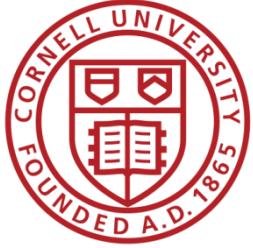


# Replication and Reproducibility in Social Sciences and Statistics: Context, Concerns, and Concrete Measures

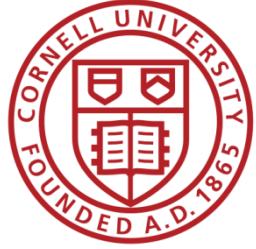
Lars Vilhuber  
Cornell University

Partial funding acknowledged under NSF-#1131848 (NCRN) and a grant from the Alfred P. Sloan Foundation. The opinions expressed in this talk are solely the authors, and do not represent the views of the U.S. Census Bureau, the American Economic Association, or any of the funding agencies.

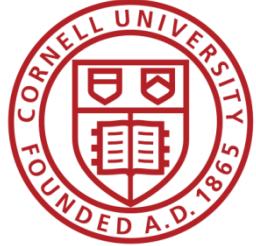


# Overview

- Crisis, what crisis?
- A brief history of reproducibility and replicability
- What seems different today?
- Economics and other social sciences
- What are we doing in Economics? (at the AEA, in general)
- What are we going to do in Economics?
- What will you be doing going forward?



# Introducing: The Reproducibility Crisis



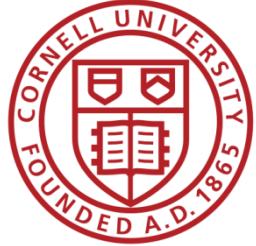
# Replicability in Economics

- Ragnar Frisch, editor of the first issue of Econometrica (1933), noted

***“the original data will, as a rule, be published, unless their volume is excessive... to stimulate criticism, control, and further studies.”***

- Publication of data, however, was discontinued early in the journal's history.

*Frisch (1933)*



# The “crisis” in the 60s and 70s

Sterling, 1959; Cohen, 1962; Lykken, 1968; Tukey, 1969;  
Greenwald, 1975; Meehl, 1978; Rosenthal, 1979

Low power

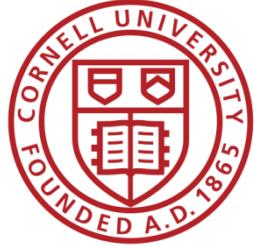
Flexibility in analysis

Selective reporting

Ignoring nulls

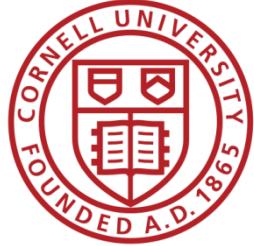
Lack of replication

Misuse of statistics



# The crisis in the 1980s in Economics

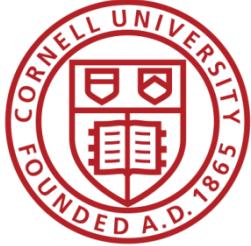
- Dewald, Thursby, Anderson (AER, 1986)
- Leamer's (1983) 'Let's take the con out of econometrics'



# Citing data sources

“Many authors cited only general sources such as *Survey of Current Business*, *Federal Reserve Bulletin*, or *International Financial Statistics*, but did not identify the specific issues, tables, and pages from which the data had been extracted.”

(Dewald Thursby Anderson **1986**, p. 591)

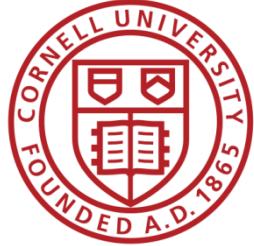


# Making USEFUL archives

- From analysis of code from 1996 to 2003 (MMH2006):

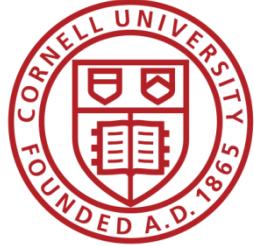
“Other authors seem to think that the entire world shares the exact same hard drive layout, with “C:\MYDATA\MYPROJECT\” **sprinkled liberally** throughout their code. Of course, a would-be replicator has to **find and change all these**.”

“The author might not realize all the data/subroutine files that his code utilizes, and **forget to include** said data/subroutine in his replication files.”



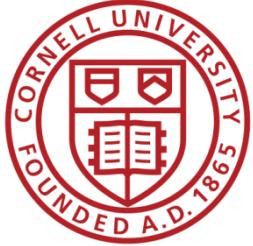
# The crisis in the 1980s ... lead to changes

- 1976- JPE Section on Confirmation and Contradictions
- 1988 Replication archive at the Journal of Applied Econometrics



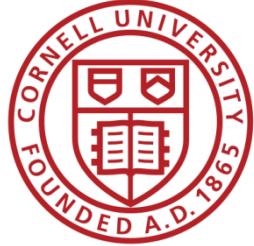
# The crisis in the 1990s

- Anderson and Dewald (1994) [Economics]
- King (1995) “Replication, replication” [Political Science]
- McCloskey and Ziliak, 1996: Do economists consider power calculations?
- McCullough and Vinod (1999): call for replication archives



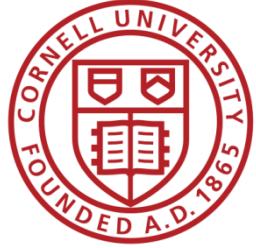
# The crisis in the 1990s... lead to changes

- Lead directly to
  - Data availability policies in Economics (1996: JCMB)
  - Replication section (Journal of Applied Econometrics, 2003)
  - More robust code archives  
(Goffe and Parks (1997), Eddelbüttel (1997): CodEC -> SSC)
- Coincidental:
  - Better public-use data (IPUMS: Sobek and Ruggles, 1999)
  - Better open-source software (R Core Team, 2000)
  - New methods of exchanging manuscripts  
(early 90s, but publications: Ginsparg, 1997: arXiv; Krichel, 1997, Halpern 1998)



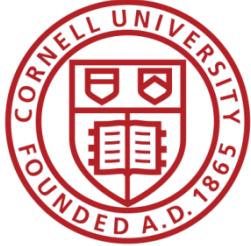
# The crisis in the 2000s

- McCullough & Vinod (2003) [Econometrics]
- Ziliak and McCloskey (2004)
- Anderson et al (2005, 2008), Vinod (2005)
- Ioannides (2005)
- McCullough et al (2006) Lessons learned JMBC, McCullough (2007)
- Hamermesh (2007)
- Koenker and Zeileis (2009)



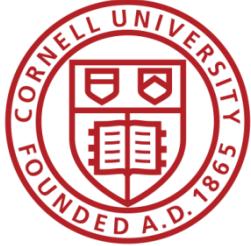
# The crisis in the 2000s... lead to changes

- Lead directly to
  - Data availability policies in Economics (2005: AER)
  - Subsequently to many other top journals
  - Social Science Registry (2011) for pre-registration



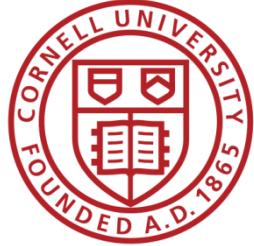
# The newest crisis

- Ioannides (2016)
  - Ioannidis, Stanley, Doucouliagos (2017): bias, low power
  - Miguel et al (2014)
  - Chang & Li (2015)
  - Zimmermann (2015)
  - Many others...
- “statistically significant, novel, and theoretically tidy results are published more easily than null, replication, or perplexing results”



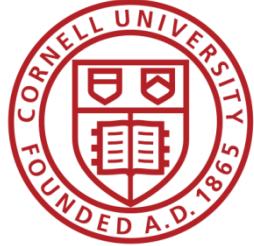
# But is it a “crisis”?

- crisis [kri'sis] (pl. *cri'ses*) (*L.*)
  1. the **turning point** of a disease for better or worse; especially a sudden change, usually for the better, in the course of an acute disease.
  2. a **sudden paroxysmal intensification** of symptoms in the course of a disease.



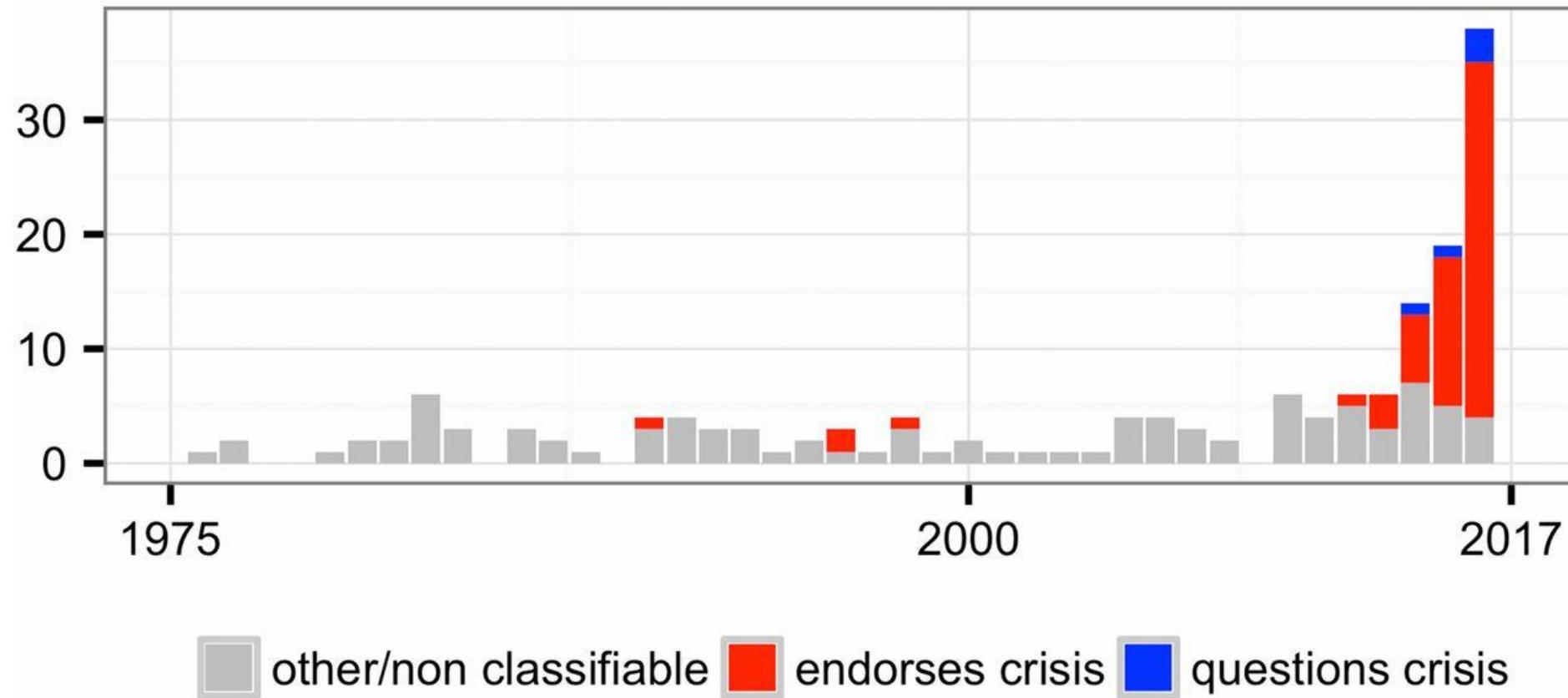
# Scholarly Communications are Meant to Communicate

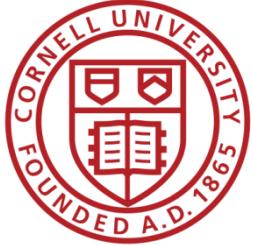
- In the 17<sup>th</sup> century, scholars communicated by letter exchange
- Collections of letters + editing became the first journals
- Peer review was incorporated
- But the review of a manuscript by an editor and two peers cannot possibly be equivalent to a full vetting
- Rather, the back and forth in the pages of academic journals IS the discourse
- In this view (also Hamermesh, 2007) bad publications (including those that do not privately replicate) die out, good publications withstand those tests



# Fanelli (2018), Is science really facing a reproducibility crisis, and do we need it to?

## Frequency of Crisis Narrative in Web of Science Records





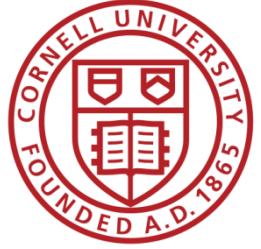
Fanelli (2018), Is science really facing a reproducibility crisis, and do we need it to?

In fact, Fanelli finds that

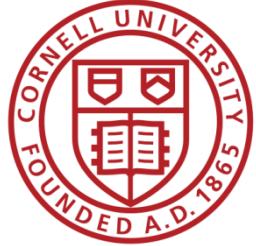
“researchers that publish at **higher frequency**, in journals with **higher impact factor**, and in countries where **pressures to publish are high**,

are

**equally or more likely** to **correct** their work, **less likely** to publish papers that are **retracted**, **less likely** to author papers that contain **duplicated** images, and **less likely** to author papers reporting **overestimated** effects”

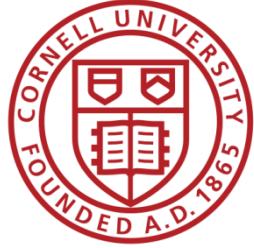


Maybe it's only a tantrum...



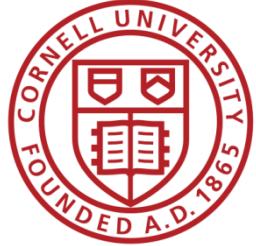
# Efficiency of scholarly discourse?

- If publishing is part of a discourse (and not the final truth), then it must be made efficient
- When publication was exchange of ideas, the exchange of papers (1600s!) included everything necessary



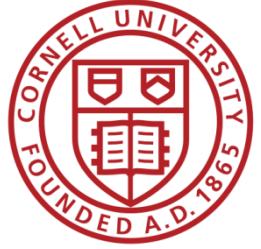
# Efficiency of scholarly discourse?

- Early publications (20<sup>th</sup> century) contained tables of data, and the math was simple (maybe)
- Data became electronic, was no longer **included** or **cited**
- Modern publications thus need the same transparency and completeness

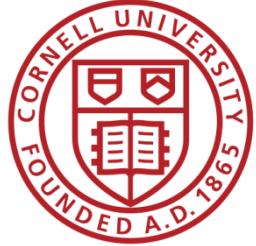


# Efficiency of scholarly discourse!

- Reproducibility helps to teach others
- Data became electronic, was no longer **included** or **cited**
- Modern publications thus need the same transparency and completeness

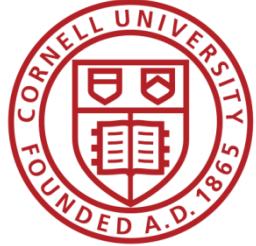


# What IS reproducibility?



# Replication continuum (Bollen et al. 2015)



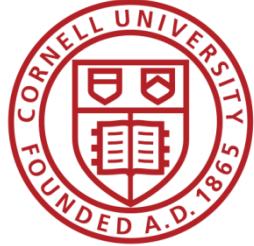


# Replication continuum



## **Reproducibility**

- Narrow Replication (Pesaran 2003)
- Pure Replication (Hamermesh  
2007)
- Verification (Clemens 2015)



# Replication continuum

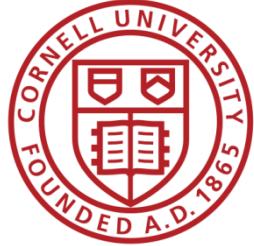


## **Reproducibility**

- Narrow Replication (Pesaran 2003)
- Pure Replication (Hamermesh 2007)
- Verification (Clemens 2015)

## **Replicability**

- Wide Replication (Pesaran 2003)
- Statistical Replication (Hamermesh 2007)
- Reproduction/Reanalysis (Clemens 2015)



# Replication continuum



## **Reproducibility**

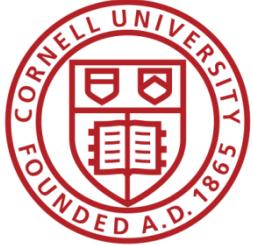
- Narrow Replication (Pesaran 2003)
- Pure Replication (Hamermesh 2007)
- Verification (Clemens 2015)

