Has the 2008 Global Financial Crisis a lasting impact on universities and public research institutes in the EU?

Marc Luwel and Thed van Leeuwen
CWTS, Leiden University

2018 Nordic Workshop on Bibliometrics and Science Policy
Borås, Sweden, November 8-9, 2018

Introduction

- Global Financial Crisis (GFC; 2007-2008) had and still has a huge impact on the public finances of (EU) countries.
- GDP contracted between 2008 and 2014, e.g. GDP of EU19 in constant prices was in 2014 still 0.6% lower compared with 2008
- Some EU countries (IE, GR, PT, ES) were more severely affected than others
- Public funding of research and innovation was affected
- In the EU universities are largely publicly funded

CWTS

Research questions

Had/s the GFC a (lasting) impact on EU countries' research capacity?

Operationalization

1. What is the evolution of the publication output of universities and publicly funded research institutes?

2. What is the impact of the changes in public research funding on the publication output of these organizations ?

CWTS

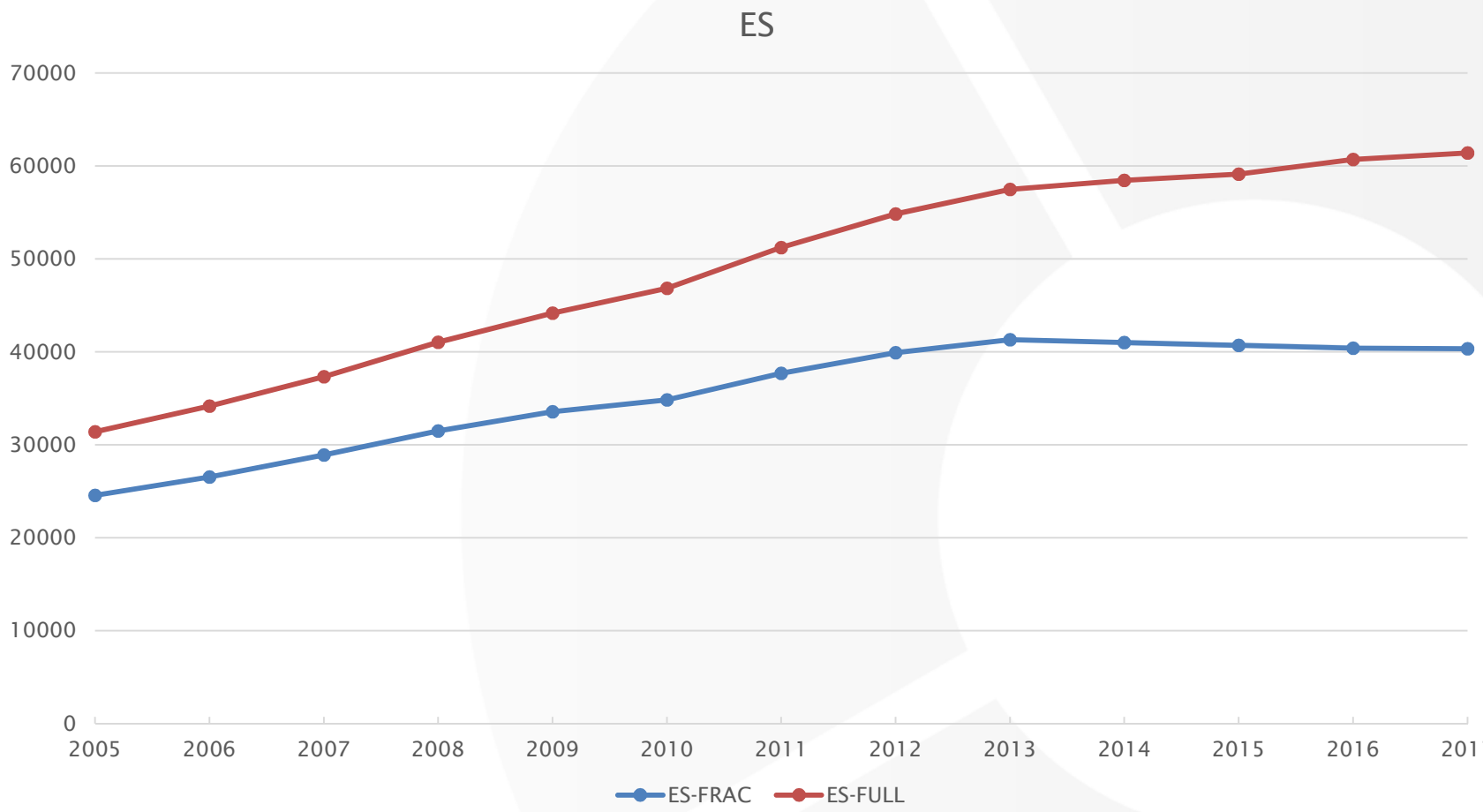
Evolution of the publication output

Publication data

- Publication output: WoS data
  - Publication types: A, R, L (weight 0.25)
  - Full and Fractional (country level) counting scheme
- Reference period: 2005-2017

