D6.1
Data Management Plan

August 2, 2023

Programme Call:	Cross-sectoral solutions for the climate transition (HORIZON-CL5-2021-D1-01)
Grant agreement ID:	101056891
Project Title:	ClimAte Policy AcceptaBiLity Economic framework
Project Acronym	CAPABLE
Partners:	CMCC (Project Coordinator), EUI, IESEG, CNRS, PIK, MCC, UAB, CUNI, RUG, E6, ETH Zurich.
Work-Package:	WP6
Deliverable #:	D6.1
Deliverable Type:	R — Document, report
Contractual Date of Delivery:	30 June 2023 (postponed to 31 August 2023)
Actual Date of Delivery:	02 August 2023
Title of Document:	Data Management Plan
Responsible partner:	CMCC
Author(s):	Johannes Emmerling, Laura Pellizzola, Silvia Pianta, Loic Berger, Niklas Döbbeling, Albert Ferrari, Kai Lessmann, Milan Ščasný, Keith E. Smith.

Content of this report: The Data Management Plan describes the data
                                                   expected to be acquired or generated during the
                                                   course of the CAPABLE project, how they will be
                                                   managed, analyzed, stored and preserved, and
                                                   how they will be made available for re-use at the
                                                   end of the project.

Dissemination Level: C-UE/EU-C - EU Classified

Document revisions		
Author(s)	*Revision content*	*Date*
Laura Pellizzola (CMCC)	Initial draft	24/05/2023
Johannes Emmerling (CMCC), Silvia Pianta (CMCC)	First revision	29/05/2023
Loic Berger (IESEG), Niklas Döbbeling (MCC), Kai Lessmann (PIK), Milan Ščasný (CUNI), Keith E. Smith (ETH), Albert Ferrari (EUI)	General revision	10/07/2023
Laura Pellizzola	Final Document	02/08/2023

Glossary

Abbreviation	Full term
CA	Consortium Agreement
CO	Project Coordinator
CB	Coordination Board
DMP	Data Management Plan
FAIR	Findable, Accessible, Interoperable and Re-usable
GA	Grant Agreement
SH	Stakeholders
WP	Work Package
WPL	Work Package Leader

Table of Contents

Executive Summary

The CAPABLE project will produce a number of data sets for internal and public use. This Data Management Plan (DMP) describes the data management life cycle for the data to be collected, processed and/or generated in the different WPs and Tasks of the project, including information regarding standards and metadata, data sharing and archiving and preservation. The aim of the plan is to ensure that data produced by the project will satisfy Horizon Europe data management guidelines. Care has been taken in the plan to describe how the data will be FAIR — Findable, Accessible, Interoperable and Re-usable. The plan also describes the data that will be generated by the CAPABLE project, whether and how they will be made accessible for verification and re-use, as well as how they will be curated and preserved.

This DMP is organized in such a way as to highlight all the data to be collected, processed and/or generated in the different WPs and Tasks of the project.

The DMP will be updated and revised throughout the project: the first update is foreseen in Month 18 and the second one in Month 30.

1   FAIR Guiding Principles

The project consortium takes measures to ensure that the data produced/collected within the project is findable, accessible, interoperable and reusable (FAIR).

Making data findable.
To aid discoverability of existing and produced data it is annotated with metadata. Where possible, it is aimed to obtain digital object identifiers (DOIs). In addition, the scientific publications in the project cross-references these data sets.

Making data open accessible.
To ensure accessibility beyond the duration of the project, the use of proprietary data formats will be avoided. Data sets will be made available to reviewers of scientific publications to aid transparency in the review process. Access will be provided through open data repositories (for example, Zenodo).

Making data interoperable.
Data produced in the project will use formats and metadata standards appropriate for the scientific community. This approach allows easy combination with data sets from other European or international projects that rely on similar topics related to environment, energy and the economy.

Increasing data re-usability.
The datasets, frameworks and tools produced in CAPABLE are intended to be re-used also outside of the project. When feasible, the publications will contain as supplementary material the link to the code and the database to ensure the replicability of the results and the reusability of data.

2 WP1 - Methodological advancement

2.1 Task 1.1 Socio-economic heterogeneity

Data Summary

This task aims to assess which dimensions of social heterogeneity are most relevant for a low-carbon transition. CUNI will study how low-carbon durable goods are distributed among households, using various measures of inequality, and examining (observed and unobserved) preference heterogeneity in the uptake of these goods.

CUNI will re-use the data from two original surveys on passenger vehicle ownership and intention to purchase (another) new passenger vehicle (one survey was conducted in 2017 within the SUPREM project, another one was conducted in 2023 within the project NOx2030, both funded from the Technology Agency of the Czech Republic, and both were carried out in the Czech Republic. The sample size of the two surveys is 2500 and 1000, respectively).

CUNI also plan to analyse thermal insulation of homes using two waves (cross-section) of ENERGO survey conducted in Czechia by the Czech Statistical Office in 2015 and 2022. Sample size of ENERGO survey is around 10,000 dwellings. Neither of these data can be used by a third party.

FAIR Data

Making data findable, including provisions for metadata

We will use secondary data that were collected within another projects (the two surveys) or that were conducted and are owned by the Czech Statistical Office. The datasets will be described but they can't be provided for any use by a third party.

Making data accessible

The data will be stored on secured CUNI hardware in such manner that allows compliance with applicable data protection and privacy legislation. The data gathered and provided by the Czech Statistical Office can be only used by persons who signed Agreement on confidentiality. The secondary data can't be made openly available since these data were collected within other projects and are owned by other bodies (CSO). The ENERGO data can be only used by people who signed Agreement on confidentiality with CSO.). No metadata will be produced in this task.

Making data interoperable

N/A

Increase data re-use

These data can be re-used only under agreement of the data owners.

Other research outputs

N/A

Allocation of resources

N/A

Data Security

The data will be stored in a secure internal repository (Google for Education cloud being licensed by CU). Only authorised persons will have access to download data, using a protected password.

Ethics

The surveys were evaluated and got an ethics approval by the Ethics Board at Charles University Environment Center.

Other issues
N/A

2.2   Task 1.2 Behavioural dimensions of climate policy performance

Data Summary
This task examines how bounded rationality and social, behavioural aspects affect the evaluation and design of climate policies. This includes attention for habit formation and myopia, serving as barriers to long term investment. In terms of social influence - connecting to Task 1.1 on socio-economic heterogeneity - we assess climate policy performance under different behavioural assumptions and models and implications for optimal policy design. IESEG will develop a tool that will be used to elicit individual preferences over several interconnected dimensions of welfare (uncertainty, time, social), with the objective to study how preferences are influenced by and determine beliefs about climate change. This includes running incentivized, online experiments eliciting detailed preference and belief measures. Always in Task 1.2, CUNI will investigate how altruistic motives might be affected when the public good is to be provided through an instrument, taxation in particular, and how acceptance, support and acceptability of policy measures might be affected when people are uncertain about their future wealth, expected benefits and incidence and environmental burden they will be responsible for in future. Specifically, CUNI investigates how distortions in stated willingness to pay might be reduced if people are placed behind a 'veil of ignorance', following a work developed within a health risk context.

For Task 1.2, primary data will be collected in two phases. The first phase will consist of the collection of various measures of individual preferences and beliefs, in the risk and uncertainty, time, and social dimensions. These will be collected by participating partners through incentivized lab and online experiments with at least 200 subjects in France, with potential contributions from other European partners. These quantitative measures will provide the basis for an extensive survey focused on preferences and beliefs (both subtasks led by IESEG and CUNI). This output will serve as a validation stage for the questions that will be later integrated into the large-scale survey in WP2.

The format of the data will be (i) quantitative measures of individual preferences, (ii) quantitative measures of subjective beliefs, and (iii) survey responses on attitudes and beliefs, (iv) socio-economic characteristics. These will be stored as CSV files and made available in other common formats where practical to increase accessibility. The purpose of the data generation is to test the survey questions that correlates best with standard incentivized tasks typically used in experimental economics.

The collected data may be useful for projects that want to re-examine relationships between the collected variables and to compare results with similar data collected at a later point in time.

FAIR Data

Making data findable, including provisions for metadata
Data, in so far as is practically possibly, given confidentiality and data protection obligations, will be FAIR following the conclusion of the project in accordance with the guiding principles outlined in the Project Data Management plan. To aid the discoverability of data sets, DOIs will be obtained for each data set released. The acquired data sets will be added to the Zenodo repository. Furthermore, the resulting scientific publications will reference these datasets. Metadata will include the questionnaire scales, method of analysis, and keywords will be provided to optimise the possibility for discovery and potential reuse. Codebook(s) detailing the nature and justification for each variable collected will be made available for each data set.

