



OPEN

DATA DESCRIPTOR

# A dataset of structural breaks in greenhouse gas emissions for climate policy evaluation

Talis Tebecis<sup>1</sup> & Jesus Crespo Cuaresma<sup>1,2,3</sup>  

The quantitative assessment of policies aimed at climate change mitigation requires rigorously identifying abnormal changes in greenhouse gas emissions. We present a new dataset of robust level changes in greenhouse gas emissions that cannot be explained by aggregate socioeconomic fluctuations. Modern methods of structural break identification based on two-way fixed effects models are employed to estimate the size of significant level changes in emissions. The resulting dataset spans information for all major greenhouse gases in OECD countries across 37 IPCC sectors, from 1995 to 2022. The data unveils large differences in abnormal changes in emissions across gases, countries and sectors, as well as over time. Our resulting data can be applied to a broad range of research questions, including the analysis of the comparative efficacy of policy instruments to mitigate climate change.

## Background & Summary

**Purpose of the study.** We provide a comprehensive dataset of structural breaks in greenhouse gas (GHG) emissions for all countries in the Organisation for Economic Co-operation and Development (OECD) between 1995 and 2022<sup>1</sup>. Structural breaks refer to persistent step changes in GHG emissions that are not accounted for by economic, demographic or technological development. Having controlled for these main determinants of emissions, we seek to isolate the persistent increases or decreases in GHG emissions that are likely attributable to policy or institutional changes. The policy and institutional changes associated with the largest (negative) structural breaks can thus be considered to be the most effective at having reduced emissions. This dataset serves as a resource for the systematic and comparative evaluation of the effects of (climate) policies on GHG emissions, and thus the effectiveness of national and supranational mitigation efforts.

Traditional policy evaluation techniques follow a forward causal approach. Typically, a researcher selects a policy of interest, then seeks to isolate the effect of that policy, controlling for other socioeconomic determinants of the outcome variable. A common statistical method used in such traditional approaches is difference-in-difference (DID) estimation. The forward causal approach requires that policies of interest are selected *ex ante*, and then these policies tend to be evaluated in isolation. This risks omitting potentially beneficial but overlooked policies, and it means that policy mixes may be ignored through attempting to isolate the impact of only single policies.

The dataset presented here can be used as the basis for reverse causal policy evaluation, which first identifies significant changes in emissions, then attributes these “effects” to “causes,” typically policies<sup>2</sup>. The reverse causal approach employs an extension of the DID estimator to systematically identify structural breaks in GHG emissions that can possibly be attributed to policies that were implemented in a period around the break date, as has already been done in a number of studies with more constrained coverage than that provided by our data<sup>3–6</sup>. Existing studies have primarily explored carbon dioxide emissions, which is only one of the major contributors to anthropogenic GHG emissions. Most of these studies have also been limited in their geographic scope and sectoral coverage. As such, as compared to existing efforts<sup>6</sup>, we provide a dataset covering structural breaks in emissions at a very high level of sectoral granularity for not only carbon dioxide emissions, but for all major GHGs.

**Study overview.** Our study provides a comprehensive dataset of structural breaks in GHG emissions for all OECD countries in 1995–2022 for all major GHGs across all sectors, excluding land use, land use change and forestry (LULUCF). We use inputs on all the gases that are reported under the common reporting format of the

<sup>1</sup>Department of Economics, Vienna University of Economics and Business (WU), Vienna, Austria. <sup>2</sup>International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria. <sup>3</sup>Austrian Institute of Economic Research (WIFO), Vienna, Austria. ✉e-mail: [jesus.crespo.cuaresma@wu.ac.at](mailto:jesus.crespo.cuaresma@wu.ac.at)

