Research Reproducibility in Political Science

Dr Nicole Janz
School of Politics and IR
More than half of psychology papers are not reproducible

Initiative to replicate findings of 100 prominent studies raises further questions about health of discipline

August 27, 2015

By Paul Jump  Twitter: @PaulJump
The Case of the Amazing Gay-Marriage Data: How a Graduate Student Reluctantly Uncovered a Huge Scientific Fraud

By Jesse Singal
How can we establish trust in science?
<table>
<thead>
<tr>
<th>Data transparency</th>
<th>Analytic transparency</th>
<th>Production transparency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Providing full access to data itself</td>
<td>Information about data analysis</td>
<td>Process of data collection</td>
</tr>
</tbody>
</table>

www.dartstatement.org
<table>
<thead>
<tr>
<th>Data transparency</th>
<th>Analytic transparency</th>
<th>Production transparency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Providing full access to data itself</td>
<td>Information about data analysis</td>
<td>Process of data collection</td>
</tr>
</tbody>
</table>

**Quantitative research**

Upload datasets used for analysis

Code for models (SPSS, STATA, R)

Good methods section in paper

Provide or describe raw data & variable codings

(see Lupia/Elman 2014; Moravcsik 2014)
University training
Replication exercises
Bringing the Gold Standard into the Classroom: Replication in University Teaching

Nicole Janz
University of Cambridge

Reproducibility is held to be the gold standard for scientific research. The credibility of published work depends on being able to replicate the results. However, there are few incentives to conduct replication studies in political science. Replications are difficult to conduct, time-consuming, and hard to publish because of a presumed lack of originality. This article sees a solution in a profound change in graduate teaching. Universities should introduce replications as class assignments in methods training or invest in new stand-alone replication workshops to establish a culture of replication and reproducibility. This article will
Why should you replicate?

Learn Statistics

- Real life data
- Author decisions
- Bugs included
- More fun than textbook

Reproducibility routine

- When are published results really reproducible?

Publish

- Add value
- Publish faster
Use terminology accepted in your field

Political Science (see King 2003)

<table>
<thead>
<tr>
<th>Duplication</th>
<th>Replication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verify research results</td>
<td>Test the robustness of the original research results</td>
</tr>
<tr>
<td>exact same data set</td>
<td>new data</td>
</tr>
<tr>
<td>exact same methods</td>
<td>new models</td>
</tr>
</tbody>
</table>
Practical steps in a replication study

1. Select paper
2. Access data & code
3. Identify each variable
4. Reproduce tables, figures
5. Compare

If you got to this point, you completed a duplication.
Practical steps in a replication study (II)

6 Add value
   • new data
   • new variables
   • new model specifications
   • theoretical contributions
   
7 Compare

8 Get feedback from peers

9 Journal submission

You now completed a full replication!
Comparing your results with the original study

Clarify with precision the extent to which you were able to replicate the author’s results.

Gary King (2006)

- Exact same data and methods: results cannot be duplicated

- New data, experiments, models, methods: describe exactly at which step the results changed and why

- Different measures of a concept can naturally yield different results

- Different lab conditions may lead to different results
Communicating failed replications

Be professional!
What replicators write

“We ... find that coding errors, selective exclusion of available data, and unconventional weighting of summary statistics lead to serious errors” (Herndon et al. 2013)

“If we cannot even reproduce the original results using the same publicly available data, there is no need for further commentary.” (Miller et al, 2001)
How original authors often respond

“less realistic”, “inconsistent with the substantive literature,” and “of limited utility” (Mansfield, Milner, and Rosendorff 2002)

“fundamentally flawed” (Peffley, Knigge, and Hurwitz 2001)

“statistical, computational, and reporting errors that invalidate its conclusions” (Gerber and Green 2005:301).
Our estimation approach builds off of the methodology and data used by Gomez et al. (2007) ..., adding measures of electoral closeness in order to focus on how the randomly assigned cost (rain) has a different impact depending on the electoral environment.
... we analyze a dyad-year data set (used by Rauchhaus 2009) to examine whether existing findings on the effect of symmetric nuclear weapons possession on conflict are robust to the improvements noted above. We find that once pre-nuclear dyadic conflict is controlled for, symmetric nuclear dyads are not more likely to experience low-level conflict.
Can inflation expectations be measured using commodity futures prices?
Rafael Saladailli1, a, b, D’Maris Coffman1

1 Department of Economics, Southern Methodist University, Dallas, TX, USA
2 Department of Economics, University of Texas at Austin, Austin, TX, USA

ABSTRACT
This paper examines the use of CT scans to measure prices of the drug.

