they offer the opportunity for direct international comparison under a standardised reporting framework. That is, since the publication of GGIs is an obligation under the Paris Agreement, there are equivalent documents published by other countries, each with their own set of citations to science. The UNFCCC library provides access to annual reports back to 2003 from over 40 countries. Cross-country comparisons can, for example, measure the similarity between citations (or the dynamics thereof) across countries, or even how often one nation sources from the other, perhaps offering a window into the quantification of international policy knowledge spillovers.

The GGIs are usually several chapters long, and each chapter has a reference list with at least 20 publications, yet Overton retrieves only 84 unique citations to published science and 36 unique citations to policy. This is less than expected for such a science-dependent set of documents. The reason for this is that much of the science input into these reports is commissioned, and these commissioned reports are often not published, even in the ministry's digital library. When they are published, there is no formal link from these publications back to the report for which they were commissioned. Figure 2 displays the number of unique citations made by each GGI in the corpus, split by their presence in Overton (i.e., whether they have been formally published) as a function of time; most citations are unpublished and thus not present in bibliometric databases.

In other words, for the majority of citations in the GGIs, all we can know about them is what we are able to garner from the citation and its associated entry in the reference list. Fortunately, we are able to use semi-automated tools to mine and retrieve as much information as possible

**Figure 2: References cited in each annual GGI report divided between those Overton matched to a DOI (grey), indicating formally peer-reviewed journal articles and similar, versus references to reports that we classify as 'unpublished' or 'commissioned work' (dotted blue line)**



about the authors and institutions responsible for the unpublished commissioned work; this work is ongoing within our broader research programme. Preliminary results from this exercise suggest that consistent commissioning of work from a set of experts within New Zealand effectively supports GGI-related policies, while also raising questions about the dependence of GGIs on grey-literature reports that remain difficult to link to more formal and systematic publications.

Such practices have clear implications for the reliability and transparency of the government's evidence-based policymaking processes in general, especially when those who set policy are also able to choose the experts who inform it. Further, while the ability to commission work that is timely and informed by local contexts is extremely valuable and should be part of effective evidence-based policymaking, these commissions and their scope have been carefully tailored to the requirements of the policymaker, which opens the door to policy-based evidence making. Finally, commissioned work is often carried out by organisations that have their own financial incentives and relationships to actors who may be affected by the policies under consideration. Even if we assume that the resulting research is of the highest quality, the fact that policymakers are able to choose the research that is commissioned – and what is not – is certainly a concern when the majority of the research informing policy is sourced in this way. These problems do not have straightforward solutions, but we suggest that independent reviews of major policy documents are a good first step. Incidentally, such a practice is already enshrined in the UNFCCC in the case of GGIs,[2] which may constitute a model on which to build more transparent evidence-based policymaking in other policy areas.

### Limitations of policy citation data

While citations extracted from policy documents clearly have some informational value, they also have many limitations, some inherent and some that may be overcome in time.

The inherent limitations are primarily due to the data generation process: that is, citation practice within government. There are several ways in which this affects the data available, including inconsistent citation practices, motivated citation, and lack of access to resources. In short, mechanisms that determine what is cited (and what is not) are generally opaque, and this fact places a natural limit on the informational content of citations.

However, there are also limitations that can be addressed with further research, standardisation and technological advances. These limitations currently place restrictions on the kinds of research that can be undertaken, at least in a systematic way. For example, the accessibility of policy documents varies by country, which hampers international comparisons of knowledge sourcing for policy. Standardisation or centralisation of policy documents (as implemented in the United Kingdom, for example, via https://www.gov.uk/) can partially address this issue.

Accessibility of policy documents also varies drastically between government agencies in the same country, and New Zealand is no exception to this. Some agencies have central repositories for their policy documents, while others do not; further, some agencies that have central repositories do not add all their policy documents to that repository or do so in a timely manner. There is also inconsistency in the different types of documents produced by agencies, which can make it difficult to determine what counts as 'policy', and how to weigh the importance of a three-page briefing against a 200-page annual report.

Empirical complications are also myriad. For example, there exist obvious statistical outliers in the citation data: 70% of the base documents that cite more than 100 DOIs were published by the Ministry for Primary Industries, many of which are serial reports such as fish assessment reports, aquatic environment and biodiversity annual reviews, and risk profiles related to potential food contaminants. Indeed, citations made by these series lead directly to the *New Zealand Journal of Marine and Freshwater Research* and the *Journal of Food Protection* being the top two cited journals across all base documents (Table 3).

