Erasmus
School of
Economics

What does Open Science mean for Econometricians?

Andreas Alfons

EI Department Meeting, November 9, 2021

Erasmus University Rotterdam

Why Open Science in methodological research?

- We wouldn't expect someone to trust a theoretical result without a mathematical proof

- Why would we expect someone to trust a table with empirical results without the code?

FAIR principle

Findable

Accessible

Interoperable

Reusable

$\longrightarrow$ Introduced for scientific data management, but also applies to other domains (e.g., code)

Low-hanging fruit

$\longrightarrow$ Make sure your research is not behind a paywall

Open Access publishing:

- VSNU agreement with many publishers
  $\longrightarrow$ Corresponding author has to be from a Dutch university
- ~~Erasmus Open Acces fund~~ (stopped in June 2021)
- Open Access Journals, e.g.:
  - Journal of Machine Learning Research
  - Journal of Statistical Software
  - Journal of Data Science, Statistics, and Visualization
- Include budget for Open Access fees in grant applications

Low-hanging fruit

$\longrightarrow$ Make sure your research is not behind a paywall

Make a preprint available, e.g.:

- arXiv: `https://arxiv.org/`
- SSRN: `https://www.ssrn.com/`
- RePub: `https://repub.eur.nl/`
- ERIM Research Report / TI Discussion Paper

Low-hanging fruit

$\longrightarrow$ Make the code for your method available (increases citations!)

Any code is better than no code:

- It doesn't have to pretty
- It doesn't have to be efficient
- You don't have to provide support

Low-hanging fruit

$\longrightarrow$ Make the code for your method available (increases citations!)

Make your code easily findable:

- Put it on a popular code sharing platform, e.g.:
  - GitHub: `https://github.com/`
  - GitLab: `https://gitlab.com/`
  - Bitbucket: `https://bitbucket.org/`
- Put it on the EUR Data Repository (figshare): `https://datarepository.eur.nl/`
  $\longrightarrow$ You get a DOI, so your code is citable!

A step further

$\longrightarrow$ Make replication files for your analyses available

Scripts that reproduce:

- Examples
- Simulation studies or benchmark experiments
- In particular all figures and tables of your paper

$\longrightarrow$ This requires some effort, but:

- Several (top) journals require this already (e.g., JASA), and this will only become more prevalent
- Grant agencies reward Open Science practices

A step further

$\longrightarrow$ Make replication files for your analyses available

Something to think about: future proofing

- Package version managers, e.g., `packrat` in R
- Docker container (virtual machine) containing the computational environment and replication files

A possible long-term project

$\longrightarrow$ Use software that is freely available and open source

- Using your method should not require an expensive license (e.g, MATLAB
- Proprietary software can be vague about important computational details of implemented algorithms (e.g., Mplus)

Open Science at EUR

EUR can help you with Open Science:

- Open Science Coordinator at Erasmus Research Services: Antonio Schettino
- Data Steward: Lizette Guzman Ramirez
- Open Science Community Rotterdam: `https://www.openscience-rotterdam.com/`

Big-picture discussion point

$\longrightarrow$ How can we avoid a replication crisis in methodological research?

$\longrightarrow$ Should there be registered reports for simulation studies and benchmark experiments?