1. Introduction

In economics, there is a growing consensus on the importance of inflation expectations. This consensus is based on the idea that inflation expectations influence the behavior of households and firms, and thereby affect the overall economy. However, there is a disagreement on the best way to measure inflation expectations. Some economists argue that consumer price indexes (CPI) and producer price indexes (PPI) are adequate measures of inflation expectations. Others argue that forward-looking indicators, such as commodity futures prices, are more appropriate measures. This paper aims to contribute to this debate by examining the usefulness of commodity futures prices as a measure of inflation expectations.

2. Literature Review

Previous studies have used commodity futures prices as a measure of inflation expectations. However, these studies have used different commodity contracts and different time periods. This paper uses the Corn futures contract to measure inflation expectations. The Corn futures contract is a fairly liquid and widely traded contract. It is also a good proxy for inflation expectations because the corn market is sensitive to changes in the overall economy.

3. Methodology

The methodology of this paper is based on the work of Bekaert and Gürkaynak (2006). They show that commodity futures prices can be used to forecast inflation. This paper extends their work by using the Corn futures contract to measure inflation expectations.

4. Results

The results show that the Corn futures contract is a good measure of inflation expectations. The contract is positively correlated with the CPI and PPI, which suggests that the contract is a good proxy for inflation expectations.

5. Conclusion

In conclusion, this paper shows that commodity futures prices, specifically the Corn futures contract, can be used to measure inflation expectations. This finding has implications for policymakers who rely on commodity futures prices as a measure of inflation expectations. It also has implications for investors who use commodity futures prices to forecast inflation.

Keywords: Inflation expectations, commodity futures prices, Corn futures contract, CPI, PPI.
Transparent workflows
4 Steps to Transparency

1. Fixed Folder Structure
2. Comment your code
3. Clear methods section in paper
4. Share materials
Fixed folder structure

Decide on a template structure for each project

Never alter raw data
Main

Readme.txt
Paper

/Data

/Code

On the fly....
/Results
/Materials
FIGURE 4.1: Example Research Project File Tree

- **Root**
  - **ExampleProject**
    - **Analysis**
      - Analysis1.R
      - Analysis2.R
    - **Paper.Rnw**
    - **Slideshow.Rnw**
    - **Website.Rmd**
    - **Main.bib**
    - **Packages.bib**
    - **README.md**
  - **Data**
    - **Makefile**
    - **Gather1.R**
    - **Gather2.R**
    - **MergeData.R**
    - **MainDataVariableDescriptions.md**
    - **MainData.csv**
# loading data

# variable transformation

# merging tables

# models for table 1
doedit "Z:\M14160\assessment\self data\data\stata operation-day 2-binary\EQ-I > Q-command-LR.do"

.do "C:\Users\lxwh\AppData\Local\Temp\STD02000000.tmp"

.use EQIQbinaryLR.dta,clear

.end of do-file

.do "C:\Users\lxwh\AppData\Local\Temp\STD02000000.tmp"

.logit BEQ2 gender representative age IQ

Iteration 0:  log likelihood = -157.28934
Iteration 1:  log likelihood = -155.59433
Iteration 2:  log likelihood = -155.59392
Iteration 3:  log likelihood = -155.59392

Logistic regression

Number of obs = 227
LR chi2(4) = 3.39
Prob > chi2 = 0.4947
Log likelihood = -155.59392 Pseudo R2 = 0.0108
load("replicationdata.Rdata") # this data file needs to be in the folder of your working directory
t.nono <- replicationdata[replicationdata$oecd==0, ] # select developing nations only (non-OECD)
t.nono <- droplevels(t.nono) # drop the unused levels (http://stackoverflow.com/questions/17217951/how-can-i-drop-unused-levels-from-a-data-frame)

# Main models for paper
# Creating table "Total FDI and Personal Integrity Rights Protection (1983-2010)" in the main part of the paper ####
# US FDI ologit with robust standard errors clustered by country
CIRI.logit.LDV <- lrm(CIRI_PHYSINT ~ Lag_lognonnegUS_fdi_totalpGDP + Lag_logtrade + Lag_logGDPPc + Lag_logpopulation + Lag_polity2 + Lag_conf1 + Lag_CIRI_PHYSINT,data=data.CIRI.PHY.LDV,x=TRUE, y=TRUE)
CIRI.logit.LDV.corr <- robcov(CIRI.logit.LDV, data.CIRI.PHY.LDV$country) #robust errors clustered by country

PTS.logit.LDV <- lrm(PTS_ai_reversed ~ Lag_lognonnegUS_fdi_totalpGDP + Lag_logtrade + Lag_logGDPPc + Lag_logpopulation + Lag_polity2 + Lag_conf1 + Lag_PTS_ai_reversed,data=data.PTS.LDV,x=TRUE, y=TRUE)
PTS.logit.LDV.corr <- robcov(PTS.logit.LDV, data.PTS.LDV$country)

