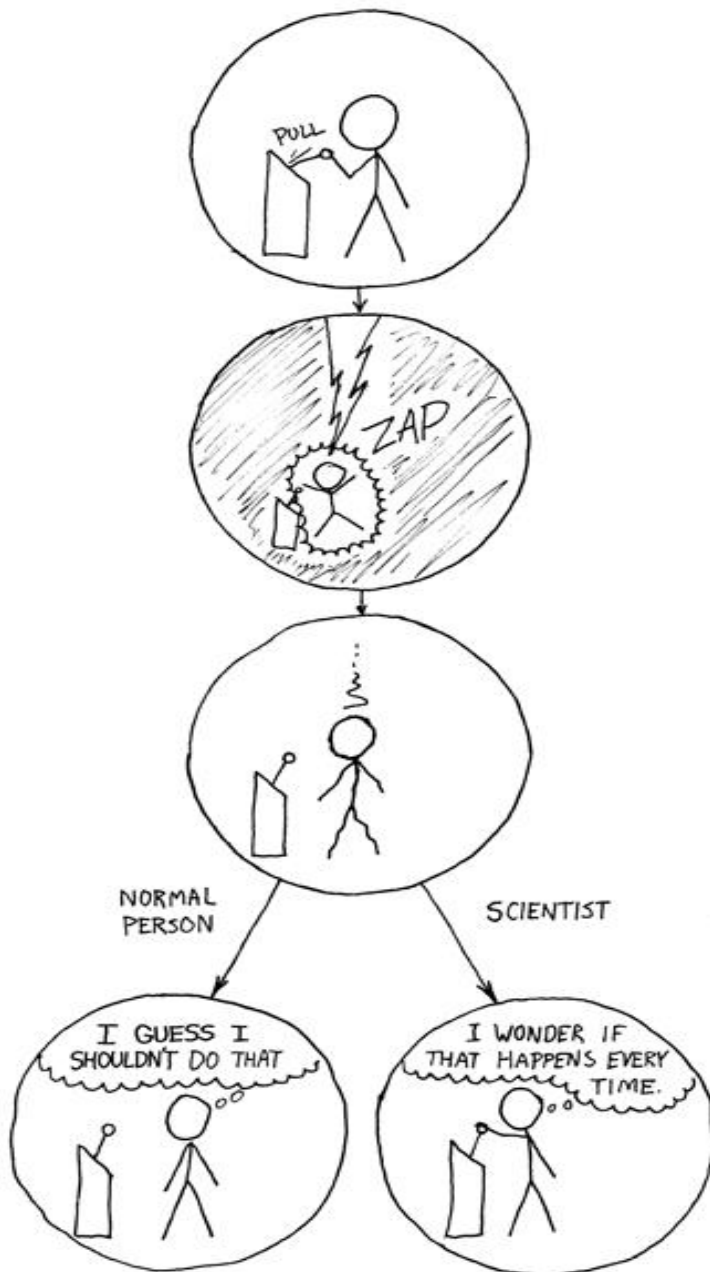


# Certifying Research Reproducibility

Prof. Christophe Pérignon

HEC Paris

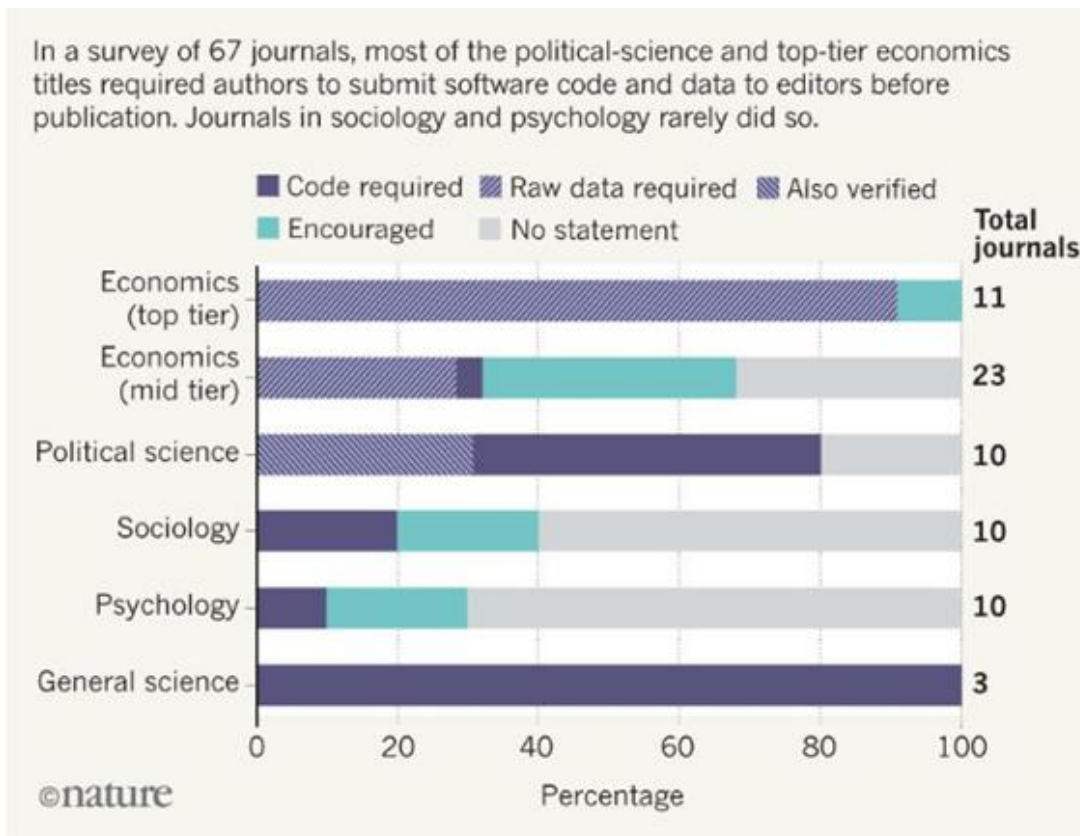


December 17, 2019

Reproducible Research and Modern Data  
Analysis: Concepts, Skills and Tools

INEXDA workshop, Porto

# Why academic research is not always reproducible



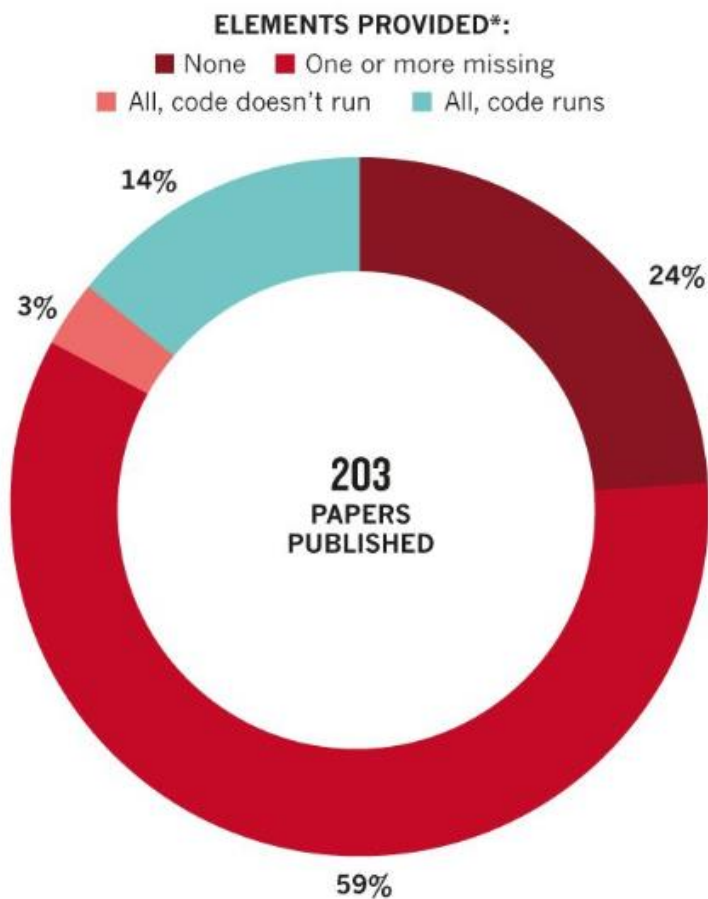
Source: P. Gertler, S. Galiani & M. Romero (unpublished data)

Numerical resources not always requested by journals  
 Numerical resources not systematically checked

# Why academic research is not always reproducible (2)

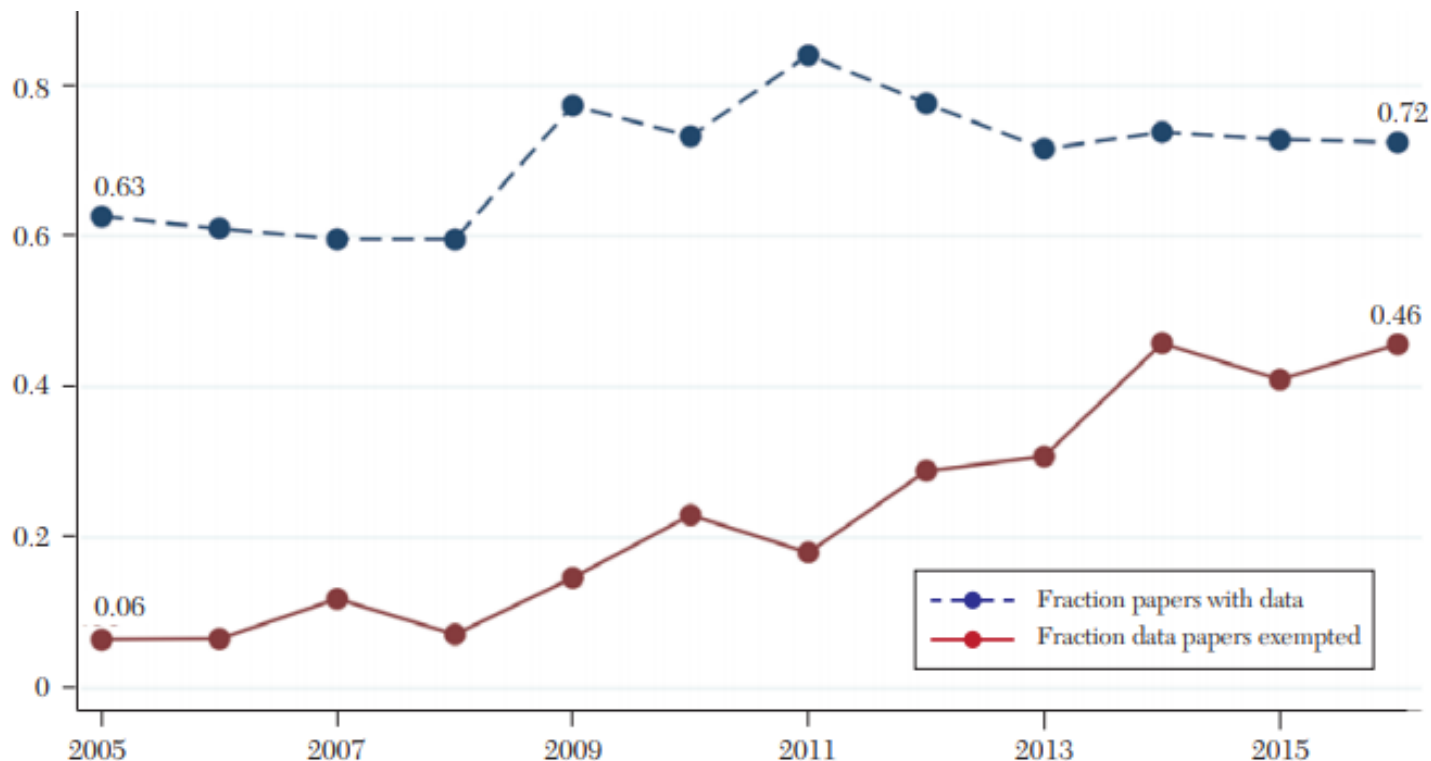
## REPLICATION RARELY POSSIBLE

An analysis of 203 economics papers found that fewer than one in seven supplied the materials needed for replication.



Many missing or non-functioning numerical resources

# Why academic research is not always reproducible (3)



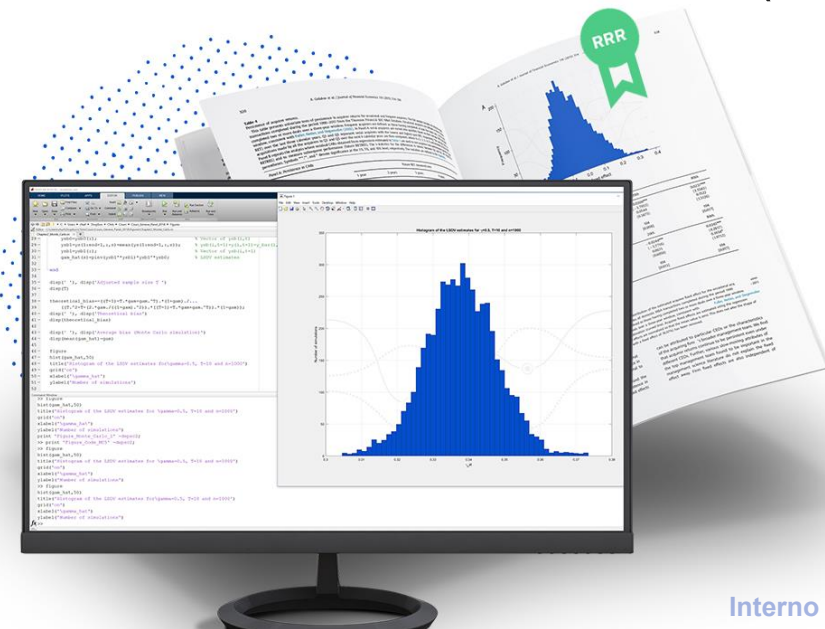
Source: Christensen and Miguel (2018)

Numerical resources cannot often be shared for confidentiality reasons

- **Funding agencies:** no disclosure/no money policy (Data Management Plans)
- **Data providers:** policy to facilitate reproducibility (CASD, ONS, central banks)
- **Academic journals:** pre-publication reproducibility reviews (*American Economic Review*)
- **External reproducibility certification:** pre-publication reproducibility reviews by a trusted third-party  
also works with confidential data

**cascad** = Certification Agency for Scientific Code and Data

- The cascad reproducibility certification attests that the numerical results (tables and figures) reported in a scientific paper can be reproduced using a set of numerical resources (code and data) used by the authors



- [www.cascad.tech](http://www.cascad.tech)
- Non-profit academic initiative
- HEC Paris & U. Orléans with the support of CNRS

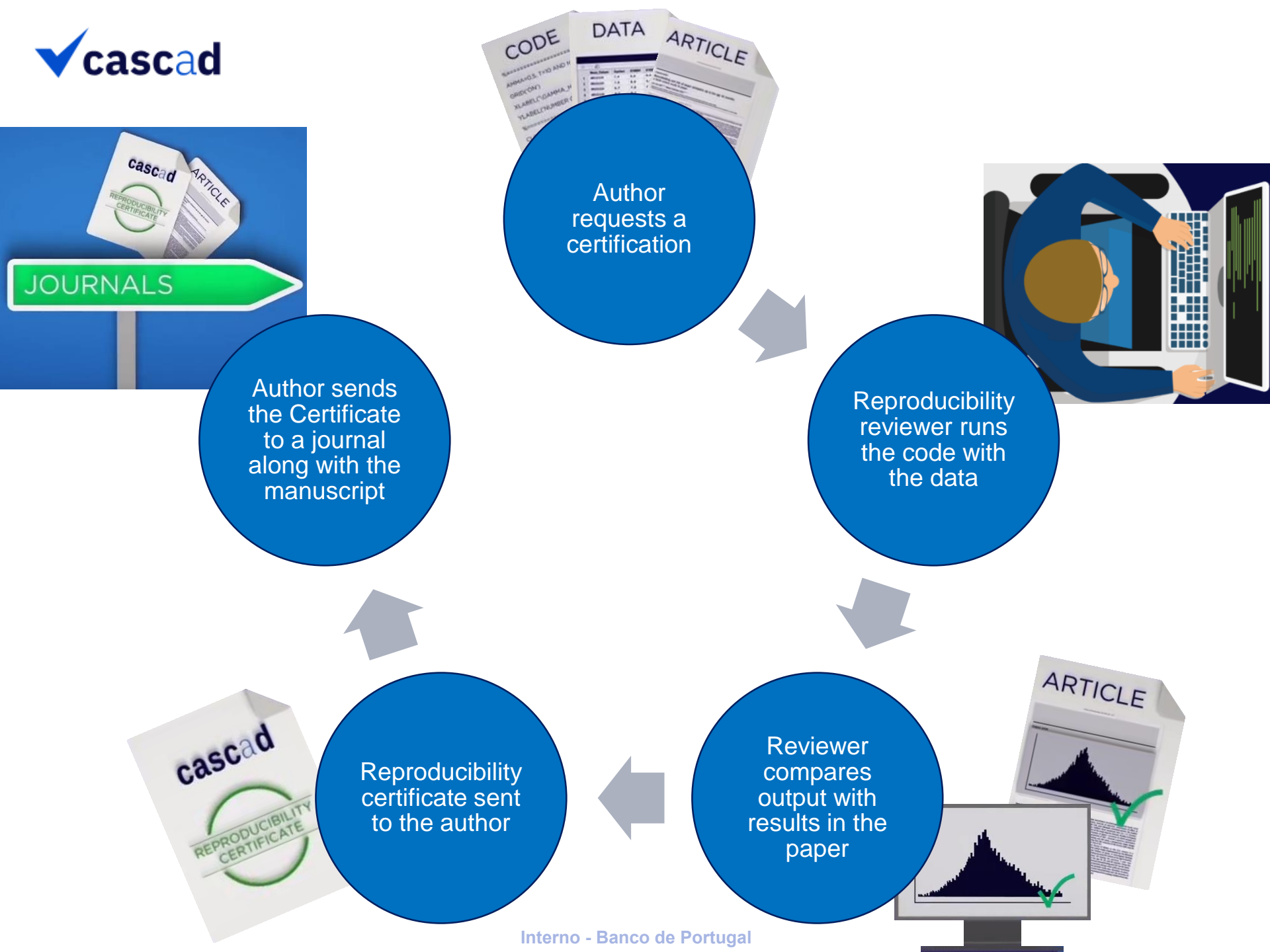




TABLE 4: QUANTILES OF WAGE GROWTH

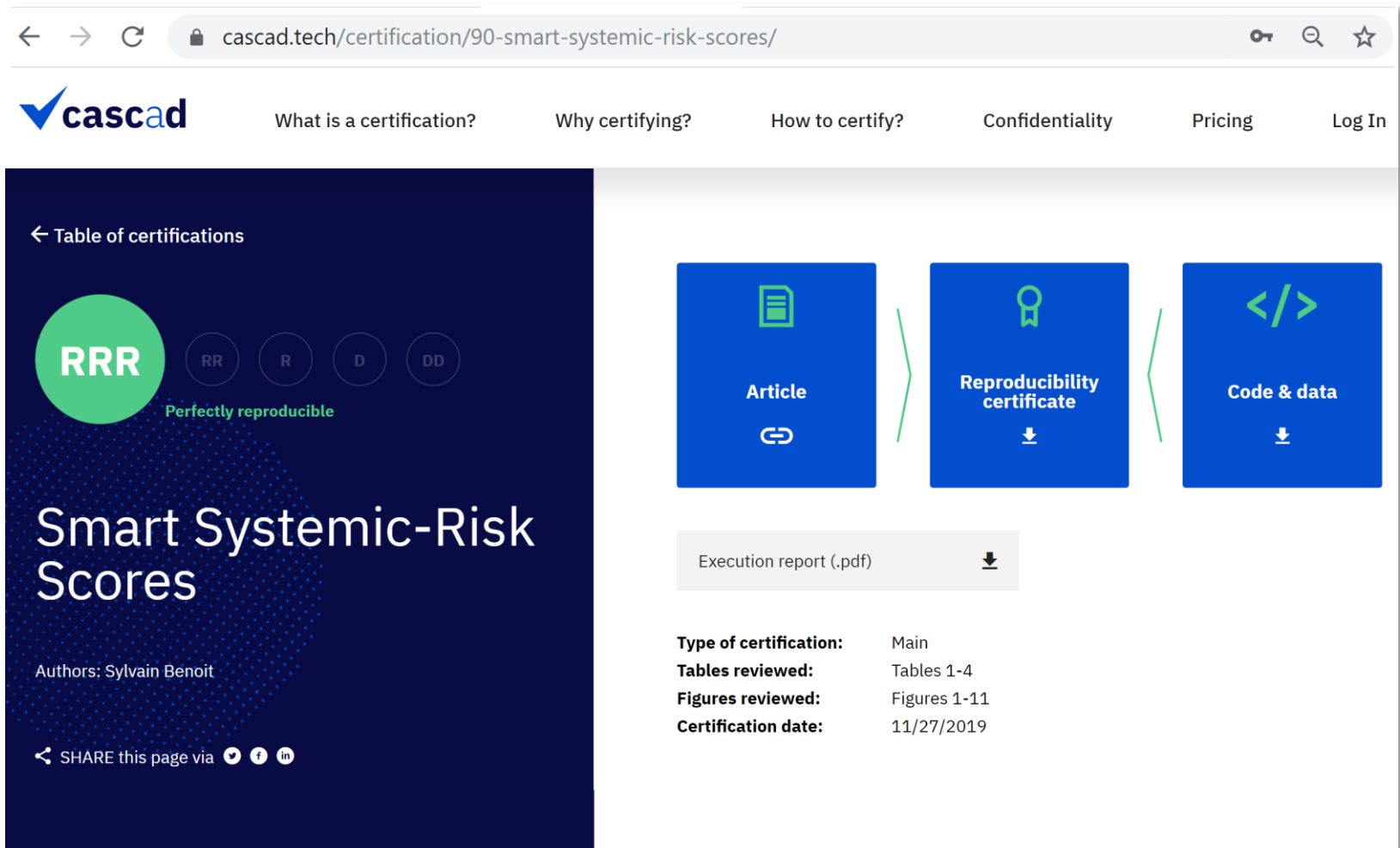
Original :

	Wage growth quantiles				
	(1)	(2)	(3)	(4)	(5)
ICT <sub>0</sub>	-.105*** (.027)	-.105*** (.018)	-.107*** (.015)	-.121*** (.018)	-.110*** (.028)
Worker Controls	✓	✓	✓	✓	✓
Observations	4,972	4,972	4,972	4,972	4,972

Reproduced :

Wage growth quantiles					
	est1	est2	est3	est4	est5
ICT <sub>0</sub>	-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





The screenshot shows a web browser at the URL [cascad.tech/certification/90-smart-systemic-risk-scores/](https://cascad.tech/certification/90-smart-systemic-risk-scores/). The page features a dark blue sidebar on the left with the cascad logo and navigation links: "What is a certification?", "Why certifying?", "How to certify?", "Confidentiality", "Pricing", and "Log In". The main content area has a header "← Table of certifications" and a row of five circular icons labeled "RRR", "RR", "R", "D", and "DD". The "RRR" icon is highlighted in green and labeled "Perfectly reproducible". Below this, the title "Smart Systemic-Risk Scores" is displayed, followed by "Authors: Sylvain Benoit" and a share button. The right side of the page shows three blue boxes with icons and labels: "Article" (document icon), "Reproducibility certificate" (certificate icon), and "Code & data" (code icon). Each box has a download icon. Below these is a grey box for "Execution report (.pdf)" with a download icon. At the bottom, a table provides certification details.

<b>Type of certification:</b>	Main
<b>Tables reviewed:</b>	Tables 1-4
<b>Figures reviewed:</b>	Figures 1-11
<b>Certification date:</b>	11/27/2019

Resources hosted on ICPSR, ZENODO (versioning)

