

<https://doi.org/10.1038/s44168-024-00160-y>

A review of national climate policies via existing databases

Yves Steinebach¹✉, Markus Hinterleitner², Christoph Knill³ & Xavier Fernández-i-Marín⁴

Various databases have been developed to track national climate policy efforts. These datasets facilitate comparisons across countries regarding policy activity, instrument choice, and policy effectiveness. This article evaluates these datasets to see whether they converge in their observations about climate policy development. Our findings reveal that all datasets agree at the aggregate level in that they show that ever-more climate policies are being adopted. However, they diverge significantly when scrutinizing more nuanced elements like policy instrument types and their stringency. The main contributions of our review are to highlight what research endeavors are already possible with existing datasets and to identify the gaps that still remain. We also provide concrete suggestions on how to enhance the existing datasets, making them more useful for social science research on climate policy. The article provides the most comprehensive and up-to-date source for scholars and practitioners interested in the comparative analysis of governmental climate policy efforts.

Climate change represents one of the most significant challenges of our time that demands immediate and concerted action by national governments. Effective government responses to climate change not only require the adoption of concrete policies (policy activity), but also the design of appropriate interventions (policy instruments) that are sufficiently ambitious (policy stringency). The urgency and global scope of the problem of climate change underline the need to comparatively assess the activity, instrumental design^{1,2}, and stringency of governmental policy responses³.

To track progress in national climate policy efforts, various databases have been developed over the last years^{4,5}. Among these repositories are the Climate Policy Database (CPDB)^{6–8}, the Climate Change Laws of the World (CCLW) (Grantham Research Institute)⁹, the European Environment Agency's (EEA) database on greenhouse gas policies and measures in Europe, the International Energy Agency's (IEA) Policies and Measures Database, the OECD Climate Actions and Policies Measurement Framework (CAPMF)¹⁰, and the Climate Policy Portfolios Dataset (CLIMAPP), which was developed in the context of the ACCUPOL project¹¹. The stated goals of these repositories are to describe and categorize countries' responses to climate change and to provide an empirical basis for evaluating their effectiveness. While there are other datasets that cover certain instrument types or sectors, the six datasets mentioned distinguish themselves by their aim to cover climate policy activity across countries comprehensively.

While tremendous work has been undertaken to track climate policy developments, we do not know whether and to which extent these datasets provide us with accurate and consistent information about countries' policy efforts and their effects. While there are some papers that assess and

compare more than one dataset⁴, they do not compare several datasets with regard to multiple policy output dimensions. Ideally, the datasets would provide consistent answers to the following questions: *Which governments are attempting to address climate change (description)? What exactly are they doing (taxonomy)? And how effective are their attempts (effects)?* However, while all datasets share the goal of tracking national climate policies, there are notable differences between them. They not only exhibit differences in geographic, sectoral, and temporal coverage but also use different analytical approaches and classificatory schemes to capture climate policies. Moreover, while some datasets only focus on description and taxonomy, others additionally provide a basis for assessing climate policies' effects by containing information about their stringency. Hence, it cannot be taken for granted that the existing databases provide us with consistent information on how governments respond to climate change and what difference it makes.

This review article addresses these knowledge gaps. It analyzes whether and how existing datasets map national climate policies with regard to three key dimensions: general policy activity, policy instruments, and policy stringency. Our analysis yields mixed results.

While the databases provide quite consistent information on countries' general enactment of climate policies over time (description), notable differences between the databases emerge once we delve into features of instrument choice (taxonomy) and stringency (effects). Specifically, we identify disparities in how the databases record the (growing) diversity of the policy instruments that countries employ to tackle climate change. These disparities become even more pronounced when examining how the

¹University of Oslo, Oslo, Norway. ²University of Lausanne, Lausanne, Switzerland. ³LMU Munich, Munich, Germany. ⁴University of Barcelona, Barcelona, Spain.

✉ e-mail: yves.steinebach@stv.uio.no

databases record policy stringency (i.e., the calibration of policy instruments).

Our findings reveal that all datasets agree at the aggregate level in that they show that ever-more climate policies are being adopted (*what is being done?*). However, they diverge significantly when scrutinizing more nuanced elements like policy instrument types (*how is it done?*) and their stringency (*to what effect?*). The analysis thus highlights that while all databases can be utilized to gauge broader developments in climate policy, the choice between these databases becomes critically important when scholars move from a more general assessment of climate policy activity to concrete inquiries on policy instrument choice and stringency. At this more intricate level, giving careful attention to the choice of databases and their inherent strengths and weaknesses is of heightened importance. Here, it becomes more crucial to validate whether the results and conclusions derived are, in any manner, influenced by the specific database selected for analysis. We demonstrate that some (pairs of) datasets are much more congruent than others. Moreover, existing databases are generally much better suited to explaining the adoption of climate policies than their effects. For evaluative studies, it is not only important to know that certain policies and policy instruments exist but also how strict these measures are. For instance, a common critique against carbon trading schemes is that they have often been too low in price or insufficient in coverage to address climate change effectively. With regard to the latter point, the CLIMAPP and especially the OECD climate policy dataset are notable examples as they not only provide information about the existence or absence of certain climate policy measures but also about (changes in) their stringency.

Our review's main contribution is to show what research endeavors are possible with the existing datasets and which are not and to provide concrete suggestions on how to enhance the reviewed datasets to make them even more useful for social science research on climate policy. Our article thus provides the most comprehensive and up-to-date source for scholars and practitioners interested in the comparative analysis of governmental climate policy efforts.

The article is structured as follows. We first delve into the criteria for evaluating climate policies and climate policy data repositories and map the existing datasets along the identified dimensions. We then compare the various datasets to determine the extent of their overall alignment, looking at general policy activity, instrument diversity, and policy stringency. In the concluding section, we reflect on the collective insights garnered from our review and tease out implications for future research in the field.

Criteria for evaluating climate policies and climate policy data repositories

At the most abstract level, any effort to fight climate change requires some *policy activity/action*. This means that *policy targets*—products or activities causing (mitigation) or being affected (adaptation) by climate change—are addressed by some governmental measures. The level of policy activity or density thus informs us about the extent to which a given sector is permeated by governmental interventions^{3,4}. In general, climate policies target multiple sources of and activities causing greenhouse gas emissions, covering both economic and private activities. At this abstract level, the essential question is whether governments address certain issues, such as emissions from the transport or industrial sector. However, these policy targets can also be more granular and differentiate between emissions derived from various industrial plants (small vs. large combustion plants) or modes of transport (passenger cars vs. heavy-duty vehicles). To structure the large variety of targets that contemporary governments address with their climate policies, targets are often grouped into sectors or areas, such as transport, industry, energy, and buildings. In this context, it is important to clarify that the term policy “targets” does not refer to *emission reduction targets* commonly debated within the framework of the Paris Agreement but rather to the question of which specific climate challenges the government addresses.

The second, more specific dimension that can be analyzed is *how* governments seek to achieve their predefined policy goals. Governments can usually choose between different *policy instruments* to address policy

targets^{1,12}. Policy research has developed various ways to categorize policy instruments. For instance, an often-used categorization in the climate domain is by Vedung², who distinguishes between regulations (“sticks”) such as emission standards and limit values, market-based instruments (“carrots”) such as emission trading or carbon taxes, and informational instruments (“sermons”) such as educational campaigns or performance labels (for other classification see¹³). A key insight from the literature is that all policy instruments have their particular strengths and weaknesses¹⁴. Consequently, governments need to choose from and craft “instrument mixes” that address the policy targets in the best possible way.

The third and most nuanced dimension is the *stringency or intensity of public policies*. This dimension refers to the exact degree of ambitiousness underlying the government's efforts to address a given issue³. Frequent examples in the area of climate policy are the stringency of emission standards (e.g., 120 vs. 95 g/CO₂ per km) or the price set by carbon taxes (20 vs. 100 Euro per tCO₂). Stringency is an important policy dimension as it is necessary for evaluating policy effects, or it can be used as a proxy to infer effectiveness if an actual evaluation of effects is not feasible for some reason. Generally, the more ambitious a particular policy, the stronger its effects can be expected to be.

Table 1 presents a compiled overview of the existing datasets related to climate policy along the key dimensions policy activity, instruments, and stringency. It also indicates the geographic coverage by stating the number of countries included in the databases. Sectoral coverage refers to the specific industries or areas of society being covered by the databases, like transportation, industry, or energy production. Temporal coverage displays the range and distribution of data entries over time. All datasets include policy targets on both mitigation and adaptation, except for the OECD dataset. This dataset does not include adaptation targets.

Policy activity, instrument mixes, and stringency across the datasets

In the previous section, we mapped the existing datasets on climate policies. In this section, we delve deeper into identifying overall patterns, correlations, and disparities that exist across these datasets. We assess the general policy activity before delving into instrument mixes and stringency.

Climate policy activity

Figure 1 displays each dataset's temporal trend of policy activity, i.e., the adoption of new climate policies over time. The figure reveals that all the datasets examined present a strong growth in the number of climate policies over time. This growth, however, does not occur “linearly” but instead comes in different surges, suggesting periodic intensifications in the production of climate policies. The datasets reveal notable peaks in climate policy production during the early 2000s, mid-2010s, and the early 2020s. The peaks observed in the 2000s and 2010s roughly coincide with (the aftermath of) major international climate agreements: the Kyoto Protocol, which was adopted in 1997 and ultimately came into force in 2005, and the Paris Agreement, which was adopted in 2015. However, it also needs to be noted that there is some variation between the peaks identified in the different datasets, indicating that despite their commonly reported growth trends, they display certain variations in the specific activities they capture. This assessment is corroborated when analyzing the correlation between the datasets at the country-year level.

A challenge for examining correlations at the country-year level is that there are significant differences across the datasets regarding the number of policy targets and instruments covered (see again Table 1). Consequently, some datasets appear to document more changes than others simply because they consider a greater number of policy targets and instruments. To address this issue, we transform all datasets into a two-dimensional portfolio space comprising all policy targets along one dimension and all instruments on the other. This “policy portfolio” approach has gained traction in recent years within public policy research to compare policy dynamics across countries and different policy sectors^{15,16}. The portfolio's value is 100 percent when every potential target-instrument combination is

Table 1 | Overview of Existing Climate Policy Databases

Name	Policy Targets	Policy Instruments	Policy Stringency	Coverage	
				Country	Time
Climate Policy Data Base (CPDB)	9 policy objectives and 6 mitigation areas	9 policy instrument types, 28 subtypes and 33 instruments	No	198 countries plus EU level	1927–2024
Climate Change Laws of the World (CCLW)	23 policy sectors	5 policy instrument types and 25 subtypes	No	199 countries plus EU level	1970–2023
EEA Database on Greenhouse Gas Policies and Measures in Europe (EEA)	9 policy sectors and 8 greenhouse gases	10 policy instrument types	No	30 countries	1930–2024
The IEA's Policies and Measures Database (IEA)	134 sectors, 16 topics and 288 technologies	190 policy instruments	No	163 countries plus EU level	1980–2023
OECD Climate Actions and Policies Measurement Framework (CAPMF)	56 policy climate actions (no systematic distinction between policy targets and instruments possible)	56 policy climate actions (no systematic distinction between policy targets and instruments possible)	Yes	52 countries plus EU level (no records for the EU currently)	1990–2022
Climate Policy Portfolio Dataset (CLIMAPP)	73 policy targets	14 policy instrument types	Yes	21 countries plus EU level	1975–2022

All datasets listed are accessible online. However, there are minor differences in how the data can be utilized. The CLIMAPP database is exclusively available for download as a data file without an accompanying web-based tool for online examination. Conversely, while the IEA dataset can be explored online, it does not offer a direct download option. Furthermore, it is important to note that the IEA dataset contains a broader range of policy topics, technologies, and instruments than those revealed by pre-selected searches. Thus, to uncover the full extent of the dataset's offerings, performing a search with no pre-set parameters is recommended, effectively running an "empty" search. To pursue the analysis presented in this paper, we utilized web scraping techniques to extract the necessary data from the IEA webpage. The code is provided in section 10 of the Supplementary Information. Please also note that the datasets use different terminologies to describe the same phenomenon, i.e., products, processes, and activities being addressed by policies. For the sake of comparability, we cluster them under the term policy targets but keep the original terminology in the table.

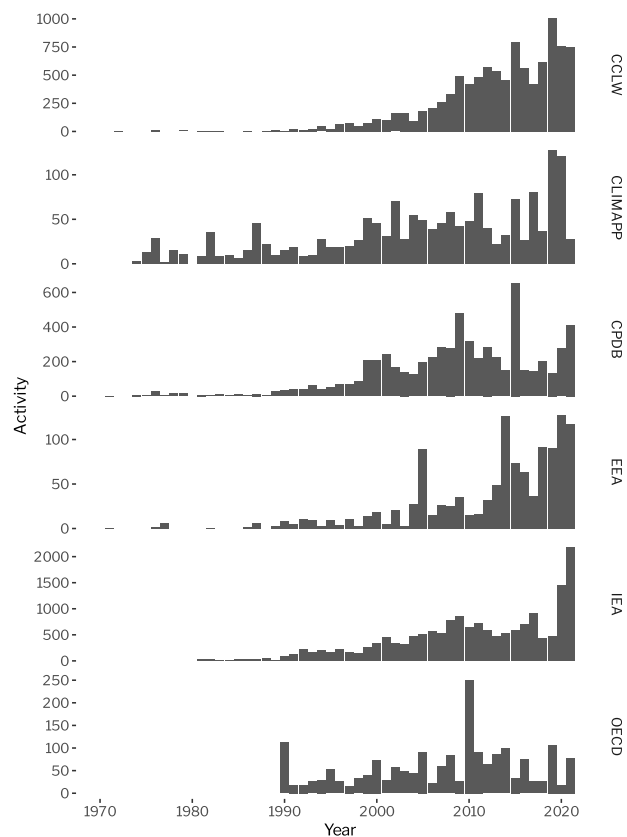


Fig. 1 | Policy activity over time and across datasets. Each bar represents the new policy activity included in the respective dataset in each year. The different number of countries and of policies in each database implies that the vertical axis is singular for each database. Please note that the figure includes all countries contained in the respective datasets. In Figure S8.1 in the Supplementary Information, we provide a figure that presents only the joint country sample across the datasets. Source: CPDB, EEA, IEA, CCLW, CLIMAPP, OECD.