Data sets will be collected independently. Persistent anonymised identifiers will be used with any identifying information held only by the collecting partner to facilitate compliance with data protection and privacy obligations.

Making data accessible

Data sets with accompanying metadata will be made available to reviewers of scientific publications to aid transparency in the review process. Access will be provided through open data repositories such as the Zenodo. Any embargo place on data accessibility will be limited to the duration of the project or if a publication is pending at the end of the project, be prolonged until the publication process is completed.

The data will be made available in accordance with Open science principles. Before data is shared, data will be assessed to ensure that no personal or sensitive data is made public and that anonymised participants are not identifiable. Only fully anonymized data and in certain cases, aggregated results may be made publicly available. If participants no longer wish for their data to be openly accessible or re-used, procedures to ensure the removal of such data from the data set will be implemented where possible.

Metadata will be made openly available on the Zenodo platform, including all necessary information to enable the user to access the data. This entails the questionnaire and method of analysis. Necessary documentation and referencing of software to access or read the data will be given.

Making data interoperable

To ensure that data collected will be interoperable to the greatest extent possible, data made available will be in simplified non-proprietary (.csv) format. It is not intended to include qualified references to previous research.

Datasets, questionnaire formats and methods of analysis will be shared openly and in case of use of uncommon ontologies or vocabularies for the project, these will be explained in the published method/appendix of analysis.

Increase data re-use

In order to increase data re-use, data will be published in a cleaned format with the cleaning code accessible for examination. The data that is being made available will be usable by third parties as soon as they are made public following any embargo period. The provenance of the data can be extracted from papers based on the referenced data set (i.e., methods/appendix section).

To ensure data quality, quality checks appropriate to the nature of the collected data will be deployed (e.g., checking for the time spent on the questionnaire, repeated answering patterns and overall response completeness). Responses that do not meet basic requirements will be excluded from the cleaned data. For individual analysis specific requirements for in/exclusion should be reported in the methods/appendix section of papers using the data as well as the code of analysis.

Other research outputs

N/A

Allocation of resources

N/A

Data Security

WP1 T1.2 data gathered and analysed for the purposes of research carried out by staff and students at the IESEG, and CUNI, respectively, and these data will be stored on secured IESEG and CUNI hardware in such manner that allows compliance with applicable data protection and privacy legislation.

Participating partners involved in data collection will be individually responsible for maintaining and observing agreed data collection and security protocols including the preservation of the anonymity of participants.

Ethics

Data gathering, analysis and storage at IESEG, and CUNI, respectively, for the purposes of research will be conducted in accordance with the French and Czech data protection law and General Data Protection Regulation (EU) 2016/679. The FR Data Protection Law in both countries complies with EU legislation and privacy/ confidentiality measures (European Directive 95/46/EC, covering collection and processing of personal data for scientific purposes).

Informed consent will be sought from each participant for the long-term preservation and sharing of anonymised respondent data with the respondents right to withdraw clearly outlined.

Finally, full compliance with ethical and human subjects' procedures and regulations as set out by the LEM research unit, accredited by the French Centre National de la Recherche Scientifique (CNRS), will be obtained before the beginning of the data collection. The same procedure will be followed for the task carried out by CUNI - CUNI will apply for an approval from the Ethics Board at The Charles University Environment Center.

Other issues
N/A

3   WP2 - Social acceptability and feasibility

3.1   Task 2.2 Large scale questionnaire studies on determinants of policy acceptance and compliance

Data Summary
These tasks will empirically assess the social and political feasibility of the EU green deal policy packages.

They aim at investigating public acceptance of different climate policies and policy instruments, assessing determinants of public support, and testing policy packaging options to maximise public support. For doing this, we consider personal factors (e.g. personal values, behavioural preferences, basic needs, knowledge), social factors (e.g. social norms, national values), and contextual factors (e.g., socio-demographics, infrastructure, pricing) that influence policy acceptance and compliance, and investigate interactions between these factors, at different societal levels, and across segments. In addition, the project assesses public perceptions and (second-order) beliefs about the goals, roles, and responsibilities of different relevant actors in policymaking, including policymakers, organizations and the general public, and how these may affect evaluations of policies and policymaking. Further, the tasks investigate public perceptions on policymaking, their current involvement in policymaking, and how they would like to participate and be involved in policymaking, considering options such as citizen assemblies.

We will collect around 35,000-40,000 survey completes of voting-age residents from 10 countries across the European Union. Likely candidate countries for data collection are CZ, DE, ES, IE, IT, FR, LT, NL, PL, SI, SE. The final country list and sample size will be determined based upon budget and sample availability. The data will be collected using online-implemented surveys. In order to develop a sample in each of the countries, we will partner with an established survey firm to recruit respondents (e.g. YouGov, Dynata, Bilendi). The primary form of data collection will be questions about individual characteristics, attitudes, climate change policy preferences, and survey experiments. Survey experiments are a common form of research design in

which treatment/control/placebo conditions and items measuring participants' responses to treatments are embedded in a survey. Survey participants are randomly assigned to treatment conditions or a control/placebo group. This data is intended to be generated in a manner that is useful for scientific and public-facing audiences. We will design the research such that it will be readily presentable as part of the project's knowledge transfer goals, as well as usable for scientific publications. We will generate new survey data for the project, but where possible, we will replicate existing survey instruments and measures for enhanced generalizability.

FAIR Data

Making data findable, including provisions for metadata

The datasets from this project's publications will be deposited on a publicly available, permanent data archive, such as Open Science Foundation (OSF) or the Harvard Dataverse. Replication materials will also be published alongside academic publications. All published files i.e. articles and datasets from this project will be/have been given a Digital Object Identifier (DOI). The DOI issued to datasets can be included as part of a data citation in publications, allowing the datasets underpinning a publication to be identified and accessed. Metadata about datasets held in OSF/Harvard Dataverse are publicly searchable and discoverable and will indicate how and on what terms the dataset can be accessed.

Making data accessible

After publication of findings, the data will be archived in a publicly accessible repository. For archival purposes we suggest either using the Open Science Foundation (OSF) or the Harvard Dataverse. Both are reputable repositories, are maintained by non-profit organisations, are primarily utilised by scientists and for scientific research, and are common archives utilised within social science studies.

Upon completion of the project, the full survey data will be made available, in a confidential form. Data that is utilised in scientific research publications will be made available at time of publication. We will not share data as it is immediately collected, rather at the end of the project, for two primary reasons. First, we need to clean, validate and make the data confidential before public publication (a common practice in survey data archiving). Second, and importantly, we need to protect the scientific integrity of the research, allowing for project partners to analyse and report the data.

The metadata will be openly available and licensed under a public domain dedication. The data will be archived in a permanent repository, and findable to the public. The survey data will be archived alongside survey instrument materials (e.g. questionnaires, consent letters, codebooks) as well as data cleaning and replication materials, fitting with common survey data archiving practices in the social sciences.

Making data interoperable

Where possible, we will adopt validated survey instruments, developed by international survey programmes (such as the European Social Survey), or within scientific publications. Particularly for individual respondent characteristics we will mirror instruments developed by the European Social Survey, while for other characteristics such as human values and trust in institutions, we will utilise established survey instruments developed within the scientific literature. This will allow for much of the single item data to be comparted, and interoperable, with existing survey data and previous scientific literature.

For the survey experimental components, these will be novel and unique from previous data collection. We will use existing methodologies and best practices common to the social and behavioural sciences, and thoroughly document the methodological approach so that these survey experiments are interoperable to other similar scientific research and understood by the scientific community.

Increase data re-use

We will provide survey documentation alongside the archiving of the survey data. This will entail documenting survey questionnaires, data cleaning, and coding for replication of analyses. These materials will be made available alongside the publication of data, on publicly available, survey data repositories. All data will be published in a confidential format, which does not include personally identifiable information about the respondents (fitting with social science survey data best practices).

Other research outputs

Scientific publications using these survey data will include analytical replication coding, allowing for re-generation and replication of these survey data findings.

