UiO : Department of Political Science
University of Oslo

Håvard Strand, Maria Seippel Bineau, Jonas Nordkvelle & Nils Petter Gleditsch

Posting your data: From obscurity to fame?

Arguments for replication policies

1. Efficiency

   Sharing is more efficient than recollection

2. Incremental research

   Building on others is a form of quality control

3. Signal effect

   Replication signals confidence in results

Arguments against replication

- Scooping
  - Young researchers spend much time collecting data
  - Replication pressure -> release of data
  - Other researchers pick up the data and write numerous articles that the Young researcher otherwise could have written.
    - These others release their replication dataset and receive all the citations
- Older, famous researchers release their replication datasets consisting of third party data assembled at little cost
  - Receives boost due to replication policy

Our argument

- We expect an independent effect of replication data
  - If the replication effect is merely a trend, then we should expect a trend variable to outperform the replication dummy
- The signal effect should mainly affect those without a reputation
  - The interaction between reputation and replication should favour the least reputable.

Research design

- 430 items from *Journal of Peace Research*,
  - Citation counts from 2002, 2013 and 2016
    - Cit2016
    - Diff 2016-2013
  - Number of item-related variables:
    - Field, North American, Gender, coauthored, lenght
  - Type of analysis:
    - Verbal argument, Illustrative figures, game theory, uni- or bivariate analyses, multivariate analyses, Special Data Feature
  - Fame:
    - The number of citations to each author at the time of publication

Model

- Unconditional fixed effects negative binomial regression model
  - Issue fixed effects through dummy variables
    - Seemingly robust to some alternative specifications
  - HC1 standard errors
    - Bootstrapping

	(1)	(2)	(3)	(4)	(5)	(6)
	Cited	Cited	Cited	diff	diff	diff
Theoretical figures	0.746***	0.720***	0.766***	0.743**	0.716**	0.755**
	(3.71)	(3.57)	(3.76)	(2.73)	(2.64)	(2.75)
Univariate/bivariate analysis	0.393	0.425*	0.400	0.401	0.400	0.441
	(1.90)	(2.09)	(1.95)	(1.54)	(1.57)	(1.65)
Game matrices	0.298	0.263	0.273	0.421	0.349	0.422
	(1.30)	(1.17)	(1.22)	(1.38)	(1.17)	(1.41)
Multivariate analysis	0.534***	0.524***	0.667***	0.253	0.242	0.490*
	(3.41)	(3.37)	(4.63)	(1.18)	(1.13)	(2.50)
Data feature	1.606***		1.848***	0.606		1.014*
	(3.62)		(4.62)	(1.17)		(2.20)
1-26 prior cit.	0.310*	0.349**	0.282	0.245	0.300	0.113
	(2.33)	(2.68)	(1.94)	(1.32)	(1.63)	(0.55)
More than 27 prior cit.	0.969***	1.098***	0.906***	0.808***	0.963***	0.620*
	(5.21)	(5.87)	(4.67)	(3.38)	(4.02)	(2.49)
Replication data	**0.805***	**0.863***		**1.222***	**1.273****	
	(2.26)	**(2.37)**		**(2.55)**	**(2.61)**	
1-26 prior cit. X Rep. Data	-0.449	-0.491		-0.421	-0.510	
	(-1.12)	(-1.19)		(-0.77)	(-0.91)	
27+ prior cit. X Rep. Data	-0.595	-0.708		-0.988*	-1.082*	
	(-1.61)	(-1.91)		(-2.06)	(-2.25)	
Replication Policy			**0.538**			**0.297**
			(1.08)			**(0.54)**
1-26 prior cit. X Rep. Pol.			-0.0613			0.330
			(-0.21)			(0.84)
27+ prior cit. X Rep. Pol.			-0.0882			0.0930
			(-0.28)			(0.25)
N	430	419	430	430	419	430