covered, while a value of 0 percent suggests that no policy actions have been taken. Changes within this portfolio space can be quantified as percentage point changes. For example, an increase from 10 to 15 percent coverage is recorded as a five-percentage point change. This approach allows us to compare standardized portfolio growth rates at the country-year level irrespective of the exact number of policy targets/instruments covered by the respective datasets¹². The underlying logic of this approach is exemplified by the policy portfolio presented in Fig. 2. Using the CLIMAPP dataset, the figure presents Germany's climate policy portfolio at two points in time, namely, the year 2005 (shown in black boxes) and the year 2022 (shown in black and grey boxes). The analysis shows that during the specified period, the portfolio expanded from 48 to 114 covered spaces, representing an increase of 6.5 percentage points (4.7% to 11.2%). While 72 target-instrument combinations were introduced between 2005 and 2022, only 6, primarily subsidies, were abolished.

We have been able to create portfolio spaces for all datasets except for the OECD database. For the CPDB, targets are the cross-combination between sectors (Table S1.1 in the Online Supplementary Document) and types (Table S1.2 in the Online Supplementary Document). For the CCLW, targets correspond to the sectors (Table S2.1 in the Supplementary Information). For the EEA, targets are the cross-combination between sectors (Table S3.1 in the Supplementary Information) and greenhouse gases affected (Table S3.2 in the Supplementary Information). For IEA, targets correspond to sectors (Table S4.1 in the Supplementary Information). For CLIMAPP, targets are shown in Table S6.1 in the Supplementary Information. For the IEA dataset, we limit our analysis to those targets (sectors) that are present in at least five countries, and to

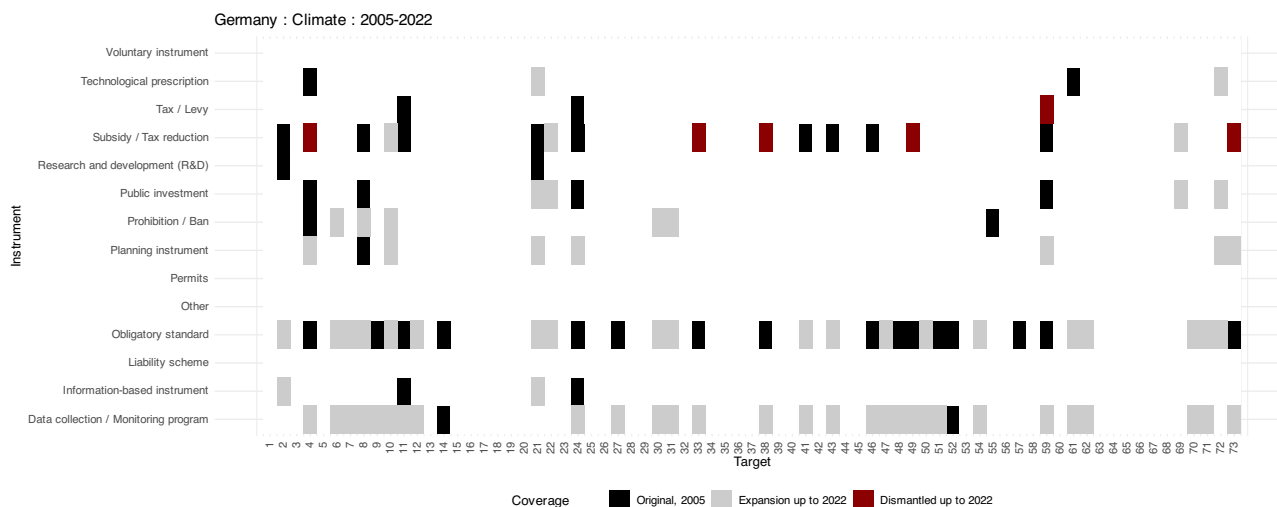


Fig. 2 | German Climate Policy Portfolio in 2005 and 2022. The horizontal axis represents the different policy targets, and the vertical axis the instruments employed. A white background means that there is no policy intervention recorded in that space. A black square identifies spaces that were covered by policy

intervention in 2005 and continue to do so in 2022. A gray square identifies new policies added between 2005 and 2022. A dark red square identifies policies in place in 2005 that are no longer part of the portfolio. Source: CLIMAPP.

instruments found in at least ten countries. This is because considering every potential combination of targets and instruments surpasses our computational capacity. By focusing on “policy actions,” the OECD dataset already blends targets and instruments from the outset; however, it does *not* consider the same instruments for all potential targets. Thus, creating a (symmetric) two-dimensional portfolio space is not possible. This does *not* imply that the quality of the OECD dataset is inferior. It merely suggests that it is more challenging to compare it with others when using the portfolio approach. By excluding the OECD dataset from parts of the following analysis, we also address the issue that the OECD dataset only includes mitigation targets, in contrast to the others. Figure 3 demonstrates the correlation between the five other datasets. Given their different sample sizes, we have only included data for which there is an overlap in the pairwise comparisons.

The strongest correlation is found between the IEA and the CPDB, with a correlation coefficient of 0.882. The weakest correlation exists between the EEA and the CPDB, with a coefficient of 0.274. On average, the datasets exhibit a correlation of 0.542. When broken down, the average correlations are as follows: 0.652 for the IEA dataset, 0.620 for CCLW, 0.547 for CPDB, 0.417 for the EEA dataset, and 0.476 for the CLIMAPP. These figures suggest that, while correlations vary across datasets, there is a general pattern of moderate to strong correlations, with some individual pairings, such as IEA and CPDB, showing particularly strong correlations. While a high level of correlation is not necessarily a quality indicator—as all datasets could potentially be inaccurate in their assessments—it does (at least) suggest some consistency across the datasets when it comes to tracking general policy activity over time.