Allocation of resources

The survey data will be publicly archived on reputable, non-profit organisation platforms. There are no direct costs associated with publicly archiving this data. ETH Zurich and RUG, as co-leaders of WP2, will be responsible for the archiving of the survey data and the data management. The data will be stored in csv., dta., or sav. formats. The data files will be stored/backuped on Network Attached Storage (NAS) as provided by the ETH IT. The data storage is backed by a service level agreement (SLA).

Data Security

Until publication of scientific results or the end of the project, only the immediate CAPABLE research team members will have access to the survey data. Internal data transfer is being/will be performed via password protected folders, housed on ETH servers.

A confidential and clean version of the data will be publicly archived at the end of the project at OSF or Harvard Dataverse servers, as mentioned above. A backup of the data will be stored on Network Attached Storage (NAS) as provided by the ETH IT.

Ethics

All survey data collection activities will undergo university institutional ethical review (e.g. IRB or Ethics Commission) before data collection will begin. Such institutional ethics review boards will review data collection, storage, and content, to ensure all survey data collection activities comply with ethical regulations. All surveys will include an informed consent, which will be the first information received by potential respondents. The informed consent will acknowledge the scientific conductors of the survey data collection, the content of the surveys, and how the data will be stored and used. Respondents will only participate in the survey if they provide voluntary consent.

Other issues

N/A

3.2   Task 2.3 Empirically testing aspects of policies, policy instruments, and policy packages that impact social feasibility and public acceptance; and Task 2.4 Public perceptions of policymaking and policy actors, and the impact of perceptions on the acceptability and feasibility of climate change-relevant policies.

Data Summary

This task will collect further experimental survey data in several EU-countries, from a smaller set of contexts (single-context studies), based on comparatively smaller sample sizes (n=500-1000 individuals in each country), but allowing to determine causal factors from the random allocation of treatment and control groups. The goal of this task is to identify mechanisms of choices instead of cross-country differences.

FAIR Data

Making data findable, including provisions for metadata

The dataset will be deposited on a publicly available, permanent data archive, such as Open Science Foundation (OSF) or the Harvard Dataverse. Replication materials will also be published alongside academic publications. All published files i.e. articles and datasets from this project will be/have been given a Digital Object Identifier (DOI). The DOI issued to datasets can be included as part of a data citation in publications, allowing the datasets underpinning a publication to be identified and accessed. Metadata about datasets held in OSF/Harvard Dataverse are publicly searchable and discoverable and will indicate how and on what terms the dataset can be accessed.

Making data accessible

After publication of findings, the data will be archived in a publicly accessible repository. For archival purposes we suggest either using the Open Science Foundation (OSF) or the Harvard Dataverse. Both are reputable repositories, are maintained by non-profit organisations, are primarily utilised by scientists and for scientific research, and are common archives utilised within social science studies.

Data that is elaborated and utilised in scientific research publications will be made available at time of publication. We will not share data as it is immediately collected, rather after two-years of the project completion, in a confidential form, for two primary reasons. First, we need to clean, validate and make the data confidential before public publication (a common practice in survey data archiving). Second, and importantly, we need to protect the scientific integrity of the research, allowing for project partners to analyse and report the data.

The metadata will be openly available and licensed under a public domain dedication. The data will be archived in a permanent repository, and findable to the public. The survey data will be archived alongside survey instrument materials (e.g. questionnaires, consent letters, codebooks) as well as data cleaning and replication materials, fitting with common survey data archiving practices in the social sciences.

Making data interoperable

Where possible, we will adopt validated survey instruments, developed by international survey programmes, or within scientific publications. Particularly for individual respondent characteristics we will mirror as much as possible instruments developed by the European Social Survey, while for other characteristics such as the attitudinal measures, we will utilise established survey instruments developed within the scientific literature. This will allow for much of the single item data to be comparted, and interoperable, with existing survey data and previous scientific literature.

Increase data re-use

We will provide survey documentation alongside the archiving of the survey data. This will entail documenting survey questionnaires, data cleaning, and coding for replication of analyses. These materials will be made available alongside the publication of data, on publicly available, survey data repositories. All data will be published in a confidential format, which does not include personally identifiable information about the respondents (fitting with social science survey data best practices).

Other research outputs

Scientific publications using these survey data will include analytical replication coding, allowing for re-generation and replication of these survey data findings.

Allocation of resources

The survey data will be publicly archived on reputable, non-profit organisation platforms. There are no direct costs associated with publicly archiving this data. CUNI, as leader of this task, will be responsible for the archiving of the survey data and the data management. The data will be stored in csv., dta., or sav. formats.

Data Security

Until publication of scientific results or the end of the project, only the immediate CAPABLE research team members will have access to the survey data. Internal data transfer is being/will be performed via password protected folders, housed on CUNI servers.

A confidential and clean version of the data will be publicly archived at the end of the project at OSF or Harvard Dataverse servers, as mentioned above. A backup of the data will be stored on Network Attached Storage (NAS) as provided by the CUNI IT.

Ethics
All survey data collection activities will undergo university institutional ethical review (e.g. IRB or Ethics Commission) before data collection will begin. Such institutional ethics review boards will review data collection, storage, and content, to ensure all survey data collection activities comply with ethical regulations. All surveys will include an informed consent, which will be the first information received by potential respondents. The informed consent will acknowledge the scientific conductors of the survey data collection, the content of the surveys, and how the data will be stored and used. Respondents will only participate in the survey if they provide voluntary consent.

Other issues
N/A

4   WP3 - Evaluation of European Climate and Environmental Policies

4.1   Task 3.1: Systematically reviewing policy evaluation studies using machine learning tools

Data Summary
This task aims to create a large and comprehensive database of ex-post policy evaluations by innovating new automated methods based on machine-learning to produce living, and systematic evidence maps that find, classify and chart out existing studies (Callaghan et al., 2021 or Lamb et al., 2019). MCC and UAB will apply and further develop supervised and unsupervised machine learning methods from natural language processing for screening, selecting and coding impacts evaluation studies of climate policies (Callaghan and Müller-Hansen, 2020; Haddaway et al., 2020). The resulting database of ex-post climate policy impact evaluation studies will serve as an automatically and continuously updated basis for comprehensive metanalysis efforts in key EU climate policy areas in Task 3.2, and the online tool developed in Task 5.4. In addition, CMCC will extend and implement a comprehensive coded database on policy evaluation based on Peñasco, Anadón, and Verdolini (2021) to create a database on environmental policy evaluation including as an underlying source for the online tool to be developed and policy maker capacity building workshops in WP5.

The data will be collected from an existing open access literature database (OpenAlex), which lists academic literature providing titles, abstracts, author information, publication year, etc. At MCC data from this database was extracted previously for a related project, with a much broader scope. For this previous project a share of the data was annotated to identify its focus on climate mitigation policies, classification of different types of policies using a fine-grained typology of climate policy measures, its sector focus as well as a broad categorisation of research methods. These annotations were used to train a machine learning classifier to extend the labels to the full set of studies. This data will fully be re-used and amended for the new research focus. The data will be in the natural language format of the data provided by OpenAlex, adding the annotations for the classification of ex-post literature. This data will be used for the research in Task 3.2 of the CAPABLE project and will be relevant for other researchers synthesising evidence on climate mitigation policies.

FAIR Data

Making data findable, including provisions for metadata

To the aim of discoverability of existing and produced data it is annotated with metadata. Where possible, it is aimed to obtain digital object identifiers (DOIs). In addition, the scientific publications in the project cross-references these data sets. To the extent possible, we will use keywords to facilitate the discovery of the data, despite the non-existence of clear standards for metadata in the field.

Making data accessible

The collected data and the codebook which explains the data collection process will be deposited in a Zenodo repository or in a repository complying with the same open access policies. The project partners have extensive experience with Zenodo as a reliable data repository. The data will be assigned a digital object identifier (DOI).

No restrictions on data publication apply. The database will be created using the OpenAlex literature database, which itself is open access. The published data will include the identifiers assigned by OpenAlex in addition to an extract of the metadata (i.e. title and authors). Together with the identifier, these data will allow for the replication of all conducted analyses and the integration of new evidence published after the completion of the project. To ensure accessibility beyond the duration of the project, the use of proprietary data formats will be avoided.

The institutional repositories of the project partners do not claim rights over the deposited publications nor preclude access. Metadata will be made openly available on a public platform, including all necessary information to enable the user to access the data.

Making data interoperable

The data will be provided in CSV format with no specific software needed to access the data.

Increase data re-use

The database will be stored in a common spreadsheet format to maximise re-use. As much of the data as possible will be obtained from public domain sources and will be shared without restrictions.

The publication of the extended spreadsheet will include documentation to allow reproducing the analysis undertaken within the CAPABLE project.

Other research outputs

N/A

Allocation of resources

N/A

Data Security

N/A

Ethics

We do currently not see any ethical issues that could have an impact on data sharing.

Other issues

N/A

4.2   Task 3.2 Policy analysis of selected Fit-for-55 policies

Data Summary

This task will carry out an empirical assessment of selected climate policies related to the recently proposed Fit-for-55 policy packages. Based on task 3.1, MCC will perform a comparative and quantitative meta-analysis on key policy areas relevant for Fit-for-55, notably the efficiency and effectiveness of various non-carbon pricing policies in the building and transportation sector (MCC).

The second part of this task focuses on the EU ETS. First, using firm-level data and econometric causal analysis (difference-in-differences), EUI will estimate the impacts of the EU ETS on regulated firms' emissions, competitiveness, and investment leakage (as proxied by reduced tangible fixed assets and increased foreign direct investments). The study will rest on the matching of installation-level data in the EU ETS registry (EU Transaction Log) to firm-level data in a suitable company database, such as ORBIS (Bureau Van Dijk). Second, EUI will investigate the impact of participation in the EU ETS on firms' investment in GHG abatement. Again, firm-level data and causal econometric analysis will be used. The study will focus on Swedish firms, for which high-quality data are known to be available. In addition, interviews with company representatives will be conducted with a view to deriving complementary information about the roles of regulatory uncertainty and carbon price uncertainty in abatement investment decisions. Special attention will be devoted to firm heterogeneity related to sector and size and relevant regulatory/policy decisions such as those concerning the Market Stability Reserve and the Carbon Border Adjustment Mechanism.

For the meta-analysis the CAPABLE partners will collect data from published scientific publications. The selection of the relevant publications will be based on transparent and reproducible criteria and will be based on the database of publications from task 3.1. Meta-data of included and excluded publications will be captured. From relevant publications estimated policy effects will be captured together with information on the study design and context of the policy. The exact data points to be captured from each publication will be set out in a codebook. The data will be captured in table format and after the end of the project, the quality controlled and approved data will be made publicly available in CSV format.

Regarding the second part of the task on the EU ETS, the study will rest on the matching of installation-level data in the EU ETS registry (EU Transaction Log) to firm-level data in a suitable company database, such as ORBIS (Bureau Van Dijk). In addition, it will use firm-level data and causal econometric analysis on Swedish firms, for which high-quality data are known to be available on investment on GHG abatement. In addition, interviews with company representatives will be conducted with a view to deriving complementary information about the roles of regulatory uncertainty and carbon price uncertainty.

The datasets on the EU ETS will come from previous work on the matchings between ORBIS and the EU Transaction Log, such as the one published on open access by the Joint Research Centre in 2022, and current work initiated at EUI under the LIFE COASE project. The dataset on Swedish forms will be purchased as foreseen in the CAPABLE budget. In addition, interviews with company representatives will be conducted with a view to deriving complementary information about the roles of regulatory uncertainty and carbon price uncertainty.

With respect the data format, the datasets and recordings of the interviews will be confidentially shared with the research team.

The data collected will be useful especially for the policymakers in order to understand the main factors to take into account if they want to introduce feasible and effective policy instruments on carbon leakage and investments for decarbonisation.

FAIR Data

Making data findable, including provisions for metadata
To the aim of discoverability of existing and produced data it is annotated with metadata. Where possible, it is aimed to obtain digital object identifiers (DOIs). In addition, the scientific publications in the project cross-

references these data sets. To the extent possible, we will use keywords to facilitate the discovery of the data, despite the non-existence of clear standards for metadata in the field.

Making data accessible

The collected data and the codebook which explains the data collection process will be deposited in a Zenodo repository or in a repository complying with the same open access policies. The project partners have extensive experience with Zenodo as a reliable data repository. The data will be assigned a digital object identifier (DOI).

No restrictions on data publication apply. The shared data will allow for the replication of all conducted analyses and the integration of new evidence published after the completion of the project. To ensure accessibility beyond the duration of the project, the use of proprietary data formats will be avoided.

The institutional repositories of the project partners do not claim rights over the deposited publications nor preclude access. Metadata will be made openly available on a public platform, including all necessary information to enable the user to access the data. The data will be provided in CSV format with no specific software needed to access the data. The codebook describing the data collection process will be in PDF format.

Making data interoperable

We will follow established systematic review guidelines by "Cochrane" and the "Collaboration for Environmental Evidence" to allow for seamless data updates and data exchange across disciplines. The data will be linked to existing systematic reviews in the field using digital object identifiers to track extracted information from primary studies across datasets.

In order to ensure a feasible interpretation of the data, data on the EU ETS and its effect on comapnies will be provided in CSV or XLSX formats as concern the survey, and in text format for the interviews.

Increase data re-use

A codebook will describe the data collection process and will be published along with the dataset. The data will be made freely available in the public domain to permit the widest re-use possible. The data will be licensed using standard reuse licenses, in line with the obligations set out in the Grant Agreement. The data produced in the project will be useable by third parties, in particular to update the analysis in the light of new evidence or to extend the scope of the analysis. All information extracted from primary studies will be verified by a team of two reviewers to ensure consistent, high-quality data.

The results will be reusable by interested parties, including researchers, civil society organisations and policymakers. The data will be shared cleaned and clear to increase the probability of re-use. In light of this, a codebook will be provided.

Other research outputs

The datasets and the survey results will be used in scientific articles.

Allocation of resources

Resources for making data FAIR and costs related to storage archiving, re-use and security are allocated as part of task 3.2.

Data Security

Data will be stored in a trusted repository for long term preservation. A local copy of the data will remain with the project partners.

On the EU ETS, the data will be contained in a hardware protected by username and password. Further, each team member will be responsible of the data preservation, including security and anonymity of the participants. Data can also be uploaded to Onedrive.

The EUI has a Data Security Officer to provide the security of information systems, to set up EUI's security policies and controls and to take appropriate technical and organisational measures for its implementation.

Ethics
We do currently not see any ethical issues that could have an impact on data sharing. The data collection will be conducted following the EU Member States (Regulation (EU) 2016/679, GDPR).

The EUI has its own Data Protection Policy (President Decision n. 10 of 18 February 2019), which establishes adequate safeguards for the protection of personal data, including a mechanism for the exercise of data subjects' rights, and the possibility for judicial review by its internal judicial system.

Other issues
N/A

4.3   Task 3.3 Policy analysis of the COVID-19 pandemic and policy response

Data Summary
In this task EUI will identify lessons from the management of the COVID-19 pandemic that are transferable to climate policy (EUI). Specifically, this task is focussed on the contribution of Scientific Advisory Boards (SAB) advising national governments. To address the COVID-19 emergency, most MSs established their own SAB. Similarly, several MSs established- or are going to establish a SAB for climate change and the ecological transition (this is also the case of the EU). The study will analyse the role and the performance of both COVID-19 and climate SABs in Europe with respect to their ability to improve the quality and legitimacy of policymaking.

Using semi-structured interviews and a survey with members of SABs in several (>6) MSs, the consortium will draw lessons for climate SABs from the experience of COVID-19 SABs and of climate SABs themselves.

FAIR Data

Making data findable, including provisions for metadata
During the project cycle, folders and files will be consistently named using the structure: CAPABLE > WP 3 > Task 3.3 > Data > Dataset ref. (theme/Taskcode) &c. > Dataset version. Technical metadata will be generated by the European University Institute (EUI) library staff to aid findability. The metadata will contain information notably on the dataset creator(s), data description, source, type and date range, codebook and supporting document, unique object identifier and related publications. Keywords using the project name (LIFE COASE) and dataset names will be used to optimise the possibility of discovery. The metadata will be offered in such way that it can be harvested and indexed.

Making data accessible
Project data outputs will be openly shared where there are no concerns pertaining to (i) data protection relating to human subjects, families, or households and (ii) database copyright provisions arising from the elaboration of pre-existing data inputs.

The research data will be published in the EUI research repository, Cadmus, which is the Repository of the project's host institution. In line with the EUI Cadmus DSpace repository policy, codebooks and documentation will be made available with data outputs which have been approved for publication. The project team will work with EUI library staff to determine whether and when a dataset can be openly shared.

Full details of EUI data repositing and open access procedures are available in the 2023 EUI Library Research Data Guide (see other issues below).