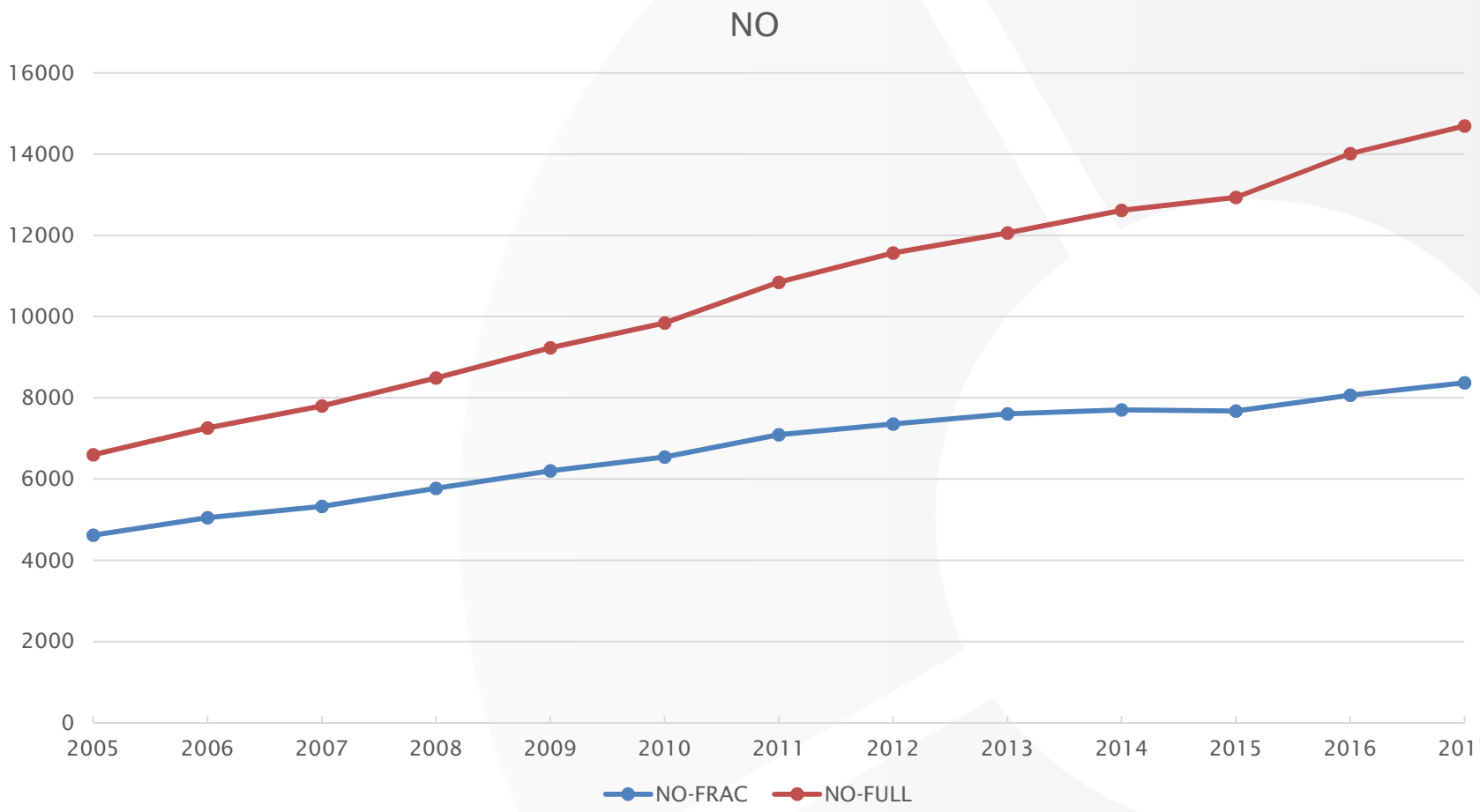
CWTS

Spain – Full / Fractional WoS publication data for [2005-2017]



CWTS

NORWAY – Full / Fractional WoS publication data for [2005-2017]



CWTS

Breakpoint in publication output – Chow breakpoint test - theory

Question:

Breakpoint in time series around 2012-2013

Are the changes in the slope statistically significant ?

CHOW TEST:

Visual inspection – fairly linear trends (with breakpoint)

Pool sample and 2 distinct subsamples (2005-2011/2012 and 2012/2013-2017)

Regression – check if improvement by using 2 subsamples is significant (F-test)

Limitations : external choice of breakpoint/small samples /homoscedasticity/...

CHOW Breakpoint test publication output - results

Country	FULL	FRAC
DK	YE	YE
FI	YE	NO/YE (-)
NO	NO	YE
SE	YE	YE
IE	YE	YE
ES	YE	YE
PT	YE	YE
IT	YE	YE
YE : breakpoint at 5%		
NO : no breakpoint at 5%		
- : 2012 no breakpoint /		
	2013 : breakpoint	

CWTS

What is the impact of the changes in public research funding on the publication output of these organizations?

CWTS

Methodology (1)

- The variables are time series: set of measures index in (equally spaced) time order
- How to measure impact?
- Correlation measures how strongly pairs of variables are related
- HOWEVER correlation DOES NOT imply causation

- **Causation indicates that one event is the result of the other event (cause and effect)**
- In econometrics time series techniques are often used to study causation.

CWTS

Methodology (2)

Granger (1969) : X causes Y by analyzing how much of the Y at t can be explained by lagged values of Y (t-1, t-2, …) and then to analyze whether adding lagged values of X can improve the explanation

