Discrete Continuous Dynamic Choice Models (C14/24/013)
FWO DMP (Flemish Standard DMP)

1. 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
		Please choose from the following options: • Generate new data • Reuse existing data	*Please choose from the following options:* • Digital • Physical	*Please choose from the following options:* • Observational • Experimental • Compiled/aggregated data • Simulation data • Software • Other • NA	*Please choose from the following options:* • .por, .xml, .tab, .csv,.pdf, .txt, .rtf, .dwg, .gml, … • NA	*Please choose from the following options:* • <100MB • <1GB • <100GB • <1TB • <5TB • <10TB • <50TB • >50TB • NA	
R package	Software code for the methods	Generate new data	Digital	Software	.R	<100MB	
Panel Study of Income Dynamics (PSID)	Data for an application	Re-use existing data	Digital	Observational	.csv	<100MB	
Current Population Survey (CPS)	Data for an application	Re-use existing data	Digital	Observational	.csv	<100MB	
American Community Survey (ACS)	Data for an application	Re-use existing data	Digital	Observational	.csv	<100MB	

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:

Panel Study of Income Dynamics (PSID): https://simba.isr.umich.edu/data/data.aspx
Current Population Survey (CPS): https://cps.ipums.org/cps/
American Community Survey (ACS): https://usa.ipums.org/usa/

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? Describe these issues in the comment section. Please refer to specific datasets or data types when appropriate.

- No

Will you process personal data? If so, briefly describe the kind of personal data you will use in the comment section. Please refer to specific datasets or data types when appropriate.

- No

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.

- No

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 in the comment section to what data they relate and what restrictions are in place.

- No

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 in the comment section to what data they relate and which restrictions will be asserted.

- No

2. 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 Notebooks, README.txt files, Codebook.tsv etc. where this information is recorded).

For the observational data: I will do a standard README.txt file + provide all the code in order to be able to reproduce the research by obtaining the original dataset (this is what is asked by the publishing journals).
For the package: I will make a package website with documentation on github, providing the code to all the simulations allowing anyone to reproduce it at home, as well as to download the package.
The package will also be made available on CRAN.

Will a metadata standard be used to make it easier to find and reuse the data? If so, please specify (where appropriate per dataset or data type) which metadata standard will be used. If not, please specify (where appropriate per dataset or data type) which metadata will be created to make the data easier to find and reuse.

- No

The R package will be uploaded on CRAN https://cran.r-project.org/ which allows to retrieve the most up to date version at all time.

3. Data storage & back-up during the research project

Where will the data be stored?

During the project, we will use the range of storage solutions provided by KU Leuven: a KUL-managed computer, desktop file storage, KUL OneDrive. These storage types are encrypted by a personal password.

How will the data be backed up?

Automatic backup by the IT department. Remark that since we use only data which is easily available online, it suffices to save the code that modifies the data in order to reproduce it by downloading the original data files on the corresponding websites. The datasets are not that big so reproducing it should take at most 20 minutes on any decent computer.

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

As said above: the datasets are not that large.

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

Only the PI have access to the repositories which are protected by strong passwords. But in any case, the original data is freely available to anyone online.

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

There are no costs as KU Leuven already provides us with the necessary storage.

4. 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...).

Data is stored on KU Leuven servers for at least 10 years.

Where will these data be archived (stored and curated for the long-term)?

On KU Leuven servers.

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

There are no cost, we use already existing KU Leuven servers.

5. Data sharing and reuse

Will the data (or part of the data) be made available for reuse after/during the project? In the comment section please explain per dataset or data type which data will be made available.

- Yes, in an Open Access repository

For the R package, there will be an online application and website for the documentation + anyone will be able to download the package.
For the data reproducibility, all the codes + readme files will be available as well (typically on the website of the journal publishing the papers).

If access is restricted, please specify who will be able to access the data and under what conditions.

-

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 in the comment section per dataset or data type where appropriate.

- No

Where will the data be made available? If already known, please provide a repository per dataset or data type.

The R package will appear on CRAN and Github. The package has not been created yet, so no existing link yet.

When will the data be made available?

As soon as possible, I already started working on the package, but it's not finished so I expect 2026.

Which data usage licenses are you going to provide? If none, please explain why.

The following MIT license, e.g., https://cbruneelzupanc.github.io/semiIVreg/LICENSE.html

Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, you have the option to provide it in the comment section.

- No

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

No costs, the data sharing will be free.

6. Responsibilities

Who will manage data documentation and metadata during the research project?

The PI: Christophe Bruneel

Who will manage data storage and backup during the research project?

The PI: Christophe Bruneel

Who will manage data preservation and sharing?

The PI: Christophe Bruneel

Who will update and implement this DMP?

The PI: Christophe Bruneel