The project team and the EUI library staff will ensure the data is assigned an identifier. The published data will have handles generated by Cadmus.

Open data generated by the project will be available under a CC-BY (International) license. As soon as the data is produced as a deliverable (public), no embargo is envisaged for data likely to be deemed sharable. No restrictions on use and access will apply to the data that will be produced as a project deliverable. The data will be made available by the end of the project, following the timeline described in the project proposal. Similar to publicly available datasets, the metadata will be reposited online following the EUI's standard procedure. It will be freely accessible. The data and associated metadata will remain available indefinitely. As of now, no update of the data is planned beyond the project duration. All data will be accessible in formats such as .csv, .txt and .xlsx or in equivalent non-proprietary formats. Technical notes and user guides to support data use and exploitation will be in .pdf, docx, /or .xlsx, xlsx or in equivalent non-proprietary.

Making data interoperable

To ease data interoperability, exchange and reuse, the datasets created will use the source data standards and vocabularies. In case uncommon or project-specific ontologies or vocabularies are used, their definition and relation to existing terminology will be described in the technical notes. This will ensure ease of reuse and extension of the newly created concepts. The data will include qualified references to other data used and matched, as well as to previous datasets upon which it will be built. Cadmus relies on an adapted Dublin Core schema infrastructure to facilitate interoperability. DSpace uses Java script, generating a web interface compatible with multiple environments.

Increase data re-use

Documentation of the generated datasets will be available in the database metadata as text files (technical notes). It will notably include methodology, variable definitions, units of measurement, data size, column statistics and information on missing values. The provenance of the data will also be documented and referenced. The purpose of these technical notes is to guarantee ease of data usage for other scholars and stakeholders. As described above, the data will be made available to other academic researchers, policymakers, partners, the media, and the general public, including DG CLIMA, DG GROW, JRC, EIB, EPRS, EEA, and national ministries, regulators and environmental agencies, and numerous academic and civil society nodes. The data will therefore remain in the EUI's Repository indefinitely after the end of the project.

Before publication on the Repository, the data and accompanying documents will be formatted and proofread by team members and reviewed by the EUI Library Staff and Communication staff of EUI.

Other research outputs

Other research outputs,will all be public and remain available to be reused in line with the FAIR Data Principles, even after the project ends. They will be published on the project's website and on the EUI Repository when applicable.

Allocation of resources

No project-line costs are associated with long-term preservation because LIFE COASE will use existing institutional infrastructures such as research data repositories. EUI will bear in-project data collection and elaboration costs. The project coordinator for EUI is responsible for data management and will work with the CAPABLE team and the Robert Schuman Centre's Data Controller to ensure correct implementation.

Data Security

All raw data will be stored on the EUI servers with institutional security set up. Data will be uploaded to OneDrive, while away from the Institute, with a two-step security measure (logging and telephone token). The EUI has also appointed a Data Security Officer with a mandate to safeguard the security of information systems at the EUI, to set up EUI's security policies and controls and to take appropriate technical and organisational measures for its implementation.

Raw data generated by the project will be stored for five years after the end of the project. After that date, they will be destroyed.

Ethics

An essential part of the DMP is to ensure the protection of the dignity, rights and welfare of people involved in the project and to promote high ethical standards of research. In the context of its methodology, the EUI does not foresee a probable negative impact of its research either before, during or after the conducting of research activities. The project has pre-identified three areas (i) involvement of experts for research; (ii) processing of personal data; and (iii) processing of special categories of personal data focused on emissions and economic activities. Research activities will be conducted on the soil of member states of the European Union, no activities will be carried out in a non-EU country, and no personal data transfer will be made outside.

The EUI, as an international organisation, has its own Data Protection Policy (President Decision n. 10 of 18 February 2019, see below), which establishes adequate safeguards for the protection of personal data, including a mechanism for the exercise of data subjects' rights, and the possibility for judicial review by its internal judicial system. The Decision was introduced following the adoption of the EU General Data Protection Regulation (GDPR). The EUI' Guide to Good Data Protection Practice in Research' provides further details of data protection at the EUI. The informed consent for data sharing and long-term preservation will be included in questionnaires dealing with personal data.

Other issues

The 11th edition of the EUI Library's Research Data Guide was published in April 2023. There are ten sections:

1. Data discovery and the EUI Library Data Portal
2. Data generation, data protection, data processing and ethical use
3. Data management plans
4. Research data management in Horizon Europe, ERC, and other EU programmes
5. Managing data during the research project cycle
6. Repositing and preserving data in the EUI research repository Cadmus
7. Open Data, FAIR Principles and Open Science
8. Qualitative and unstructured data in the humanities and social sciences
9. EUI infrastructure, software, protocols, and support
10. International research data guidelines
    (https://www.eui.eu/research/library/researchdataservices/guide)

4.4   Task 3.4 Environmental benefits and co-benefits of climate policies

Data Summary

This task explores how environmental policies are motivated by the direct (intended) and unintended effects that they create for the environment and society. To address this question, a database of external costs of externalities with a focus on environmental and pollution-related externalities will be compiled for Europe. The database will facilitate the discussion of benefits and co-benefits of policies. Secondly, the RICE50+ model will be extended to include indicators for biodiversity and natural capital.

The database on the benefits and co-benefits of climate policy will contain information on external costs of environmental and pollution-related externalities. Examples are health implications of local pollutants, land use, or fertilisation, covering sectors such as transport and agriculture. The data will be compiled from existing studies and publications and stored in common spreadsheet formats.

FAIR Data

Making data findable, including provisions for metadata
The resulting database will be made available by publication via an open data repository (such as Zenodo) and a DOI will be obtained as a persistent identifier. This will ensure that the generated data received is broadly visible and discoverable.

Making data accessible
The resulting database will be made available by publication via an open data repository (such as Zenodo) and a DOI will be obtained as a persistent identifier.

Making data interoperable
Interoperability in the sense of re-use of the data in other contexts by others will be facilitated by including references to the original sources of the compiled data, and transparent information about any conversions that may have been necessary to harmonise data for the database.

Increase data re-use
The database will be stored in a common spreadsheet format to maximise re-use. As much of the data as possible will be obtained from public domain sources and will be shared without restrictions.

The publication of the extended spreadsheet will include documentation to allow reproducing the analysis undertaken within the CAPABLE project.

IP rights to the models participating in CAPABLE remain with their host institutions. This also applies if the model development (e.g. new data or code modules) is realised by a partner organisation different from the host institution.

Other research outputs
The RICE50+ model is written in GAMS. This is a proprietary modelling language. For unlimited use, a licence needs to be bought from the GAMS company (but public servers offer a free use of GAMS, for example the NEOS server, see https://www.gams.com/blog/2020/11/running-large-models-on-neos-for-free/). The RICE50+ model code is open source and available on GitHub (https://github.com/witch-team/RICE50xmodel). Model extensions will likewise be made available at this repository.

Allocation of resources
Preparing data and model codes for public access as outline above is standard procedure and covered by the resources (person months) within the work package.

Data Security
To increase data security, the collected data and the developed model code will additionally be stored locally at the responsible institutions (PIK for the database, CMCC for RICE50+ development) according to the data backup and data security provision of the institutions.

Ethics
N/A

Other issues
N/A

4.5   Task 3.5 Implementation of the EU green deal via the Fit-for-55 package.

Data Summary

This task extends climate policy models (PRIDE, REMIND, WITCH, DSK) to address the questions: how to integrate new policies with existing EU policies for maximum synergies? The energy-economy models REMIND and WITCH will be extended to assess bounded rationality scenarios. The PRIDE model will be used to explore how to efficiently address multiple objectives with a limited set of policy instruments. DSK will be used to explore the roles of innovation funds and uncertainty. Thus, the task will produce policy scenarios and model extensions.

The policy scenarios will be collected in spreadsheet or text file format (Excel or CSV format). The files will contain the data sets as well as metadata, including a short description of source/origins and usage rights.

Individual modelling teams participating in the CAPABLE project shall retain control of their preliminary scenario data regarding external use outside the CAPABLE project. External use of preliminary scenario data by other partners than the modelling team itself is not allowed without explicit permission from the modelling team. After the end of the project, the quality controlled and approved data will be made publicly available in the form of consolidated snapshots in CSV format.