## **Replicability**

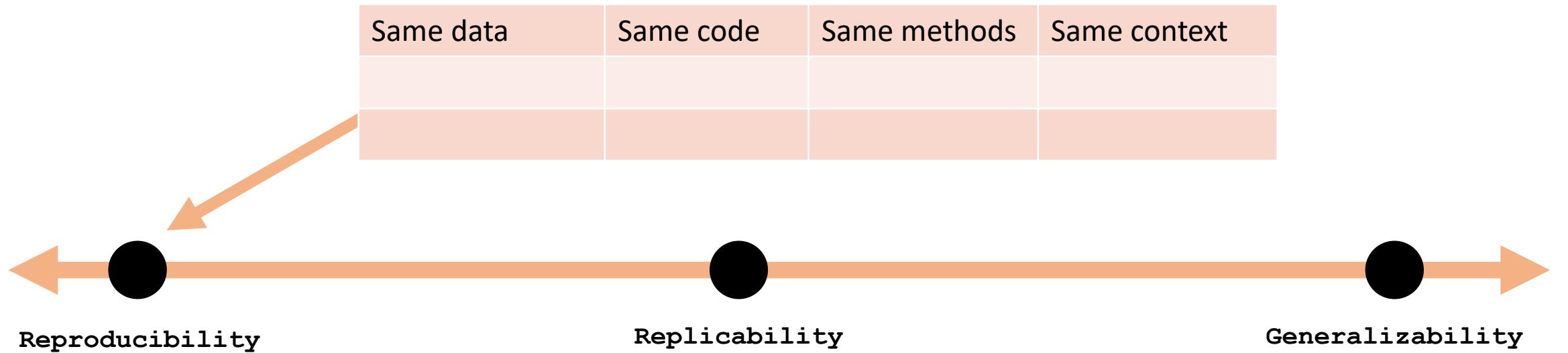
- Wide Replication (Pesaran 2003)
- Statistical Replication (Hamermesh 2007)
- Reproduction/Reanalysis (Clemens 2015)

## **Generalizability**

- Wider Replication (Pesaran 2003)
- Scientific Replication (Hamermesh 2007)
- Reanalysis/Robustness (Clemens 2015)



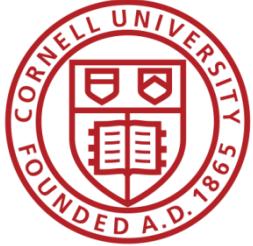
# Replication continuum



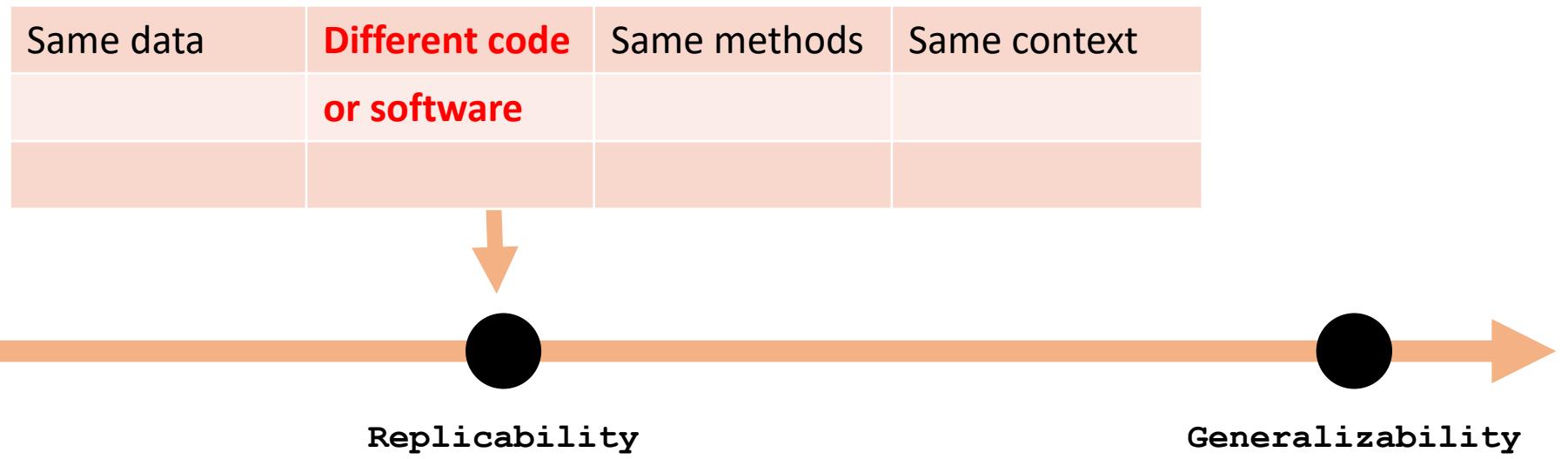
- Narrow Replication (Pesaran 2003)
- Pure Replication (Hamermesh 2007)
- Verification (Clemens 2015)

- Wide Replication (Pesaran 2003)
- Statistical Replication (Hamermesh 2007)
- Reproduction/Reanalysis (Clemens 2015)

- Wider Replication (Pesaran 2003)
- Scientific Replication (Hamermesh 2007)
- Reanalysis/Robustness (Clemens 2015)



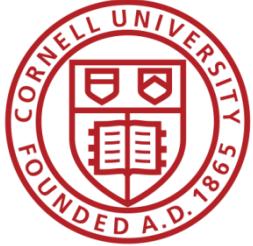
# Replication continuum



- Narrow Replication (Pesaran 2003)
- Pure Replication (Hamermesh 2007)
- Verification (Clemens 2015)

- Wide Replication (Pesaran 2003)
- Statistical Replication (Hamermesh 2007)
- Reproduction/Reanalysis (Clemens 2015)

- Wider Replication (Pesaran 2003)
- Scientific Replication (Hamermesh 2007)
- Reanalysis/Robustness (Clemens 2015)



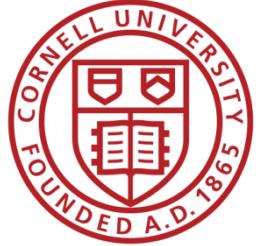
# Replication continuum



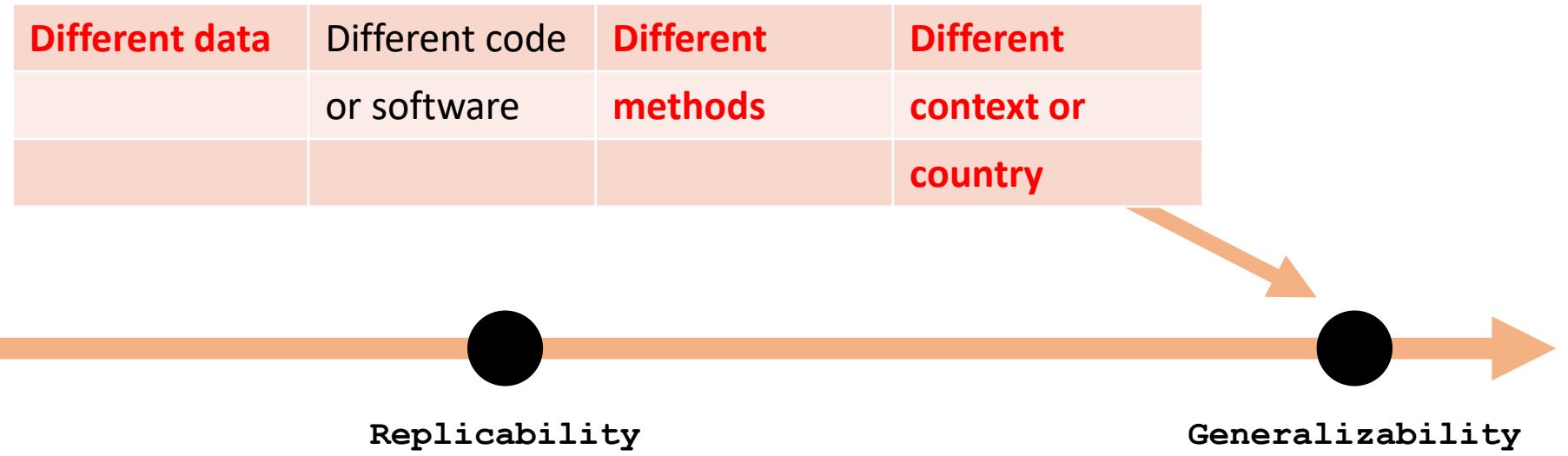
- Narrow Replication (Pesaran 2003)
- Pure Replication (Hamermesh 2007)
- Verification (Clemens 2015)

- Wide Replication (Pesaran 2003)
- Statistical Replication (Hamermesh 2007)
- Reproduction/Reanalysis (Clemens 2015)

