January 4, 2022 - cascad#214

Execution Report

Title: **The Economics of Research Reproducibility**
Authors: **Jean-Edouard Colliard, Christophe Hurlin, Christophe Pérignon**

Full reference: Colliard, Jean-Edouard, Hurlin, Christophe, and Pérignon, Christophe, "The Economics of Research Reproducibility", January 4, 2022.

The structure and contents of this execution report provided by **cascad** for the certification are similar to those recommended by the AEA Data Editor.

1. DATA DESCRIPTION

For the purpose of this study, data on authors' decisions to voluntarily make their paper reproducible were collected. The researchers used all papers published in the Journal of Financial Economics between January 2010 and September 2020, which corresponds to a total number of 1,347 papers written by 2,231 authors. For a thorough description of the data, please refer to section 1.2.2 of the paper.

2. CODE DESCRIPTION

For the purpose of this certification, we aimed to check the results displayed in Table 1 and in Figures 2 and 4. Note that Figures 1 and 3 are simply images and contain no results.

The Code is written in SAS and Mathematica. The replication materials contain one data file, *JFE data and code.xlsx*, that is first called by *Database Management.sas* to generate an Excel file, *Data_JFE.xslx*. This file is then used by *Modelisation.sas* to generate Table 1. Meanwhile, two Mathematica notebooks, *Graph_2.nb* and *Graph_4.nb*, generate their respective figure.

3. REPLICATION STEPS

The resources were downloaded from the **cascad** website and run as per readme, using SAS 9.4 on a computer with 64GB RAM, intel® Core™ i9-9900K CPU @3.60-5.00GHz, Nvidia Geforce RTX 2060, and Windows 10 OS.

We encountered no issues during the replication.

4. FINDINGS

We reproduced the table and the two figures with perfect accuracy.

4.1 TABLE 1: THE DRIVERS OF CODE AND DATA SHARING: EVIDENCE FROM THE JOURNAL OF FINANCIAL ECONOMICS.

Original:

Panel A: Summary Statistics

	Mean	Std-dev	Min	Q1	Median	Q3	Max	Obs.
International	0.318	0.466	0	0	0	1	1	1,347
Top 10	0.232	0.422	0	0	0	0	1	1,347
Citations	98.70	226.67	0	17.00	41.42	112.33	3,756.50	544
Seniority	9.14	6.62	-4	4.67	8.50	12.67	45.50	544

Panel B: Logistic Regression

	(1)	(2)	(3)	(4)	(5)	(6)
International	0.5089	0.5306	0.4470	0.4609		
	(0.0658)	(0.0579)	(0.0857)	(0.0817)		
Top 10	0.0197	0.0015	-0.0386	-0.0237		
	(0.9527)	(0.9965)	(0.9023)	(0.9406)		
Log Citations	0.1306	0.1210			0.0793	0.0797
	(0.3334)	(0.3769)			(0.3958)	(0.3951)
Seniority	-0.0098	-0.0061				
	(0.7354)	(0.8348)				
Year FE	No	Yes	No	Yes	No	Yes
Obs.	544	544	1,347	1,347	544	544

Reproduced:

Panel A: Summary statistics:

PANEL A - line 1 and 2

Variable	Label	Mean	Std Dev	Minimum	Lower Quartile	Median	Upper Quartile	Maximum	N
international	international	0.318	0.466	0	0	0	1	1	1347
Top10	Top10	0.232	0.422	0	0	0	0	1	1347

PANEL A - line 3 and 4

Variable	Label	Mean	Std Dev	Minimum	Lower Quartile	Median	Upper Quartile	Maximum	N
mean_citations	mean_citations	98.7	226.67	0	17	41.42	112.33	3756.5	544
mean_seniority	mean_seniority	9.14	6.62	-4	4.67	8.5	12.67	45.5	544

Panel B: Logistic regression:

Column 1
544 individuals

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-2.563	0.4237	36.5868	<.0001
Top10	1	1	0.0197	0.3322	0.0035	0.9527
international	1	1	0.5089	0.2765	3.3858	0.0658
log_citations		1	0.1306	0.135	0.9358	0.3334
mean_seniority		1	-0.0098	0.029	0.1142	0.7354

Column 2
544 individuals

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-2.6957	1.1229	5.7632	0.0164
Top10	1	1	0.00146	0.3354	0	0.9965
international	1	1	0.5306	0.2798	3.5953	0.0579
log_citations		1	0.121	0.1369	0.7808	0.3769
mean_seniority		1	-0.00614	0.0294	0.0435	0.8348