FAIR Data

Making data findable, including provisions for metadata

To the aim of discoverability of existing and produced data it is annotated with metadata. Where possible, it is aimed to obtain digital object identifiers (DOIs). In addition, the scientific publications in the project cross-references these data sets. Search functions and metadata are integrated in the scenario data. Individual scenarios and subgroups of scenarios can be easily identified by search-functionality, keywords and meta-categories that describe the scenarios by characteristics.

Making data accessible

Scenario data (if no limitation is specified in the consortium agreement) is collected and the quality controlled and approved final data is made openly accessible in the form of consolidated snapshots in CSV format through institutional and other, subject-specific repositories for archiving project publications and related metadata, which will allow reproducibility and transparency of the findings presented in the publications. The institutional repositories of the project partners do not claim rights over the deposited publications nor preclude access. Examples are repositories suggested by the European Commission (OpenAIRE or Zenodo).

To ensure accessibility beyond the duration of the project, the use of proprietary data formats will be avoided. Model documentations will be made openly accessible via the IAMC website, or as part of the publication of the results.

Several models are or will be made open access with the advancements and added data and be freely accessible and reusable based on GitHub repositories.

Making data interoperable

CAPABLE promotes Open Data and Open Source via publication of improved model documentations in the IAM documentation wiki. This will include descriptions of new modelling methods with their associated code (e.g. in GAMS and R). While R is a free software environment for statistical computing and graphics, GAMS needs to be purchased (but public servers offer a free use of GAMS, for example the NEOS server, see https://www.gams.com/blog/2020/11/running-large-models-on-neos-for-free/).   This   will   allow reproducibility and transparency of model findings and may incentivize further model code improvement, transfer of software development skills and collaborations within the broader scientific community.

Data produced in the project will use data formats and metadata standards established by the Integrated Assessment Modeling Consortium (IAMC). Applying community standards will allow easy combination with data sets from other European or international projects that rely on energy-economy and integrated assessment models.

For the assessment of the databases, open-source python software packages are available that allow for scenario visualisation and contain a toolbox of analytical tools. Documentation can be found at http://pyam-iamc.readthedocs.io.

Increase data re-use

Datasets produced in CAPABLE are intended to be used also outside of the project. Therefore, CAPABLE commits to open science by publicly releasing all the scenario results and providing free of charge and unrestricted online access to all related publications.

IP rights to the models participating in CAPABLE remain with their host institutions. This also applies if the model development (e.g. new data or code modules) is realised by a partner organisation different from the host institution.

Other research outputs

N/A

Allocation of resources

Resources for this are allocated as part of task 3.5.

Data Security

The final scenario database will be archived using the consolidated snapshot in CSV format on a public repository for general use at the end of the project.

Ethics

We do currently not see any ethical issues that could have an impact on the sharing of research/scenario data.

Other issues

N/A

5   WP4 - Environmental policy-making

5.1   Task 4.1 Comprehension and use of scientific knowledge by policymakers

Data Summary

This task will investigate how policymakers employ scientific evidence on climate change and other environmental issues. The task will combine an inductive approach, employing semi-structured interviews, and experimental approaches. For the semi-structured interviews, interviewees will be selected through a purposive sampling strategy and will include policymakers working at different levels (local, national, EU) and in different branches of government and organizations, as different categories possess different competencies and face different contexts and problems. The results of the semi-structured interviews will inform both the design of experiments that constitute the second part of this task and the design of the communication strategy and capacity building efforts in WP5.

For WP4 Task 4.1, primary data will be collected in two phases. The first phase will consist in collecting data from semi-structured interviews. The interviews will be conducted by participating partners and will be recorded. These insights collected through these interviews will provide the basis for the second phase, which

consists in running an (online) experimental survey with a target of 100 respondents, focused on preferences and beliefs.

The format of the data will be (i) interviews recording and transcripts, (ii) survey responses on attitudes and beliefs, (iii) other general socio-economic survey responses.

The interview recordings and transcripts will be kept confidential and accessed only by the research team. The anonymised survey responses will be stored as CSV files and made available in other common formats where practical to increase accessibility.

FAIR Data

Making data findable, including provisions for metadata

Survey data, in so far as is practically possibly, given confidentiality and data protection obligations, will be FAIR following the conclusion of the project in accordance with the guiding principles outlined in the Project Data Management plan. To aid the discoverability of data sets, DOIs will be obtained for each data set released. The acquired data sets will be added to the Zenodo repository. Furthermore, the resulting scientific publications will reference these datasets. Metadata will include the questionnaire scales, method of analysis, and keywords will be provided to optimise the possibility for discovery and potential reuse. Codebook(s) detailing the nature and justification for each variable collected will be made available for each data set.

Data sets will be collected independently. Persistent anonymised identifiers will be used with any identifying information held only by the collecting partner in order to facilitate compliance with data protection and privacy obligations.

Making data accessible

Survey data sets with accompanying metadata will be made available to reviewers of scientific publications to aid transparency in the review process. Access will be provided through open data repositories such as the Zenodo. Any embargo place on data accessibility will be limited to the duration of the project or if a publication is pending at the end of the project, be prolonged until the publication process is completed.

The data will be made available in accordance with Open science principles. Before data is shared, data will be assessed to ensure that no personal or sensitive data is made public and that anonymised participants are not identifiable. Only fully anonymized data and in certain cases, aggregated results may be made publicly available. In the event that participants no longer wish for their data to be openly accessible or re-used, procedures to ensure the removal of such data from the data set will be implemented where possible.

Metadata will be made openly available on the Zenodo platform, including all necessary information to enable the user to access the data. This entails the questionnaire and method of analysis. Necessary documentation and referencing of software to access or read the data will be given.

Making data interoperable

In order to ensure that data collected will be interoperable to the greatest extent possible, data made available will be in simplified non-proprietary (.csv) format.

It is not intended to include qualified references to previous research.

Datasets, questionnaire formats and methods of analysis will be shared openly and in case of use of uncommon ontologies or vocabularies for the project, these will be explained in the published method/appendix of analysis.

Increase data re-use

In order to increase data re-use, data will be published in a cleaned format with the cleaning code accessible for examination. The data that is being made available will be usable by third parties as soon as they are made public following any embargo period. The provenance of the data can be extracted from papers based on the referenced data set (i.e., methods/appendix section).

To ensure data quality, quality checks appropriate to the nature of the collected data will be deployed (e.g., checking for the time spent on the questionnaire, repeated answering patterns and overall response completeness. Responses that do not meet basic requirements will be excluded from the cleaned data. For individual analysis specific requirements for in/exclusion should be reported in the methods/appendix section of papers using the data as well as the code of analysis.

Other research outputs
N/A

Allocation of resources
N/A

Data Security
WP4 T4.1 data gathered and analysed for the purposes of research carried out by staff and students at the IESEG will be stored on secured IESEG hardware in such manner that allows compliance with applicable data protection and privacy legislation.

Participating partners involved in data collection will be individually responsible for maintaining and observing agreed data collection and security protocols including the preservation of the anonymity of participants.

Ethics
Data gathering, analysis and storage at IESEG for the purposes of research will be conducted in accordance with the French data protection law and General Data Protection Regulation (EU) 2016/679. The FR Data Protection Law complies with EU legislation and privacy/ confidentiality measures (European Directive 95/46/EC, covering collection and processing of personal data for scientific purposes).

Informed consent will be sought from each participant for the long-term preservation and sharing of anonymised respondent data with the respondents right to withdraw clearly outlined.

Finally, full compliance with ethical and human subjects procedures and regulations as set out by the LEM research unit, accredited by the French Centre National de la Recherche Scientifique (CNRS), will be obtained before the beginning of the data collection.

Other issues
N/A

5.2   Task 4.2 The role of political actors

Data Summary
The task will focus on political actors at the EU and Member State level, who exert a major influence on the definition of policies that can contribute to the EU long-term strategy. First, a survey will be administered to Members of the European Parliament (MEPs), to investigate their climate policy preferences - focusing on different aspects of the Fit for 55 package and on biodiversity policies - as well as their perceptions of citizens' preferences (CMCC, EUI, UAB). The survey will pay attention to opinions about distinct policies but also to weights assigned to key policy evaluation criteria (effectiveness, efficiency, equity, monitoring, political

feasibility). The survey data can be complemented by data on casted votes in the European Parliament in order to more comprehensively assess MEPs' positions. The timing of the surveys will be aligned with the public opinion surveys carried out in WP2.