→ Y is said to be Granger-caused by X if X helps in the prediction of Y

(Granger causation can be bi-directional, two way causation: X Granger causes Y and Y Granger causes X)

CWTS

Methodology (3)

Granger causality is used in:

- **Economic modelling** (e.g. energy consumption and GDP; growth and export, market prices and future prices)
- **Neuroscience** (modelling of the brain functions)

Granger worked with linear models; extensions to nonlinear cases have been developed

Caution: Bivariate time series - Assumption there is no third underlying variable

To our knowledge Granger causality has never been used in bibliometric studies

CWTS

Methodology (4)

To test the null hypothesis that X does not Granger cause Y, determine the proper lagged values of *Y* to be included in a univariate auto regression of *Y*:

$$Yt = \acute{\alpha}0 + \acute{\alpha}1 * Yt - 1 + \ldots + \acute{\alpha}m * Yt\text{-}m + \varepsilon t$$

Next the lagged values of X are included

$$Yt = \acute{\alpha}0 + \acute{\alpha}1 * Yt - 1 + \ldots + \acute{\alpha}m * Yt\text{-}m + \beta p * Xt\text{-}p + \ldots + \beta q * Xt\text{-}q * + \varepsilon t$$

with l is the shortest lag and q the longest lag for which the lagged value of X is significant

In this regression all lagged values of *X* are taken into account that are individually significant according to their t-statistics, provided that collectively they add explanatory power to the regression according to an F-test (Wald test).