###########################################################################
#Creating "Correlation matrix (I) to (III)
###########################################################################
corstars <- function(x){
x <- as.matrix(x)
R <- rcorr(x, type="spearman")$r
p <- rcorr(x)$P
# n <- rcorr(x)$n
mystars <- ifelse(p < .01, "***", ifelse(p < .05, "** ", ifelse(p < .1, "* ", " ")))
R <- format(round(R, 3))
Rnew <- matrix(paste(R, mystars, sep=""), ncol=ncol(x)) #removed n here
1 Clinical data

1.1 Data overview

First we load the data and R packages. The data file is part of the paper supplement, and we have also made a copy available online.

```r
library(survival)
library(kernlab)
library(mrs)
library(spatstat)
library(RColorBrewer)
library(gplots)
```

```r
# load data file from local copy or from URL
if (file.exists("Schwarz2015-supplement.Rdata")) {
  load("Schwarz2015-supplement.Rdata")
} else {
}
```

The first object in the .Rdata file is a table \texttt{D} with patient information:

\begin{verbatim}
# D  
#  
# 003-01 1 4.790231 1.2025074 511 271 1 1 H2OC 47 IV <1cm
# 003-02 2 NA 0.7106901 977 363 1 1 H2OC 62 IV >1cm
# 003-03 3 3.785366 1.2932829 209 153 1 1 H2OC 69 IV >1cm
# 003-04 4 5.801712 1.4705535 547 303 1 1 H2OC 63 IV >1cm
# 003-05 5 6.0558896 0.7298828 744 298 1 1 H2OC 59 IV >1cm
# 003-10 6 6.6588596 0.7298828 744 298 1 1 H2OC 63 IV >1cm
# 003-13 7 3.0000090 0.6839691 1887 358 1 1 H2OC 67 IV >1cm
# 003-17 8 3.422117 2.2387817 809 373 1 1 H2OC 51 IV >1cm
# 003-20 9 4.487826 0.6493453 1278 563 1 1 H2OC 71 IV >1cm
# 003-21 10 4.715946 0.8008039 1139 303 1 1 H2OC 60 IIIC >1cm
# 003-22 11 5.707707 0.8340685 1588 382 1 1 H2OC 58 IIIC >1cm
# 003-24 NA NA 1986 634 1 1 H2OC 58 IIIC >1cm
# 003-24 NA NA 316 76 1 1 H2OC 60 IIIC >1cm
# 003-25 13 NA 0.6212971 1166 776 1 1 H2OC 57 IIIC >1cm
# 003-26 14 4.621984 0.6033119 1513 601 1 1 H2OC 63 IIIC >1cm
# 003-21 15 0.741273 706 332 0 0 H2OC 54 IV <1cm
# 003-27 16 NA NA 1408 1408 0 0 H2OC 58 IIIC >1cm
# 003-30 17 NA 0.8591206 949 203 0 0 H2OC 60 IIIC >1cm
# 003-31 NA NA 1500 527 0 0 H2OC 60 IIIC >1cm
#```

Rownames correspond to sample identifiers. Columns indicate the patient identifier (\texttt{Nr}), tumor type (\texttt{TH}), clinical stage (\texttt{ST}), overall survival in days (\texttt{OS}), progression-free survival (\texttt{PFS}) and its status (\texttt{dead}), indicators for survival (\texttt{dead}) and progression (\texttt{dead}), covariates for age, stage (ordered factor), residual disease after debulking surgery and ordered factor (\texttt{Stage residual}).
Describe methods clearly

Name exact models with citations for statistical choices

Footnote should contain software versions

If space is restricted: Appendix
Models

For the ordered categorical outcome variables, CIRI and PTS, I estimate an ordered logit model with robust standard errors clustered on country to correct for heteroskedasticity. For the continuous outcome variable, the Latent Human Rights Scores by Fariss, I employ ordinary least squares (OLS) with panel-corrected standard errors (PCSE).\textsuperscript{17} In both models, I include a lagged dependent variable (“Past”) since countries that repressed their citizens in the past are more likely to use repressive acts in the future (Gurr 1988). A lagged dependent variable also corrects the serial correlation (Beck and Katz 1995, 2009).

I include a one-year lag between the outcome and predictors to allow the effect of FDI stock to spread in the country. This means that the accumulated FDI in a country in a given year is expected to correlate with rights protection in the following year, which establishes a time order and suggests a direction of causality from FDI\((t-1)\) to rights protection\((t)\).

The data set ranges from 1983 to 2010 and includes up to 121 non-OECD countries. The selection of these cases is limited to countries with available data on personal integrity rights and FDI measures (see a list of all countries in the online appendix). Two main sets of models are estimated: The first set of models includes total FDI to compare my results with previous work, while the second set of models replaces total FDI with investment in 10 industry sectors.