Task 4.2 analyses and produces novel evidence on policy actors' preferences and behaviors concerning feasibility of different environmental policy instruments. CMCC (leader) will design the surveys of policy makers of European Parliament (including their experimental aspects) and analyze the results. Survey design and methodology will be adapted to the size of the surveyed actors. The data collection itself will be carried out in coordination with all other surveys in the project. The interview recording and transcripts will be kept confidential and accessed only by the research team.

The format of the collected data will be in line with the requirements of survey data and will include numbers, descriptive statistics, and textual data regarding the firms' preferences and behaviors amid environmental policy issues.

The data will serve for the identification of policy's preferences and current approaches that will comply to assure a sustainable transition to a decarbonized economy. In the research, CMCC will take account of other specific studies and other surveys to compare the results and to identify the research gap related to policy issues. The results of our surveys will be used by several public and private actors, especially national public authorities that are involved in the environmental policy instruments decision and implementation, helping them in formulating more informed and evidence-based public policies.

FAIR Data

Making data findable, including provisions for metadata

The acquired data sets which are not subject to confidentiality and data protection obligations will be added to a public repository. The data sets will be added to a public repository. Furthermore, any resulting scientific publication from the process described will reference these datasets.

The data sets will be added to a public repository. Metadata will include the questionnaire scales, method of analysis, and keywords will be provided to optimize the possibility for discovery and potential reuse. Codebook(s) detailing the nature and justification for each variable collected will be made available for each data set.

The policy actors' personal data will not be identified with a persistent identifier. In this task, data will not be generated about any personal and vulnerable topics, only people's personal and subjective vision about (good) life in general and visions of a fair and just transition.

Data sets will be collected independently. Persistent anonymized identifiers will be used with any identifying information held only by the collecting partner to facilitate compliance with data protection and privacy obligations.

Making data accessible

Data sets with accompanying metadata will be made available to reviewers of scientific publications to aid transparency in the review process.

The data should be openly available in reports and working papers and the data will be made available under Open science principles. Before data is shared, data will be assessed to ensure that no personal or sensitive data is made public. All data will be anonymized and made public unless a survey respondent objects.

Personal data about policy actors collected for the survey will be stored on a server that is password protected. If personal data will be used for publication, these data will be anonymised and impossible to track down to an individual. The generated data will only contain information about opinions and visions of a

climate policy preferences and behaviors. Individual opinions or statements will not be trackable to individual persons, and no personal information about the participants involved will be given.

Data will be identified by a persistent identifier. Metadata will be provided to allow discovery, following basic standards from databases like Harvard Dataverse. Search keywords will be provided in the metadata to optimize the possibility for discovery and then potential re-use.

All the data will remain available and findable indefinitely.

Metadata will be made openly available on a public platform, including all necessary information to enable the user to access the data. This entails the survey questionnaire and method of analysis. Necessary documentation and referencing of software to access or read the data will be given in a structured manner.

Making data interoperable
To ensure that the data collected will be interoperable to the greatest extent possible, data made available will be in simplified text format (docx, etc.). However, the data resulting from the surveys will be operated in the most compatible formats (like the SPSS program, or Stata) to scientifically analyze and process the policy makers' answers.

Increase data re-use
The data will represent answers to several questions on policy maker's preferences and behaviors concerning environmental policy instruments. To increase data re-use, data will be published in a simplified text format and cleaned. For example, responses or information that do not meet basic requirements will be excluded. A codebook for the dataset will be publicly available, including variable definitions and measurement.

Other research outputs
The answers and the survey results will be used in reports and scientific articles.

Allocation of resources
There are no costs associated for making data or other research outputs FAIR in the project (e.g. direct and indirect costs related to storage, archiving, re-use, security, etc.).

Data Security
This task data gathered and analyzed for research carried out by staff will be stored on secured hardware (username, password) to assure data protection and privacy legislation.

CMCC, involved in data collection, will be individually responsible for maintaining and observing agreed data collection and security protocols including the preservation of the anonymity of participants.

Ethics
Data gathering, analysis, and storage at all organizations involved in the project for research will be conducted following the General Data Protection Regulation (EU) 2016/679.

Informed consent will be sought from each participant for the long-term preservation and sharing of anonymized respondent data with the respondents' right to withdraw outlined.

It could possible that we may not be able to make all data public. For example, if a survey respondent specifically asks that we do not make their anonymized responses publicly available.

Other issues
N/A

5.3   Task 4.3 Local governance of the ecological transition

Data Summary

Task 4.3 aims to investigate the main challenges that local and regional policymakers have to address for reaching the ecological transition. First of all, EUI will interview members of the regional government and the second step will be implementing a survey with regional assemblies' members in order to capture the strengths and the barriers of the transition in Emilia-Romagna. The structure of the survey will be adapted to the sample of interest.

The second aim of the task is to address the issue of the social impact of the ecological transition. According to this, EUI and CMCC will analyse the Territorial Just Transition Plans (TJTPs) and will use interviews with experts of TJTPs to grasp challenges at local level. To the extent possible, this analysis of the TJTPs will also include a quantitative analysis of the labour impacts of the ecological transition collecting public data from selected TJTPs. Further, it will be carried out semi-structured interviews, considering several large cities on governance challenges of city-level ecological transition. The data collocation will be carried out in collaboration with the other partners involved in this task.

With respect the data format, the recordings of the interview will be confidentially shared with the research team. The output obtained by the surveys will be anonymised, available in CSV or XLSX formats and shared only with the research team.

The data collected will be useful especially for the policymakers in order to understand the main factors to take into account if they want to introduce feasible and effective policy instruments.

FAIR Data

Making data findable, including provisions for metadata

Data should be FAIR after the publication of the project, according to the guiding principles outlined in the Project Data Management plan and where possible, DOIs will be obtained for each data set created for the tasks.

Where possible, a codebook will be provided in order to explain the nature of each variable used.

Making data accessible

Metadata will be made available for scientific purposes ana for ensuring transparency once the deliverable is made public. Further, before the publication, anonymity will be assessed, and search keywords will be provided.

The metadata will remain open access and it will be correlated to clear and transparent information about the access of the data.

Making data interoperable

In order to ensure a feasible interpretation of the data, data will be provided in CSV or XLSX formats as concern the survey, and in text format for the interviews.

Increase data re-use

The results will be reusable by interested parties, including researchers, civil society organisations and policymakers. The data will be shared cleaned and clear to increase the probability of re-use. In light of this, a codebook will be provided.

Other research outputs

The answers and the survey results will be used in reports and scientific articles.

Allocation of resources

No cost has been estimated.

Data Security

The data will be contained in a hardware protected by username and password. Further, each team components will be responsible of the data preservation, including security and anonymity of the participants. Data can also be uploaded to Onedrive.

The EUI has a Data Security Officer to provide the security of information systems at the EUI, to set up EUI's security policies and controls and to take appropriate technical and organisational measures for its implementation.

Ethics

The data collection will be conducted following the EU Member States (Regulation (EU) 2016/679, GDPR). The surveys will include an informed consent at the beginning. This consent will inform participants about the aim of the surveys and how their answers will be used.

The EUI has its own Data Protection Policy (President Decision n. 10 of 18 February 2019), which establishes adequate safeguards for the protection of personal data, including a mechanism for the exercise of data subjects' rights, and the possibility for judicial review by its internal judicial system.

Other issues

N/A

5.4   Task 4.4: Assessment of different approaches to citizen engagement

Data Summary

Task 4.4 aims to tackle the issue of social acceptability of climate policy, especially the solutions and challenges of the citizens' engagement in the climate actions. Firstly, it will be carried out a computational text analysis of public consultations on EGD-related legislative proposals by the European Commission; the purpose is to identify the peculiarities of parties taking part in the consultations.

Another purpose will be analysed the European Climate Pact, it will be mapping the projects under this initiative. It will be applied a survey and some semi-structured interviews for capturing the characteristics of these projects.

Further, also through semi-structured interview and survey, it will be analysed and compared the different citizens' experiences in the climate assemblies, in different Countries.

Finally, in collaboration with CUNI and RUG, it will be examined perceived citizens 'barriers to public climate actions engagements. The results of the large-scale surveys across different European countries mentioned in Task 2.2 will be used for that purpose.

All the data will be anonymised, available in CSV or XLSX formats and confidentially shared with the research team and collaborators.

The output that will be obtained will be useful for researchers, civil society organisations and policymakers for better understanding the weakness of the public engagement instruments as well as the way for improving such tools and increasing climate policy acceptability.

FAIR Data

Making data findable, including provisions for metadata

Data should be FAIR after the publication of the project, according to the guiding principles outlined in the Project Data Management plan and DOIs will be obtained for each data set.