X does not Granger-cause *Y* is accepted if and only if no lagged values of X are retained in the regression: ***null hypothesis***.

CWTS

Methodology (5)

For the Granger causality test a condition is that X and Y are **STATIONARY** times series

Stationary: a shift in time does not cause a change in the shape of the distribution.

Different types of **non-stationary processes**, e.g. deterministic trends; stochastic trends in a time series (unit root (process))

If a time series has unit roots a series of successive differences (d) can transform it into one with stationarity – (notation I(d))

In most cases the order of integration is I(0), I(1) or I(2)

CWTS

Methodology (6)

- Two sets of variables are **cointegrated** if a linear combination of these variables has a lower order of integration.

- E.g. Two I(1) time series Xt and Yt are cointegrated if they can be described by a stationary I(0) process Zt=Yt+ά*Xt

CWTS

Methodology (7)

How to carry out Granger causality test?

Step 1

Test the stationarity of the variables : values (level),the 1st difference and if necessary the 2nd difference, using:

1) Augmented Dickey Fuller (ADF) test and the Perron-Philips (PP) test (with deterministic trend);

2) When necessary the breakpoint unit root test.

The lag length is determinated using Akaiki's Information Criterion (AIC) and Swartz Information Criterion (SIC)

t-test is used with for statistical significance McKinnon (1996) finite sample critical values

CWTS

Methodology (8)

STEP 2 – testing cointegration of variables at level by using Johansen method (1995)

Five different trend assumptions (e.g. the level data and the cointegrating equations have linear trends)

Two statistics:

*) the trace statistics

*) the maximum eigenvalue statistics

Cointegration is accepted if BOTH statistics indicate at least 1 cointegration vector

CWTS

Methodology (9)

Step 3 – Granger causality test for bivariate time series

- If the two variables are not cointegrated **a vector autoregression (VAR) model** can be used and **VAR Granger causality** test is used to test causality

- If there is a cointegration relationship between the 2 variables **vector error correction (VECM) model** can be used (Cointegrated VAR) and causality will be tested using the **VEC Granger causality** test

TOOLS : EVIEWS (Statistical package, IHS Markit)

numXL (add-in for MS EXCEL, SPIDER Finance)

CWTS

FUNDING DATA

- Funding data: OECD database

  - Government expenditure on R&D (GOVERD) + Higher education expenditure on R&D (HERD)