United Nations Framework Convention on Climate Change (UNFCCC): carbon dioxide ( $CO_2$ ), methane ( $CH_4$ ), nitrous oxide ( $N_2O$ ) and fluorinated greenhouse gases (F-gases)<sup>7</sup>. We thus cover emissions from 23 different gases, including 19 F-gases, as well as an aggregate of all GHGs in terms of  $CO_2$  equivalents ( $CO_2-e$ ) based on their 100-year global warming potential (GWP). Emissions are disaggregated into 37 categories consistent with the 2006 Intergovernmental Panel on Climate Change (IPCC) guidelines for national greenhouse gas inventories<sup>7</sup>. This is the most granular level of detail under the 2006 IPCC guidelines, providing a very high level of detail for the analysis of climate policies in the reverse causal framework. We identify structural breaks in all OECD countries (Austria, Australia, Belgium, Canada, Chile, Colombia, Costa Rica, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Latvia, Lithuania, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Türkiye, the United Kingdom and the United States). Both positive and negative structural breaks, or step changes in emissions, are included in the dataset.

The process of identifying structural breaks follows a novel approach based on model selection within a two-way fixed effects (TWFE) model<sup>4</sup>. First, the annual panel dataset of emissions of a given gas in a given sector is regressed on year and country dummy variables and a vector of control variables which includes the gross domestic product (GDP), its square, and country population. These control variables account for affluence and population size, the two main determinants of emissions<sup>8</sup>, while the inclusion of countries in the sample with relatively homogeneous access to technology controls for technological development. The analysis has been conducted for the full sample of OECD economies and for a subsample including only the EU15 countries (Austria, Belgium, Germany, Denmark, Spain, Finland, France, United Kingdom, Ireland, Italy, Luxembourg, the Netherlands, Greece, Portugal and Sweden), as the EU15 countries were all subject to similar (supranational) climate regulations over the observation period, being within the European Union (EU).

Significant structural breaks are identified by using a saturated set of possible treatment effects. A treatment effect is a persistent (step) increase or reduction in emissions, identified after having accounted for the controls mentioned above (GDP, its square and population, as well as all non time-varying country and sector specific factors). This is essentially an extension of the common DID estimator, where variables potentially identifying every possible combination of country-year pairs are included in the model. Statistically significant effects of these variables correspond to step changes in emissions in a given country from a given year. For example, the estimated effect corresponding to “Austria-2010” would refer to a (positive or negative) step change in emissions for Austria from 2010 onward for a given gas in a given sector. As GDP and population are included as control variables, major economic or population shocks will not be identified as structural breaks. For example, a major economic downturn will likely reduce both economic activity and emissions, and so will not be identified as a break. With a fully saturated set of all possible treatment effects, the regression parameters are not identifiable, so we employ a general-to-specific (GETS) variable selection approach to keep only statistically significant treatment effects (country-year pairs) in the model. Only these statistically significant structural breaks are captured in the dataset we compile.

The reverse causal approach does not seek to replace traditional policy evaluation techniques, but it serves as a complementary approach to directly compare policies and to identify potentially overlooked policies. It overcomes some key limitations of traditional approaches to policy evaluation. First, the lack of comparability between different evaluation approaches in traditional methods is remedied by using a harmonized statistical approach to identifying effective policies. Secondly, as a single structural break can be attributed to multiple policies, the reverse causal approach allows one to identify effective policy mixes, as opposed to only exploring the effects of isolated policies. At the same time, the reverse causal approach has limitations. Structural breaks are identified as statistically significant step changes in emissions which means that effective policies that cause gradual changes in emissions over time may not be identified, or may be identified as only having delayed effects. The low target false-positive rate employed in our dataset implies a conservative approach to the identification of structural breaks and means that we tend to identify only highly significant structural breaks, potentially omitting smaller or less significant breaks.

## Methods

**Data inputs.** The Emissions Database for Global Atmospheric Research (EDGAR v8.0, [https://edgar.jrc.ec.europa.eu/dataset\\_ghg80](https://edgar.jrc.ec.europa.eu/dataset_ghg80)) database<sup>9</sup> provides annual information on GHG emissions by sector, including carbon dioxide ( $CO_2$ ), methane ( $CH_4$ ), nitrous oxide ( $N_2O$ ) and fluorinated greenhouse gases (F-gases). We use the annual time series data covering the period from 1995 to 2022 with emissions expressed as kilotons per year. Emissions are measured in  $CO_2$  equivalents ( $CO_2-e$ ), computed using the methodology in the IPCC Fifth Assessment Report (AR5) for global warming potential (GWP) over a 100-year time horizon<sup>10</sup>. The econometric model employed required data on population and GDP, which are used as control variables. GDP data is measured in constant 2015 US dollars (indicator code: NY.GDP.MKTP.KD; <https://data.worldbank.org/indicator/NY.GDP.MKTP.KD>) and population is measured as the total population per country per year (indicator code: SP.POP.TOTL; <https://data.worldbank.org/indicator/SP.POP.TOTL>). These are both sourced from the World Bank's World Development Indicators database<sup>11,12</sup>.

Two different country samples are employed as comparator groups in our analysis:

- a) All EU15 countries: Austria, Belgium, Germany, Denmark, Spain, Finland, France, United Kingdom, Ireland, Italy, Luxembourg, the Netherlands, Greece, Portugal and Sweden; and
- b) All OECD countries: the EU15 group plus Australia, Canada, Chile, Colombia, Costa Rica, Czechia, Estonia, Hungary, Iceland, Israel, Japan, Korea, Latvia, Lithuania, Mexico, New Zealand, Norway, Poland, Slovak Republic, Slovenia, Switzerland, Türkiye and the United States.

Countries	<b>Sample 1: EU15 (15 countries)</b> Austria, Belgium, Germany, Denmark, Spain, Finland, France, United Kingdom, Ireland, Italy, Luxembourg, the Netherlands, Greece, Portugal and Sweden.
	<b>Sample 2: OECD (38 countries)</b> EU15 + Australia, Canada, Chile, Colombia, Costa Rica, Czechia, Estonia, Hungary, Iceland, Israel, Japan, Korea, Latvia, Lithuania, Mexico, New Zealand, Norway, Poland, Slovak Republic, Slovenia, Switzerland, Türkiye and the United States.
Gases	1. All GHGs (in CO <sub>2</sub> -equivalents)
	2. CO <sub>2</sub>
	3. CH <sub>4</sub>
	4. N <sub>2</sub> O
	5. F-gases (aggregated emissions of c-C <sub>4</sub> F <sub>8</sub> , C <sub>2</sub> F <sub>6</sub> , C <sub>3</sub> F <sub>8</sub> , C <sub>4</sub> F <sub>10</sub> , CF <sub>4</sub> , HCFC-141b, HCFC-142b, HFC-125, HFC-134a, HFC-143a, HFC-152a, HFC-227ea, HFC-23, HFC-245fa, HFC-32, HFC-365mfc, HFC-43-10-mee, NF <sub>3</sub> , SF <sub>6</sub> )
Sectors (2006 IPCC Categories)	<b>37 Sectors:</b> Electricity and heat, Petroleum refining, Manufacturing & construction, Civil aviation, Road transport, Railways, Water-borne navigation, Other transportation, Residential, Non-specified, Solid fuels, Oil and natural gas, Cement production, Lime production, Glass Production, Carbonate uses, Chemical industry, Metal industry, Non-energy fuel products, Electronics industry, Substitutes for ODS, Other product manufacture, Enteric fermentation, Manure management, Biomass burning, Liming, Urea application, Direct N <sub>2</sub> O - managed soils, Indirect N <sub>2</sub> O - managed soils, Indirect N <sub>2</sub> O - manure, Rice cultivations, Solid waste disposal, Biological treatment - waste, Incineration of waste, Wastewater treatment, Indirect N <sub>2</sub> O - nitrogen, Fossil fuel fires.

**Table 1.** Overview of input data including countries, gases and sectors in the sample.

The two above sample groups were specifically chosen due to the relative homogeneity of countries in terms of regulations and the countries' ability to implement decarbonisation technologies. The period of 1995 to 2022 was chosen because most EU countries were subject to similar regulations under the European Single Market for this period. In principle, as the data inputs are emissions, population and GDP, for which accessible data exists for almost all countries, this approach could be implemented for almost all countries in the world. However, it is possible that the impact of population, GDP and policy changes on emissions differs between developed and developing countries, hence the decision to maintain a relatively homogeneous sample of only developed countries in the analyses.

The identification of structural breaks in the EU15 sample exploits the fact that EU countries are subject to a set of similar regulations with regards to climate policy, making the regulatory environment for implementing climate policies relatively homogeneous over the observation period. For instance, the European Climate Law establishes a legally binding commitment for the EU to achieve climate neutrality by 2050, and the EU Industrial Emissions Directive establishes common guidelines for reducing industrial emissions across the whole of the EU. While such harmonized regulations exist at the European level, countries have the freedom to implement additional national-level policies.

The use of the OECD sample allows for the identification of structural breaks that may have occurred across the entire EU15 sample due to the harmonized regulations, while still maintaining a relatively homogeneous sample group, as OECD countries are all market-based economies with relatively similar access to technologies and levels of development. The two sample groups allow for comparison and robustness, while still maintaining a relatively homogeneous sample in terms of countries' ability to decarbonize.

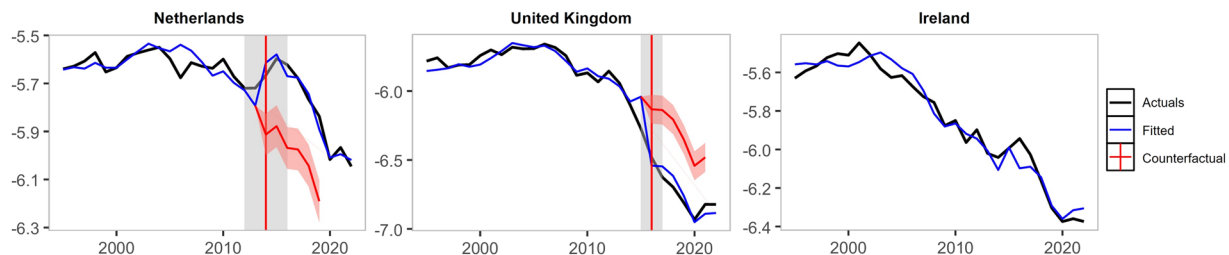
Table 1 provides an overview of the input data, including countries, gases and sectors in the sample.

**Identifying structural breaks in emissions data.** The abnormal changes in emissions are estimated using methods of structural break detection for panel models, with two-way fixed effects (TWFE) and a general-to-specific (GETS) variable selection approach, as has been implemented in existing studies<sup>4,13,14</sup>. The assumed data generating process for emissions of gas  $g$  in sector  $i$  of country  $j$  is given by

$$\log(E_{i,j,t}^g) = \alpha_j + \phi_t + \sum_{k=1}^N \sum_{s=2}^T \tau_{k,s} \mathbf{1}_{\{j=k, t \geq s\}} + x'_{j,t} \beta + \varepsilon_{i,j,t}, \quad (1)$$

where  $E_{i,j,t}^g$  denotes emissions,  $\mathbf{1}$  is the Heaviside function, which takes value one if its argument is true, and zero otherwise, and the column vector  $x_{j,t}$  includes (log) population, (log) GDP and (log) GDP-squared as control variables. The error term  $\varepsilon_{i,j,t}$  is assumed to fulfill all the assumptions of the normal linear regression model. We carry out our inference in balanced panel datasets by sector. For a sector-specific panel formed by  $N$  countries and  $T$  time periods, the most general specification includes  $N(T-1)$  potential breaks, with corresponding coefficients  $\tau_{k,s}$  on the indicator covariates. Each one of these variables represents a potential structural break, or a step function, which changes the level of the emissions variable from year  $s$  until the end of the sample. The model as presented by this regression specification cannot be estimated, since the number of unknown parameters is larger than that of sample observations. We thus employ the GETS algorithm to perform model selection in the specification<sup>14</sup>, and the estimated coefficients of the selected model allow for an interpretation similar to that of the treatment effect in a DID estimator.

