FWO DMP Template - Flemish Standard Data Management Plan

Version KU Leuven

Project supervisors (from application round 2018 onwards) and fellows (from application round 2020 onwards) will, upon being awarded their project or fellowship, be invited to develop their answers to the data management related questions into a DMP. The FWO expects a **completed DMP no later than 6 months after the official start date** of the project or fellowship. The DMP should not be submitted to FWO but to the research co-ordination office of the host institute; FWO may request the DMP in a random check.

At the end of the project, the **final version of the DMP** has to be added to the final report of the project; this should be submitted to FWO by the supervisor-spokesperson through FWO's e-portal. This DMP may of course have been updated since its first version. The DMP is an element in the final evaluation of the project by the relevant expert panel. Both the DMP submitted within the first 6 months after the start date and the final DMP may use this template.

The DMP template used by the Research Foundation Flanders (FWO) corresponds with the Flemish Standard Data Management Plan. This Flemish Standard DMP was developed by the Flemish Research Data Network (FRDN) Task Force DMP which comprises representatives of all Flemish funders and research institutions. This is a standardized DMP template based on the previous FWO template that contains the core requirements for data management planning. To increase understanding and facilitate completion of the DMP, a standardized **glossary** of definitions and abbreviations is available via the following link.

1. General Project Information	
Name Grant Holder & ORCID	Christian Pröbsting 0000-0003-4885-9327
Contributor name(s) (+ ORCID) & roles	Damien Vliegen 0009-0000-4147-0664 (contributor)
Project number [1] & title	ENERGY PRICE SHOCKS – EVIDENCE FROM U.S. REGIONS (3H230743)
Funder(s) GrantID [2]	D-2024-2882
Affiliation(s)	☑ KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel ☐ Other: ROR identifier KU Leuven: 05f950310
Please provide a short project description	The macroeconomic literature emphasizes that energy price shocks have surprisingly large effects on economic activity, despite the modest share of spending on crude oil. What could explain this strong sensitivity of economic activity to energy prices? A wide range of supply- and demand-side mechanisms have been proposed to explain this: endogenous response of monetary policy oil price shocks, time-varying markups, complementarity in energy and capital goods in the production function, disruption in households' and firms' spending on goods and services other than energy, to name a few. However the link between model predictions and empirical tests is often quite loose. There is currently a lack of reliable macro model to study energy crises and evaluate alternative policy options. The goal of this project is to foster our understanding of how energy price shocks affect the economy by confronting state-of-the-art macroeconomic models with causally identified moments that help discriminate between supply-driven or demand-driven models. To achieve this scientific objective, the project is composed of two large blocks: (i) an empirical framework to identify the partial equilibrium response of economic activity both at the state and the state-industry level to a local energy price shock, and (ii) a structural, general equilibrium model that captures the empirically identified transmission channels and serves to translate the regional responses into aggregate responses.

[1] "Project number" refers to the institutional project number. This question is optional. Applicants can only provide one project number.

[2] Funder(s) GrantID refers to the number of the DMP at the funder(s), here one can specify multiple GrantIDs if multiple funding sources were used.

2. Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data.

				Only for digital data	*Only for digital data*	*Only for digital data*	*Only for physical data*
Dataset Name	Description	New or Reused	Digital or Physical	Digital Data Type	Digital Data Format	Digital Data Volume (MB, GB, TB)	Physical Volume
US-Level Energy Prices	US Energy Information Administration (EIA) electricity, oil, and natural gas monthly price series. Fed of Saint Louis (FRED) coal monthly price series.	☐ Generate new data ☒ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	.xls	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
Industry-Level Output Prices	US Bureau of Labor Statistics (BLS) output monthly price series (SIC and NAICS code based).	☐ Generate new data ☒ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	.xls	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
Industry-Level Utilization Rates	US Federal Reserve Board (FRB) industry-level capacity utilization rate monthly data (NAICS code	☐ Generate new data ☒ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical	.xls	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB	

	based).			☐ Textual ☐ Model ☐ Software ☐ Other:		☐ > 5 TB ☐ NA	
Industry-Level Industrial Production	US Federal Reserve Board (FRB) industry-level industrial production rate monthly data (NAICS code based).	☐ Generate new data ☒ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	.xls	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
Input-Output Tables	US Bureau of Economic Analysis (BEA) 5-yearly input-output tables (1972-2017). IO industry code based.	☐ Generate new data ☒ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	.txt .csv .xls .fmt .DAT	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
Oil Supply News Shocks	Monthly oil supply news shocks and oil price series from Känzig (2023) replication package.	☐ Generate new data ☒ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	MATLAB data	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
SIC/IO Code Concordance Tables	Industry SIC/IO code concordance tables by Nekarda and Ramey	☐ Generate new data ☒ Reuse	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound	.xls	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB	

