Codebook

We capture two levels of information: study level and effect level information. We do the first by using the NACSOS software[1] which automatically records study level information concerning authors, publication year, publication name and publication type, where available from the searched databases. We then use the NACSOS software for relevance coding on the title-abstract level. Our search strategy is described in detail in our protocol.

This document describes the coding at the study level and the effect level done in Excel. For studies which appeared relevant on the title-abstract level we first record whether the study fulfils our inclusion criteria after reading the full text. At the study level we further assess the risk of bias, informed by the ROBINS-I framework[2]. We amend the ROBINS-I assessment criteria in order to be applicable to our set of non-randomised observational studies on the impact of an intervention at the aggregated level (e.g. country-level, sector-level). The criteria for our risk of bias assessment are described below. We capture the relevant information on the carbon pricing schemes from the primary studies. Further information on the schemes can be added later from other sources.

Study level information

Field Name	Explanation	Choices or Examples
Include	The inclusion/exclusion criteria are specified in the protocol.	*Options* include exclude
Exclusion reason	If you exclude the article, please specify why you exclude it.	*Options* No relevant effects, Effect information incomplete/missing, Other
Exclusion reason description	Please describe your decision for exclusion further. You have put some thoughts into this so it is the right time to verbalise them. If someone "challenges" your decision later, it will be very helpful to have these thoughts recorded. Especially if you had to exclude the study because of missing or incomplete information, all pieces of information provided here will be	e.g. "The study only analyses leakage effects of the policy intervention, but does not provide any estimates on the development of carbon emissions within the territory of its application.", "The study only provides descriptive information on the carbon emission path but does not estimate any causal effects regarding the implementation of the intervention."

[1] More information on the NACSOS software can be found at https://github.com/mcallaghan/nacsos.
[2] More information on the ROBINS-I framework can be found at http://www.riskofbias.info.

	very helpful in the subsequent process of trying to recapture the missing information. Also provide as accurately as possible information about in which table or paragraph and on which page you would expect the missing information.	
The subsequent labels cover the **risk of bias assessment**. The assessment is inspired by the ROBINS-I tool which is based on the Cochrane risk of bias tool for randomized trials. All studies in our sample are non-randomized intervention studies working with real-world policies and real-world data. As such, study results are at risk not to report the true effect of the intervention on the outcome variable but to report biased estimates influenced by confounding factors, selective reporting, or other biases. The subsequent questions assess how severe this risk is. The benchmark should always be an idealised study, which can perfectly single out the effect of the intervention on the dependent variable. You can select from five different answers: Yes, Probably Yes, Probably No, No, No Information. Not applicable can be selected for some questions if the question does not apply to the study design. The answers should be selected according to the following judgement. Yes: Critical risk of bias (the study design is too problematic in this domain to provide any useful evidence on the effects of intervention); Probably Yes: Serious risk of bias (the study has some important problems in this domain); Probably No: Moderate risk of bias (the study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial); No: Low risk of bias (the study is comparable to a well-performed randomized trial with regard to this domain); No information: There is no information on which to base a judgement about risk of bias for this domain.		
RoB – selection of treated study objects	Is there a risk of bias in the selection of treated study objects? *The treatment in this context is independent of the research design and subject to political decisions of the jurisdictions. The study should cover all treated study objects or a representative sample.* *We judge the risk of bias with respect to the country-sector-fuel analysed by the authors. So if the study e.g. focuses on the energy sector, this does not lead to an increased risk of bias.*	*Options* Yes, Probably Yes, Probably No, No, No Information
RoB – selection of control group	Is there a risk of bias in the selection of the control group (if relevant for the study design)? *In the idealized setting the control group would perfectly mimic the development of the emissions in the treatment group in absence of the carbon pricing policy.* *Risk of bias can be reduced if the control group is chosen based on regional, institutional, cultural, or economic arguments (e.g., comparing neighbouring countries or regions of the same country). It should be reasonably*	Yes, Probably Yes, Probably No, No, No Information, Not applicable

	outlined how the control group compares with the treatment group in decisive factors. In particular pre-intervention GHG pathways and other descriptive statistics should be compared. Risk of bias can be further reduced through the application of study designs which increase the comparability of the control group with the treatment group, like matching or synthetic controls.	
RoB – selection description	Describe or use quotes from the paper which illustrate how the authors deal with the risk of bias from the selection of treatment and control group.	
RoB – confounding factors	<u>Is there a risk of confounding factors, which are not appropriately controlled for?</u> *The risk can be reduced if the estimation procedure reasonably controls for confounding factors. Such confounding factors could either be differences in characteristics between the treatment and control group which could, independent of the treatment, have an effect on the emissions (dependent variable), like initial GHG emissions, environmental awareness, country factors like GDP, governance or other environmental regulation/policy. Other confounding factors could be other events, like the global financial crisis or the implementation of other relevant policies, happening during the study period, which could have a significant effect on the emissions.* *The selection of control variables (and exclusion of others) should be reasonably justified and discussed with reference to the relevant literature as well as with robustness analyses of the findings with respect to the inclusion/exclusion of critical control variables.* <u>If instrumental variables (IV) are used, is there a risk of bias due to an invalid IV?</u> *The validity of the IV should be reasonably justified. In particular the relevance condition and the exclusion restriction should hold. For the relevance condition, tests should be provided,*	Yes, Probably Yes, Probably No, No, No Information, Not applicable

	which show that the instrument is a relevant predictor for the observed variation in the explanatory variable (the policy intervention). For the exclusion restriction it should be credibly justified that the instrument only affects the emissions (dependent variable) through the explanatory variable (is it credible that the (changes in) greenhouse gas emissions/fuel consumption were only affected by the instrument through the variation in carbon pricing?). <u>If a regression discontinuity in time design is used, is there a risk of bias due to the identifying assumptions underlying the design?</u> *For the justification of the identifying assumptions, in particular the political and economic context should be discussed as well as (graphical) evidence provided for the smoothness of trends in observed control variables around the cutoff. Robustness checks with respect to a varying bandwidth around the cutoff as well as with respect to the inclusion of higher order polynomials in the outcome (emissions/fuel consumption) trends before and after the cutoff further reduce the risk of bias.*	
RoB – confounding factors Description	Describe or use quotes from the paper which illustrate how the authors deal with the risk of bias caused by confounding factors.	
RoB – other	Are there any other biases in the study? If the validity of the study is restricted by any other factors, please describe the present biases in the next label. *One possible bias could arise in price elasticity studies from the endogeneity of carbon prices and carbon emissions. This can particularly be the case if prices are formed on markets and the emission levels of single study objects have a significant influence on the price.*	Yes, Probably Yes, Probably No, No
RoB – other description	Please describe any other present biases here and/or paste in quotes from the paper.	

Effect level information

For included studies we will subsequently record information at the effect size level, i.e. the estimated effects of carbon pricing on one of the outcome variables of interest. The risk of bias assessment does not influence the inclusion/exclusion of the study and all further information shall be captured regardless of the risk of bias assessment.

Each study may report multiple estimates of the effect. This can be due to different model specifications, different control groups, different periods, separately calculated effects for different sectors, etc. You are asked to capture all the *relevant* effect sizes and corresponding information reported by the study, this includes effects reported in supplementary information such as the appendix. Obviously, model specifications which do not contain relevant effect sizes (e.g. if the carbon pricing variable was excluded from a model specification) can and should not be captured.

Field Name	Explanation	Choices or Examples
Effect size type	Introduction (treatment) effects specify the effect of the presence of the policy compared to its absence. This would usually be done using a dummy variable. A price elasticity is estimated using the carbon price level as an explanatory variable for the emission level.	Introduction effect Price elasticity
Effect number	Please assign numbers to the effects. Start with the first reported effect size in the main text. Do not use the numbers assigned to columns in tables of results as these may be ambiguous (there may be several tables of results).	1, 2, 3,…
Location of the information	Please note down where the effect size (and related uncertainty measures etc.) was found. Was it captured in a table, chart or a text paragraph? If available, please use the page numbers of the article, not the pdf. If no other page numbers are available, use the page numbers of the pdf. Please keep in mind, that it might be necessary to go back to individual studies at some point in a later project phase. If any additional information on the location of the values is	page 683; table 4; column 1, page 89; table 4; column 5 and page 90 second paragraph, page 9; paragraph 16 appendix B; page 5; table A.3; column 4;

Field Name	Explanation	Choices or Examples
	needed to find the information again, please capture that here.	
Study design type	Choose one or multiple of the options. Please contact us, if you discover a study design which does not fit the predefined options, or if you are unsure which category to choose.	DiD, DiDiD, Matching, synthetic control, 2SLS, Regression discontinuity/structural break, Panel without fixed effects, Panel with fixed effects, time series OLS (pure time series), OLS (pure cross section), other
Dependent variable co2/GHG	How is the dependent variable defined? If the dependent variable originally is not in CO2 emissions or CO2 equivalents but can be recalculated with information provided by the authors, perform the calculation into CO2 equivalents on your own and select Greenhouse gases.	CO2, Greenhouse gases (in CO2 equivalents)
Dependent variable total/capita	How is the dependent variable defined?	Total, Per capita
Dependent variable levels/growth rates	How is the dependent variable defined?	Levels, Growth rates
Log-transformations of dependent and independent variables	Capture whether the dependent and/or independent variables are transformed to log values. e.g., choose log-level if the dependent variable is in logs and the independent variable is in levels.	Level-level, Log-Level, Log-Log, Level-Log, relative change in percent/100

Field Name	Explanation	Choices or Examples
Larger risk-of-bias dummy	Here we are interested in the risk of bias of individual estimates relative to the risk of bias assessment on the study level. Does the estimate have a larger risk of bias then the best estimates (i.e. the estimates with the lowest risk of bias, included in the study). E.g., this can be estimates which have been estimated without the control variables deemed necessary by the authors or without fixed effects etc. Some robustness checks may also fall into this category.	Yes/No
Larger risk-of-bias - Explanation	If you assessed the estimated effect size to have a larger risk of bias, please explain what causes the higher risk of bias.	e.g. “The effect size was estimated without fixed effects.", “The effect size was estimated leaving out the control variables for GDP."
Complete effect size information	For the meta regression we need information on the effect size as well as complete information on at least one uncertainty measure.	yes/no
Effect size – statistical estimate	Capture the statistical estimate of the effect of the explanatory variable (introduction/existence of carbon pricing/carbon price) on the outcome variable (emissions etc.) as given by the authors.	
Effect size – statistical estimate standardized	We want to standardize effect sizes to make them comparable across studies. If the effect sizes in a study require recalculations to fit the required format, carry out the necessary recalculations and capture the standardized effect size. If the effect is given in percent, divide it by 100. If the effect size is given in CO2 (equivalent) capture it in tons.	

Field Name	Explanation	Choices or Examples
Effect size - direction	Capture the direction of the effect carbon pricing had on the outcome variable. If the effect suggests that carbon pricing decreased the outcome variable relative to a baseline/control scenario choose “decrease”. Otherwise choose “increase”.	increase, decrease
Cumulative-periodic effect size	Capture in which form the effect size is provided. If a cumulative emission reduction is provided for the full time period from the time of introduction until the end of the observation period, select "cumulative". If the reduction is reported as the reduction in each observation point (e.g., for annual observations, the annual reduction), select “periodic”. If the reduction in each period is estimated choose “periodic”.	Cumulative, Periodic, other
Uncertainty measure	Capture the uncertainty measure provided for the effect size. Fill in all provided uncertainty measures. If only a confidence interval is provided capture the confidence interval in the following form: confidence interval; lower bound; upper bound.	standard error t-statistic p-value confidence interval (upper and lower bound)
Uncertainty measure standardized	If the statistical estimate had to be recalculated for standardization, the uncertainty measure needs to be recalculated accordingly.	
Sample Size	Capture the observed entities/individuals with as much detail as possible (i.e. if all three options are given, record all). Total – Full sample size (for synthetic control estimations, capture the treatment group and the full donor pool) Treatment – Capture sample size of entities with carbon pricing Control – Capture the sample size of entities without carbon pricing.	In a clinical study, this would be the number of participants in the trial.

Field Name	Explanation	Choices or Examples
Number of observations	If panel data is analysed the number of observations in a regression differs from the sample size of the study.	
Significance level	If the study indicates the level of significance at which the effect size is significant, capture the highest level of indicated significance. Often this indication of significance is provided by asterisks/stars in the results table or it is mentioned in the text.	<0.90, 0.90, 0.95, 0.99, >0.99, Insignificant, not reported,
Main specification	Choose 'main specification' to indicate if the authors describe this estimate as their main specification.	main specification
Mean carbon price	Capture the mean carbon price level of the treated countries as specified in the study. This is manly important, if the study estimates a price elasticity. For price semi-elasticities the carbon price entered here should correspond to the semi-elasticity recorded in the "Effect size – statistical estimate" field such that captured semi-elasticity can be transformed to an introduction effect by multiplying it with the mean carbon price. Only consider observation periods where a carbon price was introduced i.e. the average carbon price between the introduction date of the carbon tax and the end of the treatment period.	
Mean carbon price - currency	Capture the currency of the explanatory variable. If you do not indicate otherwise we assume that prices are current prices. It is preferrable to capture the mean price level in constant prices, if given in the paper.	Euro US dollar Canadian dollar Chinese Yuan New Zealand dollar Norwegian krone

Field Name	Explanation	Choices or Examples
	If the carbon price is not given per ton of CO_2-equivalent, specify the currency and the quantity of fuel the carbon price is referring to (see example).	Swedish krona UK pound sterling Canadian cents per litre of gasoline
Mean carbon price – currency base year	Please specify the base year if a paper reports constant prices.	
Mean carbon price - **standardized**	Capture the mean carbon price in currency unit per ton of CO2. For price semi-elasticities the carbon price entered here should correspond to the semi-elasticity recorded in the "Effect size – statistical estimate **standardized**" field such that captured semi-elasticity can be transformed to an introduction effect by multiplying it with the mean carbon price standardized.	
Mean carbon price – currency **standardized**	As specified above but for the standardized mean carbon price.	
Mean carbon price – currency base year **standardized**	As specified above but for the standardized mean carbon price.	
Explanatory variable - currency	Capture the currency of the explanatory variable. If you do not indicate otherwise we assume that prices are current prices. This is manly important, if the study estimates a price elasticity.	Euro US dollar Canadian dollar Chinese Yuan New Zealand dollar Norwegian krone Swedish krona UK pound sterling
Explanatory variable – currency base year	Please specify the base year, if papers use constant prices.	
Introduction date	Capture the year when the carbon price analysed in the study was introduced.	1999, 2005, 01.07.2015