We implement the GETS algorithm using the "getspanel" package in R<sup>15</sup>, using the block search algorithm of the "gets" package<sup>13</sup>. We calibrate the level of target significance, so as to control the expected false-positive rate of the selected indicators<sup>16</sup>, and employ three different levels of target significance: 5%, 1% and 0.1%. The use of different target significance levels provides a test of the robustness of the identified breaks, with low target significance levels implying that the identified breaks are likely to constitute large (significant) changes in emissions. Our conservative approach effectively implies that we identify minimum effect sizes, and thus lower-bound estimates of abnormal changes in emissions.



**Fig. 1** Example plots of a positive structural break (Netherlands), a negative structural break (United Kingdom), and a time series without any identified structural break (Ireland). The black line shows actual values, the blue line shows fitted values, and the red lines shows the counterfactual, in the absence of a structural break, for a period of five years after the break date. Vertical red lines depict the break date, and grey bars around these lines depict a 95% confidence interval around the break date. Data is from the “Main Activity Electricity and Heat” sector, and structural breaks are identified at a 5% significance level.

	Negative			Positive		
	Mean	SD	Obs.	Mean	SD	Obs.
<b>Gases</b>	732.6	373.8	3663	776.0	399.7	3880
<b>Countries</b>	96.4	61.3	3663	102.1	51.4	3880
<b>Sectors</b>	99.0	65.2	3663	104.9	66.9	3880
<b>Years</b>	135.7	65.9	3663	143.7	78.7	3880
<b>Samples</b>	1831.5	450.5	3663	1940.0	620.0	3880

**Table 2.** Descriptive statistics. Mean values refer to the mean number of total structural breaks identified in each category.

Figure 1 presents three examples plots from the “Main Activity Electricity and Heat” sector for a positive structural break (identified in the Netherlands), a negative structural break (United Kingdom), and a time series without any identified structural break (Ireland). The graphs depict also the counterfactual dynamics of emissions in the absence of the structural break, which allows us to visualize the magnitude of the abnormal change in emissions.

The data input of the method employed for the creation of our dataset is thus composed by the emissions data sourced from EDGAR v8.0., as well as the GDP and population data from the World Bank’s World Development Indicators. The output is a panel dataset of country/sector structural breaks in emissions identified by the model, as well as their magnitude. These structural breaks are to be interpreted as persistent step (level) changes in emissions for a given sector in a given country, as compared to the benchmark region (EU15 or OECD).

### Data Records

The dataset “all\_structural\_breaks.csv” is available in an online Zenodo database (<https://doi.org/10.5281/zenodo.13325884>)<sup>1</sup>. Each row in the dataset corresponds to an identified structural break, with the 14 columns containing details about the structural break including the GHG, sample, sector, magnitude, country, year and direction of the break.

The column “gas” details which type of greenhouse gas the structural break was identified for, and can take the following values: “all\_ghg” which is the aggregate of all GHGs measured in CO<sub>2</sub>-e, “ch4” which is methane, “co2” which is carbon dioxide, “fgas” referring to all fluorinated gases, and “n2o” which is nitrous oxide. The “sample” column indicates whether a break was identified in the EU15 or OECD sample. The “IPCC code” column details the sector code according to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories, with the corresponding “sector group” and “sector description” in the subsequent columns. Sectors are grouped into five categories: Energy; Industrial processes and product use (IPPU); Agriculture, forestry, and other land use; Waste; and Other. The “target p-value” column indicates which target p-value (or false positive rate) the break was identified for. The columns “country” and “country code” list the country in which the break was identified, both using the full name of the country and its standardised 2-letter ISO country code. The column “year” corresponds to the break date. The “coef”, “std.error”, “t-stat”, and “p-value” refer to the estimated coefficient on the break, and its standard error, test statistic and p-value, respectively. Finally, the “break direction” column details if the break is negative or positive.

