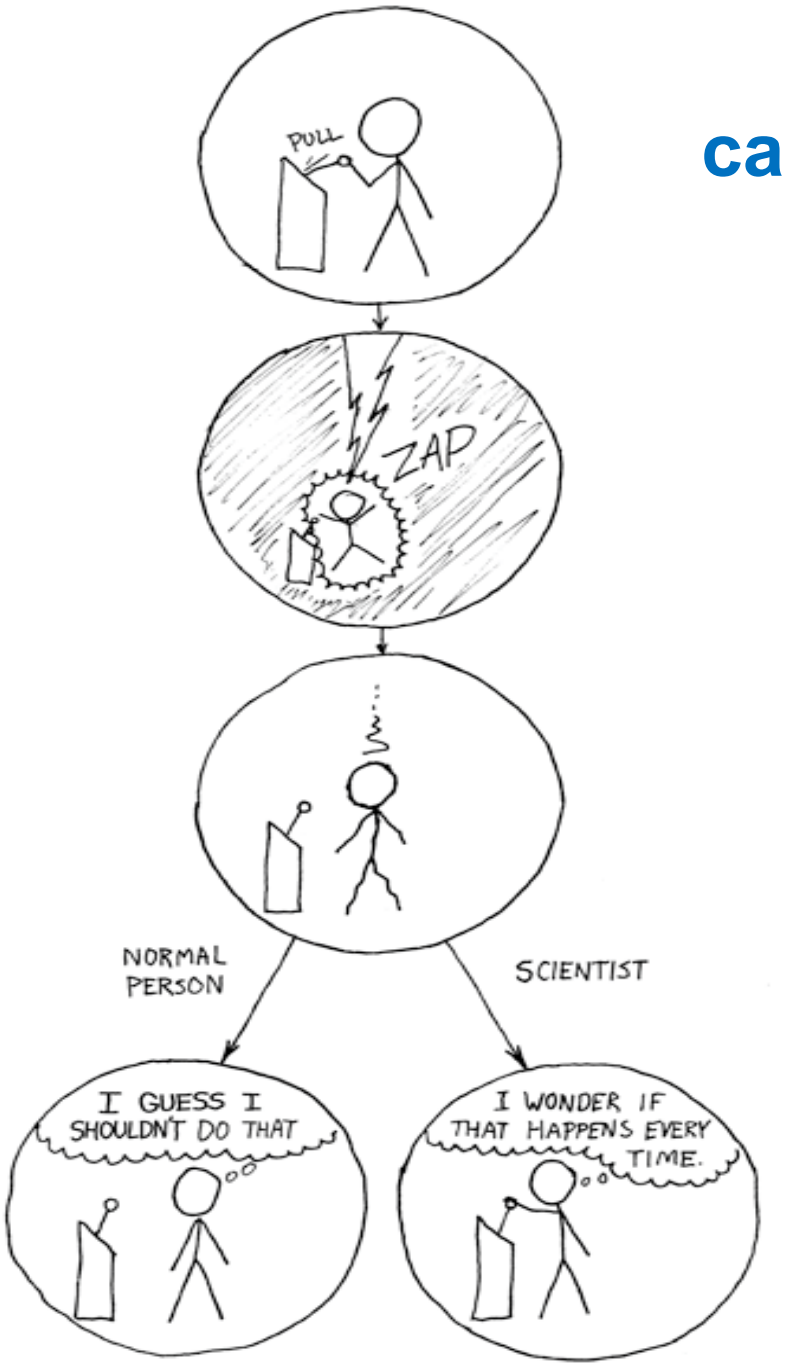


cascad: a trusted third-party to verify research reproducibility

Prof. Christophe Pérignon

HEC PARIS

cascad

October 1, 2021

DataFest

@RSM

cascad

1. What is reproducibility? What is NOT reproducibility?

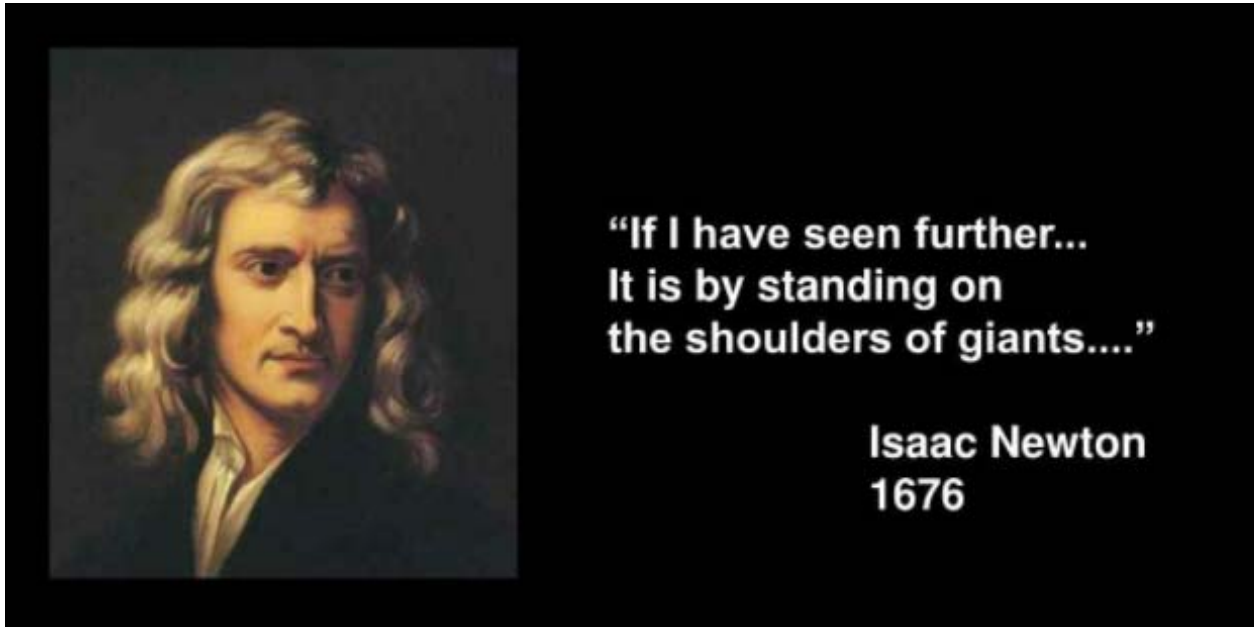
Welcome to the world of Open Science!

Science aims to push the boundary of knowledge.