Thus far, we have evaluated the datasets’ consistency by checking for cross-correlations (external validity). Another way to analyze and compare the datasets is to see whether climate policies are consistently coded *within* them (internal validity). For this purpose, we can take advantage of the fact that all EU policies being adopted through EU Directives—as opposed to EU Regulation—must also be transposed into national-level policies. One could expect that in a perfectly consistent dataset, all the policy measures coded at the EU level should also have been coded at the member state level. Put differently, when “laying” the EU policy portfolio over those of the member states, there should be a substantive, if not exhaustive, overlap between EU Directives and member state transpositions. To test for this, we created a distinct EU portfolio for the CPDB, the IEA, and the CCLW. At the time of writing, the OECD’s CAPMF lists the EU as a separate unit for analysis, but all the entries are, unfortunately, empty. As depicted in Fig. 4, there is

generally a high level of consistency between the EU and the member states’ coding. The average consistency reaches an impressive 98% for the IEA dataset and 94% for the CCLW. The coding in the CPDB is also quite good, with an overall consistency of 89%. However, this consistency decreases when we shift focus from the older to the newer member states. For most EU member states that acceded post-2004, the consistency between the EU and national coding drops to around 80%. Note that due to the time lag between the adoption of an EU policy and its transposition into national law, it might be the case that some of the EU policies are not yet represented in the national policy portfolios. To control for this eventuality, we present an additional analysis in Figure S8.2 in the Supplementary information that sets the year 2018 as the upper time limit for policies included in the EU portfolio. This way, we avoid that our analysis of internal validity is biased by possible delays in transposition. This adjustment improves the overlap by about 2% for the CPDB and the CCLW and remains almost unchanged for the IEA.

In sum, the aggregate insight from the first part of our analysis on climate policy databases is that—despite their pronounced conceptual variation (see again Table 1)—they are rather consistent regarding their assessment of general trends in climate policy activity across countries.

Climate policy instrument mixes

But does this picture also hold if we climb one step down on the “ladder of analytical abstraction” and focus on policy instruments? A key insight from public policy research is that instrument choice and combinations are crucial for the government’s ability to tackle societal challenges. Within this framework, the more quantitative policy design literature has developed concepts such as “instrument diversity”¹⁷ or “instrument balance”¹⁸ to evaluate instrument mixes. Essentially, both approaches involve assessing the variety of the policy instruments being utilized.

In the following, we use the concept of average instrument diversity (AID) developed by Fernández-i-Marín et al.¹⁷ to assess and contrast the information on the composition of instrument mixes provided by the datasets. The AID measure can be calculated using the *R PolicyPortfolio* package once the respective data is structured in policy targets and instruments. Put simply, AID is calculated by selecting instruments dedicated to specific targets (the vertical axis in Fig. 2), and then determining the average probability that the same instrument is picked when making “draws” across other policy targets (the horizontal axis in Fig. 2). For a more detailed explanation of this approach, please consult Chapter 7 of the Supplementary Information.

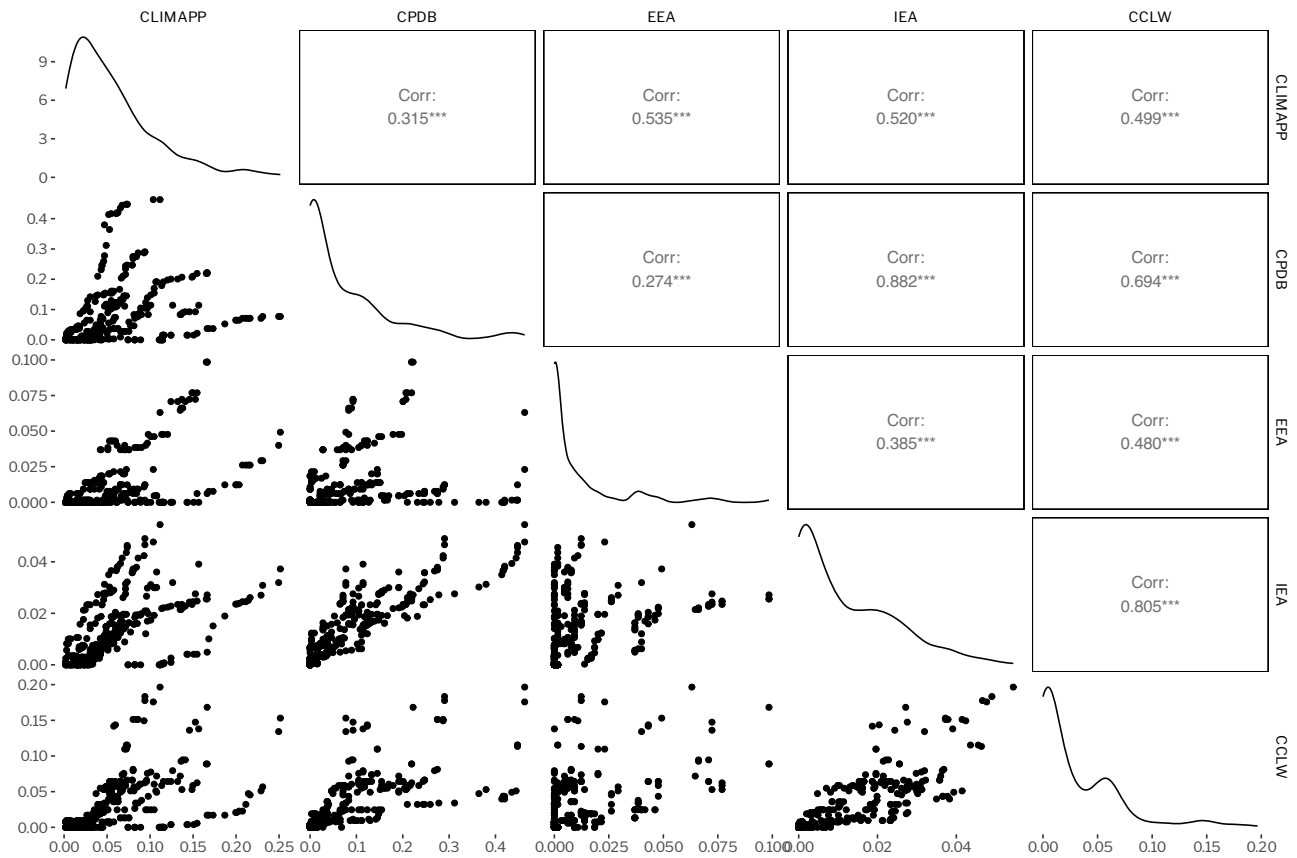


Fig. 3 | Cross-correlations in general policy activity. The figure depicts the correlation between the datasets under study (except the OECD dataset). The diagonal contains the distribution of the portfolio sizes in each database. The upper diagonal

shows the cross-correlations between databases, and the lower diagonal shows the scatterplots. Source: CLIMAPP, CPDB, EEA, IEA, CCLW.

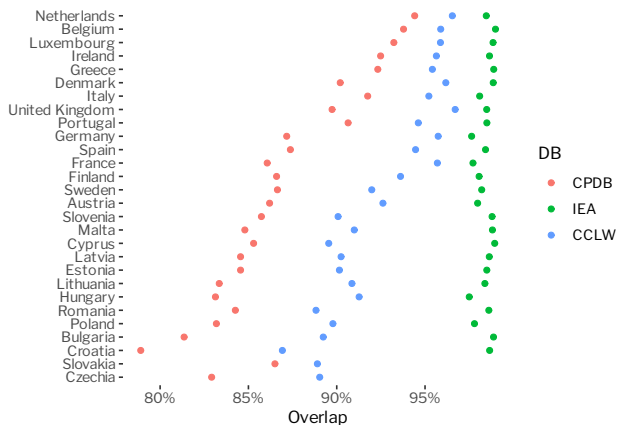


Fig. 4 | Coding accuracy using EU policies as benchmark. The figure depicts the consistency of different datasets for the coding of the EU and the member states. The horizontal axis displays the percent of overlap between the EU policies and the national (in the vertical axis) policies. Each color represents a different database. Source: CPDB, IEA, CCLW.

Figure 5 illustrates the correlations in AID among the five previously evaluated datasets, focusing on the years 2000 to 2023. This time span is chosen because the instrument diversity metric is quite sensitive to fluctuations when the portfolio size is small. The analysis suggests a substantially lower correlation level when examining the composition of instrument mixes rather than general policy activity. The average correlations for each database, when examined separately, are as follows: 0.419 for CPDB, 0.386

for CCLW, 0.350 for CLIMAPP, 0.323 for the IEA, and 0.208 for the EEA. The calculated average correlation across all datasets is 0.349, indicating a strength about two-third of that observed for general policy activity. Hence, while the datasets provide consistent information on whether climate action is taken in a particular country, they provide much more conflictive information on what governments are doing to tackle climate change.

An obvious explanation for these discrepancies is that the datasets rely on different—and more or less detailed—classifications of policy instruments. Datasets that exhibit more parsimonious categorizations of policy instruments automatically display smaller levels of instrument diversity. However, our analysis indicates that the identified discrepancies may also stem from coding inconsistencies *within* the individual datasets, i.e., that not all policy instruments in the respective datasets are coded at the same level of “detail”.

One way to examine the consistency with which a dataset classifies policies across countries and areas of intervention is to leverage the datasets’ distinction between instrument types and subtypes (among the six datasets analyzed, three differentiate between various types and subtypes of policy instruments), and to use alluvial diagrams to analyze their relations. Figure 6 presents an alluvial diagram picturing the various instrument types and subtypes considered in the CPDB. In the diagram, green flows symbolize instances where comprehensive information is available. This means that, for the policy in question, the coding effectively captures the full spectrum of detail: from the instrument category to the subcategory and the exact policy tool used. Conversely, red flows indicate coded policies where only the broad instrument category has been coded, without further detail on the subcategory or the specific tool. The diagram also features two shades of yellow, representing instances in the data where the available information is only partially complete. One shade of yellow highlights instances where detail is

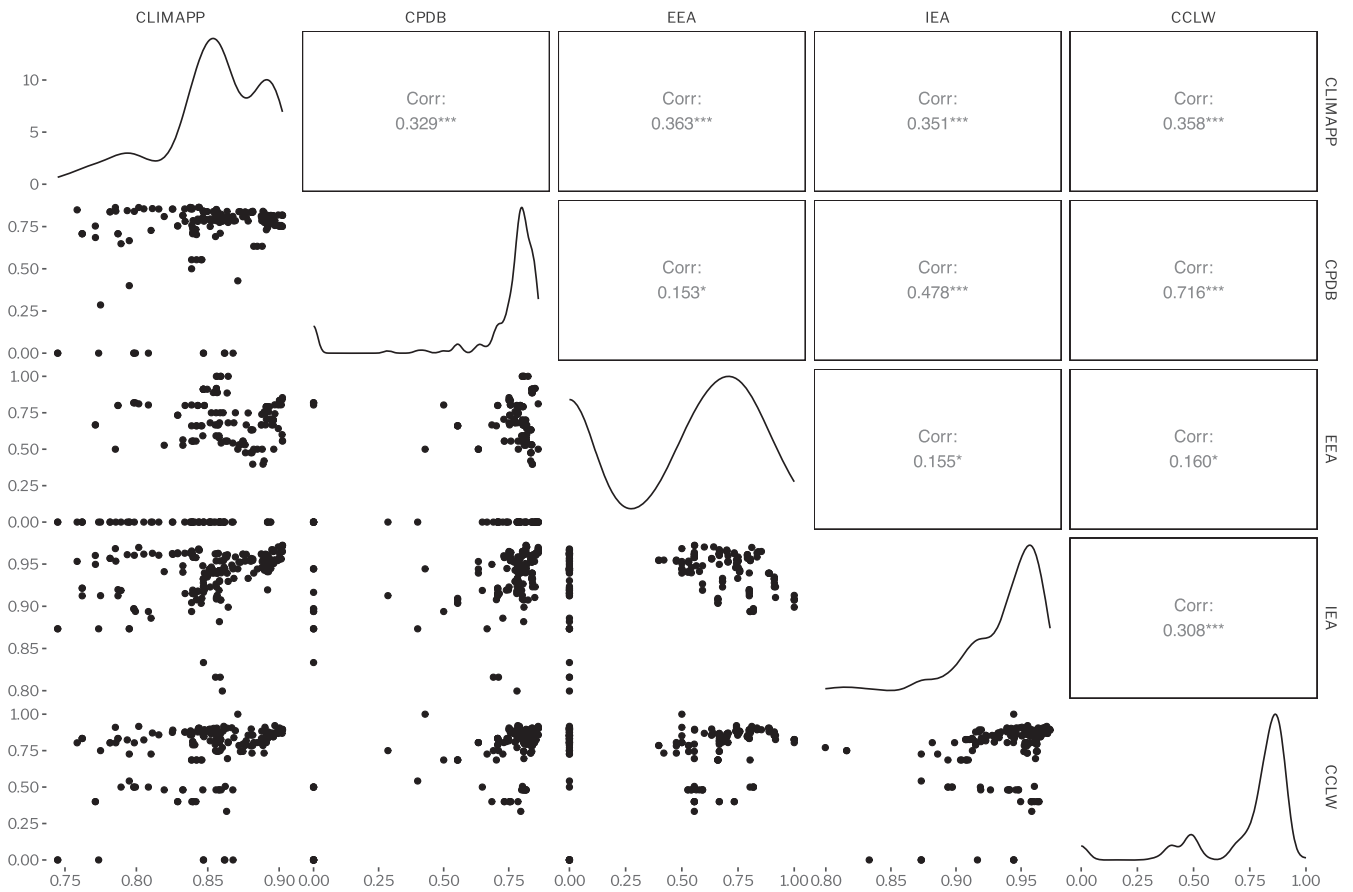


Fig. 5 | Cross-correlation for average instrument diversity. The figure depicts the correlations in AID among the five previously evaluated datasets. The diagonal contains the distribution of the average instrument diversity. The upper diagonal shows the cross-correlations between databases, and the lower diagonal shows the scatterplots. Only countries shared across databases are included. Covers the 2000-2020 period. Source: CLIMAPP, CPDB, EEA, IEA, CCLW.