- Wider Replication (Pesaran 2003)
- Scientific Replication (Hamermesh 2007)
- Reanalysis/Robustness (Clemens 2015)



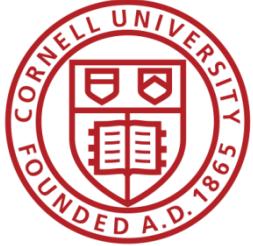
# Replication continuum



- Narrow Replication (Pesaran 2003)
- Pure Replication (Hamermesh 2007)
- Verification (Clemens 2015)

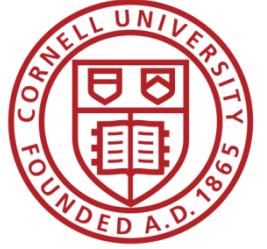
- Wide Replication (Pesaran 2003)
- Statistical Replication (Hamermesh 2007)
- Reproduction/Reanalysis (Clemens 2015)

- Wider Replication (Pesaran 2003)
- Scientific Replication (Hamermesh 2007)
- Reanalysis/Robustness (Clemens 2015)

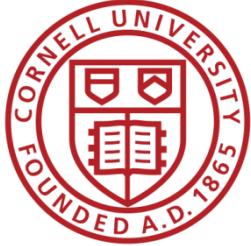


# What is a “Data Availability Policy”?

- It is the policy of the American Economic Association to publish papers only if the data used in the analysis are clearly and precisely documented and are readily available to any researcher for purposes of replication.
- Authors of accepted papers that contain empirical work, simulations, or experimental work must **provide**, prior to publication, the **data, programs, and other details of the computations sufficient to permit replication**. These will be posted on the AEA website. The Editor should be notified at the time of submission if the data used in a paper are proprietary or if, for some other reason, the requirements above cannot be met.

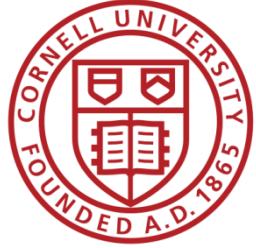


Today is different



# What is different?

- Technical infrastructure supporting openness
  - More and easier archives with self-service options (Figshare, Zenodo, OpenICPSR)
  - Github & Co.
  - Making science more easily reproducible
- Institutions
  - Center for Open Science / Open Science Framework (OSF)
  - BITSS, J-PAL, etc.
  - Registries
- New methods to do replications at scale
  - Collaborative efforts
  - Use of Mechanical Turk, Odesk, and other mechanisms



# Nosek: “Why ineffective?”

## *“Strategy”*

Study the issue

Knowing what to change

Education & training

Knowing how to change

Inculcate the values

Wanting to change

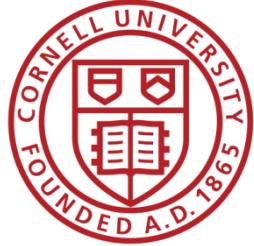
## Additions

Infrastructure

Technology to enable change

Reward system

Incentives to embrace change

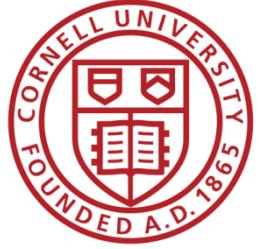


# Fanelli (2018): Opportunities

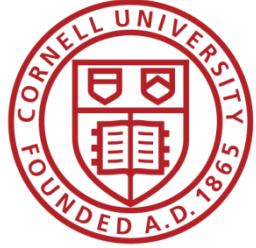
“These [information and communication] technologies promise to make research

- **more accurate, powerful,**
- **open, democratic,**
- **transparent, and**
- **self-correcting**

than ever before. At the same time, this technological revolution creates **new expectations and new challenges** that metaresearchers are striving to address.”

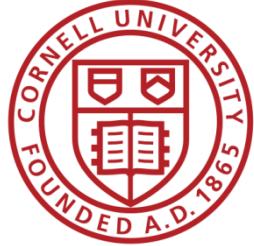


What does progress look like?



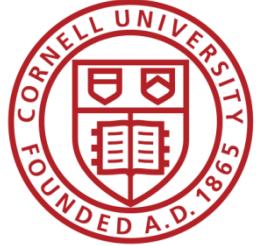
# First round (2005-2015)

- Adding data availability requirements
  - AEA, 2005 others follow
  - Which have led to code availability as well
- Experiments in post-publication discussion
  - AEA: Comments on articles



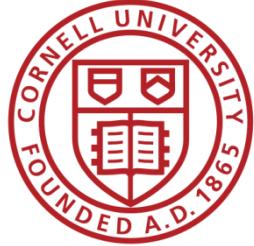
# Broad adherence to AEA policy

Journals (Publisher)	Type of policy	Archive	Confidential data
AER and Journals (self)	AEA	Journal website	Exemption
QJE (OUP)	AEA	Dataverse	Exemption
ReStud (OUP)	Generic + assistance	Journal website	Exemption
ReStat (MIT)	Own	Dataverse	“... way to apply for data...”
J Applied Econometrics	Own	Own (Queens, 1988-)	Exemption
Econometrica	Own	Journal website	Exemption with “...reasonable effort...”
JOLE (Chicago)	AEA	Journal website	Exemption
JPE (Chicago)	AEA	Journal website	Exemption
JMCB	Own (barebones)	Journal website	--



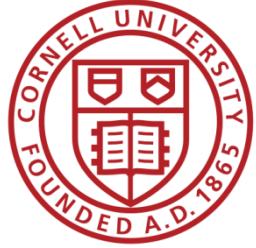
# Second round (2012-)

- Broad replication projects in Psychology
  - Open Science Collaboration (2015, others)
- Replication projects in Behavioral Economics
  - Camerer et al (2016)



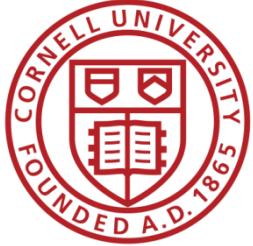
# Second round (2013-)

- New data (and code) availability policies
  - 2012, AJ Political Science
  - 2016, QJE (last top journal)



## Second round (2013-)

- Greater enforcement of data (and code) availability
  - 2015, AJ Political Science
  - 2016, Data Editor for ASA Software Section
  - 2016, Statistical review added Science
  - 2017: AEA appoints Data Editor, with mandate to do similar activities

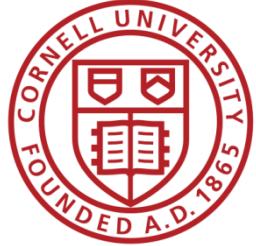


# Second round (2013-)

- Registered Reports <https://cos.io/rr>
  - Chambers (2014)
  - Nosek & Lakens (2014)

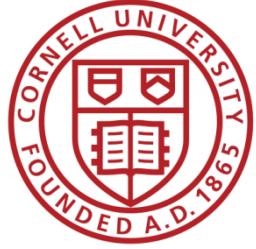


- Close cousin: Results-blind review



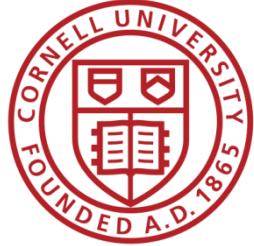
# Second round (2013-)

- Preprints in other sciences
  - bioXxiv (2013)
  - PsyArXiv (2016)

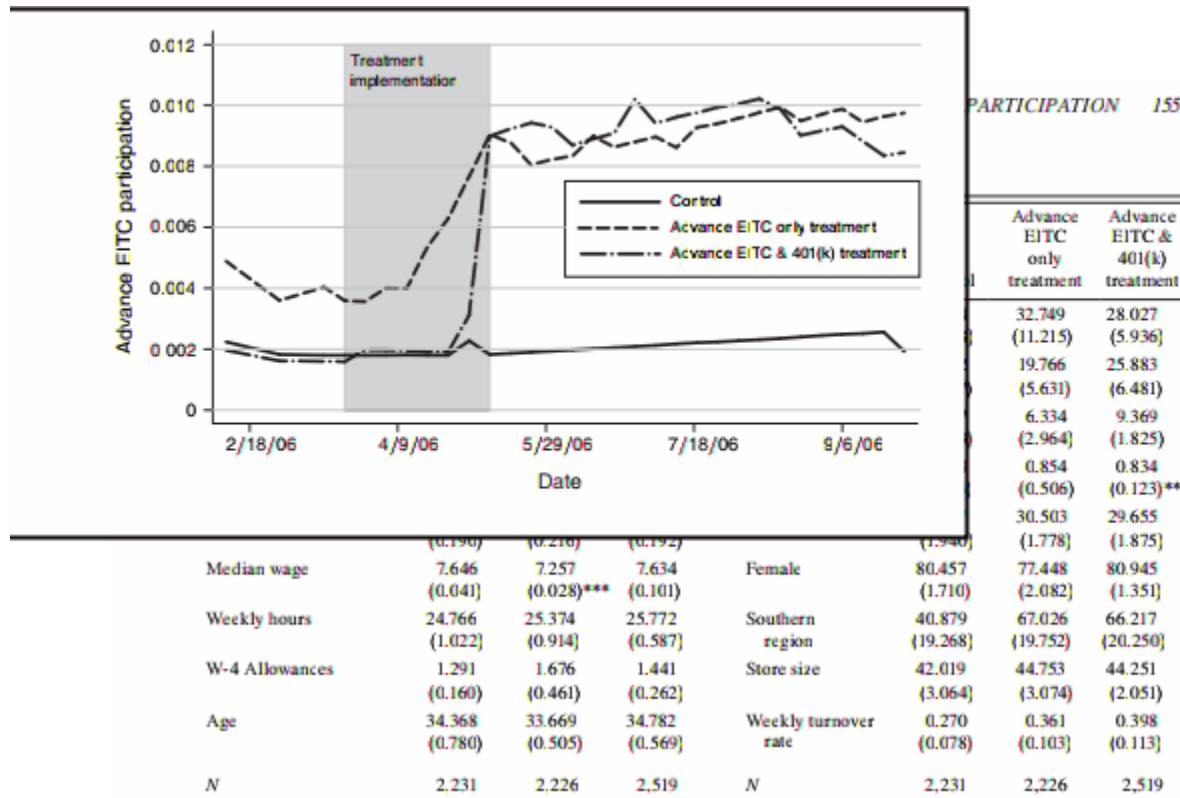


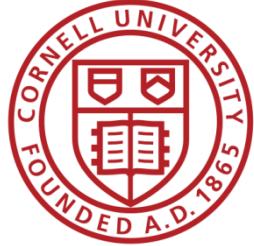
# Economics and other social sciences

A brief retrospective



# Analysis in Economics





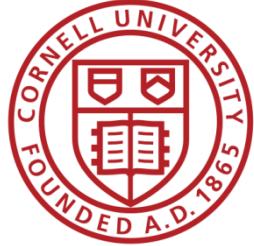
# Wide use of public-use data

- **Macrodata:**

“We use data downloaded from  
the Bureau of Economic Analysis...”

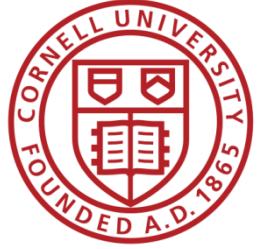
- **Microdata:**

“... this paper uses data from  
the Current Population Survey...”

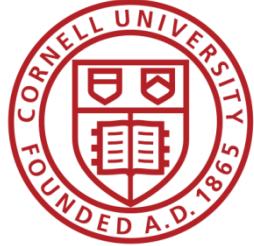


# Positive attitudes?

- Often liberal sharing of data and code  
(via email, networks)
- Sharing of articles prior to publication  
(RePEc)



And yet, there are issues



# An example

J Econom. Author manuscript; available in PMC 2012 Mar 1.

Published in final edited form as:

J Econom. 2011 Mar 1; 161(1): 82–99.

doi: [10.1016/j.jeconom.2010.09.008](https://doi.org/10.1016/j.jeconom.2010.09.008)

PMCID: PMC3079891

NIHMSID: NIHMS246950

## National Estimates of Gross Employment and Job Flows from the Quarterly Workforce Indicators with Demographic and Industry Detail

[John M. Abowd](#) and [Lars Vilhuber](#)

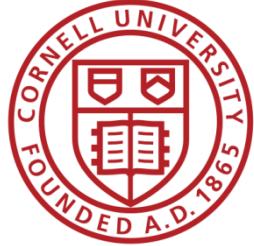
[Author information ▶](#) [Copyright and License information ▶](#)

### Abstract

[Go to:](#)

The Quarterly Workforce Indicators (QWI) are local labor market data produced and released every quarter by

No confidential data were used in this paper. All public-use Quarterly Workforce Indicators data can be accessed from <http://www.vrdc.cornell.edu/news/data/qwi-public-use-data/>. The national indicators developed in this paper can be accessed from <http://www.vrdc.cornell.edu/news/data/qwi-national-data/>. We are grateful for the comments and suggestions of many of our colleagues, past and present, too numerous to list here and thus listed at the website above and in the working paper version of this article. The opinions expressed in this paper are those of the authors and not the U.S. Census Bureau nor any of the research sponsors.



# An example: not cited...

J Econom. Author manuscript; available in PMC 2012 Mar 1.

Published in final edited form as:

J Econom. 2011 Mar 1; 161(1): 82–99.

doi: [10.1016/j.jeconom.2010.09.008](https://doi.org/10.1016/j.jeconom.2010.09.008)

PMCID: PMC3079891

NIHMSID: NIHMS246950

## National Estimates of Gross Employment and Job Flows from the Quarterly Workforce Indicators with Demographic and Industry Detail

[John M. Abowd](#) and [Lars Vilhuber](#)

[Author information ▶](#) [Copyright and License information ▶](#)

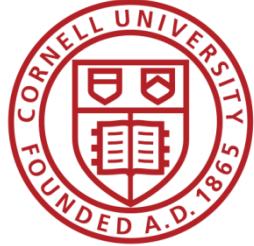
### Abstract

[Go to:](#)

The Quarterly Workforce Indicators (QWI) are local labor market data produced and released every quarter by

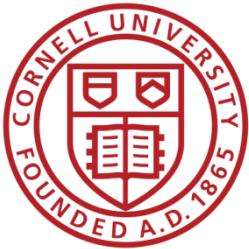
Press for the NBER; 2009. pp. 149–230.

5. Abowd JM, Vilhuber L. The sensitivity of economic statistics to coding errors in personal identifiers. *Journal of Business and Economic Statistics*. 2005;23(2):133–152. 
6. Abowd JM, Zellner A. Estimating Gross Labor Force Flows. *Journal of Business and Economic Statistics*. 1985;3:254–283.



# Data not attached to article

- J of Econometrics Data Policy at the time could not accommodate 50MB file
  - Data was not attached to paper.
- Today's J of Econometrics policy suggests using third-party repositories
  - We will get to that later



We went back, archived it

Lars Vilhuber Dataverse (Cornell University)

Harvard Dataverse > Lars Vilhuber Dataverse >

**Replication data for: National estimates of gross employment and job flows from the Quarterly Workforce Indicators with demographic and industry detail**

Metrics 4 Downloads

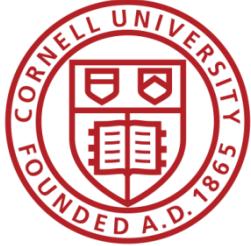
**Replication data for: National estimates of gross employment and job flows from the Quarterly Workforce Indicators with demographic and industry detail**

Abowd, John M.; Vilhuber, Lars, 2014, "Replication data for: National estimates of gross employment and job flows from the Quarterly Workforce Indicators with demographic and industry detail", <http://dx.doi.org/10.7910/DVN/27923>, Harvard Dataverse, V2

If you use these data, please add this citation to your scholarly resources. Learn about [Data Citation Standards](#).

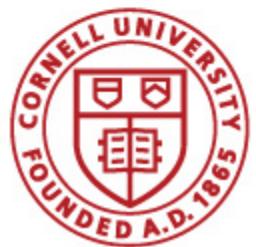
**Description**

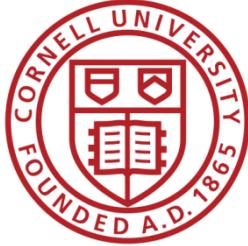
The Quarterly Workforce Indicators are local labor market data produced and released every quarter by the U.S. Bureau of Labor Statistics. Unlike any other local labor market series produced in the U.S. or the rest of the world, the QWI provide detailed information on job flows for workers (accession and separations), jobs (creations and destructions) and earnings (by race and sex), economic industry (NAICS industry groups), and detailed geography (county, County Metropolitan Statistical Area, and the nation's 400+ Workforce Investment Areas, as well as experimental, unreleased block-level estimates). Job flows are the key innovation of the QWI, allowing the existing public-use data (and only those public-use data) to construct the first national estimates of gross employment and job flows. This is an important enhancement to existing series because they include demographic and industry detail that have been compiled from data that have been integrated at the micro-level by the Longitudinal Employment Dynamics Survey.



# We went back, archived it, linked it back

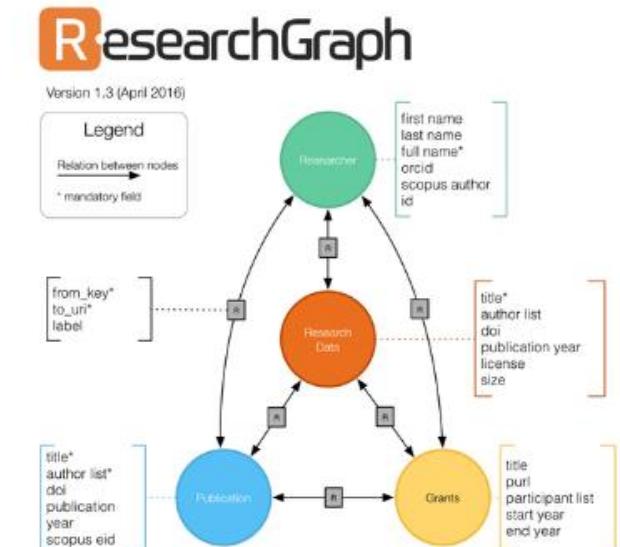
	Dataverse	About
<b>Keyword</b>	Employment Dynamics	
<b>Topic Classification</b>	Economics	
<b>Related Publication</b>	John M. Abowd and Lars Vilhuber, "National estimates of gross employment and job flows from the Quarterly Workforce Indicators with demographic and industry detail," Journal of Econometrics, vol. 161, iss. 1, pp. 82-99, 2011. doi: <a href="https://doi.org/10.1016/j.jeconom.2010.09.008">10.1016/j.jeconom.2010.09.008</a> <a href="http://www2.vrdc.cornell.edu/news/data/qwi-national-data/">http://www2.vrdc.cornell.edu/news/data/qwi-national-data/</a>	
	John M. Abowd and Lars Vilhuber, "National estimates of gross employment and job flows from the Quarterly Workforce Indicators with demographic and industry detail," Journal of Econometrics, vol. 161, iss. 1, pp. 82-99, 2011. doi: <a href="https://doi.org/10.1016/j.jeconom.2010.09.008">10.1016/j.jeconom.2010.09.008</a> <a href="http://www2.vrdc.cornell.edu/news/data/qwi-national-data/">http://www2.vrdc.cornell.edu/news/data/qwi-national-data/</a>	
	John M. Abowd and Lars Vilhuber, "National estimates of gross employment and job flows from the Quarterly Workforce Indicators with demographic and industry detail (with color graphs)," Center for Economic Studies, U.S. Census Bureau, Washington, D.C., 2011. <a href="http://ideas.repec.org/p/cen/wpaper/10-11.html">http://ideas.repec.org/p/cen/wpaper/10-11.html</a>	
<b>Producer</b>	Labor Dynamics Institute (Cornell University) (LDI) <a href="http://www2.vrdc.cornell.edu/news/data/qwi-national-data/">http://www2.vrdc.cornell.edu/news/data/qwi-national-data/</a>	

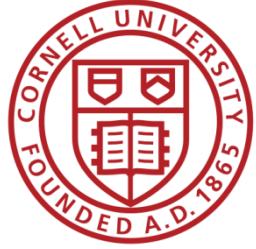




# But journal and data infrastructure are incomplete

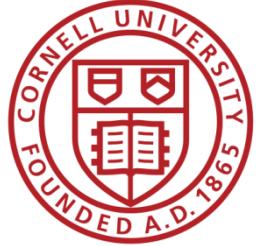
- While Dataverse allows to manually link back...
- ... the article itself reveals none of that
- True for most journals, and most data archives
  - ICPSR (manual linking to articles)
  - RePEc (no linkage possible)
- Infrastructure starting to emerge
  - If article cites data (DOI!)
  - If archive and/or journal leverages infrastructure





# Using personal websites

Don't even think about  
using personal websites!

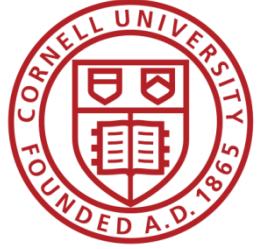


# Google

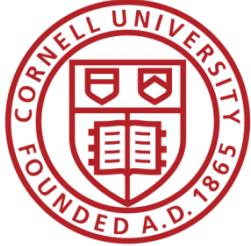
**404.** That's an error.

The requested URL /a\_cool\_website was not found on this server. That's all we know.





[source](#)



# Yielding to the temptation of “administrative”data

- **State-level data**

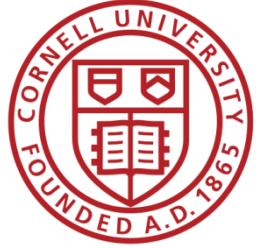
“Our analysis draws on administrative records from the Detroit Work First program linked with unemployment insurance (UI) wage records for the State of Michigan”

Autor/Houseman doi:10.1257/app.2.3.96

- **School-district data**

“confidential student-level panel dataset provided by the School Board of Alachua County in Florida”

Carrel and Hoekstra doi:10.1257/app.2.1.211

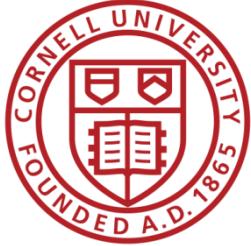


... yielding...

- **Proprietary data**

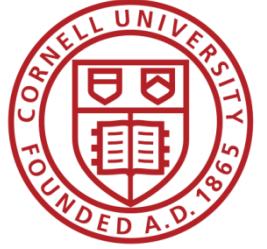
“This field experiment was made possible by the collaboration of a large-scale, nationwide firm in the retail sector.”

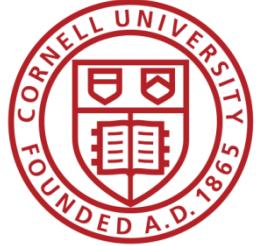
Damon doi:10.1257/app.2.2.147



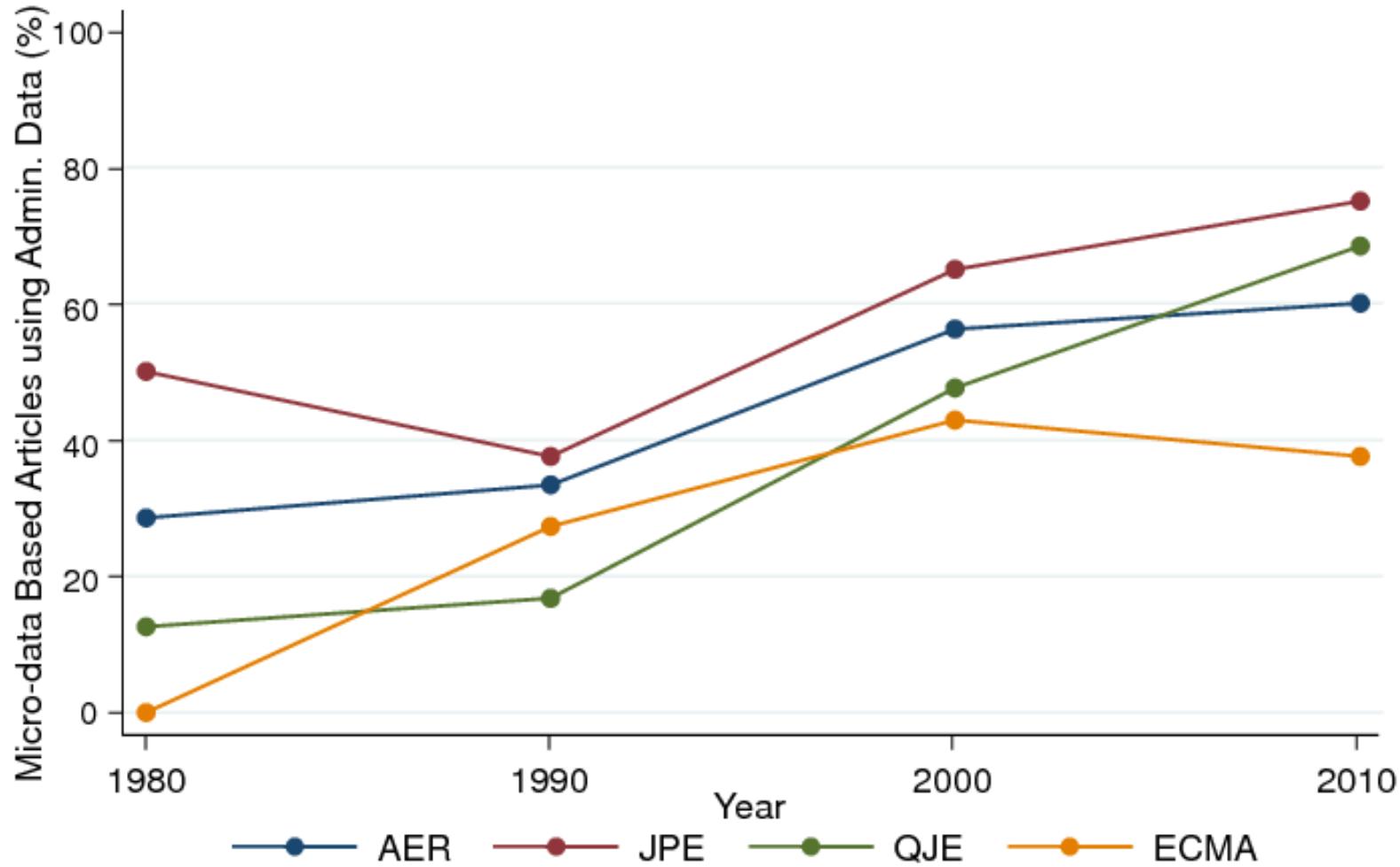
# Death Knell for Public-use Data?