	(2011).	existing data		☒ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:		☐ < 5 TB ☐ > 5 TB ☐ NA	
US Census SIC/NAICS Concordance Tables	Industry SIC/NAICS concordance tables by the US Census.	☐ Generate new data ☒ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	.xls	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	

GUIDANCE:
The data description forms the basis of your entire DMP, so make sure it is detailed and complete. It includes digital and physical data and encompasses the whole spectrum ranging from raw data to processed and analysed data including analysis scripts and code. Physical data are all materials that need proper management because they are valuable, difficult to replace and/or ethical issues are associated. Materials that are not considered data in an RDM context include your own manuscripts, theses and presentations; documentation is an integral part of your datasets and should described under documentation/metadata.
RDM Guidance on data

If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL etc.) per dataset or data type.	EIA Energy Price data: https://www.eia.gov/outlooks/steo/realprices/ FRED Coal Price data: https://fred.stlouisfed.org/series/WPU0512 BLS Output Price data: https://data.bls.gov/cgi-bin/srgate (enter "PCU + 6-digit industry code (NAICS) twice" or "PDU + 4-digit industry code (SIC) twice") FRB Industrial Production and Capacity Utilization data: https://www.federalreserve.gov/datadownload/Choose.aspx?rel=G17 BEA Input-Output data (1972-2002): https://www.bea.gov/industry/historical-benchmark-input-output-tables BEA Input-Output data (2007-2017): https://www.bea.gov/industry/input-output-accounts-data Känzig (2023) Oil Supply News shock: https://github.com/dkaenzig/replicationOilSupplyNews Nekarda and Ramey (2011) SIC/IO concordance tables: https://www.aeaweb.org/articles?id=10.1257/mac.3.1.36 US Census SIC/NAICS concordance tables: https://www.census.gov/naics/?68967
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, refer to specific datasets or data types when appropriate and provide the relevant ethical approval number.	☐ Yes, human subject data; provide SMEC or EC approval number: ☐ Yes, animal data; provide ECD reference number: ☐ Yes, dual use; provide approval number: ☒ No Additional information:
Will you process personal data[3]? If so, please refer to specific datasets or data types when appropriate and provide the KU Leuven or UZ Leuven privacy register number (G or S number).	☐ Yes (provide PRET G-number or EC S-number below) ☒ No Additional information:

[3] See Glossary Flemish Standard Data Management Plan

Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, …)? If so, please comment per dataset or data type where appropriate.	☐ Yes ☒ No If yes, please comment:
Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material/Data transfer agreements, research collaboration agreements)? If so, please explain to what data they relate and what restrictions are in place.	☐ Yes ☒ No If yes, please explain:
Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain to what data they relate and which restrictions will be asserted.	☐ Yes ☒ No If yes, please explain:

3. Documentation and Metadata	
Clearly describe what approach will be followed to capture the accompanying information necessary to keep **data understandable and usable**, for yourself and others, now and in the future (e.g. in terms of documentation levels and types required, procedures used, Electronic Lab	Documentation and metadata will be added to all raw and processed datasets. A README file for each dataset will contain all the necessary information about the context of the data generation, the research project to which it belongs, and an interpretation of the data. A data dictionary/code book will be created as specified in the following section.

Notebooks, README.txt files, Codebook.tsv etc. where this information is recorded).
 	
RDM guidance on documentation and metadata.
	The code used for data analysis and transformation will be documented in-file with detail, adding key comments at each particular step. Every code file will include a documentation section on top indicating:
	• Goal: a short description of the code.
	• Input: name, location, and format of all datasets/files needed to execute the code, clearly specifying their key identifiers.
	• Transformations: short bullet points detailing the critical steps/transformations conducted in the code.
	• Output: name, location, and format of all the datasets/files produced by the code.
	
	All the documentation (data, corresponding documents, methodologies) will be kept in the folder where the dataset is stored. The project will use the following folder structure:
	• Literature: relevant papers and bibliography related to the research.
	• Data: contains the data dictionary and an individual folder for each dataset. In each dataset folder, the structure will be:
	  ○ code: the code files used for data transformation, which follow standard naming conventions.
	  ○ docs: relevant documentation/methodology.
	  ○ input: raw datasets.
	  ○ output: processed datasets.
	  ○ tmp: temporary data files generated by the code.
	  ○ A README file.
	• Archive: obsolete data/code which merits storage.
	• Notes: meeting notes relating to the project, with participants and dates clearly stated.
	• Presentations: external presentations made about the project.
	• Tasks: contains folders (with order of execution) indicating the transformations/analysis made in the project. In each task folder, the structure will be:
	  ○ code: the code files used for data transformation, which follow standard naming conventions and indicate with numbers the order of execution.
	  ○ docs: relevant documentation/methodology.
	  ○ output: output generated by the task (tables, derived datasets).
	  ○ fig: figures generated by the task.

	○ tmp: temporary data files generated by the code. ○ report: a LaTex document indicating the results/analysis conducted in the particular task. All good practices in terms of file and variable naming will be followed. Every variable will be stored together with a descriptive label. The variable type (string, float, integer,…) will be clearly indicated and kept consistent across datasets. GitHub will be used as a task organization and versioning system. All standard naming conventions for variables, controlled vocabularies, and ontologies will be followed (main reference: Gentzkow, M., & Shapiro, J. M. (2014). Code and data for the social sciences: A practitioner's guide. Chicago, IL: University of Chicago)
Will a metadata standard be used to make it easier to **find and reuse the data**? If so, please specify which metadata standard will be used. If not, please specify which metadata will be created to make the data easier to find and reuse. *REPOSITORIES COULD ASK TO DELIVER METADATA IN A CERTAIN FORMAT, WITH SPECIFIED ONTOLOGIES AND VOCABULARIES, I.E. STANDARD LISTS WITH UNIQUE IDENTIFIERS.*	☒ Yes ☐ No If yes, please specify (where appropriate per dataset or data type) which metadata standard will be used: A data dictionary (.xslx) of the derived datasets used for analysis is to be kept and regularly updated. It contains all file names, variables, and descriptions. It clearly indicates the data keys (= variables defining dimensions of dataset/unique identifiers). If no, please specify (where appropriate per dataset or data type) which metadata will be created:

4. Data Storage & Back-up during the Research Project

Where will the data be stored? *Consult the interactive KU Leuven storage guide to find the most suitable storage solution for your data.*	☐ Shared network drive (J-drive) ☐ Personal network drive (I-drive) ☒ OneDrive (KU Leuven) ☐ Sharepoint online ☐ Sharepoint on-premis ☐ Large Volume Storage ☐ Digital Vault ☐ Other:
How will the data be backed up? *WHAT STORAGE AND BACKUP PROCEDURES WILL BE IN PLACE TO PREVENT DATA LOSS?*	☒ Standard back-up provided by KU Leuven ICTS for my storage solution ☒ Personal back-ups I make (specify) ☐ Other (specify) Code files documenting the transformation of the raw data to processed data will be version-controlled using GitHub, stored online, and made publicly available together with the published papers.
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 ☐ No If no, please specify:
How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons? *CLEARLY DESCRIBE THE MEASURES (IN TERMS OF PHYSICAL SECURITY, NETWORK SECURITY, AND SECURITY OF COMPUTER SYSTEMS AND FILES) THAT WILL BE TAKEN TO ENSURE THAT STORED AND TRANSFERRED DATA ARE SAFE.* *Guidance on security for research data*	The data will be stored using KU Leuven OneDrive. As such, it is subject to KU Leuven's network security structure. The GitHub accounts of the coauthors will be protected by multi-factor authentication.

What are the expected costs for data storage and backup during the research project? How will these costs be covered?	The OneDrive license we will use is free for KU Leuven staff members.

5. Data Preservation after the end of the Research Project	
Which data will be retained for at least five years (or longer, in agreement with other retention policies that are applicable) after the end of the project? In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...). *Guidance on data preservation*	☒ All data will be preserved for 10 years according to KU Leuven RDM policy ☐ All data will be preserved for 25 years according to CTC recommendations for clinical trials with medicinal products for human use and for clinical experiments on humans ☐ Certain data cannot be kept for 10 years (explain)
Where will these data be archived (stored and curated for the long-term)? *Dedicated data repositories are often the best place to preserve your data. Data not suitable for preservation in a repository can be stored using a KU Leuven storage solution, consult the interactive KU Leuven storage guide.*	☒ KU Leuven RDR ☐ Large Volume Storage (longterm for large volumes) ☐ Shared network drive (J-drive) ☐ Other (specifiy):
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	There are no costs expected for the data preservation.

6. Data Sharing and Reuse	
Will the data (or part of the data) be made available for reuse after/during the project? Please explain per dataset or data type which data will be made available. *NOTE THAT 'AVAILABLE' DOES NOT NECESSARILY MEAN THAT THE DATA SET BECOMES OPENLY AVAILABLE, CONDITIONS FOR ACCESS AND USE MAY APPLY. AVAILABILITY IN THIS QUESTION THUS ENTAILS BOTH OPEN & RESTRICTED ACCESS. FOR MORE INFORMATION: HTTPS://WIKI.SURFNET.NL/DISPLAY/STANDARDS/INFO-EU-REPO/#INFOEUREPO-ACCESSRIGHTS*	☒ Yes, as open data ☐ Yes, as embargoed data (temporary restriction) ☐ Yes, as restricted data (upon approval, or institutional access only) ☐ No (closed access) ☐ Other, please specify:
If access is restricted, please specify who will be able to access the data and under what conditions.	NA
Are there any factors that restrict or prevent the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)? Please explain per dataset or data type where appropriate.	☐ Yes, privacy aspects ☐ Yes, intellectual property rights ☐ Yes, ethical aspects ☐ Yes, aspects of dual use ☐ Yes, other ☒ No If yes, please specify:
Where will the data be made available? If already known, please provide a repository per dataset or data type.	☒ KU Leuven RDR ☐ Other data repository (specify) ☐ Other (specify)

When will the data be made available?	☒ Upon publication of research results ☐ Specific date (specify) ☐ Other (specify)
Which data usage licenses are you going to provide? If none, please explain why. *A DATA USAGE LICENSE INDICATES WHETHER THE DATA CAN BE REUSED OR NOT AND UNDER WHAT CONDITIONS. IF NO LICENCE IS GRANTED, THE DATA ARE IN A GREY ZONE AND CANNOT BE LEGALLY REUSED. DO NOTE THAT YOU MAY ONLY RELEASE DATA UNDER A LICENCE CHOSEN BY YOURSELF IF IT DOES NOT ALREADY FALL UNDER ANOTHER LICENCE THAT MIGHT PROHIBIT THAT.* *Check the RDR guidance on licences for data and software sources code or consult the License selector tool to help you choose.*	☒ CC-BY 4.0 (data) ☐ Data Transfer Agreement (restricted data) ☒ MIT licence (code) ☐ GNU GPL-3.0 (code) ☐ Other (specify)
Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, please provide it here. *INDICATE WHETHER YOU INTEND TO ADD A PERSISTENT AND UNIQUE IDENTIFIER IN ORDER TO IDENTIFY AND RETRIEVE THE DATA.*	☒ Yes, a PID will be added upon deposit in a data repository ☐ My dataset already has a PID ☐ No
What are the expected costs for data sharing? How will these costs be covered?	There are no expected costs for the data sharing.

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	The responsible for supervising the process of data documentation and metadata is prof. Christian Pröbsting. An authorized staff member is PhD student Damien Vliegen.

Who will manage data storage and backup during the research project?	The responsible for supervising the process of data storage and backup is prof. Christian Pröbsting. An authorized staff member is PhD student Damien Vliegen.
Who will manage data preservation and sharing?	The responsible for supervising the data preservation and sharing is prof. Christian Pröbsting. An authorized staff member is PhD student Damien Vliegen.
Who will update and implement this DMP?	The responsible for supervising the updating and implementation of this DMP is prof. Christian Pröbsting. An authorized staff member is PhD student Damien Vliegen.