Journals in Economic Sciences: Paying lip-services to reproducible Research?

The Welfare Effects of Bundling in Multichannel Television Markets (pp. 643-85)
Gregory S. Crawford and Ali Yurukoglu
Abstract/Tools | Full-Text Article | Download Data Set | Online Appendix

The Use of Full-Line Forcing Contracts in the Video Rental Industry (pp. 665-719)
Katherine Ho, Justin Ho and Julie Holland Mortimer
Abstract/Tools | Full-Text Article | Download Data Set | Online Appendix

Slow to Anger and Fast to Forgive: Cooperation in an Uncertain World (pp. 729-49)
Drew Fudenberg, David G. Rand and Anna Dreber
Abstract/Tools | Full-Text Article | Download Data Set | Online Appendix

Source: Economics-E-Journal
Source: American Economic Review
Source: ckan.org

Sven Vlaeminck
ZBW – Leibniz Information Centre for Economics
IASSIST 2016 Conference | May 31 – June 3 | Bergen | Norway

Funded by
DFG

EDaWaX
European Data Watch Extended

Max Planck Institute for Innovation and Competition
DIWOESEPRatSWD ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics
Agenda:

> 1. Introduction

> 2. Data Policies in Economic Sciences’ Journals

> 3. Do Journals enforce their own Data Policies?

> 4. Take home Messages
1. Introduction:
The Beauty and the Beast

Changing Patterns of published economic Research and the Replication Crisis in the Social Sciences
Changing Patterns of published Research

> Hamermesh analysed 728 articles published in economics top-journals between 1963 and 2011.

> Results: Methodology used changed fundamentally.
  - Research based on borrowed or self-collected datasets, experimental designs or simulations reached 80% in 2011.
  - In 1963 this share was below 50%.*

> More than ever, research in economics is ‚data-based‘

The Replication Crisis in Economic Sciences

> Dewald et al. (1986) tried to replicate 54 articles of the *Journal of Money, Credit, and Banking (JMCB)*.
  - They succeeded two times (3.7%).
> Hubbard & Armstrong (1994) analysed 1,120 papers in marketing journals.
  - None of these were replications; 1.8% were extensions.
  - Of these extensions, 60% found conflicting results.
> McCullough et al. (2008) tried to replicate 117 articles of the *Federal Reserve Bank of St. Louis Review*.
  - They were able to replicate 9 articles (7.7%).
> Chang & Li (2015) successfully replicated 29 of 59 papers (49%) from seven journals with authors’ support.

➔ ...most often, economic research is not replicable.
2. Data Policies in Economic Sciences’ Journals

“The economics profession has made some attempt to require researchers to make their data available to others, in part because replication would seem to be a useful check on research and in part because of studies showing that economists had a habit of having their data immediately eaten by their dogs upon publication of their articles.” (Liebowitz, 2007)
Research Questions:

Among other things, we investigated:

- The share of journals equipped with data policies.
- The type of the data policy.
- The specifications of these data policies.
- Differences and commonalities of journals’ data policies in economics and business studies.
- ...

Data Policies of Journals in Economic Sciences

346 Journals in our Sample

Primary Subject of Journals in our Sample
(n=346)

- 160; 46.2% primarily economics
- 132; 38.2% primarily business studies
- 34; 9.8% economics & business sciences in equal parts
- 20; 5.8% other

Other logos and affiliations: Max Planck Institute for Innovation and Competition, DIW Berlin, SOEP, RatSWD, ZBW Leibniz-Informationszentrum Wirtschaft.
Data Policies of Journals in Economic Sciences

346 Journals in our Sample

71 Journals with Data Policies
Data Policies of Journals in Economic Sciences

346 Journals in our Sample

\[ \text{71 Journals with a Data Availability Policy} \]

\[ \text{49} \]
Data Policies of Journals in Economic Sciences

346 Journals in our Sample

71

49

30

Mandatory Policies
Data Policies of Journals in Economic Sciences

346 Journals in our Sample

71
49
30
25

Data Policies that require the program code. (7.2% of the sample)
Data Policies of Journals in Economic Sciences

346 Journals in our Sample

71 Policies with rules for papers based on restricted data (5.8%)
3. Do journals enforce their own data policies?

“If the empirical basis for an article cannot be reproduced, of what use to the discipline are its conclusions? What purpose does an article like this serve?”

(Gary King, 1995)
**Methodology:**

> 39 journals equipped with a Data Availability Policy have been investigated (mostly two issues - in total 628 articles).

> We determined whether an article is 'data-based' (-> the paper contains empirical work, simulations or experimental work) and whether it employed restricted data.

> We investigated for each data-based article whether and which replication files were available (-> datasets / program code / references on datasets?).

> We determined the publication history for each article (if available) AND the date a data policy became effective (-> long publication cycles in economic sciences!).