- Sounded by young scholars pursuing research programs that mandate inherently identifiable data:
  - Geospatial relations,
  - Exact genome data,
  - Networks of all sorts,
  - Linked administrative records
- These researchers acquire authorized, generally unfettered, restricted access to the confidential, identifiable data and perform their analyses in secure environments.
- But: generally don't leave behind the scientific trail that has made public-use files so important.

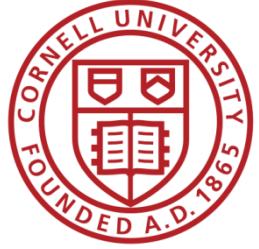




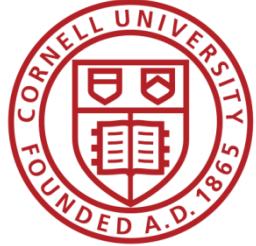
## Use of Administrative Data in Publications in Leading Journals, 1980-2010



Note: Administrative data refers to any data that was collected with the primary purpose of surveying individuals (e.g., scanner data, stock prices, school district records, social security records).  
Sample excludes studies whose primary data source is from developing countries.

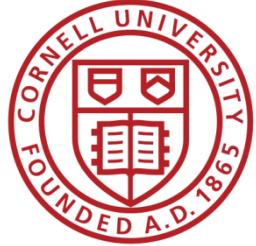


# Progress in Economics

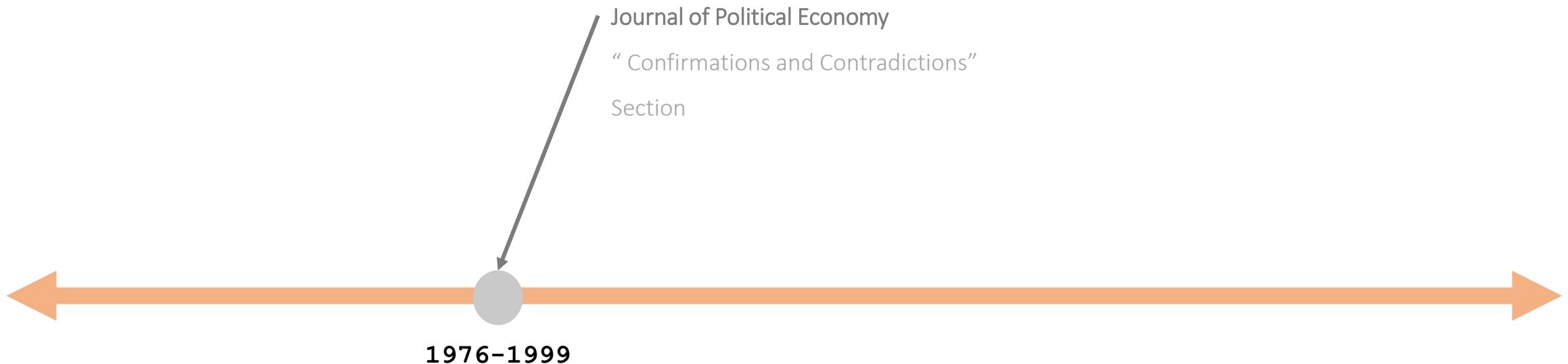


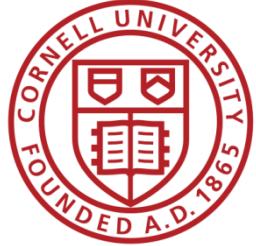
# Increasing transparency in economics journals



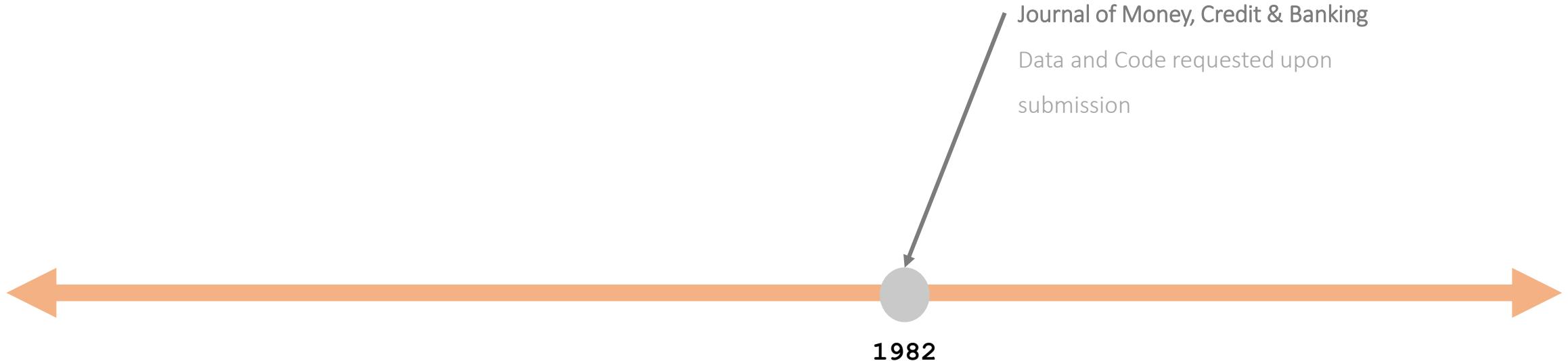


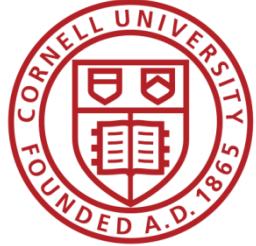
# Increasing transparency in economics journals