Column 3
1347 individuals

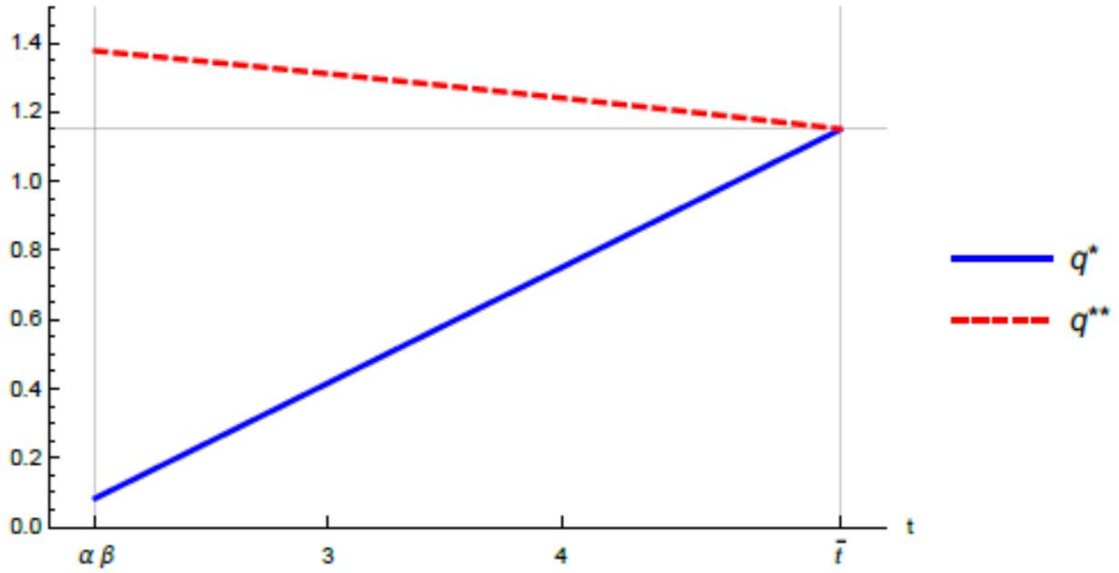
Analysis of Maximum Likelihood Estimates						
Parameter		**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr > ChiSq**
Intercept		1	-3.1044	0.1855	279.9371	<.0001
Top10	**1**	1	-0.0386	0.3143	0.0151	0.9023
international	**1**	1	0.447	0.2602	2.9526	0.0857

Column 4
1347 individuals

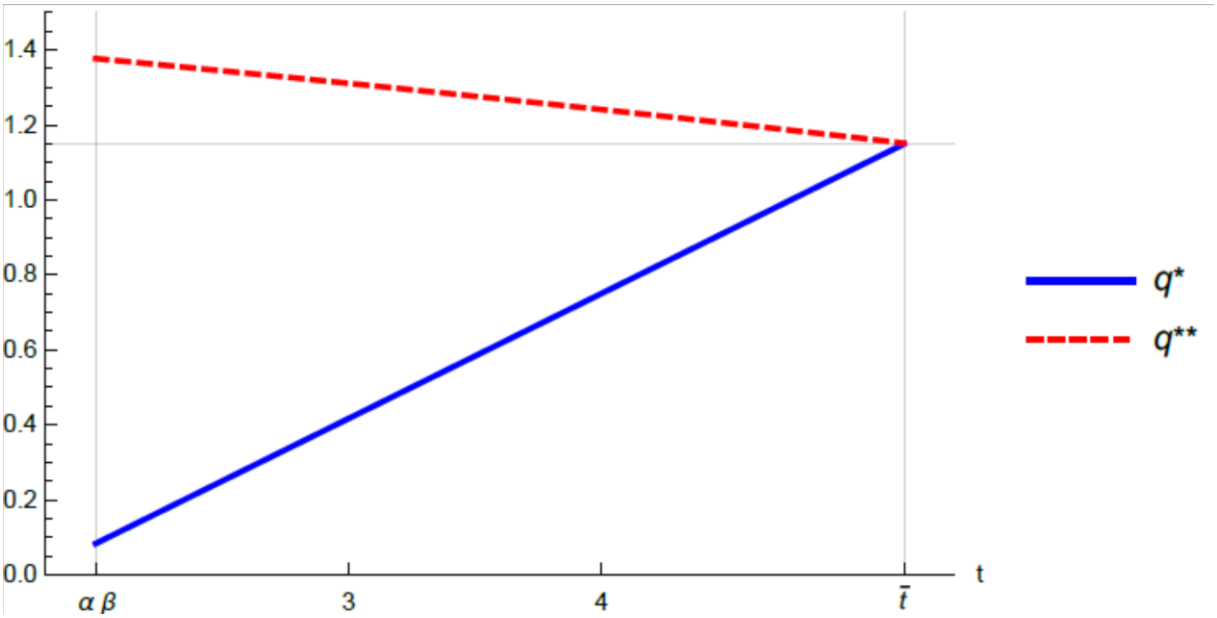
Analysis of Maximum Likelihood Estimates						
Parameter		**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr > ChiSq**
Intercept		1	-4.724	1.0122	21.7796	<.0001
Top10	**1**	1	-0.0237	0.3178	0.0055	0.9406
international	**1**	1	0.4609	0.2648	3.0299	0.0817

4.2 FIGURE 2: REPRODUCIBILITY LEVEL UNDER COMPETITION (Q*) AND UNDER A SOCIAL PLANNER (Q**)

Original:

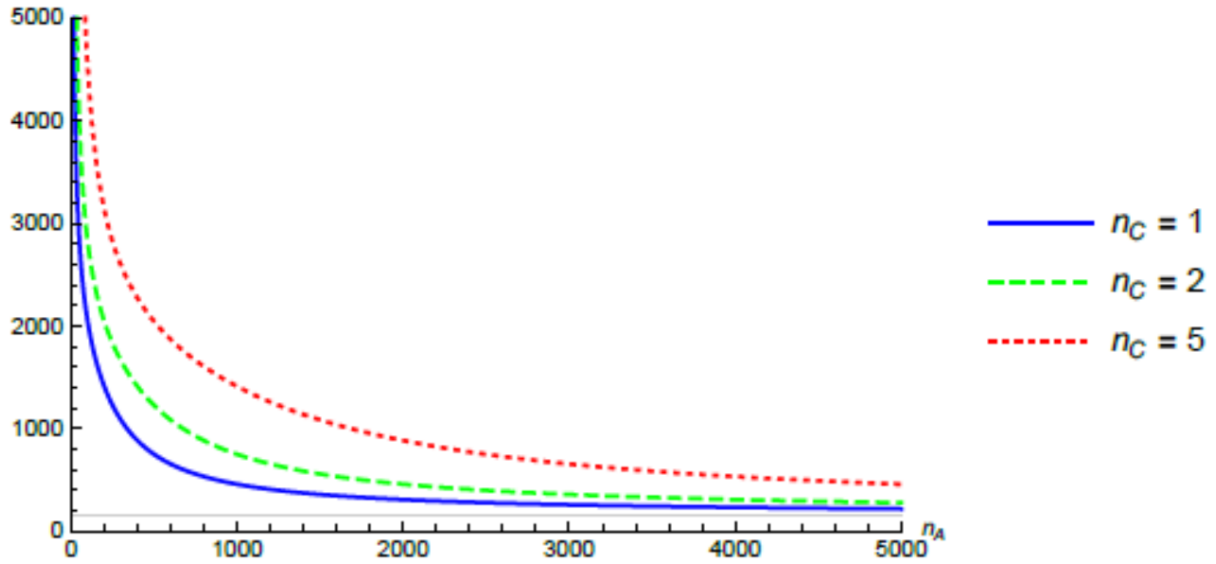
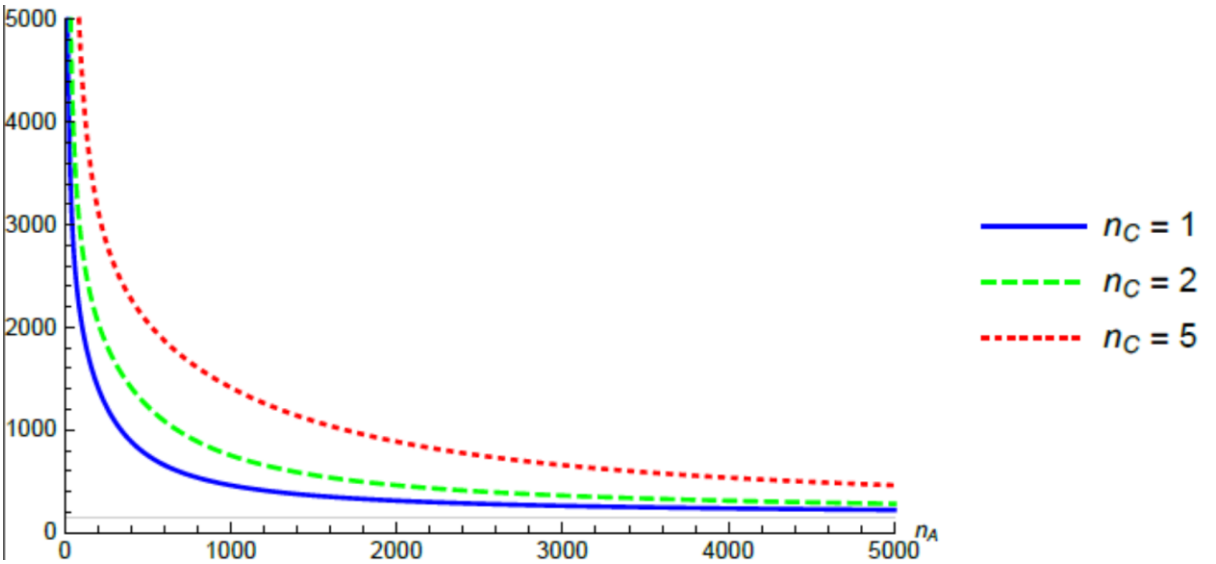


Reproduced:



4.3 FIGURE 4: AVERAGE COSTS PER ARTICLE, AS A FUNCTION OF THE NUMBER OF ARTICLES NA AND NUMBER OF VERIFICATION TEAMS NC

Original:



Reproduced: