## ERRATUM

In an earlier version of the article "Journals in Economic Sciences: Paying Lip Service to Reproducible Research?" by Vlaeminck and Podkrajac, the following appearred uncited on page 2:

> "While in other sciences replicability is regarded as a fundamental principle of research and a prerequisite for the publication of results, in economic sciences it is not treated as a top priority.  In 2006..."

This should instead have been cited as follows:

> "According to Höffler (2017), replicability does not take a high priority in economics. In his opinion, this is in sharp contrast to other sciences where replicability is "regarded as a fundamental principle of research and a prerequisite for the publication of results" (Höffler, 2017, p.1). Already in 2006..."

# Journals in Economic Sciences: Paying Lip Service to Reproducible Research?

Sven Vlaeminck[1] and Felix Podkrajac[2]

## Abstract

The findings of numerous replication studies in economics have raised serious concerns regarding the credibility and reliability of published applied economic research. Literature suggests that economic research often is not replicable because (i) only a small proportion of journals in the field have implemented functional policies on the disclosure of employed datasets and program code, (ii) authors frequently do not comply with these data policies and (iii) editorial offices do not ensure that these policies are enforced. In this paper, we focus on the aspect last mentioned. We empirically evaluate 599 articles published in 37 journals with a data availability policy. We present the share of articles that fall under a data policy, because replication data is needed to verify the published results. Afterwards, we check the journal data archives and supplemental information section of each article for the availability of replication files. For a reduced sub-sample of 245 data-based articles, we check in depth whether the replication files we found are compliant with the requirements of the journal's respective data policy. Thereby, we are able to determine how much journals in economic sciences enforce their data policies. Our findings suggest a mixed picture: While some journals achieve high compliance rates, a significant share of journals only sporadically provides replication files for data-based research papers.

## Keywords

reproducible research, economics, journals, data policies, data archives,

## 1. Introduction

In economics, data-based[3] research has become increasingly important. According to the US economist Hamermesh (2013), the number of contributions to top journals in which authors utilised self-collected or borrowed datasets, employed experimental designs, or used real data for simulations of theoretical models has massively increased over the last decades. While in 1963 the share of publications in economic journals with a solely theoretical orientation was 50.7%, this percentage dropped to 19.1% in 2011.

With the growing relevance of data-based publications, new questions and challenges for academic publishing emerge. One of the most pressing challenges is the validation of published data-based research, which is closely connected to the principle of replicable research, as a cornerstone of the scientific method[4]:

'Replication ensures that the method used to produce the results is known. Whether the results are correct or not is another matter, but unless everyone knows how the results were produced, their correctness cannot be assessed. Replicable research is subject to the scientific principle of verification; non-replicable research cannot be verified. Second, and more importantly, replicable research speeds scientific progress. […] Third, researchers will have an incentive to avoid sloppiness. […] Fourth, the incidence of fraud will decrease' (McCullough, 2009, p.118f.).

Replication refers to the duplication of scientific findings, but literature shows that the common understanding of replication differs widely: Clemens, for instance, lists several terms used by

researchers (e.g. reproduction, verification, extensions, robustness tests,...) which all refer to replication. He concluded that 'there is no consensus meaning of the term replication' (Clemens, 2015, 7). For our paper, we employ the definition of Hamermesh's 'pure replication', which he defines as to 'duplicate, repeat, as in a statistical experiment' and 'to make or do something again in exactly the same way' (Hamermesh, 2007).

According to Höffler (2017), replicability does not take a high priority in economics. In his opinion, this is in sharp contrast to other sciences where replicability is "regarded as a fundamental principle of research and a prerequisite for the publication of results" (Höffler, 2017, p.1). Already in 2006, McCullough et al. criticised the status quo in economics: 'Results published in economics journals are accepted at face value and rarely subjected to the independent verification that is the cornerstone of the scientific method. Most results published in economics journals cannot be subjected to verification, even in principle, because authors typically are not required to make their data and code available for verification.' (McCullough, McGeary & Harrison, 2006, p.1093 f.)

McCullough's negative appraisal is based on the findings of his own studies and those of other researchers in the field, who systematically tried to replicate published applied economic research. One of the first of these studies was released by the economists Dewald, Thursby and Anderson in 1986. In their seminal paper, the researchers collected programs and data from authors of the *Journal of Money, Credit and Banking* (JMCB) with the aim to replicate the results of 54 published articles. They were able to replicate the key findings of only two papers (3.7%) – a result that has led to an ongoing debate on reproducible research in economics (Dewald, Thursby & Anderson, 1986).

The data policy of JMCB at that time was based on the willingness of researchers to cooperate in cases of requests for data and code – which we label an 'author responsibility policy' (ARP). But this willingness does not conform to the incentive structures in science and research, as Mirowski & Sklivas (1991) and Feigenbaum & Lewi (1993) point out.

One important reason for the replication crisis in economics is based on such missing incentives for researchers to share their data. As Duvendack, Reed and Palmer-Jones (2015) argue, authors of original studies are concerned about the costs of compiling data and program code into (re-)usable forms. Original authors may expect that the benefits of providing well documented, easily usable code are small or even negative. Little credit is given to the original author if the replicating authors confirm the original results, while the damage to reputation may be large, if the original results cannot be confirmed (Hamermesh, 2007). This situation is aggravated by the fact that replication is still viewed as a 'parasitic activity' by some researchers (Hamermesh 1997, c.f. Longo & Drazen, 2016). Also the wish to retain exclusive rights to data that had taken a lot of time to produce is an important reason for researchers not to share their data (Savage & Vickers, 2009 and Mueller-Langer & Andreoli-Versbach, 2014). Other concerns comprise legal issues and the fear of a misinterpretation/misuse of data (Tenopir et al, 2009).

These missing incentives to share data have also been noticed by Dewald, Thursby and Anderson: The researchers reported that one-third of the authors never replied to their repeated requests, and an additional one-third replied that they could not furnish their programs or data. Dewald, Thursby and Anderson concluded: 'Our findings suggest that the existence of a requirement that authors submit to the journal their programs and data along with each manuscript would significantly reduce the frequency and magnitude of errors' (Dewald, Thursby & Anderson, 1986, p.588).