> For each paper, we contrasted the replication files we found to the requirements of journal’s data policy.

> It was not of interest, whether the policy was robust or weak.
Our research Samples at a Glance

- Full Sample (n=628)
- Data-based papers (n=467)
- Compliance Analysis (n=259)
Compliance Rates for Papers based on (non)-restricted Data:

Table 5. The availability of replication files, information on datasets and the compliance rate for articles employing (non)-restricted data (n=259)

<table>
<thead>
<tr>
<th>Type of article</th>
<th>Data available</th>
<th>Code available</th>
<th>Information on restricted data</th>
<th>Compliant to data policy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Articles using restricted data</td>
<td>---*</td>
<td>44 (55%)</td>
<td>50 (62.5%)</td>
<td><strong>41 (54.7%)</strong></td>
</tr>
<tr>
<td>(n=80/30.9%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Articles using non-restricted data</td>
<td>86 (48%)</td>
<td>84 (46.9%)</td>
<td>--</td>
<td>81 (45.3%)</td>
</tr>
<tr>
<td>(n=179/69.1%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* A few articles that use both restricted data and freely available datasets also provide some non-restricted datasets.

** Five cases have been removed from the compliance analyses, because these journals exempted papers based on restricted data from their data policy.

> The majority of all investigated ‘data-based’ articles are NOT in compliance with the requirements of the respective data policy.

> Surprisingly, the compliance rate of articles employing restricted data is higher than for those employing non-restricted data.
Journals’ Compliance to its own Data Policies

Compliance Rate: Numbers and Shares for the 18 Journals in our Sample
(n=18 journals / 259 articles)

<table>
<thead>
<tr>
<th>Range</th>
<th>No. of journals</th>
<th>Share of all journals in sample</th>
<th>Share of all articles in sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%-20%</td>
<td>10</td>
<td>55.6</td>
<td>47.5</td>
</tr>
<tr>
<td>&gt;20%-40%</td>
<td>0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>&gt;40%-60%</td>
<td>1</td>
<td>5.6</td>
<td>5.4</td>
</tr>
<tr>
<td>&gt;60%-80%</td>
<td>2</td>
<td>11.1</td>
<td>7.7</td>
</tr>
<tr>
<td>&gt;80%-100%</td>
<td>5</td>
<td>27.8</td>
<td>39.4</td>
</tr>
</tbody>
</table>
Compliance Rates for Journals with mandatory / voluntary DAPs:

**Table 6.** Replication files and compliance with own data policy for mandatory and voluntary data policies (n=259)

<table>
<thead>
<tr>
<th>Type of article</th>
<th>Data available</th>
<th>Code available</th>
<th>Information on restricted data</th>
<th>Compliant to data policy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Articles w. mandatory DAP using restricted data</td>
<td>---*</td>
<td>42 (70%)</td>
<td>44 (73.3%)</td>
<td>41 (68.3%)</td>
</tr>
<tr>
<td>(n=60)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Articles w. voluntary DAP using restricted data</td>
<td>---*</td>
<td>2 (10%)</td>
<td>6 (30%)</td>
<td>**0 (0%)</td>
</tr>
<tr>
<td>(n=20)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Articles w. mandatory DAP using non-restricted data</td>
<td>82 (73.2%)</td>
<td>81 (72.3%)</td>
<td>---</td>
<td>63 (70.5%)</td>
</tr>
<tr>
<td>(n=112)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Articles w. voluntary DAP using non-restricted data</td>
<td>4 (6.0%)</td>
<td>3 (4.5%)</td>
<td>---</td>
<td>2 (3.0%)</td>
</tr>
<tr>
<td>(n=67)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* A few articles that use both restricted data and freely available datasets also provide some non-restricted datasets.** Five cases have been removed from the compliance analyses, because these journals exempted papers based on restricted data from their data policy.
Take home messages:

> Only a small minority of journals in economic sciences (5.8%) have implemented suitable data availability policies.

> Big disparities: In economics every fourth journal has a DAP, while in business studies only one in 18 journals has one.

> We found light and shadow – both on the level of policies and on the level of journal’s compliance to its own data policies.

> Even journals with a DAP fail to enforce its data policies in large parts (only 48% of all articles honoured journal’s data policy).

> A small group of (top-)journals really enforce its data policies.

> A mandatory DAP is the only type of a data policy that works!
Further reading:

> List of all publications and presentations of the EDaWaX-project: www.edawax.de/downloads


> List of all journals investigated: http://www.edawax.de/wp-content/uploads/2016/05/Alle-untersuchten-Journals.pdf

Thank you very much for your attention!

...do you have questions or comments?

Contact:
Sven Vlaeminck | s.vlaeminck@zbw.eu | www.edawax.de
ZBW – Leibniz Information Centre for Economics
Neuer Jungfernstieg 21
D- 20354 Hamburg | Germany