Other limitations are more technical, such as those related to the automated extraction of the citations from the documents. This process is complicated by differences in citation style, whether a bibliography is present, citations only found in footnotes, etc. Finding solutions to these issues is mostly up to the data providers – the scale of extraction at the global level presents unique challenges.

There is … inconsistency in the different types of documents produced by agencies, which can make it difficult to determine what counts as 'policy', and how to weigh the importance of a three-page briefing against a 200-page annual report.

Finally, one limitation stands out as being inherently problematic and particularly relevant to the New Zealand context: commissioned research. These documents do not have a DOI and thus cannot be linked to standard bibliometric databases and their rich metadata. Further, most are not accessible to the public. These documents cannot be ignored, however, as they represent key evidentiary inputs into the policymaking process and the most direct interactions between researchers and policymakers, while often also making up a large fraction of the citations made in the resulting policy document. When these commissioned reports are co-produced by researchers and

policymakers, we also miss valuable and increasingly common interactions that do not neatly fit a linear model of policymaking (Strassheim and Kettunen, 2014).

In summary, while policy documents contain much valuable information about the research that informs them, we must be aware of and account for their numerous limitations. For researchers, this means transparency about analytical methods is vital to ensure appropriate and consistent interpretation of results. From a policymaking point of view, should governments want to make the best use of their policy documents as a tool for evaluating policy development, accessibility of these documents and some minimal standardisation of citation practice would make extraction and analysis of data significantly easier and more reliable.

### Looking forward

Tools such as Overton make efforts to follow the use of evidence in policymaking dramatically more transparent. This matters in areas of complex evidence, such as environmental science, which may require input from many disparate fields of research and cross-agency collaboration. Our preliminary analyses find that patterns of policy citations vary widely between public environmental policy agencies in New Zealand, in ways that suggest that evidence-based policymaking can be made more uniform, transparent and effective. The Greenhouse Gas Inventory reports provide a case study detailing the imperfect process of weighing and improving the quality of evidence while limiting sources of bias, and may serve as a model for other policy challenges. This model may also be used to understand the use of evidence in policy domains that experience more unstable and contested policy cycles, such as freshwater management, and also for larger-scale analyses to assess the historic role of new public management in the evolution of public sector capability in the context of knowledge sourcing.

Further, the variation in citation patterns across different government agencies (see, e.g., Figure 1 and Table 2) might reflect genuine differences in policy domains; however, these differences also raise the possibility of sector-specific biases or even regulatory capture. While our quantitative analysis alone cannot confirm such biases, it suggests that complementary qualitative investigation is warranted to assess decision-making processes surrounding the search for, and citation of, relevant evidential inputs into evidence-based policymaking, particularly to determine the extent to which this search is framed by predetermined policy positions.

Current and future work within the present research programme explores the extraction and use of grey literature citations (i.e., commissioned research) and applies network analytic tools to policy citation networks to explore the flow of knowledge at the science–policy interface in detail. We also note that new tools, such as the New Zealand Research Information System that is currently in development, will significantly augment the amount of information available to understand how policy-relevant knowledge is produced in New Zealand by linking this knowledge production to research funding, and make transparent the valuable social contributions of the researchers who undertake this work.

1 For example, https://votecompass.tvnz.co.nz/
2 https://unfccc.int/resource/docs/cop8/07a02.pdf#page=15

### References


Albert, M.A., M.M. Lalu and A. Grudniewicz (2025) 'Investigating the trustworthiness of research evidence used to inform public health policy: a qualitative interview study on the use of predatory journal citations in policy documents', *Health Research Policy and Systems*, 23 (1), 7

Arts, S., B. Cassiman and J.C. Gomez (2018) 'Text matching to measure patent similarity', *Strategic Management Journal*, 39 (1), pp.62–84

Ash, E. and S. Hansen (2023) 'Text algorithms in economics', *Annual Review of Economics*, 15 (1), pp.659–88

Banuri, S., S. Dercon and V. Gauri (2019) 'Biased policy professionals', *World Bank Economic Review*, 33 (2), pp.310–27

Barabási, A.-L., H. Jeong, Z. Néda, E. Ravasz, A. Schubert and T. Vicsek (2002) 'Evolution of the social network of scientific collaborations', *Physica A: Statistical mechanics and its applications*, 311 (3–4), pp.590–614

Bornmann, L., R. Haunschild and W. Marx (2016) 'Policy documents as sources for measuring societal impact: how often is climate change research mentioned in policy-related documents?', *Scientometrics*, 109, pp.1477–95