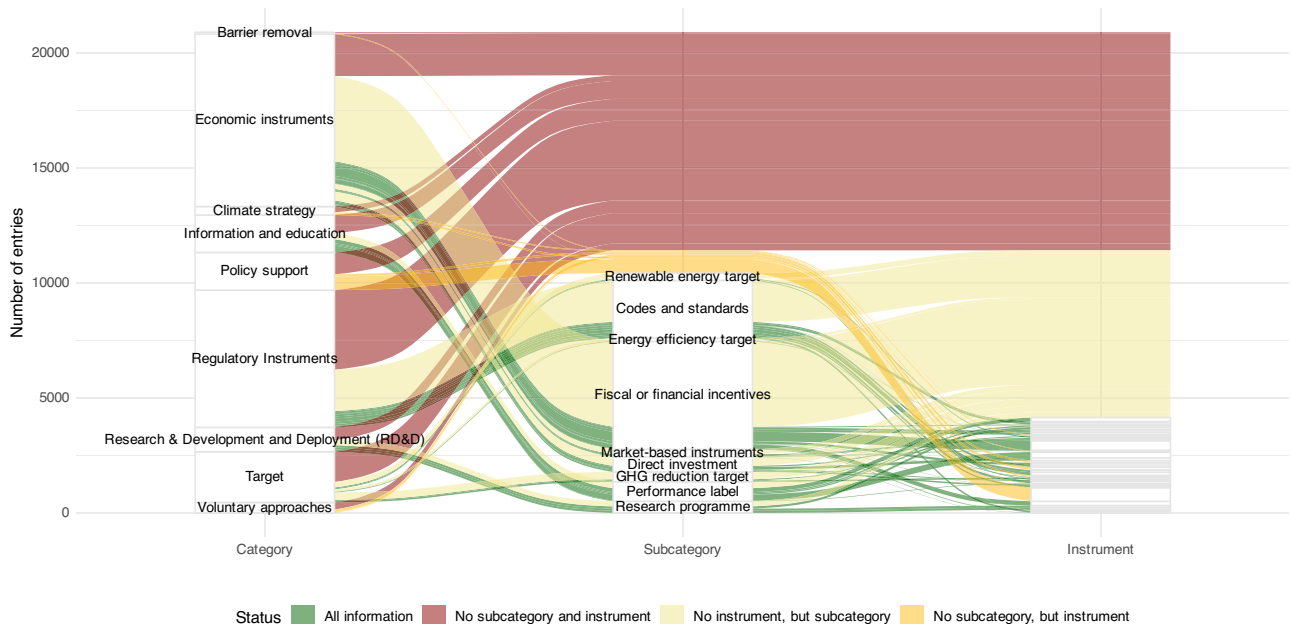


Fig. 6 | Policy instrument and subtypes considered in CPDB. Alluvial figure depicting instrument types and subtypes in the CPDB, showing the correspondence between the entries with a “Category” (first column), the “Subcategory” (second, middle column), and the “Instrument” (third category). Colors correspond to the status of each policy. Source: CPDB.

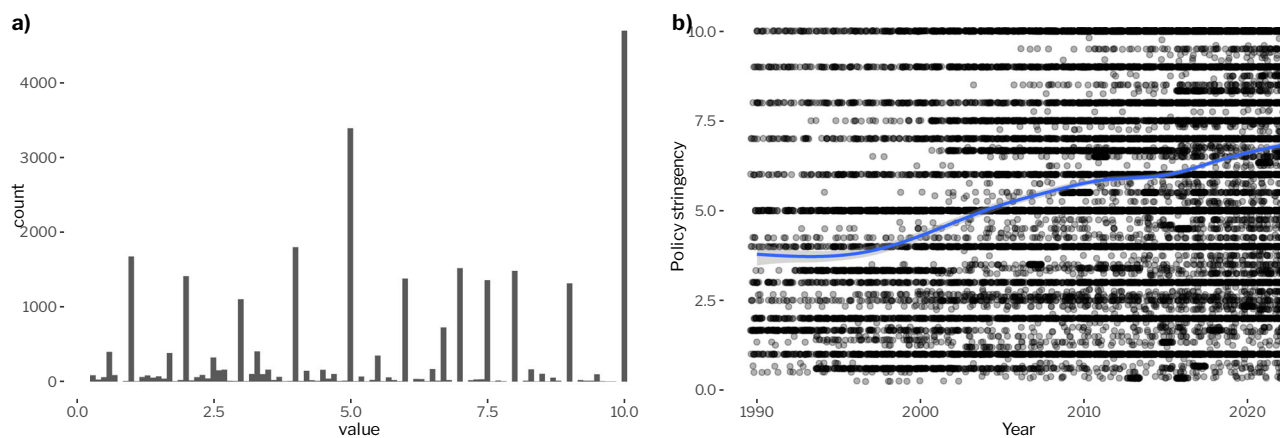


Fig. 7 | Policy stringency in the OECD dataset. The figure shows the policy stringency of climate actions in the OECD dataset. **a** presents the distribution of stringency (between 0 and 10) on the dataset; **b** presents their temporal distribution, with a temporal evolution. Source: OECD.

given for the subcategory, yet the instrument category is not specified. The other shade shows the inverse: detail is provided for the instrument category but not for the subcategory.

In essence, this figure suggests that a complete coding between the instrument type and the first subcategory is observed in only 45 percent of the policies contained in the dataset. Furthermore, merely about 34 percent consistency is found between the first and the second instrument subcategory. This issue is most pronounced for the substantial group of regulatory instruments, for which only about 15 percent are coded across all three levels: category, subcategory, and individual instruments. The analysis thus indicates that the level of “detail” or the completeness with which policy instruments are described is not uniform across all the codings included in the dataset.

These inconsistencies have implications for analyzing instrument mixes and diversity, as the extent and quality of information provided can lead to significantly different conclusions about a country’s policy instrument diversity. Consider an example where three policy instruments are documented. In one instance, the information merely indicates that all three are regulatory instruments. In another instance, the policies are explicitly identified as auditing, building codes, and obligation schemes. In the first scenario, the AID would present a value of zero, suggesting no diversity, because it appears as though the same type of instrument is being applied repeatedly. However, in the second example, where specific types of regulatory instruments are identified, the AID would indicate very high diversity, as it shows a range of different instruments being employed.

Climate policy stringency

The main takeaway from the preceding sections is that there is broad agreement between the datasets regarding general policy activity, but notable differences become visible when we transition to the analysis of policy instruments. Even greater discrepancies appear when we assess these instruments’ stringency level. Of the six datasets analyzed, only two include the analytical dimension of policy stringency. What is more, both datasets measure stringency changes in very different ways.

The OECD dataset measures policy stringency by averaging different “policy variables.” For example, the stringency of emission trading schemes is assessed using two variables. The first variable captures the cost of the annual allowances measured in USD per tCO₂e. This value is then scaled between 0 and 10, with 10 being the highest and 0 the lowest empirically observed value. The second variable captures the coverage of GHG, differentiated by CO₂, CH₄, N₂O, and all other GHG. The stringency value of each GHG reflects the contribution of each gas to global GHG emissions: CO₂ gets a stringency value of 6, CH₄ a value of 2, and N₂O a value of 1. The final stringency value for the CO₂ Emissions Trading Schemes is then the *mean* of the two observed values.