Field Name	Explanation	Choices or Examples
	If the carbon price was not introduced on January 1, capture the full introduction date in the form DD.MM.YYYY	
Start date	We would like to capture the period which the calculation of the effect is based on. Capture the start of the observation period. If the observation period does not start on January 1, use the format DD.MM.YYYY	*e.g.* 2005, 01.03.2003 In a synthetic control study, capture the first year of the period on which the calculation of the weights is based on.
End date	Fill in the last year of the treatment period here. If the observation period does not end on December 31, use the format DD.MM.YYYY	2017, 01.12.2018
Time aggregation level	Capture the level of aggregation for the emission data on a time dimension. Choose one of the options.	Larger than year, Year, Quarter, Month, Lower than a month
Entity aggregation	Capture the entity aggregation level at which the data is being analysed. Choose one option. If, for example, the observations are aggregated by sector and country, choose both.	block, country, state/region, city, firm, plant, sector
Mean emissions	EITHER 1.) Capture the mean annual emissions in tonnes of CO2 (or CO2 equivalents), if provided or if they can be calculated. Then standardize the effect size to match the emission reduction in tonnes and enter it in the field effectSize.statisticalEstimate.standardized OR 2.) Capture in the same unit as the effectSize.statisticalEstimate such that the ratio between effectSize.statisticalEstimate/mean(emissions) can be interpreted as the percentage change in emissions.	The mean emissions should have the same base, as the effect, e.g., if the effect is based on a per capita calculation, capture the mean emissions per capita. If you have to perform recalculations: 1 year = 365 days

Field Name	Explanation	Choices or Examples
	Preference order for the recording of mean emissions: (1) mean annual counterfactual emissions during the treatment period (2) mean annual emissions of the treatment group during the treatment period (3) mean annual emissions of the full sample during the treatment period (4) mean annual emissions of the full sample over the full observation period Specify any thoughts you put into your selection in the notes.	
Geographic location - treatment group	Capture the geographic location for which emission data was analysed. Here we do not look for the scope of the intervention. For the effect of the EU ETS on emission from the German electricity sector the value will be *Germany*. Note: Try and stick to the format city/state;country/block. In case that there is more than one entity analysed connect them with a comma. city/state;country/block,city/state;country/block If a block of locations was analysed and there exists an unambiguous name, put the name. Remove all white spaces. If country names require a whitespace, replace it with a lower bar. As ambiguity in spelling names is common, please consider the following partial list: British_Columbia Be**iji**ng China Hubei New_Z**ea**land USA	*e.g.* city/state;country/block Berlin;Germany, France, Pennsylvania;USA, Shenzhen;China. British_Columbia;Canada EU-25, EU-27, China_pilots (if **all** China pilot jurisdictions are included), Regex: \S+

Field Name	Explanation	Choices or Examples
Multiple interventions	Capture if a single carbon pricing intervention was studied or whether multiple carbon pricing interventions are studied. (e.g. does the study analyse the effect of the EU ETS only or does the dataset contain multiple jurisdictions with multiple carbon pricing policies.) Choose one of the options.	single, multiple
Emission sector	Capture the emission sector which the analysis is performed on. Choose one or multiple options. Note: if sectors are not specified but it is indicated that all sectors covered by an intervention are analysed choose *all covered sectors.*	*Options (sector definition of the IPCC)* Energy, Industry, Transport, Buildings, AFOLU, International Aviation and Shipping, all covered sectors, economy,
Fuel type	If only a specific fuel type is analysed choose the type accordingly. Leave unanswered otherwise.	Coal, Natural gas, Petrol, Gasoline, Diesel
Sub-category	If only a subset of the sector or fuel is analysed, please specify the sector/fuel here.	Residential buildings only
Used data	Capture the data source used for the dependent variable. In particular, which database was used to extract (or calculate) the carbon/GHG emissions. Note: *Try* and stick to the format: dataset provider;dataset name;version/publication year	e.g. International Energy Agency;CO2 Emissions from Fuel Combustion; 2019, EU;European Union Transaction Log
Notes	Any notes.	
Coding completed	The date when the coding was completed. Stick to the format: DD.MM.YYYY	*15.11.2021*