Boston, J. (2017) 'Protecting long-term interests in a short-term world: an agenda for better governmental stewardship', *New Zealand Journal of Public and International Law*, 15 (1), pp.93–121

Brandts-Longtin, O., M.M. Lalu, E.A. Adie, M.A. Albert, E. Almoli, F. Almoli, G.L. Bryson, C. Dony, D. Dunleavy, A. Grudniewicz et al. (2022) 'Assessing the impact of predatory journals on policy and guidance documents: a cross-sectional study protocol', *BMJ Open*, 12 (4), e059445

Burt, R.S. (2004) 'Structural holes and good ideas', *American Journal of Sociology*, 110 (2), pp.349–99

Cairney, P. (2016) *The Politics of Evidence-based Policy Making*, Springer

Callen, S., Gulzar, A. Hasanain, M.Y. Khan and A. Rezaee (2020) 'Data and policy decisions: experimental evidence from Pakistan', *Journal of Development Economics*, 146, 102523

Coscia, M. (2021) 'The atlas for the aspiring network scientist', *arXiv preprint*, arXiv:2101.00863

Davila, K., S. Setlur, D. Doermann, B.U. Kota and V. Govindaraju (2020) 'Chart mining: a survey of methods for automated chart analysis', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43 (11), pp.3799–819

De Solla Price, D.J. (1965) 'Networks of scientific papers: the pattern of bibliographic references indicates the nature of the scientific research front', *Science*, 149 (3683), pp.510–5

Devlin, J., M.-W. Chang, K. Lee and K. Toutanova (2018) 'BERT: pre-training of deep bidirectional transformers for language understanding', *arXiv preprint^*, arXiv:1810.04805

Farahani, A.M., P. Adibi, M.S. Ehsani, H.-P. Hutter and A. Darvishy (2023) 'Automatic chart understanding: a review', *IEEE Access*, 11, pp.76202–21

Foray, D. (2014) *Smart Specialisation: opportunities and challenges for regional innovation policy*, Routledge

Furnas, A.C., T.M. LaPira and D. Wang (2024) 'Partisan disparities in the use of science in policy', working paper, Institute for Policy Research, Northwestern University

Ganguli, I., J. Lin, V. Meursault and N.F. Reynolds (2024) 'Patent text and long-run innovation dynamics: the critical role of model selection', working paper 32934, National Bureau of Economic Research

Gentzkow, M., B. Kelly and M. Taddy (2019) 'Text as data', *Journal of Economic Literature*, 57 (3), pp.535–74

Grimmer, J. and B.M. Stewart (2013) 'Text as data: the promise and pitfalls of automatic content analysis methods for political texts', *Political Analysis*, 21 (3), pp.267–97

Haaland, I., C. Roth and J. Wohlfart (2023) 'Designing information provision experiments', *Journal of Economic Literature*, 61 (1), pp.3–40

Head, B.W. and J. Alford (2015) 'Wicked problems: implications for public policy and management', *Administration and Society*, 47 (6), pp.711–39

Hidalgo, C.A., B. Klinger, A.L. Barabási and R. Hausmann (2007) 'The product space conditions the development of nations', *Science*, 317 (5837), pp.482–7

Hjort, J., D. Moreira, G. Rao and J.F. Santini (2021) 'How research affects policy: experimental evidence from 2,150 Brazilian municipalities', *American Economic Review*, 111 (5), pp.1442–80

Huang, J., H. Chen, F. Yu and W. Lu (2024) 'From detection to application: recent advances in understanding scientific tables and figures', *ACM Computing Surveys*, 56 (10), 261

Kertzer, J.D. and J. Renshon (2022) 'Experiments and surveys on political elites', *Annual Review of Political Science*, 25 (1), pp.529–50

Krimsky, S. (2012) 'Do financial conflicts of interest bias research?', *Science, Technology, and Human Values*, 38 (4), pp.566–87

Kusner, M., Y. Sun, N. Kolkin and K. Weinberger (2015) 'From word embeddings to document distances', in F. Bach and P. Blei (eds), *Proceedings of the 32nd International Conference on Machine Learning*, Lille: JMLR

Larivière, V., N. Desrochers, B. Macaluso, P. Mongeon, A. Paul-Hus and C.R. Sugimoto (2016) 'Contributorship and division of labor in knowledge production', *Social Studies of Science*, 46 (3), pp.417–35

Le, Q. and T. Mikolov (2014) 'Distributed representations of sentences and documents', in E. Xing and T. Jebara (eds), *Proceedings of the 31st International Conference on Machine Learning*, JMLR