The graph on the right side of Fig. 7 displays an increasing trend in the average policy stringency over time, effectively showing an almost doubling of stringency values from around 3.8 to roughly 7. Yet, an important aspect to consider in this context is that each instrument has its own “method” for calculating stringency, which makes direct comparisons of stringency across different types of policy instruments difficult, if not impossible, or even worse, misleading. This issue is *not* indicative of a shortcoming in the dataset; rather, it is a consequence of the inherent diversity in climate policy instruments and their respective settings.

The left-hand side of Fig. 7 shows that the observed data points tend to cluster around the 5 and the 10 marks. This clustering pattern emerges due to the scale employed in measuring most policy variables. Most policy variables score as either 0 or 10. When these values are then aggregated at the (higher) policy level, an average score of 5 often appears, indicating that one policy variable scored 0 and the other scored 10. These findings underscore that even with a sophisticated analytical approach, capturing the nuances of policy stringency can be quite challenging.

This is also evident in the approach adopted by the CLIMAPP dataset, which captures policy stringency in *relative* rather than absolute terms. It provides insights into any modifications in existing policy instruments, registering whether they have become (1) stricter or (2) less strict over time. Thus, the dataset captures shifts in policy stringency, rather than providing an absolute measurement of policy stringency. Although the granularity of the dataset might be less than that of the OECD, it provides the benefit that it systematically distinguishes between changes in the level and the scope of policy stringency. Specifically, the data allows scrutiny of whether modifications imply stricter requirements such as emission limits, technologies, taxes, etc., or an expansion of the policy’s reach, i.e., inclusion of more categories, such as additional automobiles or industrial plants. Thus, the dataset provides a quite nuanced understanding of the shifts in policy stringency over time.

For instance, Fig. 8 (left side) shows that countries not only significantly increased their general policy activity since the 1980s, they also increasingly fine-tuned the adopted policies by adapting the stringency dimension. Moreover, as shown on the right side of Fig. 8, it seems that level changes are more widespread than scope changes.

Discussion and conclusion

In this paper, we conducted an in-depth review of all existing climate policy datasets. We evaluated the datasets regarding their overall level of agreement (or divergence) concerning broader policy growth dynamics, the employment of different instrument types, and policy stringency. In this final section, we summarize and discuss the key takeaways from our analysis.

To begin, it is essential to acknowledge the exceptional depth of data available; to our knowledge, the richness of datasets in the realm of

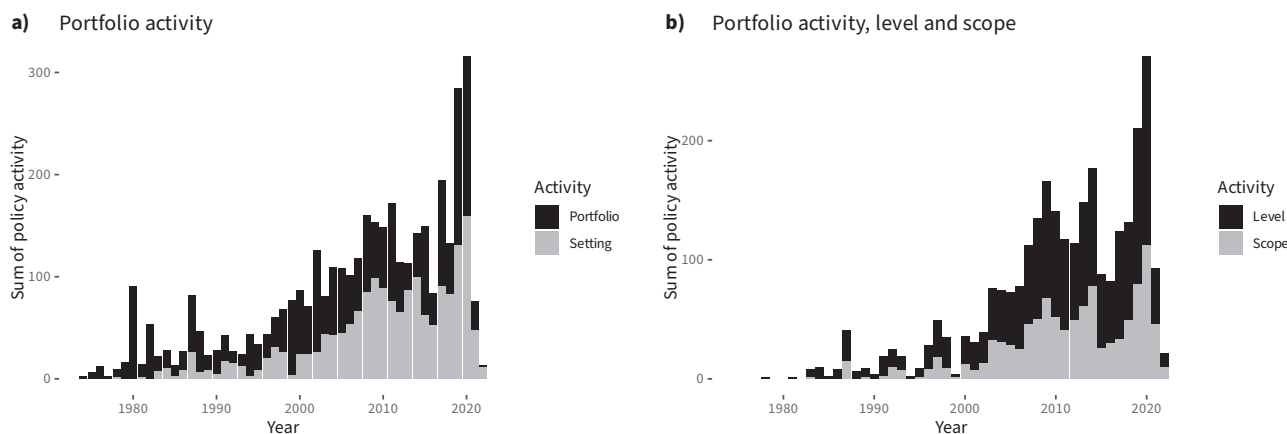


Fig. 8 | Changes in settings dimensions in comparison to general policy activity. **a** shows the sum of policy activity over time, dividing between portfolio activity (in black) and settings (in grey); **b** shows the level (dark grey) and scope (light grey). Source: CLIMAPP.

climate policy is unprecedented, surpassing even long-established research areas such as welfare and environmental studies^{19,20}. The sheer volume and variety of public policy datasets in this field represent a commendable achievement.

That said, our analysis suggests five key points to consider when using the available climate policy datasets. On the upside, we found a notable level of agreement between the datasets regarding general trends in climate policy development. Researchers who aim to examine and explain climate policy growth patterns over time can confidently choose from all the available datasets. The first key takeaway is thus that (1) all the existing datasets are well-suited to examine questions such as why some countries have adopted particular climate policies while others have not. In this case, the dataset choice may be predominantly driven by the specific research scope, including the geographic regions, sectors, and time frame of interest.

However, when the focus turns to the analysis of policy instruments, we observe much more variability between the datasets. The second key takeaway from our analysis is thus that (2) researchers diving into this area should proceed with caution, acknowledging that their findings could be influenced by their choice of dataset. In this analysis, we observe that the CPDB has the strongest correlation with all other databases examined, followed by the CCLW, which also shows a significant degree of agreement. Nevertheless, regardless of the dataset used, it is advantageous for researchers to validate their findings with multiple datasets whenever possible. In this regard, Fig. 4 serves as a valuable resource for researchers, offering an efficient way to identify datasets that are most aligned with or, conversely, present the greatest challenge for cross-checking or validating a certain finding.

Finally, when examining policy stringency, the divergence among datasets is most pronounced. The third takeaway is that (3) only two of the six datasets available provide some metrics on policy stringency, and they measure it in quite distinct ways. The OECD dataset offers a robust measure of policy stringency levels, while the CLIMAPP dataset identifies whether there have been any changes in stringency but not the specific levels of change. When selecting a dataset for studying policy stringency, researchers should thus carefully consider what aspect of stringency they wish to investigate. The OECD dataset is preferable if the analysis focuses on the impact of policies, such as their emissions reduction potential. In contrast, the CLIMAPP dataset offers more detailed insights into the diverse strategies policymakers employ to extend or modify existing policy instruments.

When contrasting this assessment with the information displayed in Fig. 1, particularly the number of countries covered, it becomes evident that (4) there is an inherent trade-off involved in shifting from a predominantly descriptive or taxonomic approach to more evaluative datasets. It is remarkable that the two datasets involving policy stringency (OECD, CLIMAPP) also come with the smallest geographical coverage. This

underscores the considerable data collection efforts required to transition from more descriptive efforts towards a more evaluative focus on governments' climate policy efforts.