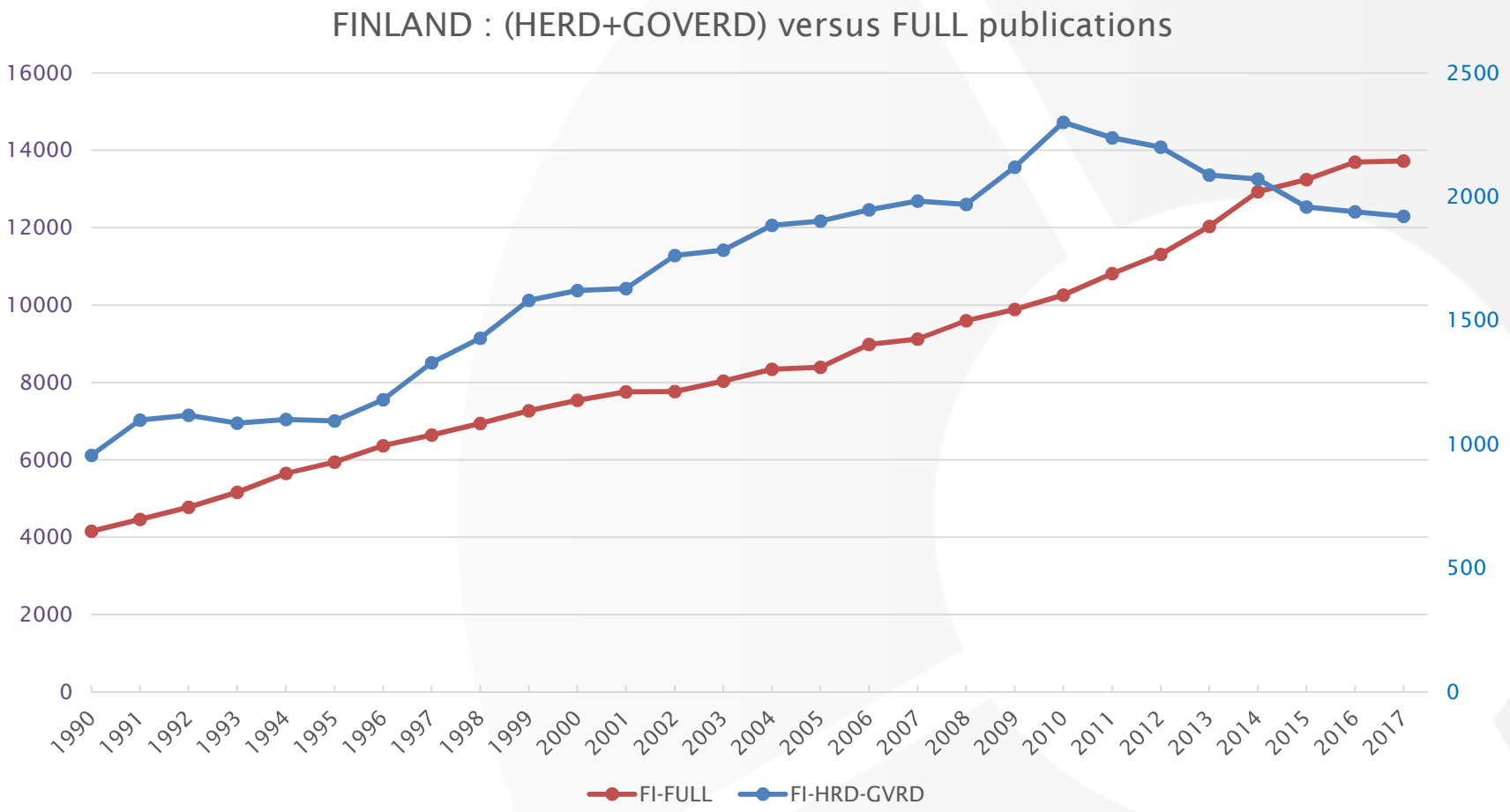
  - Constant Prices (US$, million, 2010) and PPP$

  (Aksnes et al, 2017)

- Reference period: 1990-2017

CWTS

RESULTS : Finland



FI – Unit root test level publications

Unit Root with Break Test on FI_FULL

Null Hypothesis: FI_FULL has a unit root
Trend Specification: Trend and intercept
Break Specification: Trend and intercept
Break Type: Innovational outlier

Break Date: 2007
Break Selection: Minimize Dickey-Fuller t-statistic
Lag Length: 3 (Automatic - based on Schwarz information criterion, maxlag=6)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-4.135247	0.4353
Test critical values:	1% level	-5.719131	
	5% level	-5.175710	
	10% level	-4.893950	

*Vogelsang (1993) asymptotic one-sided p-values.

FI – Unit root test 1st difference publications

Unit Root with Break Test on D(FI_FULL)

Null Hypothesis: D(FI_FULL) has a unit root
Trend Specification: Trend and intercept
Break Specification: Trend and intercept
Break Type: Innovational outlier

Break Date: 2013
Break Selection: Minimize Dickey-Fuller t-statistic
Lag Length: 0 (Automatic - based on Schwarz information criterion, maxlag=6)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-5.864962	< 0.01
Test critical values:	1% level	-5.719131	
	5% level	-5.175710	
	10% level	-4.893950	

*Vogelsang (1993) asymptotic one-sided p-values.