Table 2 outlines basic descriptive statistics on negative and positive structural breaks by gas, country, sector, year and sample. The mean values refer to mean number of total breaks identified within category. For example, a mean of 96.4 negative structural breaks were identified per country across the entire dataset.

Tables 3–6 provide a breakdown of the total number of negative and positive identified structural breaks by gas, country, sector and break year, respectively.

Gas (chemical symbol)	Negative	Positive	Total
All GHGs (in CO <sub>2</sub> -e based on 100 year GWP)	1218	1293	2511
Methane (CH <sub>4</sub> )	717	693	1410
Carbon Dioxide (CO <sub>2</sub> )	540	639	1179
Fluorinated gases (F-gases)	154	147	301
Nitrous oxide (N <sub>2</sub> O)	1034	1108	2142
Total	3663	3880	7543

**Table 3.** Total number of breaks by type of gas and break direction.

Country	Negative	Positive	Total
Australia	47	36	83
Austria	62	131	193
Belgium	126	118	244
Canada	31	36	67
Chile	46	118	164
Colombia	49	112	161
Costa Rica	85	141	226
Czechia	85	35	120
Denmark	167	85	252
Estonia	89	89	178
Finland	132	95	227
France	149	95	244
Germany	131	116	247
Greece	175	232	407
Hungary	59	84	143
Iceland	74	78	152
Ireland	261	147	408
Israel	67	80	147
Italy	133	145	278
Japan	52	34	86
Latvia	70	153	223
Lithuania	58	144	202
Luxembourg	178	129	307
Mexico	3	144	147
Netherlands	154	66	220
New Zealand	26	98	124
Norway	15	50	65
Poland	58	59	117
Portugal	88	194	282
Republic of Korea	82	86	168
Slovak Republic	95	32	127
Slovenia	98	97	195
Spain	109	235	344
Sweden	237	114	351
Switzerland	48	23	71
Turkiye	34	140	174
United Kingdom	230	64	294
United States	60	45	105
<b>Total</b>	<b>3663</b>	<b>3880</b>	<b>7543</b>

**Table 4.** Total number of breaks by country and break direction.

### Technical Validation

We validate the identified structural breaks using the block search algorithm of the GETS variable selection process. An iterative process selects significant structural breaks to reduce the general model to a specific model, with only statistically significant structural breaks included in the final specification. To ensure a high level of

Sector	Negative	Positive	Total
Biological treatment - waste	64	106	170
Biomass burning	129	118	247
Carbonate uses	42	70	112
Cement production	8	4	12
Chemical industry	71	126	197
Civil aviation	119	91	210
Direct N <sub>2</sub> O - managed soils	126	128	254
Electricity and heat	205	217	422
Electronics industry	19	13	32
Enteric fermentation	102	107	209
Fossil fuel fires	9	7	16
Glass Production	6	36	42
Incineration of waste	140	153	293
Indirect N <sub>2</sub> O - nitrogen	126	126	252
Indirect N <sub>2</sub> O - managed soils	126	94	220
Indirect N <sub>2</sub> O - manure	140	140	280
Lime production	2	16	18
Liming	26	40	66
Manufacturing & construction	221	202	423
Manure management	175	200	375
Metal industry	68	75	143
Non-energy fuel products	150	146	296
Non-specified	94	84	178
Oil and natural gas	150	167	317
Other product manufacture	155	156	311
Other transportation	117	152	269
Petroleum refining	119	133	252
Railways	72	68	140
Residential	127	161	288
Rice cultivations	14	25	39
Road transport	277	306	583
Solid fuels	54	54	108
Solid waste disposal	26	34	60
Substitutes for ODS	95	67	162
Urea application	22	24	46
Wastewater treatment	153	157	310
Water-borne navigation	114	77	191
Total	3663	3880	7543

**Table 5.** Total number of breaks by sector and break direction.

confidence in the identified breaks and a low false positive rate, we calibrate the model with three levels of target significance: 5%, 1% and 0.1%. The conservative significance levels mean that only highly significant structural breaks are identified. The GETS method is known to have very good performance in terms of false positive rate and predictive error as compared to LASSO and adaptive LASSO alternatives<sup>13</sup>. The GETS method proceeds in three steps:

- (i) it starts with a general unrestricted model, which includes as many variables as possible and that passes standard specification tests,
- (ii) backwards elimination of variables in the model is carried out along several paths, making use of significance tests and ensuring validity using a set of diagnostic tests,
- (iii) a final specific, identifiable model is chosen maintaining only those variables that are still identified as statistically significant.

Such a methodological framework ensures that the chosen model only contains statistically significant structural breaks and, thus, that the effects found are strongly supported by the available data.

Further, validation of specific breaks is achieved through comparison of breaks across samples and visual inspection of the model fit with the actual time series of emissions. As shown in Fig. 1, the fitted models align

Year	Negative	Positive	Total
1996	5	14	19
1997	43	29	72
1998	112	86	198
1999	115	83	198
2000	203	177	380
2001	166	169	335
2002	120	171	291
2003	188	212	400
2004	170	223	393
2005	210	196	406
2006	213	230	443
2007	150	207	357
2008	185	195	380
2009	203	241	444
2010	225	182	407
2011	219	229	448
2012	151	193	344
2013	166	246	412
2014	95	165	260
2015	190	129	319
2016	121	162	283
2017	124	172	296
2018	83	102	185
2019	76	22	98
2020	122	37	159
2021	1	5	6
2022	7	3	10
Total	3663	3880	7543

**Table 6.** Total number of breaks by year and break direction.

closely with the actual time series, suggesting that the specified model captures a significant proportion of the variation in the true data generating process, thus validating the chosen specification of the model.

### Usage Notes

This dataset serves as the foundation for analysing climate policies using the reverse causal approach across all major GHGs and all sectors. The reverse causal approach has already been implemented in a number of studies to evaluate the effectiveness of climate policies, but these have been limited in their sectoral coverage, scope of GHGs or geographic coverage<sup>3-6</sup>. These studies serve as a useful basis for implementing this method, and our dataset serves as a foundation for further exploration of the effectiveness of climate policies across different GHGs and countries. Further, the richness of the dataset allows for comparisons between countries across a range of sectors, a dimension of heterogeneity which has not yet been systematically implemented in literature to date. Specifically, the dataset covers 37 different sectors as defined by the 2006 IPCC Guidelines for National Greenhouse Gas Inventories, with an aggregation of sectors provided into the following five groups: Energy; Industrial processes and product use (IPPU); Agriculture, forestry, and other land use; Waste; and Other. The data do not include emissions from land use, land use change and forestry (LULUCF).

The approach of the above studies involves first identifying structural breaks, then attributing these breaks to relevant policies. Policies are generally attributed to breaks in the same sector in a confidence interval around the year of the break, accounting for lags in policy effects on emissions or anticipatory effects of upcoming policies. For example, a negative break in emissions in the Austrian transport sector from 2010 onward may be attributed to a transport emissions tax implemented in 2009. A break could be attributed to a single policy, or a mix of policies implemented in the interval. In fact, in previous studies, many breaks are associated with multiple policies and the interaction effects between policies is starting to be explored. The approach to identifying policies differs in the aforementioned studies, but the use of policy databases, such as those of the OECD or the International Energy Agency (IEA) serve as a good starting point. These attributed policies can further be analysed to draw conclusions about which kinds of policies were most effective in a particular country or sector.

Further analysis of these structural breaks and attributed policies could help to identify “role model countries” or “best practice policies” based on those countries or policies with the most or largest negative structural breaks. This would then inform decision making about the most effective climate policies going forward, based on a systematic and holistic approach to measuring climate policy effectiveness.

Another point of interest not yet explored in the literature is the presence of positive structural breaks, that is, persistent increases in emissions not attributed to GDP, population or technological change. Analysis of such breaks could shed light on potential unintended consequences of policies or rebound effects.