Lee, N. (2022) 'Do policy makers listen to experts? Evidence from a national survey of local and state policy makers', *American Political Science Review*, 116 (2), pp.677–88

Lindblom, C.E. (1959) 'The science of muddling through', *Public Administration Review*, 19 (2), pp.79–88

Manning, M., J. Lawrence, D.N. King and R. Chapman (2015) 'Dealing with changing risks: a New Zealand perspective on climate change adaptation', *Regional Environmental Change*, 15, pp.581–94

Morris, J., X.V. Kuleshov, V. Shmatikov and A.M. Rush (2023) 'Text embeddings reveal (almost) as much as text', *arXiv preprint*, arXiv:2310.06816

Newman, J., A. Cherney and B.W. Head (2017) 'Policy capacity and evidence-based policy in the public service', *Public Management Review*, 19 (2), pp.157–74

Nicholson, J.M., M. Mordaunt, P. Lopez, A. Uppala, D. Rosati, N.P. Rodrigues, P. Grabitz and S.C. Rife (2021) 'scite: a smart citation index that displays the context of citations and classifies their intent using deep learning', *Quantitative Science Studies*, 2 (3), pp.882–98

Oliver, K., S. Innvar, T. Lorenc, J. Woodman and J. Thomas (2014) 'A systematic review of barriers to and facilitators of the use of evidence by policymakers', *BMC Health Services Research*, 14

Parkhurst, J. (2017) *The Politics of Evidence: from evidence-based policy to the good governance of evidence*, Taylor and Francis

Phelps, C., R. Heidl and A. Wadhwa (2012) 'Knowledge, networks, and knowledge networks: a review and research agenda', *Journal of Management*, 38 (4), pp.1115–66

Powell, W.W., D.R. White, K.W. Koput and J. Owen-Smith (2005) 'Network dynamics and field evolution: the growth of interorganizational collaboration in the life sciences', *American Journal of Sociology*, 110 (4), pp.1132–205

Priem, J., H. Piwowar and R. Orr (2022) 'OpenAlex: a fully-open index of scholarly works, authors, venues, institutions, and concepts', *arXiv preprint*, arXiv:2205.01833

Rose, D.C., N. Mukherjee, B.I. Simmons, E.R. Tew, R.J. Robertson, A.B. Vadrot, R. Doubleday and W.J. Sutherland (2020) 'Policy windows for the environment: tips for improving the uptake of scientific knowledge', *Environmental Science and Policy*, 113, pp.47–54

Sanderson, I. (2002) 'Evaluation, policy learning and evidence-based policy making', *Public Administration*, 80 (1), pp.1–22

Sandström, A. and L. Carlsson (2008) 'The performance of policy networks: the relation between network structure and network performance', *Policy Studies Journal*, 36 (4), pp.497–524

Saul, J.E., C.D. Willis, J. Bitz and A. Best (2013) 'A time-responsive tool for informing policy making: rapid realist review', *Implementation Science*, 8, pp.1–15

Shibayama, S., D. Yin and K. Matsumoto (2021) 'Measuring novelty in science with word embedding', *PLOS One*, 16 (7), e0254034

Sorenson, O., J.W. Rivkin and L. Fleming (2006) 'Complexity, networks and knowledge flow', *Research Policy*, 35 (7), pp.994–1017

Strassheim, H. and P. Kettunen (2014) 'When does evidence-based policy turn into policy-based evidence? Configurations, contexts and mechanisms', *Evidence and Policy*, 10 (2), pp.259–77

Szomszor, M. and E. Adie (2022) 'Overton: a bibliometric database of policy document citations', *Quantitative Science Studies*, 3 (3), pp.624–50

Toma, M. and E. Bell (2024) 'Understanding and increasing policymakers' sensitivity to program impact', *Journal of Public Economics*, 234, 105096

Vijaymeena, M. and K. Kavitha (2016) 'A survey on similarity measures in text mining', *Machine Learning and Applications*, 3 (2), pp.19–28

Weiss, C.H. (1979) 'The many meanings of research utilization', *Public Administration Review*, 39 (5), pp.426–31

Yan, E. and Y. Ding (2009) 'Applying centrality measures to impact analysis: a coauthorship network analysis', *Journal of the American Society for Information Science and Technology*, 60 (10), pp.2107–18

Yu, H., B. Murat, J. Li and L. Li (2023) 'How can policy document mentions to scholarly papers be interpreted? An analysis of the underlying mentioning process', *Scientometrics*, 128 (11), pp.6247–66