Where possible, a codebook will be provided in order to explain the nature of each variable used.

Making data accessible

Metadata will be made available for scientific purposes and for ensuring transparency. Further, before the publication, anonymity will be assessed, and search keywords will be provided.

The metadata will remain open access and it will be correlated to clear and transparent information about the access of the data.

Making data interoperable

In order to ensure a feasible interpretation of the data, data will be provided in CVS format as concern the survey and the computational text analysis, and in text format for the interviews.

Increase data re-use

The data will be shared cleaned and clear to increase the probability of re-use. In light of this, a codebook will be provided.

Other research outputs

The computational text analysis, answers and the survey results will be used in reports and scientific articles.

Allocation of resources

No cost has been estimated.

Data Security

The data will be contained in a hardware protected by username and password. Further, each team components will be responsible of the data preservation, including security and anonymity of the participants. Data can also be uploaded to Onedrive.

The EUI has a Data Security Officer to provide the security of information systems at the EUI, to set up EUI's security policies and controls and to take appropriate technical and organisational measures for its implementation.

Ethics

The data collection will be conducted following the EU Member States (Regulation (EU) 2016/679, GDPR). The survey will include an informed consent at the beginning. This consent will inform participants about the aim of the survey and how their answers will be used.

The EUI has its own Data Protection Policy (President Decision n. 10 of 18 February 2019), which establishes adequate safeguards for the protection of personal data, including a mechanism for the exercise of data subjects' rights, and the possibility for judicial review by its internal judicial system.

Other issues

N/A

6 WP5 - Co-design, Capacity Building, Dissemination and Communication

6.1 Task 5.1: Assessment of different approaches to citizen engagement

Data Summary

EUI will coordinate the interaction with the Advisory Board (AB) and ensure their involvement in the project activities and elicit their advice on the direction of the project in periodic intervals.

The advisory board (AB will be composed of experienced policymakers and policy experts engaged first hands-on evidence-based policymaking. The number of members will be limited to about 10-15 people. As per Consortium Agreement (CA), the Advisory Board (AB) is appointed and steered by the Project Steering Committee.

The members of the AB were invited to joined mostly during the proposal phase of the project, and confirmed on the first few month of the project. Their personal details were obtain from previous interactions with partners of the consortium. For the sake of data privacy, only CMCC, EUI and possibly the project meeting organisers will handle their personal details. Their personal data will be stored and preserved in the secured folders of the project shared drive, and the internal EUI servers.

The AB provides scientific and policy relevant feedback on the Project activities, thus facilitating and supporting with information the decisions made by the Consortium Bodies. More specifically, the role of the AB is:

- to follow the progress of the CAPABLE project
- to provide strategic advice for the project's future directions during its three-year duration
- to give advice on the stakeholder involvement and policy relevance of the project
- to support its dissemination and exploitation activities

Membership of the AB is voluntary, unremunerated and by invitation, coordinated by EUI on behalf of the Consortium. The AB will meet at least once a year or at any time upon request of the CO or 1/3 of the Members of the CB. The AB members will be also invited to attend (some of) the Project Meetings, which will take place on a regular basis. The meetings are programmed where some of the Consortium partners are located, Remote participation will also be available in case travel to these locations is not convenient or not possible. AB members can be reimbursed by the consortium partners for their travel and subsistence costs in line with the CAPABLE reimbursement rules and with the partner institutions internal rules that will be shared with the AB members.

During the project meetings, AB members in attendance will have the opportunity to provide feedback on the progress of the CAPABLE project to the CAPABLE Project Steering Committee and the consortium partners. Public information (final documents, tools, etc.) will be explicitly identified as such and the AB is encouraged to disseminate them as widely as possible. Upon their agreement, the composition and role of the AB members could be publicly acknowledged on the project website unless an AB member explicitly asks for her/his membership to AB to remain confidential. The project management team at CMCC and the consortium partner EUI will support the AB in its activities, including convening of teleconferences and keeping AB members updated on the proceedings of the project.

The AB members will sign a Non-Disclosure Agreement with Parties of the Project as detailed in the CAPABLE CA (Attachment 5: NDA for Advisory Board agreed under Section 6 "Confidentiality Undertaking"). Any sensitive data will be collected (i.e. name, surname, affiliation and email which are also publicly available on the web).

FAIR Data
Making data findable, including provisions for metadata
N/A

Making data accessible
Name, surname and affiliation of the AB members will be published on the CAPABLE website upon their agreement, All the other personal details of the members of the AB will remain confidential.

Making data interoperable
N/A

Increase data re-use
N/A

Other research outputs
N/A

Allocation of resources
Resources have been allocated to pay travels for the AB members. No costs associated with the management of the Advisory Board data.

Data Security
All data will be stored on the EUI servers, which have institutional security set up. Data can also be uploaded to Onedrive, while away from the Institute, with a two-step security measure (logging and telephone token).

The EUI has also appointed a Data Security Officer with a mandate to safeguard the security of information systems at the EUI, to set up EUI's security policies and controls and to take appropriate technical and organisational measures for its implementation.

Raw data generated will be stored for five years after the end of the project. After that date, they will be destroyed.

Ethics
The participation to the Advisory Board is on a voluntary basis and the involvement of the members is not foreseen to raise issues that require approval of the ethics committee. If necessary, a higher level of data protection will be applied.

Other issues
N/A

6.2 Task 5.2: Communication and dissemination

Data Summary
CMCC, with the important support of E6, will propose and coordinate a range of C&D&E activities. A GDPR-compliant project website will be created and will provide project information such as partners, objectives, upcoming events, news, and major project developments. All dissemination documents and public deliverables will be accessible for download. The website will be maintained during the project lifetime and also beyond.

FAIR Data
Making data findable, including provisions for metadata
All the scientific materials collected for Task 5.2 will be uploaded to a CAPABLE website (set up by the E6 partner with the support of the CO).

Making data accessible
All the scientific materials collected for Task 5.2 will be made publicly available on the CAPABLE website (set up by the E6 partner with the support of the CO).

Metadata
N/A

Making data interoperable

N/A

Increase data re-use

N/A

Other research outputs

All the scientific materials can be used in reports and scientific articles.

Allocation of resources

All the scientific materials will be validated by the CO internally. There are no costs associated for making data or other research outputs FAIR in the project (e.g. direct and indirect costs related to storage, archiving, re-use, security, etc.).

Data Security

To ensure lawfulness of the processing of personal data, informed consent procedures (cf. art 6 (1a) EU GDPR) will be implemented for the participation of humans and concerning data processing i.e. in the context of project meetings, the webinars and stakeholder workshops. The informed consent forms will be based on validated consent templates. An exemplary privacy policy is attached in the Annex 1 of this document.

The scientific material from speakers, experts and stakeholders who will participate to the CAPABLE events are collected through personal emails by the consortium partner hosting the specific event.

The local partners will adapt the available consent forms to the different activities and workshops, including the appropriate language; EUI is the leading partner for stakeholder workshops. If needed, EUI will be supported by the CMCC Data Protection Officer.

The data subject will be provided with information about the envisaged data processing in an intelligible and easily accessible form (that is, in language and term intelligible to the participants) to ensure a transparent and voluntary choice. This includes:

- the identity of the data controller;
- the specific purpose(s) of the processing for which the personal data will be used;
- the subject's rights as guaranteed by the GDPR, in particular the right to withdraw consent or access their data;
- information as to whether data will be shared with or transferred to third parties and for what purposes;
- and the time data will be retained before they are destroyed.

Once finalized, the forms will be stored and archived at the responsible institution. We will also use the project website to inform on the participant rights (with respect to their given consent) incl. right of access (art 15 EU GDPR), right to rectification (art 16 EU GDPR), right to restriction of processing (art 18 EU GDPR), right to erasure (art 17 EU GDPR), right to withdraw (art 7 (3) sentence 1 EU GDPR), right to lodge a complaint with a supervisory authority (art 77 EU GDPR).

Ethics

The participation to the events/dissemination activities (i.e. Webinar, if any) will be on a voluntary basis and the involvement of the speakers and moderators is not foreseen to raise issues that require approval of the ethics committee. If necessary, a higher level of data protection will be applied.

Other issues

N/A

ANNEX 1 - Information on the processing of personal data in accordance with article 13 of the Regulation 2016/679 (GDPR) regarding Webinar & Events hosted or organized by the CMCC Foundation