### Code availability

The code used to identify the structural breaks as well as the underlying data sources are available in an online database (<https://doi.org/10.5281/zenodo.13325884>)<sup>1</sup>. The statistical programming software R was used to generate the results.

Received: 20 August 2024; Accepted: 17 December 2024;

Published online: 10 January 2025

### References

1. Tebecis, T. & Crespo Cuaresma, J. Structural breaks in greenhouse gas emissions for OECD countries in 1995–2022 and 37 sectors. *Zenodo*. <https://doi.org/10.5281/zenodo.14223997> (2024).
2. Gelman, A. & Imbens, G. Why ask Why? Forward Causal Inference and Reverse Causal Questions. *National Bureau of Economic Research*, (19614) (2013).
3. Tebecis, T. Have climate policies been effective in Austria? A reverse causal analysis. *WU Vienna University of Economics and Business*, (346), August (2023).
4. Koch, N., Naumann, L., Pretis, F., Ritter, N. & Schwarz, M. Attributing agnostically detected large reductions in road CO<sub>2</sub> emissions to policy mixes. *Nature Energy* 7(9), 844–853 (2022).
5. Yao, J. & Zhao, Y. Structural Breaks in Carbon Emissions: A Machine Learning Analysis. *IMF Working Paper* (2022).
6. Stechemesser, A. *et al.* Climate policies that achieved major emission reductions: Global evidence from two decades. *Science* 385(6711), 884–892 (2024).
7. Eggleston, H.S., Buendia, L., Miwa, K., Ngara, T. & Tanabe, K. IPCC guidelines for national greenhouse gas inventories. *Intergovernmental Panel on Climate Change (IPCC)* (2006).
8. Hamilton, C. & Turton, H. Determinants of emissions growth in OECD countries. *Energy Policy* 30(1), 63–71 (2002).
9. European Commission, JRC (Datasets). EDGAR (Emissions Database for Global Atmospheric Research) Community GHG Database, a collaboration between the European Commission, Joint Research Centre (JRC), the International Energy Agency (IEA), and comprising IEA-EDGAR CO<sub>2</sub>, EDGAR CH<sub>4</sub>, EDGAR N<sub>2</sub>O, EDGAR F-GASES version 8.0., 2023. Online; accessed 26 October (2023).
10. The Intergovernmental Panel on Climate Change (IPCC). Climate Change 2014 Synthesis Report. *Intergovernmental Panel on Climate Change (IPCC)* (2014).
11. The World Bank. GDP (constant 2010 US\$). *World Bank Open Data*. Online; accessed 26 October (2023).
12. The World Bank. Population, total. *World Bank Open Data*. Online; accessed 26 October (2023).
13. Pretis, F., Reade, J. & Sucarrat, G. Automated general-to-specific (GETS) regression modeling and indicator saturation for outliers and structural breaks. *Journal of Statistical Software* 86, 1–44 (2018).
14. Pretis, F. & Schwarz, M. Discovering what mattered: answering reverse causal questions by detecting unknown treatment assignment and timing as breaks in panel models. *SSRN 4022745*, (2022).
15. Pretis, F. & Schwarz, M. *getspanel*, *GitHub repository* (2021).
16. Nielsen, B. & Qian, M. Asymptotic properties of the gauge of step-indicator saturation. *University of Oxford Working paper* (2018).

### Acknowledgements

We thank the editor and two anonymous reviewers for helpful comments on the original submission. We thank Felix Pretis and Moritz Schwarz for their consultation on the approach and review of the initial results. We acknowledge financial support from the eXplore! initiative, funded by the B&C Privatstiftung and Michael Tojner, under the grant “Ein Klimaplan für Österreich”, as well as from the Vienna Science and Technology Fund under the grant ESS22-040.

### Author contributions

Conceptualisation: J.C.C., T.T. Methodology: J.C.C., T.T. Visualization: T.T. Data collection: T.T. Validation: T.T. Writing: J.C.C., T.T. Editing: J.C.C., T.T.

### Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to J.C.C.

**Reprints and permissions information** is available at [www.nature.com/reprints](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) 2025