\textsuperscript{17} For the ordered logit models, I use the functions lrm() and robcov() from the R package “rms” Version 4.3–0, which produces the same results as the corresponding STATA command ologit with the cluster() modification (STATA Version 13.0). For OLS with PCSE, I use the functions plm() and vcovBK() from the R package “plm” Version 1.3.1, which produces the same results as the STATA command xtpcse with the pairwise specification.
Share your materials

Readme file

Data, code, variable codebook

Information to reconstruct data from original sources
Data sharing platforms

- Github
- OSF
- UK Data Service
- Inter-university Consortium for Political and Social Research
- ICPSR
- Harvard Dataverse
- CISSER
Reproduction Data for: "Foreign Direct Investment and Repression: An Analysis Across Industry Sectors" (Version 1.0)


Description

The impact of foreign direct investment (FDI) on repression in developing nations is still disputed. Some argue that FDI improves economic development and exports human rights values. Others criticize the exploitation of cheap labor and resources, which may lead to tensions and government oppression. Previous studies have employed aggregate FDI data with conflicting results. Alternatively, I propose that the effects depend on what kind of FDI enters a country. I build a sectoral framework to discuss how skills and technology levels, as well as the motivation for FDI, can mediate the impact. I then examine the link in a panel data analysis (1983-2010) in 121 countries, integrating sectoral FDI in several resource, manufacturing, and service industries. The results show that investment in high-skilled and high-tech sectors has positive effects. The results are robust across several measures for repression, and when accounting for sector size, regional and time effects.

Subject

Social Sciences

Keyword

human rights, foreign direct investment, repression, sectors
<table>
<thead>
<tr>
<th>Code</th>
<th>Data</th>
<th>Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>JHR_Replciliation.R</td>
<td>replicationdata.tab</td>
<td>Variable_Codebook.xlsx</td>
</tr>
<tr>
<td>JHR_Replciliation_Notes.pdf</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Replication Notes for
Foreign Direct Investment and Repression:
An Analysis Across Industry Sectors

Nicole Janz University of Nottingham
School of Politics and International Relations
University Park
Nottingham, NG7 2RD
Email: nicole.janz@nottingham.ac.uk

September 14, 2016
Instructions for Replication

- First, load the data. This is a panel data set with country-years. Each table or figure is identified by its header as in the article or online appendix. If you are looking for a specific table, search the (admittedly long) Rscript for that particular heading.

- Run the code chunk by chunk. Many variables are created 'on the go' and re-used at a later point. Make sure to run the code in the same order, and run all the code even if you only want to replicate e.g. the last table.

- In order to preserve the largest possible sample size, I have created data sets for each model separately (based on the original table that you have loaded in R). This way, depending on the particular model, slightly different country-years are included. Make sure to create all these data sets (again, in order of the code) to run the models.
Citation

If you work with the data for your own study (replication or original work based on these data), please cite my article as well as the data set. It would be great if you could let me know about your results. A suitable citation of the data is provided by Dataverse where you downloaded the data.

Contact

Please contact me if you have any questions about the study or replication files at: nicole.janz@nottingham.ac.uk or nicolejanz@gmail.com. Any feedback on your replication attempts is more than welcome.

Acknowledgements

Gu Li and Sergio Cuesta have verified that the provided Rscript produces the tables and figures for the study. All errors remain my own.
Quantitative research

Data transparency
Providing full access to data itself

Analytic transparency
Information about data analysis

Production transparency
Process of data collection

Upload datasets used for analysis
Code for models (SPSS, STATA, R)
Provide or describe raw data & variable codings

Good methods section in paper

(see Lupia/Elman 2014; Moravcsik 2014)
Replication Exercises

Transparent workflows
Five Selfish Reasons to Share Data
<table>
<thead>
<tr>
<th>Quality</th>
<th>Establish trust &amp; credibility in your work</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reputation</td>
<td>Be known as a transparent researcher in your field</td>
</tr>
<tr>
<td>Citation</td>
<td>Your data will be cited</td>
</tr>
<tr>
<td>Consistency</td>
<td>Transparent workflow makes it easier to re-use your own data later</td>
</tr>
<tr>
<td>Practicality</td>
<td>Meet journal &amp; funder standards</td>
</tr>
</tbody>
</table>

(Markowetz 2015)
Thank you!

nicole.janz@nottingham.ac.uk

@polscireplicate

Political Science Replication Blog
Qualitative research

Provide (partial)
- Interview transcripts
- Field notes
- Videos…

Describe which evidence supports which claims
- Discursive footnotes / supplement

Explain how data were collected:
- Interviewee selection
- Participants
- Documents…
Literature on Replication


• Open Science Framework. *Transparency and Openness Promotion (TOP) Guidelines* [https://cos.io/top/](https://cos.io/top/)

• TIER Documentation Protocol [https://www.haverford.edu/project-tier/protocol-v2](https://www.haverford.edu/project-tier/protocol-v2)