To be useful, science must be novel, trustful and available

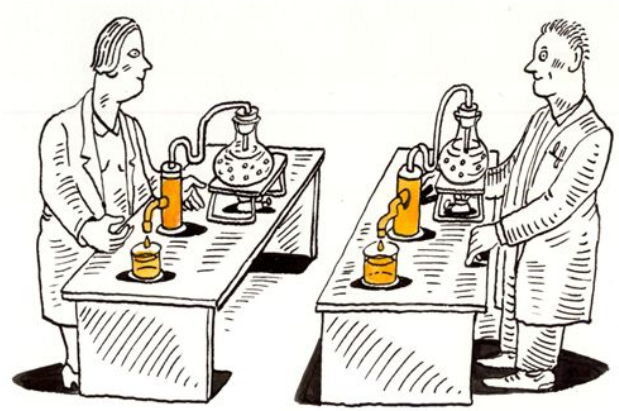
Open science is the movement to make scientific research (publications, data* and code*) accessible to all levels of society:

- researchers
- policy makers
- companies
- general public



cascad

Definition



Results are considered reproducible if:

another researcher, including the original researcher in the future, can regenerate the results from a set of files (code and data) and written instructions (Kitzes, Turek and Deniz, 2017).

cascad

Important distinction

(Computational) Reproducibility is defined as:

same data + same method = same results

Chang and Li (AER 2017) aim to reproduce the results of AER papers from the code and data used & shared by the authors

Replication is defined as:

new data + same method = same results

same data + **new** method = same results

Camerer et al. (Science 2016) evaluate the replicability of lab experiments in economics by re-running them independently

cascad

2. Why do results need to be reproducible?

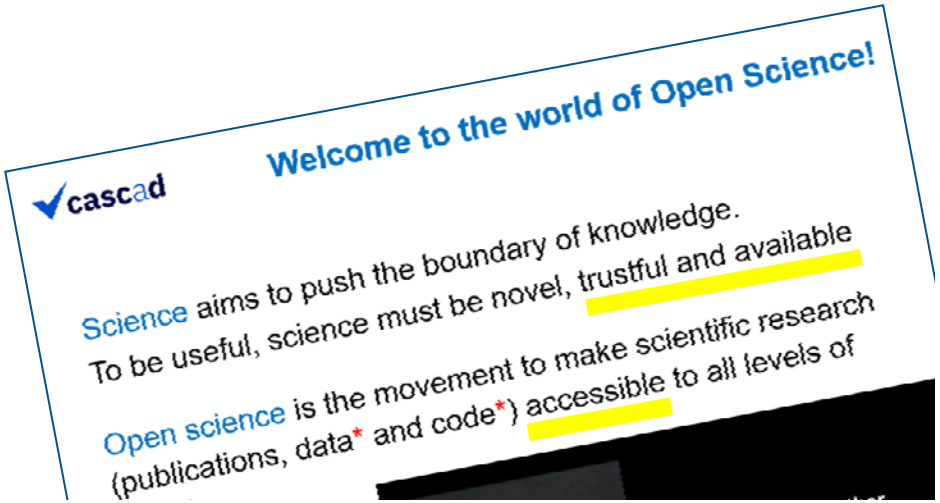
cascad

Why do results need to be reproducible?

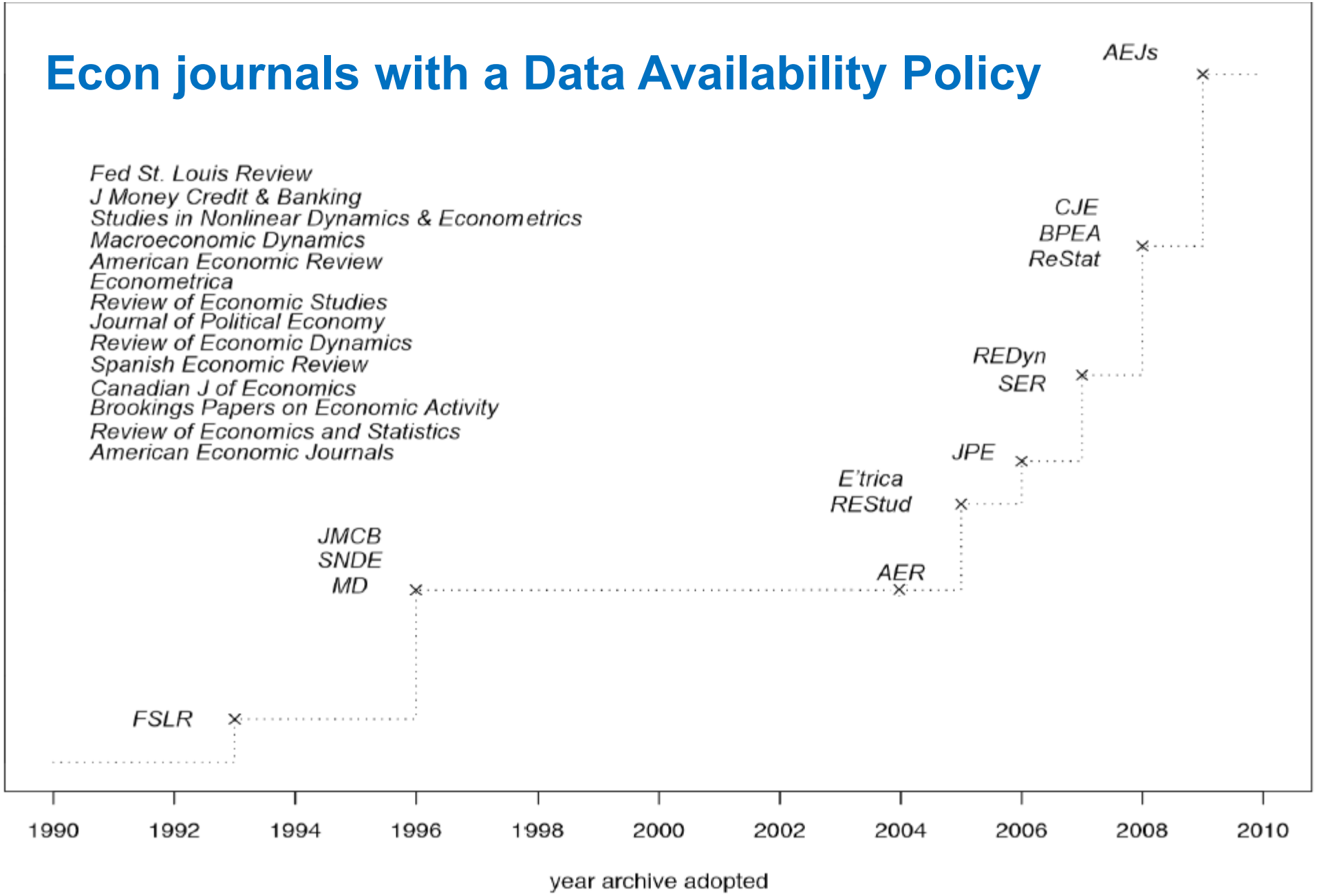
- Serving as a minimum requirement for judging scientific claims
- Facilitating the diffusion of new ideas and replication (less p-hacking)
- Encouraging researchers to exert more effort in their analyses (fewer mistakes)

Why do results need to be reproducible?