When moving from descriptive to more evaluative purposes of climate policy data, another aspect that needs consideration is the assessment of administrative capacities and structures that influence the implementation of these measures²¹. Even very stringent climate policies may fail to deliver if the administration lacks the resources to implement them. However, (5) all existing databases touch on the aspect of climate administrative capacity in a rather superficial way: the CPDB has one instrument on “institutional creation”, the IEA has one on “institutional mandates”, the IEA has one on “institutions for sustainable finance”, and the OECD has a variable on “climate advisory bodies”. Yet, these codings are far from capturing the complex administrative arrangements and capacities required to implement climate policies^{22–24}. Future work on climate policies should thus not only aim to expand in terms of geographical scope or in terms of capturing policy nuances (stringency), but also consider expanding the analysis to include more administrative aspects. A first step, following, for example, ref. 25, might be to code whether the legal acts that bring about the respective climate policies involve a statement about implementation procedures, specifically assigning actors and rules, or setting up a rigorous monitoring regime.

In conclusion, our findings highlight a prevailing challenge in comparative public policy research, namely to capture and compare government actions across different policy measures and contexts and to identify and compare their effects³. We believe that there are clear advantages in the existence of multiple climate policy datasets and the various approaches they offer to examine climate policy actions and their ambitions. It is the diversity in datasets that enables researchers to conduct various types of research and to answer different research questions. Nevertheless, further discussion is necessary to understand how the selection of datasets may influence and possibly bias specific research findings. We hope that our review paper contributes to a thoughtful conversation on this important topic, encouraging awareness and transparency in data selection choices in future research.

Received: 30 April 2024; Accepted: 31 August 2024;
Published online: 30 September 2024

References

1. Howlett, M. Governance modes, policy regimes and operational plans: a multilevel nested model of policy instrument choice and policy design. *Policy Sci.* **42**, 73–89 (2009).
2. Vedung, E. Policy instruments: typologies and theories. In M.-L. Bemelmans-Videc, R. C. Rist, & E. Vedung (Eds.), *Carrots, sticks, and*

- sermons: policy instruments and their evaluation* (5th ed., pp. 21–58). New Brunswick, NJ: Transaction Publishers (2010).
3. Knill, C., Schulze, K. & Tosun, J. Regulatory policy outputs and impacts: exploring a complex relationship. *Regul. Gov.* **6**, 427–444 (2012).
 4. Schaub, S., Tosun, J., Jordan, A. & Enguer, J. Climate policy ambition: exploring a policy density perspective. *Politics Gov.* **10**, 226–238 (2022).
 5. Zheng, S. et al. Global climate policy effectiveness: a panel data analysis. *J. Clean. Prod.* **412**, 137321 (2023).
 6. NewClimate Institute, Wageningen University and Research & PBL Netherlands Environmental Assessment Agency. Climate Policy Database. <https://doi.org/10.5281/zenodo.10869734> (2023).
 7. Nascimento, L. & Höhne, N. Expanding climate policy adoption improves national mitigation efforts. *Npj Clim. Action* **2**, 12 (2023a).
 8. Nascimento, L. et al. Twenty years of climate policy: G20 coverage and gaps. *Climate Policy* (2021).
 9. Fankhauser, S., Gennaioli, C. & Collins, M. Do international factors influence the passage of climate change legislation? *Clim. Policy* **16**, 318–331 (2016).
 10. Nachtigall, D. et al. The climate actions and policies measurement framework: a database to monitor and assess Countries' mitigation action. *Environ. Resour. Econ.* **87**, 191–217 (2024).
 11. Knill, C. *Datasets*. https://publicpolicy-knill.org/?post-type=research&post_id=2019 (2024).
 12. Fernández-i-Marín, X., Knill, C., Steinbacher, C. & Steinebach, Y. Bureaucratic quality and the gap between implementation burden and administrative capacities. *Am. Polit. Sci. Rev.* 1–21 (2023).
 13. Capano, G. & Engeli, I. Using instrument typologies in comparative research: conceptual and methodological Trade-Offs. *J. Comp. Policy Anal. Res. Pract.* **24**, 99–116 (2021).
 14. Henstra, D. The tools of climate adaptation policy: analyzing instruments and instrument selection. *Climate Policy*, 16 (2016).
 15. Adam, C., Knill, C. & Fernandez-i-Marín, X. Rule growth and government effectiveness: Why it takes the capacity to learn and coordinate to constrain rule growth. *Policy Sci.* **50**, 241–268 (2017).
 16. Limberg, J., Steinebach, Y., Bayerlein, L. & Knill, C. The More The Better? Rule Growth And Policy Impact From A Macro Perspective. *Eur. J. Political Res.* **60**, 438–454 (2021).
 17. Fernández-i-Marín, X., Knill, C. & Steinebach, Y. Studying Policy Design Quality in Comparative Perspective. *Am. Political Sci. Rev.* **115**, 931947 (2023).
 18. Schmidt, T. S. & Sewerin, S. Measuring the temporal dynamics of policy mixes An empirical analysis of renewable energy policy mixes' balance and design features in nine countries. *Res. Policy* **48**, 103557 (2019).
 19. Scruggs, L., Jahn, D. & Kuitto, K. Comparative Welfare Entitlements Dataset Version 2017-09. University of Connecticut & University of Greifswald. (2017).
 20. Van Vliet, O. & Caminada, K. Unemployment replacement rates dataset among 34 welfare states 1971–2009: An update, extension and modification of the Scruggs' Welfare State Entitlements Data Set, NEUJOBS Special Report No. 2 (Leiden University, 2012).
 21. Steinebach, Y. Instrument choice, implementation structures, and the effectiveness of environmental policies: a cross-national analysis. *Regul. Gov.* **16**, 225–242 (2022).
 22. Meckling, J. & Nahm, J. The power of process: state capacity and climate policy. *Governance* **31**, 741–757 (2018).
 23. Basseches, J. A. et al. Climate policy conflict in the U.S. states: a critical review and way forward. *Climatic Change* **170**, 32 (2022).
 24. Schaffrin, A., Sewerin, S. & Seubert, S. Toward a comparative measure of climate policy output. *Policy Stud. J.* **43**, 257–282 (2015).
 25. Wagner, P. M., Torney, D. & Ylä-Anttila, T. Governing a multilevel and cross-sectoral climate policy implementation network. *Environ. Policy Gov.* **31**, 417–431 (2021).

Acknowledgements

We acknowledge funding from the European Research Council (Grant No: 788941 – ACCUPOL) and by the Norwegian Research Council (Grant No. 335073 – ACCELZ Project).

Author contributions

Y.S., M.H., C.K., and X.F.M. have made substantial contributions to the conception or design of the work or the acquisition, analysis, or interpretation of the data; drafted the work or revised it critically for important intellectual content; approved the completed version, and are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. Y.S., M.H., C.K., and X.F.M. have read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s44168-024-00160-y>.

Correspondence and requests for materials should be addressed to Yves Steinebach.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024