CWTS

FI – VAR – lag test

VAR Lag Order Selection Criteria
Endogenous variables: D(FI_FULL) D(FI_HRD_GVRD)
Exogenous variables: C
Date: 11/01/18 Time: 14:09
Sample: 1990 2017
Included observations: 21

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-264.0181	NA*	3.45e+08	25.33506	25.43454*	25.35665
1	-261.0124	5.152696	3.81e+08	25.42975	25.72819	25.49452
2	-257.4822	5.379316	4.04e+08	25.47450	25.97189	25.58244
3	-256.3989	1.444451	5.52e+08	25.75228	26.44862	25.90340
4	-249.2601	8.158596	4.37e+08	25.45334	26.34865	25.64765
5	-241.1256	7.747173	3.30e+08*	25.05958*	26.15384	25.29706*
6	-239.8116	1.001137	5.14e+08	25.31539	26.60861	25.59605

* indicates lag order selected by the criterion
LR: sequential modified LR test statistic (each test at 5% level)
FPE: Final prediction error
AIC: Akaike information criterion
SC: Schwarz information criterion
HQ: Hannan-Quinn information criterion

CWTS

FI – Cointegration – Johanssen method

Johansen Cointegration Test Summary

Date: 11/09/18 Time: 07:11
Sample: 1990 2017
Included observations: 22
Series: FI_FULL FI_HRD_GVRD
Lags interval: 1 to 5

Selected (0.05 level*) Number of Cointegrating Relations by Model

Data Trend:	None	None	Linear	Linear	Quadratic
Test Type	No Intercept No Trend	Intercept No Trend	Intercept No Trend	Intercept Trend	Intercept Trend
Trace	1	1	2	0	0
Max-Eig	1	1	0	0	0

*Critical values based on MacKinnon-Haug-Michelis (1999)

FI - VAR Granger causality
H0 : indep. var. does NOT GC dep. var.

VAR Granger Causality/Block Exogeneity Wald Tests
Date: 11/01/18 Time: 14:11
Sample: 1990 2017
Included observations: 22

Dependent variable: D(FI_FULL)

Excluded	Chi-sq	df	Prob.
D(FI_HRD_GVRD)	18.49642	5	0.0024
All	18.49642	5	0.0024

Dependent variable: D(FI_HRD_GVRD)

Excluded	Chi-sq	df	Prob.
D(FI_FULL)	7.077566	5	0.2149
All	7.077566	5	0.2149

CWTS

Granger causality : (HERD+GOVERD) and Full publ. numbers

CO	Fund⟶Pubs	Pubs⟶Fund
DK	Y (3,E)***	N(3,E)
FI	Y (5,V)***	N(5,V)
NO	Y(2,V)*	N(2,V)
SE	N(2,V)	N(2,V)
ES	-	-
IE	-	-
PT	N(3,V)	N(3,V)
IT	N(2,V)	N(2,V)
**** ά=0.01; *ά=0.1		
Period: [1990,2017]		
V : VAR Granger Causality test		
E: VECM Granger Causalit test		
- : no results - multiple breakpoints in time series		

Conclusions

- Work in progress: multiple breakpoints test, additional countries, ….
- BE CAUTIOUS WITH STATISTICS – no black box – many underlying assumptions made and tests on data necessary; relatively small sample
- Break in publication output around 2012 (more pronounced for fractional publication numbers)

  (not a database effect!)
- There are reductions in public funding for universities and research institutes starting around 2008.
- For 3 out of 4 Nordic countries (changes in) funding Granger cause (changes in) publication output BUT model is rather simplistic for both the input and output data (however similar models frequently used in econometrics)

CWTS

Thank you
… also on behalf of Thed

Luwel@cwts.leidenuniv.nl

CWTS