- Serving as a minimum requirement for judging scientific claims
- Facilitating the diffusion of new ideas and replication (less p-hacking)
- Encouraging researchers to exert more effort in their analyses (fewer mistakes)

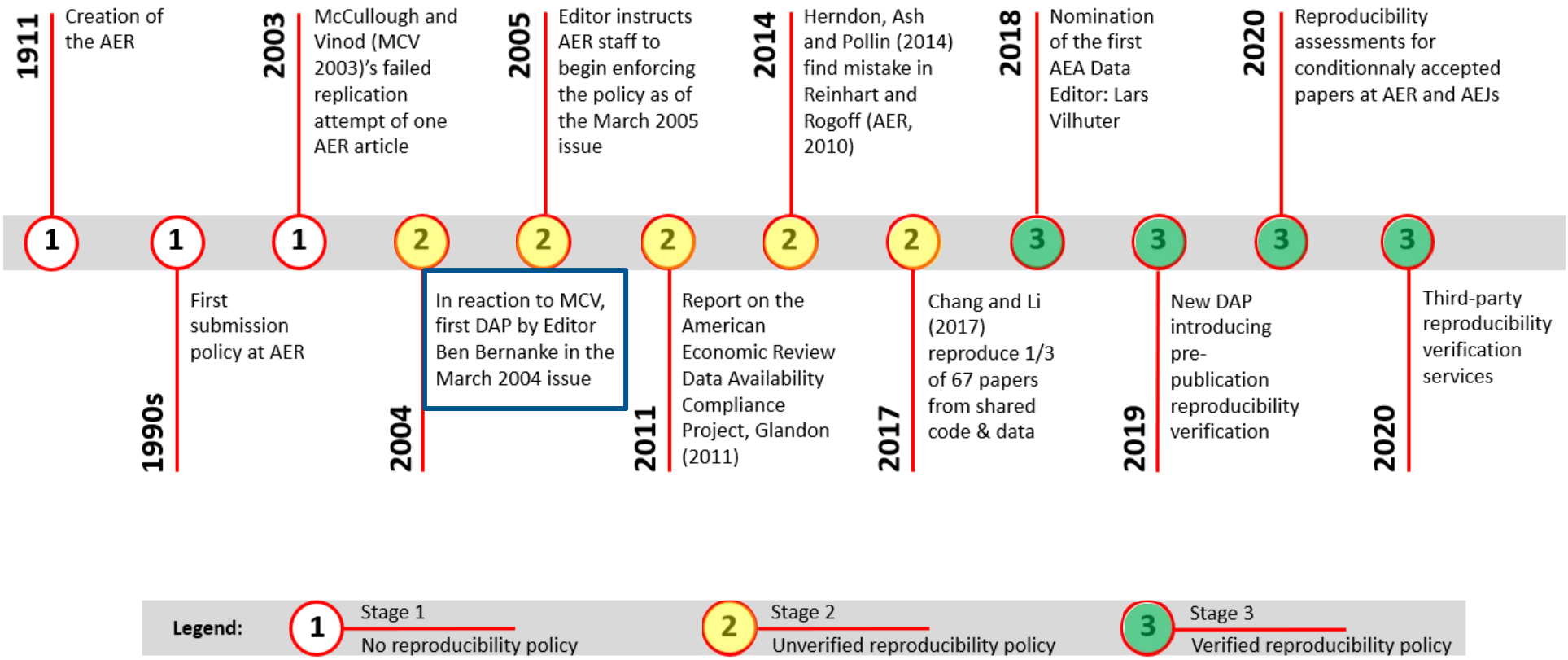


Econ journals with a Data Availability Policy



Source: McCulllough (2009)

Making Research Reproducible: The *American Economic Review* Timeline



Source: Colliard, Hurlin and Pérignon (2021)

cascad

3. What is the current level of reproducibility in economics?

cascad

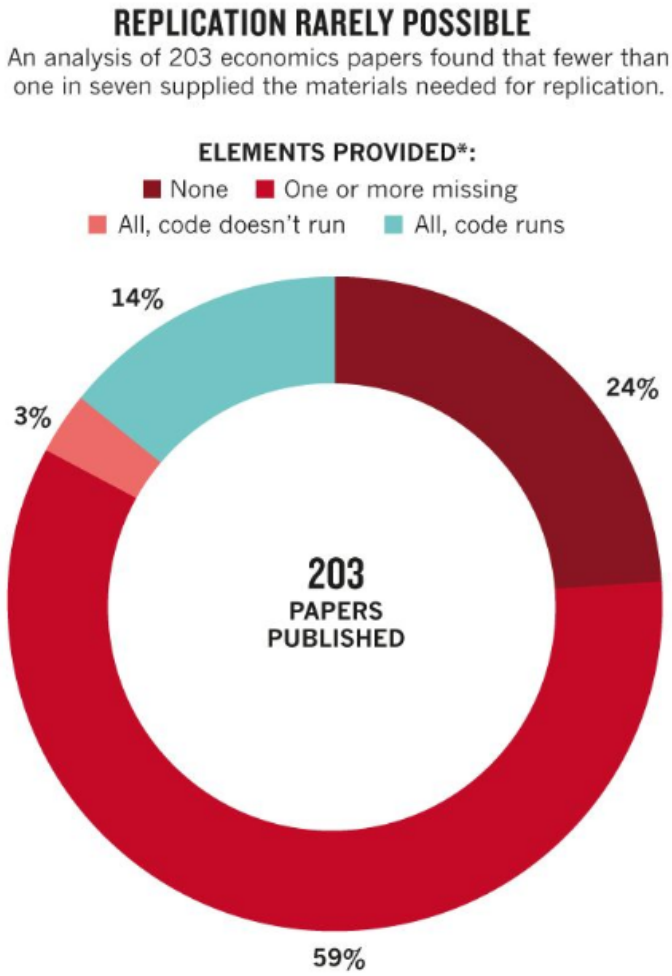
Code and data repositories

Höffler (*AER*, 2017) reports that 54% of a sample of 343 economics journals included in the Thomson Reuters Social Science Citation Index have a Data Availability Policy (DAP)

Vlaeminck and Hermann (*EconStor Open Access*, 2015) show that 84% of their sample journals with a DAP have their own archive website, whereas 14% suggest their authors use external code and data repositories

New Studies: Reproducibility rate is low

Gertler et al. (Nature, 2018)



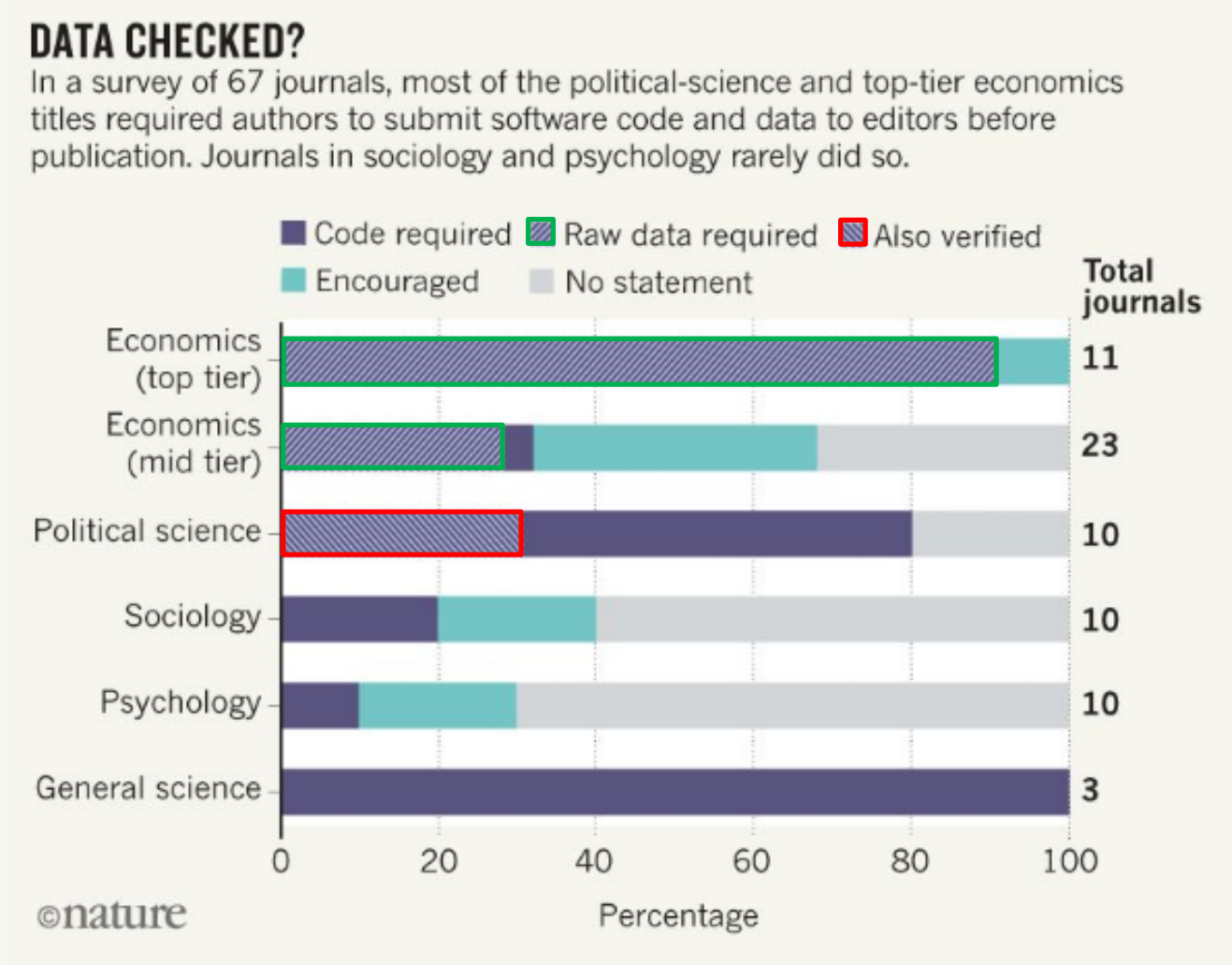
Chang and Li (AER, 2017)

Only able to reproduce the results for one-third of 67 papers published in top economic journals from the code and data available on the journals' repositories.

Another 10% of the papers were reproduced using the help of the authors

cascad

Data are not checked
Code are not run



Source: P. Gertler, S. Galiani & M. Romero (unpublished data)

cascad

4. What can we do to improve the situation?

cascad

Pre-publication verification of the reproducibility

(1) Checking whether the authors have complied with the guidelines of the verificator.

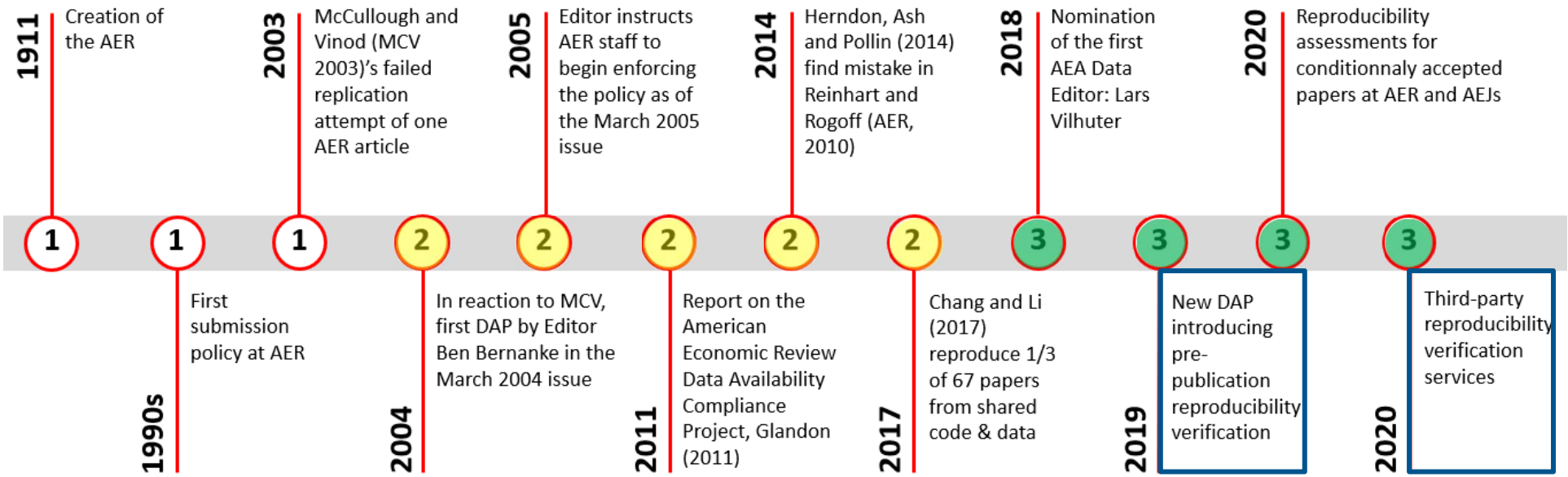
(2) Checking whether the numerical results (tables and figures) of the paper correspond to the numerical results regenerated from the code and data of the authors.

cascad

Who should verify?

- Journals themselves: Regular editors and referees of the journal: may not have the time, skills, and have a conflict of interest
- Journals themselves (2): Dedicated team with reproducibility reviewers, e.g. AER, REStud, EJ
- A trusted third party, e.g. cascad certification agency. Economies of scale

Making Research Reproducible: The *American Economic Review* Timeline

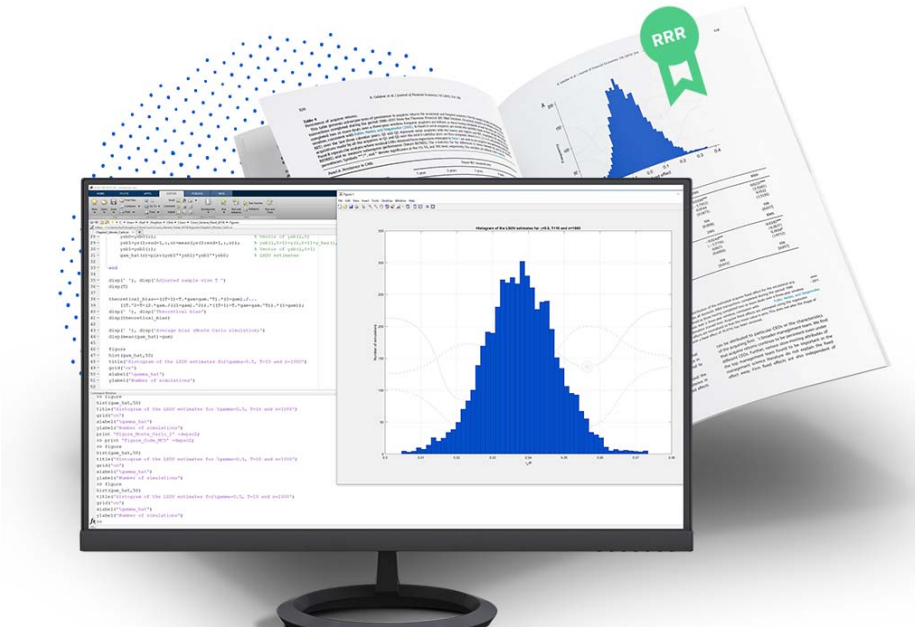


Source: Colliard, Hurlin and Pérignon (2021)

Example of a trusted third party

cascad = Certification Agency for Scientific Code and Data

The cascad reproducibility certification attests that the numerical results reported in a scientific article can be reproduced from a set of numerical resources (code and data) provided by the authors.



- CNRS, HEC Paris, U. Orléans
- 4 staff
- www.cascad.tech
- Founded by researchers for researchers (and journals)

Video

Link:

https://www.youtube.com/watch?v=i17UI2bKh0E&feature=youtu.be

cascad

Author requests a certification

Reproducibility reviewer runs the code with the data

Reviewer compares output with results in the paper

Reproducibility certificate sent to the author

Author sends the Certificate to a journal along with the manuscript

cascad

Execution report

TABLE 4: QUANTILES OF WAGE GROWTH

Original :

	Wage growth quantiles				
	(1)	(2)	(3)	(4)	(5)
ICT_0	-.105***	-.105***	-.107***	-.121***	-.110***
	(.027)	(.018)	(.015)	(.018)	(.028)
Observations	4,972	4,972	4,972	4,972	4,972

Reproduced :

Wage growth quantiles					
	est1	est2	est3	est4	est5
ICT_0	-0.105***	-0.105***	-0.107***	-0.121***	-0.110***
	(0.025)	(0.017)	(0.015)	(0.016)	(0.031)
Observations	4,972	4,972	4,972	4,972	4,972

cascad

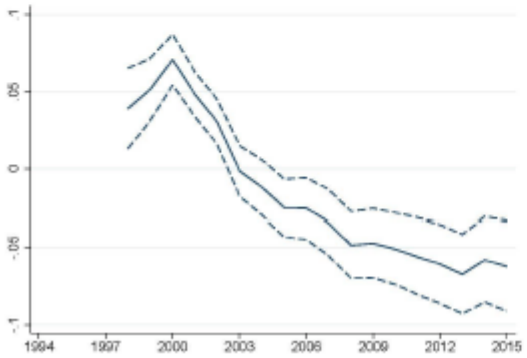
Execution report

4.2. FIGURE 2: WAGE DYNAMICS OF THE ICT BOOM COHORT

The figure displays the t coefficient of the wage regression $log(wi,t) = t + \delta_t ICT_{i,0} + \gamma_t X_i + \varepsilon_{i,t}$ where $ICT_{i,0}$ is a dummy variable equal to one if worker i's first employment spell is in a firm in the ICT sector and X_i collects control variables listed in Section 4.1. Dashed lines represent the 95% confidence interval. The regression is estimated over the cohort of skilled workers whose first full-time job was in 1998-2001.

Original:

Reproduced:

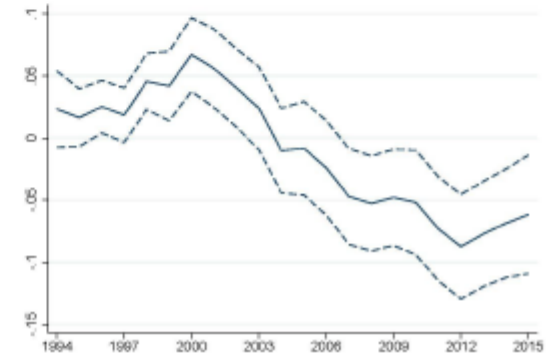


4.3. FIGURE 3: WAGE DYNAMICS OF THE PRE-BOOM COHORT

The figure displays the t coefficient of the wage regression $log(wi,t) = t + \delta_t ICT_{i,0} + \gamma_t X_i + \varepsilon_{i,t}$ where $ICT_{i,0}$ is a dummy variable equal to one if worker i's first employment spell is in a firm in the ICT sector and X_i collects control variables listed in Section 4.1. Dashed lines represent the 95% confidence interval. The regression is estimated over the cohort of skilled workers whose first full-time job was in 1994-1996.

Original:

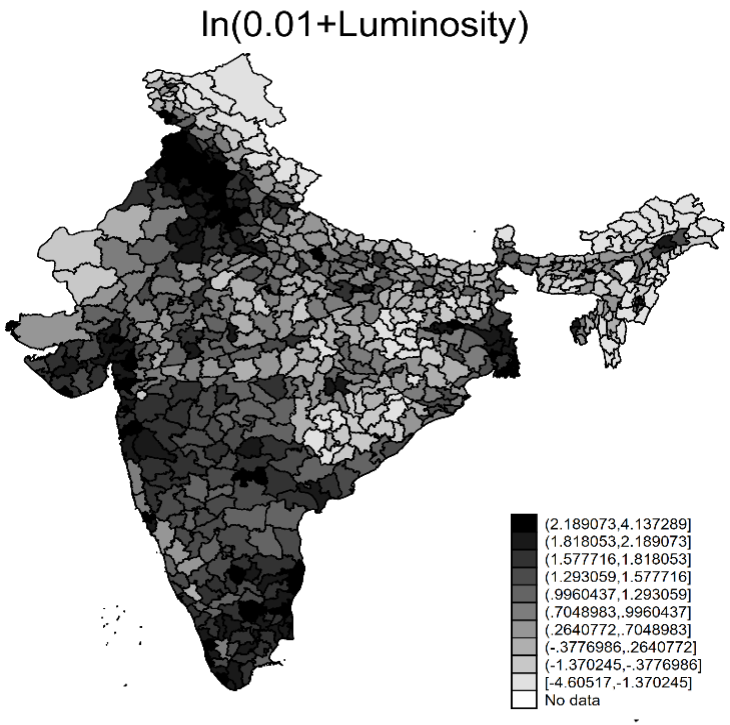
Reproduced:



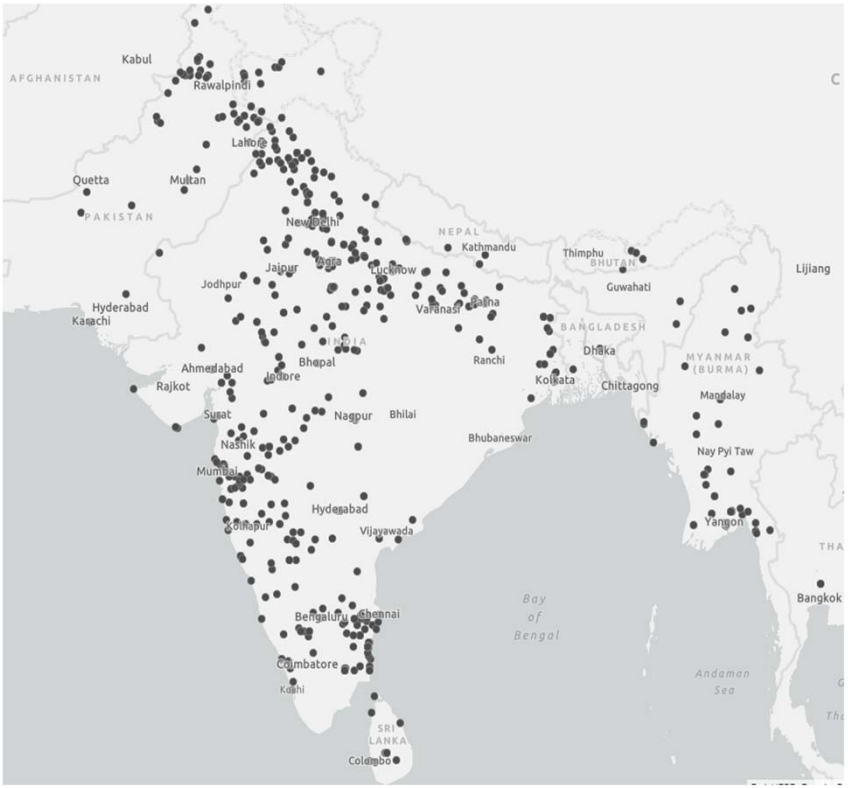
cascad

Execution report

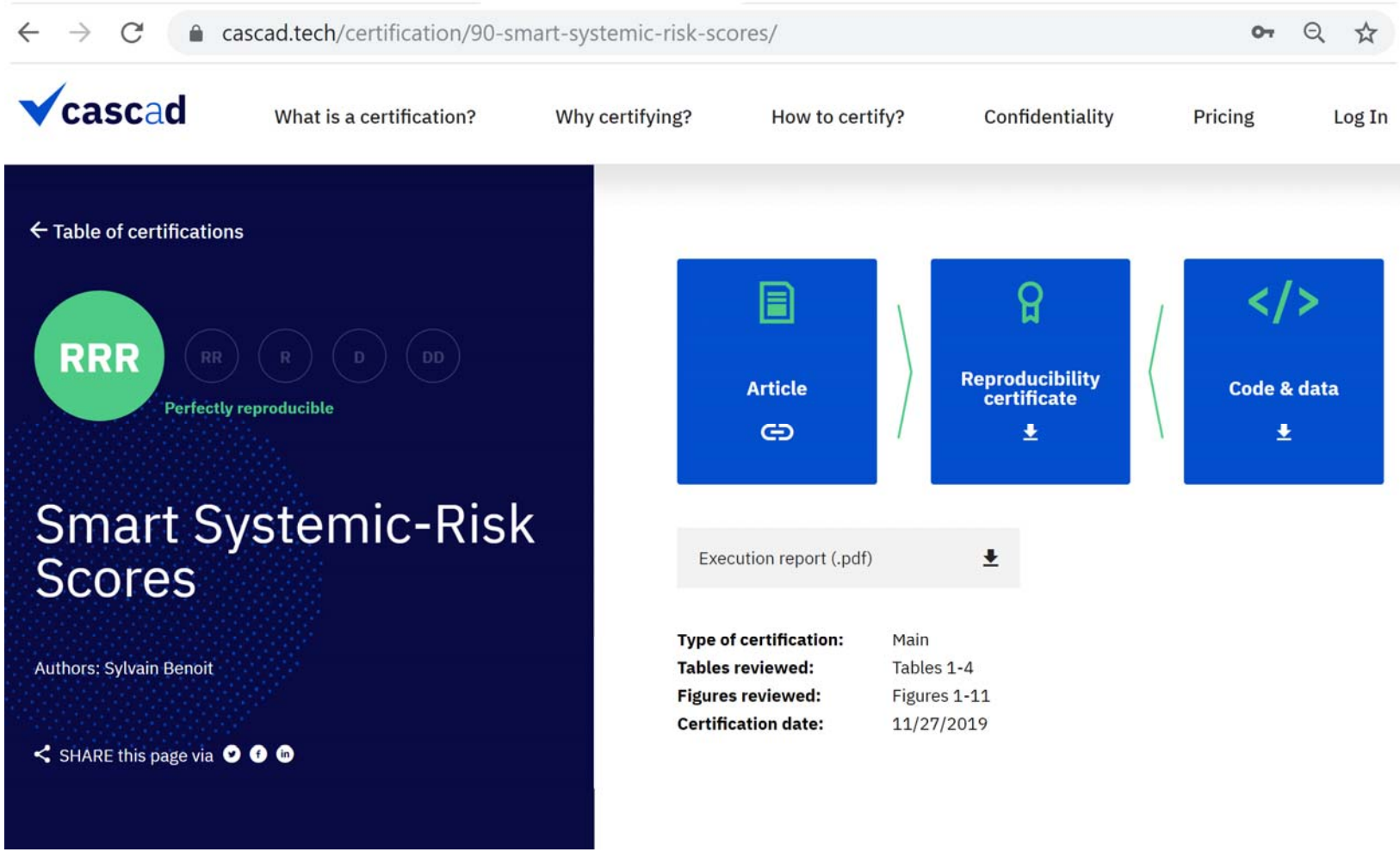
Reproduced :



Reproduced :

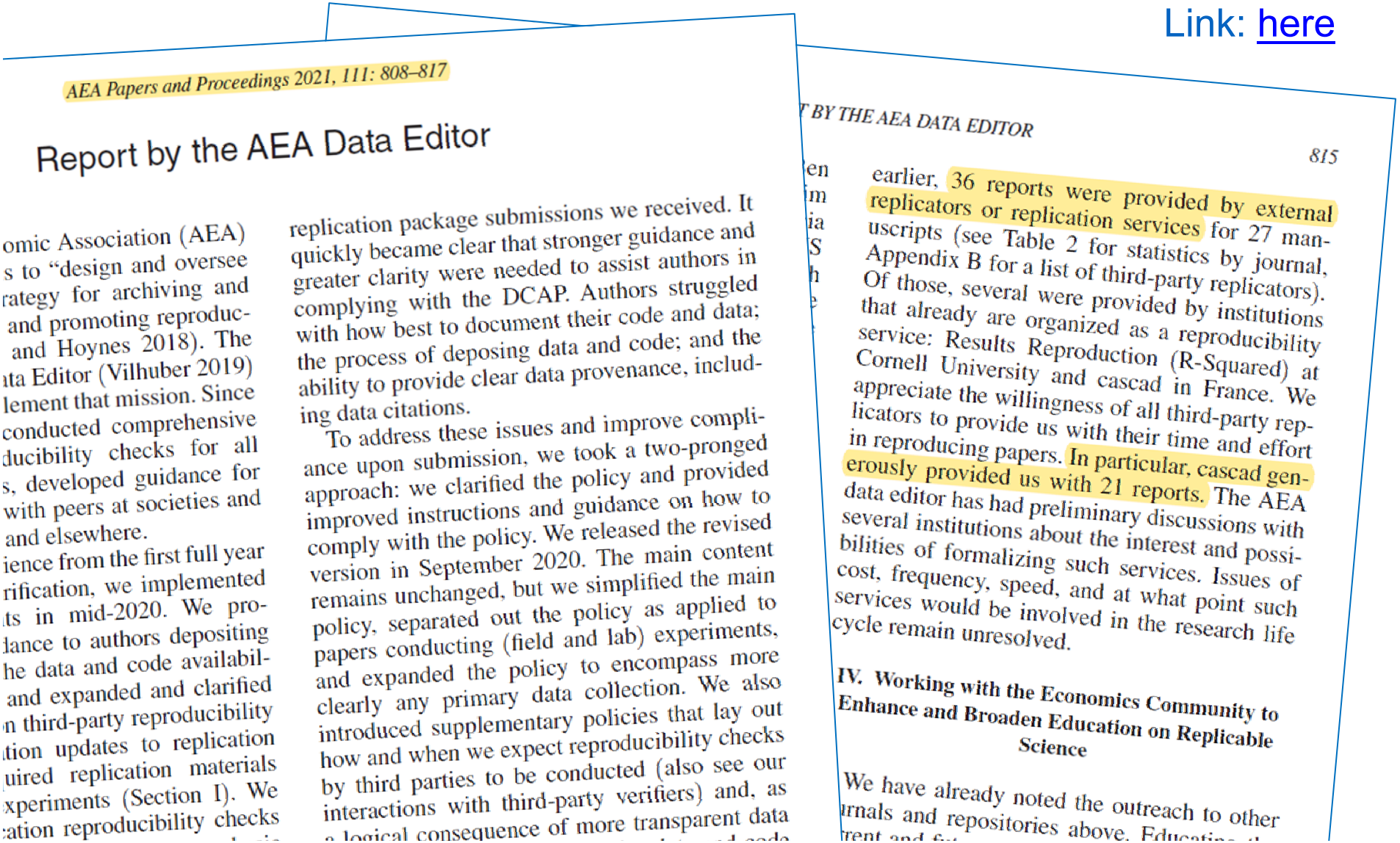


Online certified resources



Collaboration with the American Economic Association (AEA)

Link: here



AEA Papers and Proceedings 2021, 111: 808–817

Report by the AEA Data Editor

omic Association (AEA)
s to "design and oversee
rategy for archiving and
and promoting reproduc-
and Hoynes 2018). The
ıta Editor (Vilhuber 2019)
lement that mission. Since
conducted comprehensive
ducibility checks for all
s, developed guidance for
with peers at societies and
and elsewhere.
ience from the first full year
rification, we implemented
ts in mid-2020. We pro-
lance to authors depositing
he data and code availabil-
and expanded and clarified
n third-party reproducibility
tion updates to replication
uired replication materials
xperiments (Section I). We
ation reproducibility checks

replication package submissions we received. It quickly became clear that stronger guidance and greater clarity were needed to assist authors in complying with the DCAP. Authors struggled with how best to document their code and data; the process of deposing data and code; and the ability to provide clear data provenance, including data citations.

To address these issues and improve compliance upon submission, we took a two-pronged approach: we clarified the policy and provided improved instructions and guidance on how to comply with the policy. We released the revised version in September 2020. The main content remains unchanged, but we simplified the main policy, separated out the policy as applied to papers conducting (field and lab) experiments, and expanded the policy to encompass more clearly any primary data collection. We also introduced supplementary policies that lay out how and when we expect reproducibility checks by third parties to be conducted (also see our interactions with third-party verifiers) and, as a logical consequence of more transparent data

en
im
ia
S
h

earlier, 36 reports were provided by external replicators or replication services for 27 manuscripts (see Table 2 for statistics by journal, Appendix B for a list of third-party replicators). Of those, several were provided by institutions that already are organized as a reproducibility service: Results Reproduction (R-Squared) at Cornell University and cascad in France. We appreciate the willingness of all third-party replicators to provide us with their time and effort in reproducing papers. In particular, cascad generously provided us with 21 reports. The AEA data editor has had preliminary discussions with several institutions about the interest and possibilities of formalizing such services. Issues of cost, frequency, speed, and at what point such services would be involved in the research life cycle remain unresolved.

IV. Working with the Economics Community to Enhance and Broaden Education on Replicable Science

We have already noted the outreach to other urnals and repositories above. Educating

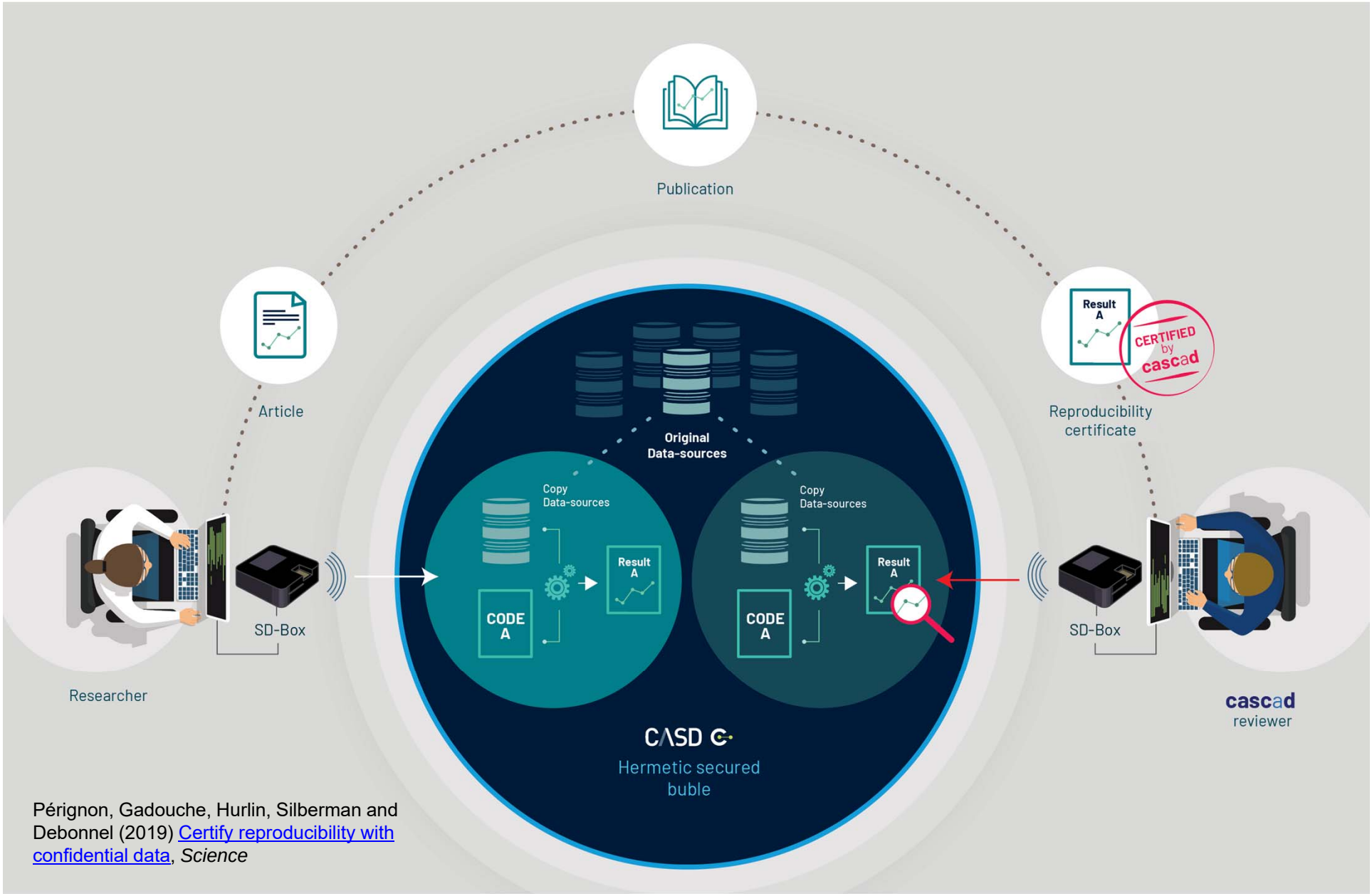
CASD – cascad partnership

CASD is a restricted data access center allowing researchers to access confidential data (tax, health) on virtual machines

CASD is a certified health data host (ISO 27001, GDPR, HDS): hospital data (PMSI), cohorts (Constances)

CASD-cascad partnership allows research based on confidential data to be reproduded and certified

Partnership with CASD



Pérignon, Gadouche, Hurlin, Silberman and Debonnel (2019) Certify reproducibility with confidential data, *Science*

Extra reading on research reproducibility

Date:
December 2020

Authors:
Baker, Lee
Lusoli, Wainer
Jaśko, Katarzyna
Parry, Vivienne
Pérignon, Christophe
Errington, Timothy
Cristea, Ioana Alina
Winchester, Catherine
MacCallum, Catriona
Šimko, Tibor

Link: here

Thank you for your attention!

Contact:

Christophe Pérignon (perignon@hec.fr, website)

cascad Certification Agency