1112122N De Canniere DMP (FWO FR)

Project Name Fueling the Road to Zero: Environmental Policy Evaluation with Multiple Market Failures - 1112122N De Canniere DMP (FWO FR)

Grant Title 1112122N

Principal Investigator / Researcher Charlotte De CanniÃ¨re

Description The project evaluates the distributional and efficiency aspects of environmental policies in a context of multiple market failures. The first work package studies the demand for (company) cars in Belgium. I focus on the differential effects of taxing fossil fuels, or imposing fuel efficiency standards. On top of that, I allow for a large extent of heterogeneity by allowing different types of consumers (company car vs no company car) . The second project focuses on anticompetitive behavior in the world oil market. This part aims to understand how the global markup found in the oil industry would interact with a global carbon tax, and its effect on total emission and redistributional consequences. The findings of this analysis can be linked to the Paris Agreement goals to reduce global fossil fuel emissions. Third, I study the Belgian electricity market. More precisely, I aim to understand how the lack of international integration of the EU electricity market affects the overall efficiency of the sector, and how it has affected the impact of a carbon tax on electricity prices.

Institution KU Leuven

1. General Information

Name applicant

Charlotte De Cannière

FWO Project Number & Title

FWO Project Title: Fueling the Road to Zero: Environmental Policy Evaluation with Multiple Market Failures

FWO Project Number: 1112122N

Affiliation

- KU Leuven

2. Data description

Will you generate/collect new data and/or make use of existing data?

- Reuse existing data

Describe in detail the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a table (see example) or as a data flow and per WP or objective of the project. If you reuse existing data, specify the source of these data. Distinguish data types (the kind of content) from data formats (the technical format).

The different work packages of the project use different types of (confidential) data sets from marketing organizations (JATO Dynamics, Rystad energy) and government agencies (Belgian Commission for the Regulation of Electricity and Gas (CREG)). The table below gives an overview.

Type of data	Format	Size	How created
automobiles	.csv	5-15 GB	Obtained from JATO
oil production	.csv	10-25 GB	Obtained from Rystad Energy
electricity production	.csv	5-15 GB	Obtained from CREG

The first dataset consists of demand volumes and prices in different countries over multiple periods (years or quarters). The second dataset covers production and cost data from worldwide oil assets at the yearly level. The third dataset contains production and cost data from the Belgian production fleet of electricity.

The datasets will not combined across the three parts, but within each part they will be merged

with publicly available variables (such as population, income, PPI, or energy forecast variables).
Statistical software packages, such as Stata and Matlab, are used to manage the data and estimate econometric models. This generates tables with aggregate summary statistics and model parameter estimates, as well as figures and graphs to illustrate the findings.

3. Legal and ethical issues

Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to your file in KU Leuven's Register of Data Processing for Research and Public Service Purposes (PRET application). Be aware that registering the fact that you process personal data is a legal obligation.

- No

Privacy Registry Reference: NA
Short description of the kind of personal data that will be used:
None of the three work packages use individual-level personal data. For the automobile project, I will use data on car purchases on the town-level in Belgium. This dataset will be complemented by publicly available datasets with town-level characteristics (average household income, employment rate, etc.). Given this aggregation, the researcher will be unable to identify household on an individual basis.

All other project involve production data, rather than personal data.

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, add the reference to the formal approval by the relevant ethical review committee(s)

- No

Does your work possibly result in research data with potential for tech transfer and valorisation? Will IP restrictions be claimed for the data you created? If so, for what data and which restrictions will be asserted?

- No

Do existing 3rd party agreements restrict dissemination or exploitation of the data you (re)use? If so, to what data do they relate and what restrictions are in place?

- Yes

The confidentiality agreements with JATO Dynamics, Rystad Energy and CREG do not allow me to publicly share the data sets. They can only be accessed by collaborating researchers who obtained permission.
All collaborators are aware of the confidentiality agreements, and work on a secure server.

4. Documentation and metadata

What documentation will be provided to enable reuse of the data collected/generated in this project?

1. The underlying data are not created by myself. Therefore, they are documented by the organizations that created them.
2. I documents the statistical code with comments. Furthermore, I keep track of a readme file that describes how the different programs are related in the various steps of the analysis (data management including sample selection and variable creation; statistical and econometric analysis; production of tables and figures).