The implementation of such mandatory data policies and corresponding data archives that require both data and program code from authors of data-based papers could have several benefits, as highlighted by Anderson et al.: With mandatory data and program code archives, checking robustness of a data-based paper should be a comparatively simple matter. Data and specifically program code should provide a record of how the published research was produced. Researchers who wish to build on previous research therefore no longer have to program everything from scratch. Consequently, not only can data policies and data archives increase the accuracy of the published results, but they could also be better examined for robustness and can be more easily extended, thus also increasing the quality of research (Anderson, Greene, McCullough & Vinod, 2008).[5]

In fact, some journals altered their data policy after publication of Dewald's, Thursby's and Anderson's article towards a mandatory data availability policy (abbreviated in the following to 'DAP'). A DAP normally asks authors of data-based papers to provide their replication files to the editorial office prior to an article's publication. Which data and files are exactly needed to facilitate replications is exemplarily outlined by King (1995a), but also depends on the methodology utilised.[6]

The *Federal Reserve Bank of St. Louis Review (1993)*, the *Journal of Money, Credit and Banking* (1996), *Studies in Nonlinear Dynamics & Econometrics (*1996) and *Macroeconomic Dynamics* (1996) were among those journals who implemented or altered their data policy after Dewald's, Thursby's and Anderson's findings till the turn of the millennium (McCullough, 2009)*.

But despite changes in journals' data policies and implementations of data archives, replication studies still reported that only a minority of applied research papers were replicable: McCullough, McGeary and Harrison (2006) again tried to replicate papers published in JMCB by using data from the journal's data archive. 193 of 266 articles published between 1996 and 2003 should have had entries in the data archive, but only 69 did. Of these, the researchers tried to replicate 62, but succeeded only 14 times (22.6%). As a consequence, McCullough (2007, 327ff.) published recommendations for managing a journal's data archive and listed files and conditions required for successful replications.

In a subsequent study, McCullough, McGeary and Harrison (2008) tried to replicate 117 articles published by the *Federal Reserve Bank of St. Louis Review*. Only 9 (7.7%) papers could be reproduced. Also a more recent study reported ongoing difficulties with the replication of published economic research: Chang and Li (2015) could replicate qualitative key results from 33% of 67 papers published in 13 well-regarded economics journals. With support from the authors, the share increased to 49% – which is still below half. The authors found that their replication success rate was significantly higher when they attempted to replicate papers from periodicals that have a mandatory replication data and program code submission policy.

But how can it be explained that so many replication attempts in economic sciences fail, despite more journals starting to implement data policies and corresponding data archives? To answer this question, McCullough, McGeary and Harrison (2008) examined the data archives of four economics journals (*Federal Reserve Bank of St. Louis Review*, *Journal of Business and Economic Statistics*, *Journal of Applied Econometrics* and *Economic Journal*) and compared the number of data-based articles with the entries in the journals' data archives. First, the researchers examined all articles in these journals, classifying each as requiring an entry in the archive or not. An article was classified as requiring an archive entry if the article displayed or otherwise represented numbers. Therefore they included classically 'empirical' articles, as well as computational economics articles. Subsequently, the authors

checked the data archives for entries of the respective articles. For each issue investigated, they show the number of articles that utilised data and therefore should have an archive entry, the number of articles that actually had an archive entry, and the 'compliance' percentage of articles that have an entry. As a result, they found compliance rates reaching from 13% (*Economic Journal*) to 99% (*Journal of Applied Econometrics*).

With this paper we provide an updated evaluation of journals' practises when it comes to the enforcement of their data policies. Because the availability of data and program code can be regarded as a prerequisite to validate the findings of published research, it is important to monitor the current practises of journals in economic research. Especially against the background of the replication crises in the social sciences, this aspect is of special interest.

Our approach follows the study of McCullough, McGeary and Harrison (2008) in several aspects: we regard the practise of 37 journals in economic sciences – all of them have a DAP. We give the number of articles that are data-based and therefore should have an entry in journal's data archive (or if no separate data archive exists, the replication files should be available in the supplemental information section of the article). Afterwards, we investigate for how many of these articles replication files are available. We distinguish between papers which are using restricted and non-restricted data, because DAP often contain varying rules for research based on proprietary or confidential data (Vlaeminck & Herrmann, 2015).[7]

Furthermore, we compare the share of articles with supplemental replication data of journals with a voluntary data policy to those with a mandatory data policy. By doing so we determine empirically how well voluntary and mandatory data policies perform against the background of missing incentives for data sharing.

With a reduced sample of journals we subsequently evaluate in depth whether journals follow and enforce their data policies. For this purpose, we compare the requirements of the journal's respective data policy to the replication files we found for each published data-based article.

Our paper is organised as follows: Section 2 describes the methodology to collect the data for our analysis. Section 3 presents the results of our study, while section 4 summarises and discusses the findings of the analyses.

## 2. Data and Methodology

In the following subsections, the approach of our study is described. Section 2.1 describes previous research and a dataset we utilised as a starting point for our analyses. Section 2.2 sets out how we determine the share of data-based articles and expounds how we ascertain the share of accompanying replication files. Section 2.3 explicates the methodology for the analysis of journals' compliance rates with their data policies.

### 2.1 Datasets available

For our study, we used a publicly available dataset of Vlaeminck and Herrmann (2015b) as a basis. The corresponding research paper (Vlaeminck & Herrmann, 2015a) presents the results of an analysis of journals' data policies in economic research. The authors evaluated the data policies in a sample of 346 journals and found 49 journals with a DAP (14.2%). To compile the sample for their analyses, the

authors employed two lists of academic journals assembled by economic associations. One list (JOURQUAL 2.1) is maintained by the German Academic Association for Business Research (VHB). Vlaeminck and Herrmann chose all journals from the JOURQUAL list ranked A+, A or B and added each 60 journals ranked C, D and E to their sample. Besides, a list of journals analysed by Bräuninger, Haucap and Muck (2011) has been added to the research sample. In a next step, Vlaeminck and Herrmann removed double entries and specified the primary subject category of these periodicals according to their classification in Thomson Reuters Journal Citation Reports and ZBW's indexing guidelines. 262 out of 346 (75.7%) journals investigated were listed in Thomson Reuters Journal Citation Reports (JCR) 2013 and therefore have an Impact Factor.

In their dataset, Vlaeminck and Herrmann also set out which of the 49 DAP are mandatory and which are voluntary for authors. Also they stated for each journal in their dataset what the data policies require from journals' authors and how these journals provide replication files.

For our analysis, we include 37 (75.5%) of these 49 journals with a DAP in our research sample. The remaining journals have been excluded because either the subject category is not primarily located in economics (e.g. *science, nature, PNAS*) or because we were not able to access the articles of these journals due to licensing restrictions (e.g. *Journal of Law and Economics, Journal of Labor Economics, European Accounting Review*,...)

## 2.2 Approach to determine the share of data based articles

To determine the share of data-based articles in our sample, we investigate 599 articles published by 37 journals in the issues 1/2014 (323) and 1/2013 (276). We choose only regular issues. In the event that one of the issues is a special issue, we use the following issue. This approach is necessary, because special issues tend to focus on single research questions or topics which might result in a bias towards a certain methodology.

For each journal, we examine the whole issue. This leads to disparate numbers of articles per periodical in our sample: While some journals only publish four articles per issue, others publish 20 and more. Therefore, journals like the *Journal of the American Statistical Association* (9.8% of all articles within our sample), the *Review of Economics and Statistics* (6.3%) or *the Journal of Economic Dynamics and Control* (5.8%) have the highest numbers of articles in our sample, while journals like the *Journal of Political Economy* (0.8%), the *Econometrics Journal* (0.8%) or the *Jahrbuecher fuer Wirtschaftswissenschaften/Review of Economics* (0.8%) only have a few articles in our sample.[8]

Subsequently, we examine the content and methods used by each of the 599 articles. First of all, we study the title and the abstract of each to determine whether an article is data-based. If we are still unsure, we go through the article and manually search for specific terms or keywords like 'data', 'simulat*', 'significant*', 'experiment*', 'evidence', 'empirical*' and read the specific paragraph.[9] We only consider original research articles and remove from the survey literary material like review articles, editorials, comments and rejoinders.

We define an article to be 'data-based' when the methods include creating, reusing, and analysing research data, other empirical methods (e.g. experiments) or self-compiled program code (e.g. simulations). For two reasons, our approach is comparatively strict: While a few data policies[10] consider papers not to be empirical if they use data only for case studies (for instance, some simulations use real data to demonstrate the effects of a theoretical model), we consider these articles

to be data-based. First, would-be replicators need the program code of the simulation to verify the computation. Second, when real data applications have been used, we also expect these case studies to be replicable – even though the case study is not the main result of the paper. Our approach is in line with the specifications of the most robust data policies, like the one of the American Economic Review (AER) (c.f. Glandon, 2010).

In addition, we determine whether an article uses proprietary or confidential data (for instance, purchased datasets, firm data or microdata) to estimate the share of articles that rely on restricted data. We examine these articles for references on datasets employed by the authors. This information is of interest for our approach described in section 2.3, because several data policies possess special rules for research based on restricted data.

Afterwards, we check whether the data-based articles are accompanied by replication files (e.g. datasets and/or program code). In case of restricted data, we look for program codes and further references to these datasets. For this purpose, we check both the publisher's web page and the separate data archive, if applicable,[11] for replication files. For each data-based article, we note which files we found.

## 2.3 Approach on journals' data policies compliance rate

To examine journals' compliance with their own data policies, we collect some additional information for all of the data-based articles. Because we have to ensure that a journal's data policy was already in effect at the time of article's submission, we note the publication history for each and the implementation date of journal's data policy, if available.

This approach is necessary because normally a paper only falls under a data policy, when the policy is effective at the time of submission. Also, some journals (e.g. the *Journal of the American Statistical Association*) demand authors to submit their data already with the submission of their manuscripts.

At that point, certain difficulties arise because not many journals provide the publication history of their articles and also the implementation date of a data policy often is not easy to confirm, because only few journals explicitly mention the date when the policy took effect.

To solve these challenges, we first determine how long the publication process in economics usually lasts. Literature shows that the publishing process in economic journals can easily exceed more than a year: Ellison (2002) ascertained the average review time (from submission to acceptance) of 16.5 months for 25 top economics journals in 1999. Björk and Solomon (2013) found that the whole publication period from submission to publication lasts 18 months, on average. Besides, by reusing a list containing the full text of journals' data policies compiled by the EDaWaX project in the spring of 2012, we made certain that there are at least 29 journals with a DAP at that time. For some other journals, we could determine the date the policy took effect, either because it was mentioned in the literature or because the policies specified the point in time at which they became effective.

As a consequence, we only keep articles in the sample for the compliance analysis, if either we are able to determine the exact submission date of the article and know the journal's data policy was effective at that point in time *or* if we know the journal's data policy was implemented more than 24 months prior to the article's publication. According to Björk's and Ellison's findings on the average

publication times in economic journals, it is thus very likely that a journal's data policy had already taken effect at the time the article was submitted.

In a next step, we contrast the replication files we found in the journal's data archive or in the supplemental information section of a paper to the specifications of the journal's respective data policy. For instance, if the data policy requires the disclosure of data and program code but only the data has been provided, we consider such an article to be not compliant with journal's data policy. Special handling is needed for research based on restricted data. For this article type, we initially consult the data policy of the respective journal. If it mentions specific requirements (e.g. to post the program code and some contact information to access the dataset), we use this requirement to check whether the policy is fulfilled. In cases of data policies without such a procedure, we exempt the article from our analysis, because such an article does not fall under the journal's data policy.

By applying this approach, we are able to determine for any article and any data policy whether the policies' requirements have been fulfilled – regardless of whether we consider the policy itself to be robust or weak, whether it is mandatory or voluntary.

## 3. Findings of our study

In the following subsections, we present the empirical findings of our study. First, we show the share of data-based papers in our full sample. Subsequently, we depict the amount of data-based articles for which we find replication data and present the respective outcomes for journals with mandatory and voluntary DAPs. Second, we show the findings of the compliance analysis. We illustrate these results both on article-level and on journal-level. Again, we pay attention to potential differences between mandatory and voluntary data policies. Figure 1 graphically illustrates the different sample sizes of our analyses.
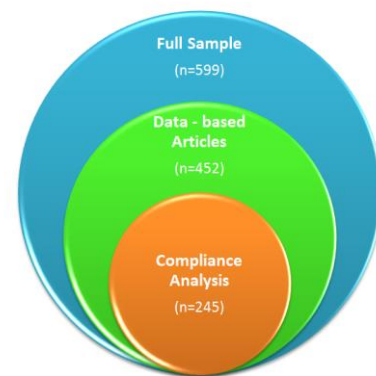


Figure 1: The research samples at a glance

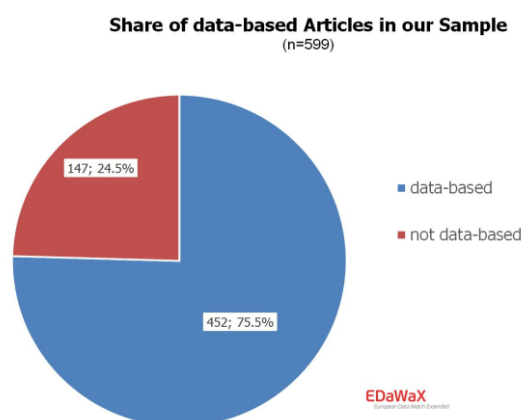### 3.1 The share of data-based articles



Figure 2: The share of articles categorised as 'data-based.'

In total, we examine 599 articles published by 37 journals. Because the sample of our study builds on journals with a DAP, we expect the share of data-based articles to be higher than normally anticipated for journals in economic sciences. Based on our approach described in section 2.2, we classify 452 (75.5%) of all articles to be data-based (see figure 2). Hence, these papers (at least those who fall under a mandatory DAP) should have an entry in journal's data archive or in the supplemental information section of the journal.

When further examining these 452 data-based articles, we find that 284 of them (62.8%) employed non-restricted data for their analyses. After checking journals' data archives and the supplemental

information section of these papers, we find accompanying replication datasets for 104 (36.6%) papers and program code for 103 (36.3%).

**Table 1.** Replication files and references to restricted data within our sample of data-based articles (n=452)

| Type of article | Data available | Code available | Information on restricted data |
|---|---|---|---|
| Articles using non-restricted data (n=284 / 62.8%) | 104 (36.6%) | 103 (36.3%) | n.a. |
| Articles using restricted data (n=168 / 37.2%) | n.a. | 51 (30.4%) | 75 (44.6%) |

Another 168 articles (37.2%) employ proprietary or confidential data to generate their findings. Of these, 75 (44.6%) provide some references on the data utilised (see table 1). Most often this information is available from the article only. We believe this information is frequently not sufficient to exactly determine which dataset was used for the analyses (although we did find some well documented examples[12]). In many cases, only the name of the dataset is mentioned, but no further references. To identify a dataset based on this information *might* work for some, but for the majority it will not (e.g. because there are several waves or corrections of a survey/dataset). For 51 (30.4%) of these papers, the program code is also available.

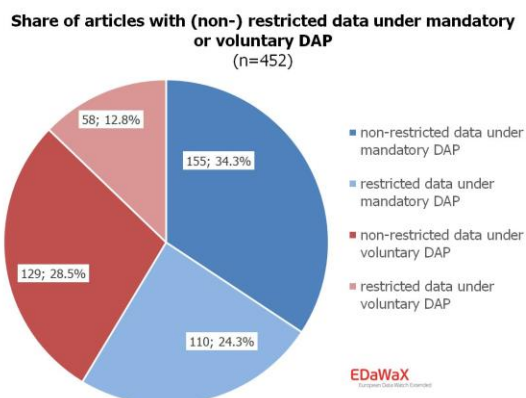**The importance of mandatory data policies for data disclosure**



*Figure 3: The sample of all data-based articles by type of data employed (restricted / non-restricted data) and type of data policy (mandatory /voluntary DAP)*

In our sample, 265 out of 452 (58.6%) articles examined fall under a mandatory DAP. 155 of these (equates 34.3% of the sample) rely on non-restricted data. 98 of these (63.2%) possess accompanying datasets, and the program code is available for 96 (61.9%) articles (see figure 3). Also for the 110 articles (equates 24.3% of the sample) based on restricted data, we ascertain that the program code exists for 46 papers (41.8%). 59 (53.6%) also give some information on restricted datasets employed in the research process.

The remaining 187 articles have been published by journals with a voluntary DAP. 129 of these (equates 28.5% of the sample) rely on non-restricted data. For these articles, replication datasets are available for six (4.7%) papers. Program code is available for seven articles (5.4%). The shares for the 58 articles based on restricted data in journals with a voluntary data policy are only slightly higher: Program code is available for five articles (8.6%) and 16 (27.6%) provide some information on the data employed (see table 2).

**Table 2.** Availability of replication files and information on restricted data by type of article and data policy (n=452)

| Type of article | Data available | Code available | Information on restricted data |
|---|---|---|---|
| Articles using restricted data in journals with mandatory data policy (n=110 / 24.3%) | n.a. | 46 (41.8%) | 59 (53.6%) |
| Articles using restricted data in journals with voluntary data policy (n=58 / 12.8%) | n.a. | 5 (8.6%) | 16 (27.6%) |
| Articles using non-restricted data in journals with mandatory data policy (n=155 / 34.3%) | 98 (63.2%) | 96 (61.9%) | n.a. |
| Articles using non-restricted data in journals with voluntary data policy (n=129 / 28.5%) | 6 (4.7%) | 7 (5.4%) | n.a. |

These results indicate why it is crucial to obligate authors to comply with a DAP. Also our findings illustrate that voluntary data policies do not sufficiently work.

## 3.2 Analysis of journals' compliance rates

Of the 245 articles in the sample for the compliances analysis, 99 (40.4%) were published in 2013, and 146 (59.6%) in 2014. The biggest share of articles again comes from the *Journal of the American Statistical Association* (53; 21.6%), followed by the *Review of Economics and Statistics* (37; 15.1%), the *American Economic Review* (20; 8.2%) and the *American Economic Journal: Applied Economics* (19; 7.8%). Therefore, these four high-ranked journals represent more than half of the articles in the sample.[13]

167 out of 245 articles (68.2%) employed non-restricted data. For 77 (46.1%) of these, datasets for replication purposes are available, and the program code is provided for 75 (44.9%). Of the 78 articles (31.8%) that used restricted data, 50 (64.1%) give some information on the dataset(s) employed. The program code is available for 44 (56.4%) papers (see table 3).

**Table 3.** Availability of replication data, information on restricted datasets employed and compliance rate on article-level (n=245)**

| Type of article | Data available | Code available | Information on restricted data | Compliant with data policy |
|---|---|---|---|---|
| Articles using non-restricted data (n=167/68.2%) | 77 (46.1%) | 75 (44.9%) | n.a. | 73 (43.7%) |
| Articles using restricted data (n=78/31.8%) | n.a. | 44 (56.4%) | 50 (64.1%) | 41 (56.2%)** |

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

These numbers are comparatively higher than for the sample of all 452 data-based articles – for both research based on restricted data (for comparison: program code 30.4%; information on datasets 44.6%) and on non-restricted data (datasets 36.6%; program code 36.3%). Possibly, this higher share reflects that only articles remain in the sample for which the journal's data policy was in effect at time of manuscript's submission.

When examining the overall share of all articles that are compliant with journals' DAP, we find a compliance rate of 47.5% on article-level for both papers based on both restricted and non-restricted data (see figure 4).[14] This percentage indicates that more 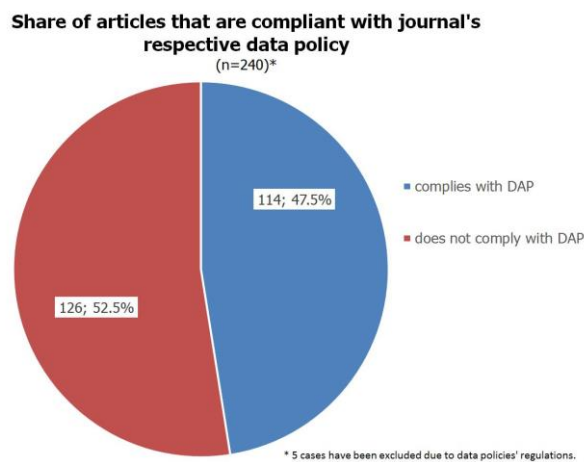than half of all articles did not honour the regulations of the respective data policy. Interestingly, when subdividing the sample into papers utilising restricted data and those using non-restricted data, we find that articles employing restricted data are more often compliant to journal's data policy than those using non-restricted data.

**Share of articles that are compliant with journal's respective data policy**
(n=240)*

- complies with DAP
- does not comply with DAP

114; 47.5%
126; 52.5%

* 5 cases have been excluded due to data policies' regulations.

*Figure 4: Share of articles that comply with the DAP of the respective journal.*

We also find that journals can be divided into two groups (see figure 5): one group enforces its data policy (albeit to varying degrees) while another group apparently does not care much about its data policy.

Journals with the highest compliance rates are the *American Economic Journal: Applied Economics* (100% compliance rate), the *American Economic Review* (100% compliance rate), *Review of Economics and Statistics* (91.9%) and the *Review of Economic Studies* (90% compliance rate).

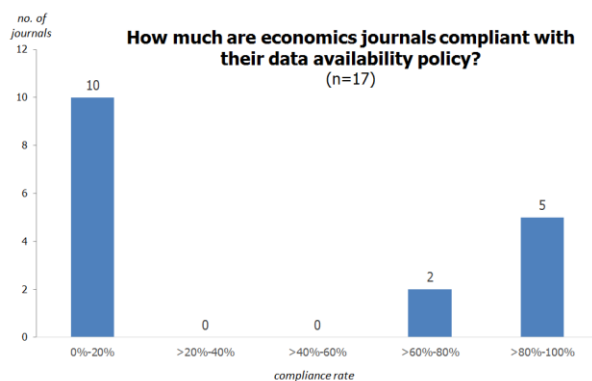On the opposite side 10 journals reach only low compliance rates up to 20 percent. In eight journals not a single data-based article of the investigated issues is compliant with the journal's data policy. On journal-level, in total 10 out of 17 periodicals (58.8%) investigated fall into this group (see figure 5). This distribution is also apparent by the median and the mean of the journals' compliance rates: whereas the mean compliance rate on journal-level is 36.9%, the median of the sample's distribution reaches only 3.8%.

no. of journals

**How much are economics journals compliant with their data availability policy?**
(n=17)

compliance rate

0%-20%: 10
>20%-40%: 0
>40%-60%: 0
>60%-80%: 2
>80%-100%: 5

*Figure 5: How much journals in economic sciences comply with their DAP (means)*

However, the reservation must be made that for some journals only a few data-based papers have been analysed, so that these findings should be interpreted with caution.

**Compliance rates and (non-) mandatory data policies**

158 out of 245 articles (64.5%) were published in journals with a mandatory DAP, while 87 (35.5%) were published in journals with a voluntary DAP. Articles published in periodicals with a mandatory DAP reach a compliance rate of 70.9%. In contrast, the overall compliance rate of papers published by journals with a voluntary DAP is only 2.4%.[15] This discrepancy underlines the importance to obligate authors to honour the data policy of a journal. Table 4 details these findings.

**Table 4.** Availability of replication data, information on restricted datasets employed and compliance rate for voluntary / mandatory data policies on article-level (n=245)

| Type of article | Data available | Code available | Information on restricted data | Compliant with data policy |
|---|---|---|---|---|
| Articles w/ mandatory DAP using restricted data (n=58/24.2%) | n.a. | 42 (72.4%) | 44 (75.9%) | 41 (70.7%) |
| Articles w/ voluntary DAP using restricted data (n=20/6.3%) | n.a. | 2 (10%) | 6 (30%) | 0 (0%)** |
| Articles w/ mandatory DAP using non-restricted data (n=100/41.7%) | 73 (73%) | 72 (72%) | n.a. | 71 (71%) |
| Articles w/ voluntary DAP using non-restricted data (n=67/27.9%) | 4 (6.0%) | 3 (4.5%) | n.a. | 2 (3.0%) - |

** Five cases were removed from the compliance analysis because these journals exempt papers based on restricted data from their data policy.

## 4. Summary and discussion

When analysing the share of articles published in 37 journals with a DAP, we notice that almost three-fourths (452 out of 599) of all articles investigated are data-based and therefore fall under journals' data policies. More than a third (37.2%) of these 452 articles employs restricted data. These numbers underline the importance to implement suitable data policies for this type of papers (e.g. by obligating authors to provide useful references of the dataset(s) employed and to post the program code of their statistical analyses or simulations). For articles utilising non-restricted data, we find replication datasets for 36.6% of the articles and program code for 36.3%.

When more deeply examining 245 of the 452 data-based articles for compliance with the particular data policy of the journal, we find an overall compliance rate on article-level of 47.5%. Thus, the share of articles that are compliant with journals' respective data policy was less than half.

On article-level, we also ascertain that the compliance rate for papers utilising restricted data is almost 9 percent higher than for articles employing non-restricted datasets (52.6% to 43.7%). At first glance, this seems to be surprising. We suggest two possible explanations for this finding: The data policies' requirements for research based on restricted data are easier to fulfil than for articles using non-restricted data. Most often, only the program code and some information on the dataset(s) employed have to be submitted. Also authors might feel more confident to provide these files and some references on employed datasets, because they know their published findings will be replicated in exceptional circumstances only due to the unavailability of the data.

On journal-level, the average compliance rate is 38% - and therefore significantly lower than the rate reported by McCullough, McGeary and Harrison for their study of four top journals in economics in 2008. Our sample diverges into two groups: A minority of journals strictly enforces their data policies and achieves high compliance rates, while a larger group of journals does little in terms of ensuring data availability for published articles and fostering reproducible research. This majority of 10 out of 17 journals (58.8%) has a compliance rate of less than 20%. For eight journals, not a single article is compliant to the respective journal's data policy. However, for the interpretation of these findings we have to bear in mind that for some journals only a few data-based articles have been investigated.

Hence, when including more issues and articles of these journals in the analyses, we presume that the share of periodicals which do more or less pay lip-service to reproducible research diminishes.

In sharp contrast, four periodicals in our sample reach compliance rates of 100% or slightly less. These journals publish many articles per issue[16] and this is the only reason why our sample eventually reaches an average compliance rate of 47.5% on article-level. All of these journals have DAP that are in effect since at least spring 2012. Also, these journals are among the higher-ranked journals in economic research.

Our findings also provide evidence on the importance of mandatory data policies for journals when it comes to reproducible research. For 73 out of 100 articles (73%) that employed non-restricted data and have been published in journals with a mandatory DAP, accompanying datasets are available. For 72 (72%) papers we find program code. The overall compliance rate for these papers reaches 71%. In contrast, for papers published in journals with a voluntary DAP, we find only four out of 67 papers with accompanying datasets (6%). For three (4.5%) papers the programme code is available and the overall compliance rate reaches only 2%.

These findings reconfirm the outcome of theoretical studies that emphasise the missing incentives for researchers to support the replication of their own work. Also they illustrate why researchers like Chang and Li (2015) reported that articles published in journals with a mandatory DAP can be much easier replicated. Therefore, voluntary data policies do not seem to be a feasible instrument to foster replicable research. Nevertheless, also for journals with a mandatory DAP there is room for improvement: According to our findings, not all of these journals do continuously demand the replication files from their authors. For journals with a 'mandatory' data policy we would assume a compliance rate close to 100% - not 71%, as our findings suggest.

Considering that our investigation focusses only on the basic prerequisites for replications (because we did not try to reproduce the findings of the papers investigated), our predictions for the outcome of such replication attempts are not optimistic. If the fundamental prerequisites for replications are missing, how could a noteworthy share of these papers meet a basic criterion of the scientific method? Furthermore, we only checked whether authors followed the policies' requirements - regardless of whether we consider the policy itself to be robust or weak. Because not all of these policies can be regarded as 'robust' (cf. Vlaeminck & Herrmann, 2015a), the sheer compliance rate does not necessarily equate with the portion of articles that can be replicated.

We also would like to discuss some limitations of our analyses. A first one covers the data base of our sample for the compliance analysis. To investigate one or two issues of 17 journals is not sufficient to determine robust results for journals' general compliance with their data policies. Nevertheless our analysis provides a useful snapshot of journals' commitment towards reproducible research for the years 2013/2014. A second limitation refers to the fact that several journals in our sample publish much more articles per issue than others what results in disparate numbers of articles per journal in our sample. This also has consequences for the interpretation of journals' compliance rates. The overall compliance rate of 47.5% on article-level is only reached, because some journals with a high compliance rate publish many articles per issue. Future research on these topics should incorporate more issues per periodical to gain results which are based on a broader sample of articles. At the time of such a follow-up study, journals' data policies will be much longer in effect and journals will have

more experiences with enforcing their data policies. Whether journals' commitment towards reproducible research also grows over time due to these experiences remains an open question.

To conclude, we would like to give some policy recommendations for journals and suggest some measures for editorial offices to foster reproducible research. Based on the outcome of our analyses, we recommend that editorial offices of economic journals tighten their data policies towards mandatory DAPs that requires both data and program code. Also, journals should broaden their data policies on research which is based on restricted data. As our findings illustrate, more than a third of all papers investigated employ such data. To exempt these articles from journal's data policy would result in excluding research based on proprietary or confidential data from basic scientific requirements. Beyond, journals should be stricter in enforcing their data policies, because replicability of published research is a cornerstone of the scientific method. In the first place editors are accountable for enforcing journals' data policies, but also the reviewers should feel a responsibility to take care of a periodical's data policy. Both, editors and reviewers, invest time in ensuring that authors comply with journal's style sheet. To also invest efforts in ensuring that replication files are available according to journal's data policy is a task that would strengthen the scientific reputation of the periodical. This is even more important when we take into account, that first and foremost scientific journals have a central place as a quality instance in science and research. Thereby they are also crucial for promoting a culture of research integrity because published articles are the most visible output of research.

# References

ANDERSON, R.G., GREENE, W.H., MCCULLOUGH, B.D. & VINOD, H.D. (2008) The role of data/code archives in the future of economic research. *Journal of Economic Methodology*. 15. p. 99-119.

BJÖRK, B. & SOLOMON, D. (2013) The publishing delay in scholarly peer-reviewed journals. *Journal of Informetrics.* 7 (4). p. 914-923.

BRÄUNINGER, M., HAUCAP, J. & MUCK, J. (2011) Was lesen und schätzen Ökonomen im Jahr 2011?, *DICE - Ordnungspolitische Perspektiven* [Online] 18. Available from http://hdl.handle.net/10419/49023 [Accessed 2017-01-16]

CHANG, A.C. & LI, P. (2015) Is Economics Research Replicable? Sixty Published Papers from Thirteen Journals Say „Usually Not"'. *Finance and Economics Discussion Series* [Online] 2015 (83). p. 1-26. Available from http://doi.org/10.17016/FEDS.2015.083. [Accessed 2017-01-16]

CLEMENS, M.A. (2015) The Meaning of Failed Replications: A Review and Proposal. *IZA Discussion Paper* [Online]No. 9000. Available from http://hdl.handle.net/10419/110735. [Accessed 2017-01-16]

DEWALD, W.G., THURSBY J.G. & ANDERSON, R.G. (1986) Replication in Empirical Economics: The Journal of Money, Credit and Banking Project. *American Economic Review*. 76 (4). p. 587-603.

DUVENDACK, M., PALMER-JONES, R.W. & REED, W.R. (2015) Replications in Economics: A Progress Report. *Econ Journal Watch* [Online] 12 (2). p.164-191. Available from https://econjwatch.org/file_download/866/DuvendackEtAlMay2015.pdf. [Accessed 2017-01-16]

ELLISON, G. (2002) The Slowdown of the Economics Publishing Process. *Journal of Political Economy* [Online] 110 (5). p.947-993.

FEIGENBAUM, S. & LEVY, D.M. (1993) The market for (ir)reproducible econometrics. *Social Epistemology.* 7 (3). p.215-232.

GLANDON, P.J. (2011) Appendix to the Report of the Editor: Report on the American Economic Review Data Availability Compliance Project. A*merican Economic Review*. 101(3). p.695-699. Available from http://pubs.aeaweb.org/doi/pdfplus/10.1257/aer.101.3.684. [Accessed 2017-01-16]

HAMERMESH, D.S. (1997) Some thoughts on replications and reviews. *Labour Economics*. 4(2). p. 107-109. Available from https://doi.org/10.1016/S0927-5371(97)00015-8. [Accessed 2017-01-16]

HAMERMESH, D.S. (2007) Viewpoint: Replication in economics. *Canadian Journal of Economics/Revue canadienne d'économique* [Online] 40(3). p. 715-733. Available from http://doi.org/10.1111/j.1365-2966.2007.00428.x. [Accessed 2017-01-16]

HAMERMESH, D.S. (2013) Six Decades of Top Economics Publishing: Who and How? *Journal of Economic Literature*. 51 (1). p. 162-172.

HÖFFLER, J.H. (2017) ReplicationWiki: Improving Transparency in Social Sciences Research. *D-Lib Magazine* [Online] 23 (3/4). Available from https://doi.org/10.1045/march2017-hoeffler. [Accessed 2022-09-09]

KING, G. (1995a) Replication, Replication. *PS: Political Science and Politics.* 28 (3). p. 444-452.

KING, G. (1995b) A Revised Proposal, Proposal. *PS: Political Science and Politics.* 28 (3). p. 494-499.

LONGO, D.L., DRAZEN, J.M. (2016) Data Sharing. *The New England Journal of Medicine [Online]* 374. p. 276-277. Available from http://doi.org/10.1056/NEJMe1516564. [Accessed 2017-01-16]

MCCULLOUGH, B.D. (2007) Got Replicability? The Journal of Money, Credit and Banking Archive. *Econ Journal Watch*. 4 (3). p. 326-337.

MCCULLOUGH, B.D. (2009) Open Access Economics Journals and the Market for Reproducible Economic Research. *Economic Analysis and Policy*. 39 (1). p. 117-126.

MCCULLOUGH, B.D., MCGEARY, K.A. & HARRISON, T.D. (2006) Lessons from the JMCB archive. *Journal of Money, Credit, and Banking*. 38 (4), 1093-1107.

MCCULLOUGH, B.D., MCGEARY, K.A. & HARRISON, T.D. (2008) Do economics journal archives promote replicable research? *Canadian Journal of Economics/Revue canadienne d'économique*. 41(4). p. 1406-1420.

MIROWSKI, P. & SKLIVAS, S. (1991) Why econometricians don't replicate (although they do reproduce). *Review of Political Economy* [Online] 3(2). p. 146–163. Available from http://doi.org/10.1080/09538259100000040. [Accessed 2017-01-16]

MUELLER-LANGER, F. & ANDREOLI-VERSBACH, P. (2014) Open Access to Research Data: Strategic Delay and the Ambiguous Welfare Effects of Mandatory Data Disclosure. *Max Planck Institute for Innovation and Competition Research Paper* [Online] No. 14-09. Available from http://hdl.handle.net/10419/98845. [Accessed 2017-01-16]

SAVAGE, C.J. & VICKERS A.J. (2009) Empirical Study of Data Sharing by Authors Publishing in PLoS Journals. *PLoS ONE* [Online]. 4 (9). p. e7078. Available from http://doi.org/10.1371/journal.pone.0007078. [Accessed 2017-01-16]

TENOPIR, C., ALLARD, S. DOUGLASS, K., AYDINOGLU, A.U., WU, L., READ, E., MANOFF, M. & FRAME, M. (2011) Data Sharing by Scientists: Practices and Perceptions. *PLoS ONE* [Online] 6 (6). p. e21101. Available from http://doi.org/10.1371/journal.pone.0021101. [Accessed 2017-01-16]

VLAEMINCK, S. & HERRMANN, L.K. (2015a) Data Policies and Data Archives: A New Paradigm for Academic Publishing in Economic Sciences? In: Schmidt, B. & Dobreva, M. (eds). *New Avenues for Electronic Publishing in the Age of Infinite Collections and Citizen Science: Scale, Openness and Trust*. *Proceedings of the 19th International Conference on Electronic Publishing*. Amsterdam: IOS Press. p. 145-155. Available from http://hdl.handle.net/10419/121278. [Accessed 2017-01-16]

VLAEMINCK, S. & HERRMANN, L.K. (2015b) Data Policies and Data Archives: A New Paradigm for Academic Publishing in Economic Sciences? (Replication Data). Version: 1. Dataset. Available from http://journaldata.zbw.eu/dataset/data-policies-and-data-archives-a-new-paradigm-for-academic-publishing-in-economics. [Accessed 2017-01-16]

## Acknowledgements:

## Appendix: Replication files and documentation

The replication files for this paper are available here: http://journaldata.zbw.eu/dataset/journals-in-economic-sciences-paying-lip-service-to-reproducible-research-replication-data.

# End-notes

[1] Sven Vlaeminck was the project manager of the project 'European Data Watch Extended' (EDaWaX) and works for ZBW – German National Library for Economics / Leibniz Information Centre for Economics in Hamburg/Germany. Since 2011, he is active in the field of research data management. He can be reached by email: s.vlaeminck@zbw.eu.

[2] Felix Podkrajac is currently trainee in academic subject librarianship at the Library and Information System of the Carl von Ossietzky University of Oldenburg. Email: felix.podkrajac@uni-oldenburg.de.

[3] We define an article to be 'data-based' when the methods include creating, reusing, and analysing research data, other empirical methods (e.g. experiments) or self-compiled program code (e.g. simulations).

[4] The growing importance of research data, its management and potentially also its integration in the academic publishing process is reflected in numerous statements and requirements of universities, learned societies, funding agencies and even political bodies. The NSF (US), the ESRC (UK) and the RCUK (UK) all emphasise that data collected with public funds belongs in the public domain and highlight the advantages of data sharing: 'Data sharing strengthens our collective capacity to meet scientific standards of openness by providing opportunities for further analysis, replication, verification and refinement of research findings' (National Science Foundation, 2010). Beside policies on data sharing for publicly funded research, others also explicitly mention publication-related research data. For instance, the European Commission (EC) recommends that EU member states ought to implement policies to ensure that 'datasets are made easily identifiable and can be linked to other datasets and publications through appropriate mechanisms' (European Commission, 2012).

[5] According to King (1995b) the primary purpose of a data archive is not necessarily to ensure replicability (this would be a high demand) but to enhance extensibility of published research (this presumes replicability).

[6] The data availability policy of the American Economic Review (AER), for instance, lists varying requirements for econometric and simulation papers and for experimental papers. For econometric and simulation papers, the 'minimum requirement should include the data set(s) and programs used to run the final models, plus a description of how previous intermediate data sets and programs were employed to create the final data set(s).' For experimental research, 'we normally expect authors of experimental articles to supply the following supplementary materials […] 1. The original instructions. […] 2. Information about subject eligibility or selection. […] 3. Any computer programs, configuration files, or scripts used to run the experiment and/or to analyze the data. […] 4. The raw data from the experiment.'

[7] According to the UC Berkley, 'restricted information describes any confidential or personal information that is protected by law or policy'. Also data provided by third parties, that 'may not be redistributed or reused without the consent of the original provider' can be characterised as restricted data, as the World Bank points out.

[8] The median of articles published by journals in our full sample is 13 (equates 2.2%) (means: 16.2/2.7%).

[9] Nevertheless, nine articles remain for which we cannot finally determine whether they are data-based. These articles are removed from the sample and are no longer regarded in our analysis.

[10] For instance, this applies to the data availability policy of the Journal of Money, Credit and Banking (JMCB).

[11] For an overview on how journals provide replication data to would-be replicators, please consult Vlaeminck & Herrmann (2015a).

[12] Exemplary, we provide two good examples for a useful documentation and well-structured replication files for research based on restricted data: (i) http://restud.oxfordjournals.org/content/suppl/2013/09/28/rdt030.DC1/Supplementary.zip (Supplement to doi: 10.1093/restud/rdt030) (ii) http://repec.wirtschaft.uni-giessen.de/~repec/RePEc/jns/Datenarchiv/v234y2014i1/y234y2014i1p5_22/ (Supplement to doi: 10.1515/jbnst-2014-0103)

[13] The median of articles by journal in the 'compliance sample' was 9 or 3.7% (means: 14.4/5.9%).

[14] Five cases are removed from the compliance analysis's sample because the journals exempt papers based on restricted data from their data policy. Therefore, only 73 articles that rely on restricted data have been examined.

[15] Five cases are removed from the compliance analysis's sample because the journals exempt papers based on restricted data from their data policy.

[16] 130 out of 245 (53.1%) articles in the compliance analysis have been published by just four journals: The Journal of the American Statistical Association (53), Review of Economics and Statistics (37), Review of Economic Studies (20) and American Economic Review (20).