Will a metadata standard be used? If so, describe in detail which standard will be used. If no, state in detail which metadata will be created to make the data easy/easier to find and reuse.

- No

There is no need for this because each part of the project makes use of existing datasets from external sources and has a relatively simple structure.

I already carefully document the data management, and statistical and econometric analysis in the code and readme files, as decribed earlier.

5. Data storage and backup during the FWO project

Where will the data be stored?

The data sets are stored on the KU Leuven servers, except for the part on electricity production, which is stored on the CREG servers. The FEB ICT department provides regular backups.

For collaboration with other researchers from other research untis, we will use DropBox for active use of the data during the project.

How is backup of the data provided?

The data will be stored on the university's central servers with automatric daily back-up procedures. Servers and PC's are managed by ICT department of FEB according to latest security norms, to rule out loss of data.

Is there currently sufficient storage & backup capacity during the project? If yes, specify concisely. If no or insufficient storage or backup capacities are available then explain how this will be taken care of.

- Yes

The total storage is estimated around 60GB. There is sufficient storage capacity on servers and PCs both during and after the research.

What are the expected costs for data storage and back up during the project? How will these costs be covered?

The core storage is provided by the readily available powerful server computer of FEB ICT.

To provide the active use of the datasets in Dropbox, I will use the FWO benchfee (Dropbox Plus- 3 TB storage - 120€/year).

Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

Servers and PCs are managed by the ICT department of FEB, according to latest security norms. Data can be stored encrypted. Data sets van only accessed by researchers that obtained permission.

6. Data preservation after the FWO project

Which data will be retained for the expected 5 year period after the end of the project? In case only a selection of the data can/will be preserved, clearly state the reasons for this (legal or contractual restrictions, physical preservation issues, ...).

1. The raw datasets will be retained by JATO Dynamics, Rystad Energy and CREG indefinitely, and will be kept on KU Leuven servers for at least 10 years, except for

- The electricity data, as it is stored on the CREG servers, instead of the KU Leuven servers because of contractual restrictions.

2. The statistical programs/code and the documentation will also be maintained for at least 10 years, and in principle indefinitely, as they are very small in storage size. The statistical programs/code will also be made publicly accessible.

Where will the data be archived (= stored for the longer term)?

If contractually allowed (as mentioned in the previous question), the data and statistical programs/code will be stored on the KU Leuven servers (with automatic back-up procedures) for at least 10 years, conform the KU Leuven RDM policy.

What are the expected costs for data preservation during the retention period of 5 years? How will the costs be covered?

The costs are negligible because the infrastructure is already available.

7. Data sharing and reuse

Are there any factors restricting or preventing the sharing of (some of) the data (e.g.

as defined in an agreement with a 3rd party, legal restrictions)?

- Yes. Specify:

The agreements with JATO Dynamics, Rystad Energy and CREG do not allow me to publicly share the data sets. They can only be accessed by researchers that obtained their permission.

Which data will be made available after the end of the project?

1. The code will be made publicly accessible.

2. The agreements with Rystad Energy, JATO Dynamics, and CREG do not allow me to share the data sets publicly. They can only be accessed by researchers who obtained permission.

Where/how will the data be made available for reuse?

- Upon request by mail
- Other (specify):

Data sets can be accessed by researchers who obtained permission, and the code is made publicly accessible through journals or author's website, or on request.

When will the data be made available?

- Immediately after the end of the project
- Upon publication of the research results

Data sets can only be accessed by researchers who obtained permission, and will not be shared by the researcher. The code and results, however, is made publicly accessible without embargo through journal or author's website, or on request.

Who will be able to access the data and under what conditions?

Data sets can be accessed by researchers who obtained permission from JATO Dynamics, Rystand Energy and CREG respectively.

The code is made publicly accessible.

What are the expected costs for data sharing? How will the costs be covered?

Not applicable.

8. Responsibilities

Who will be responsible for data documentation & metadata?

Charlotte De Canniere (PhD candidate): implementation, updating, co-ordination and legal responsability

Who will be responsible for data storage & back up during the project?

Charlotte De Canniere

Who will be responsible for ensuring data preservation and reuse ?

Charlotte De Canniere (PhD candidate) and Frank Verboven (supervisor)

Who bears the end responsibility for updating & implementing this DMP?

Charlotte De Canniere bears the end responsibility of updating & implementing this DMP.