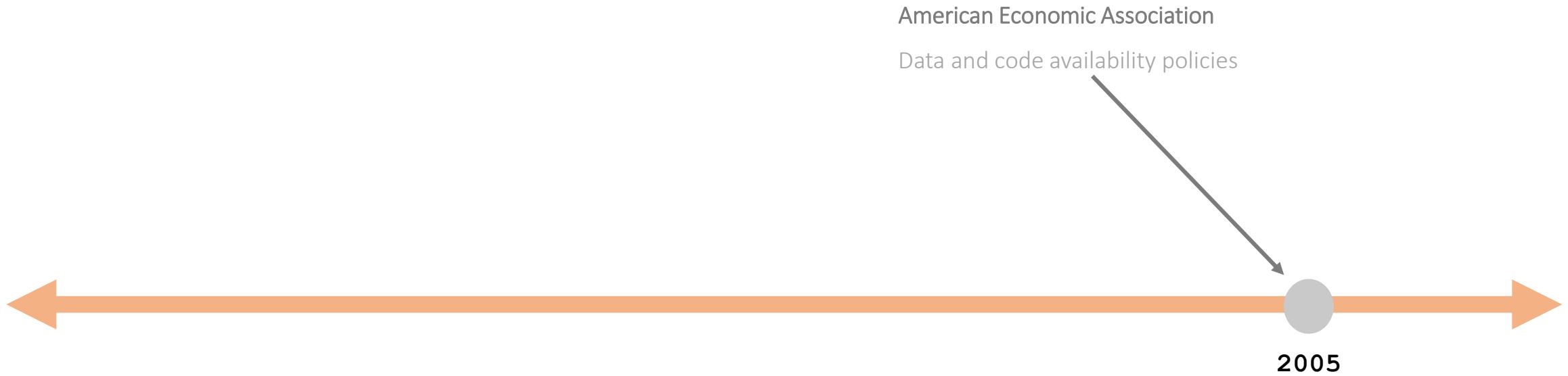


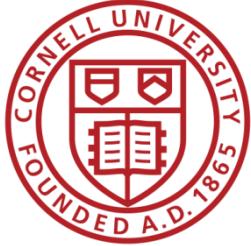
# Increasing transparency in economics journals





# Increasing transparency in economics journals





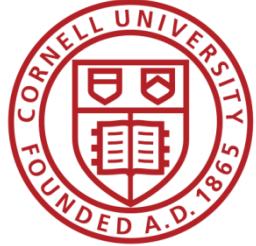
# Increasing transparency in economics journals



2014

Promoting Transparency in Social  
Sciences  
Miguel et al (Science)  
[10.1126/science.1245317](https://doi.org/10.1126/science.1245317)

- Greater transparency
- Registration and pre-analysis plans
- Open data and materials

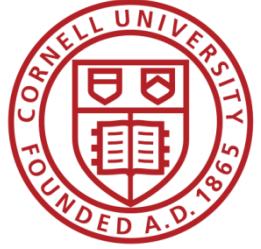


# Increasing transparency in economics journals

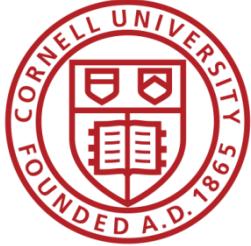
Quarterly Journal of Economics

Data and code availability policies



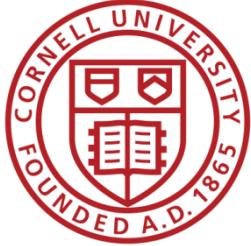


Explicit replication remains rare



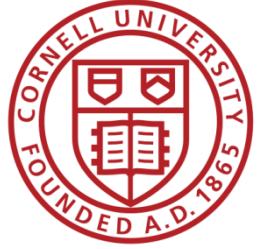
# How common is replication in economics?

- **0.1%** of 126,505 “top 50” published articles 1974-2014 are replications
  - Mueller-Langer et al. (2018)
- **6.2%** of 1,138 “top 10” development articles 2000-2015 were subjects of replications
  - Sukhtankar (2017)

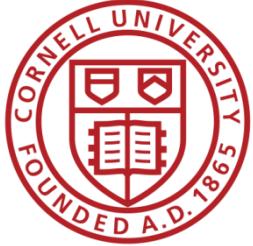


# Replications and Reproducibility Checks Remain Rare

- No significant replication journal
- ReplicationWiki identifies **44** “Comments” in the AER as “replications” of some sort, out of **~200\*13 = 2600** articles (2004-2016) = **1.7%**



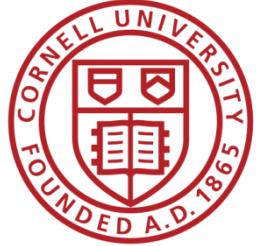
Explicit replication remains rare  
Or are they?



# “Dark Web” of Replications

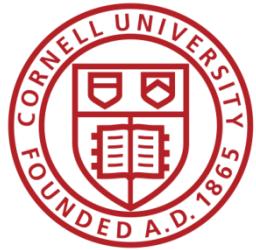
- Personal communication: lots of micro, labor, macro Ph.D. classes have students run (successful) replications
- Hamermesh (2007): papers that do not replicate do not get cited (... over time)
- Projects (...) to provide replications (in the broad sense) by “swapping country”

*(Do the results of Paper A  
– run on data from country B –  
work in country C)*

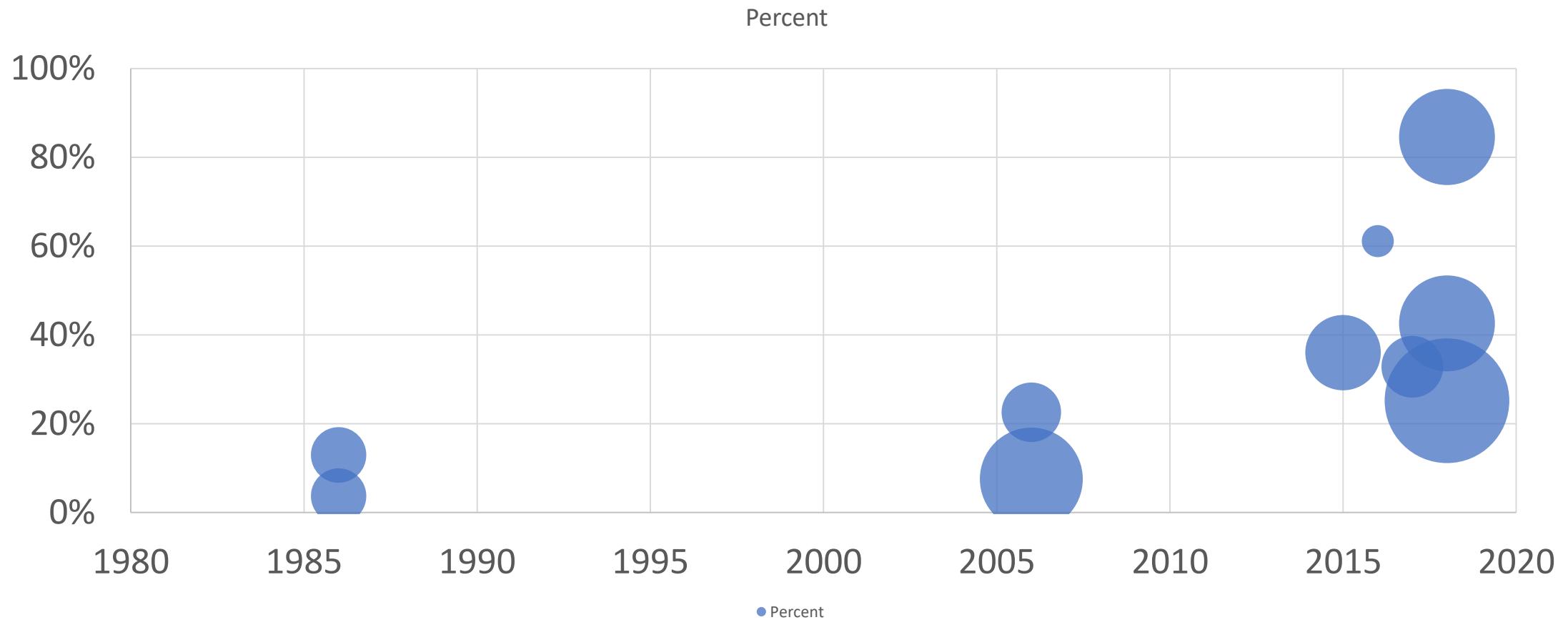


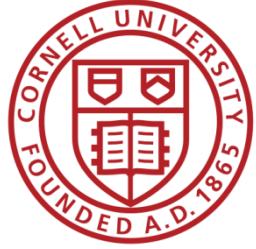
# Explicit Replications and Reproduction Attempts

- Vary in how they assess the exercise
  - Replication or Reproducibility
  - Denominator: Data available, or all articles?  
*Prevalence of restricted-access data!*
  - Numerator: Partial or full reproduction/  
replication?  
*Who makes the assessment? How?*
  - Quantity/ domain

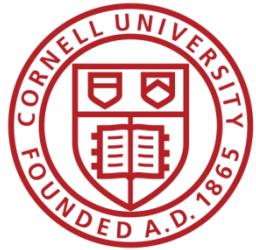


# Some key statistics

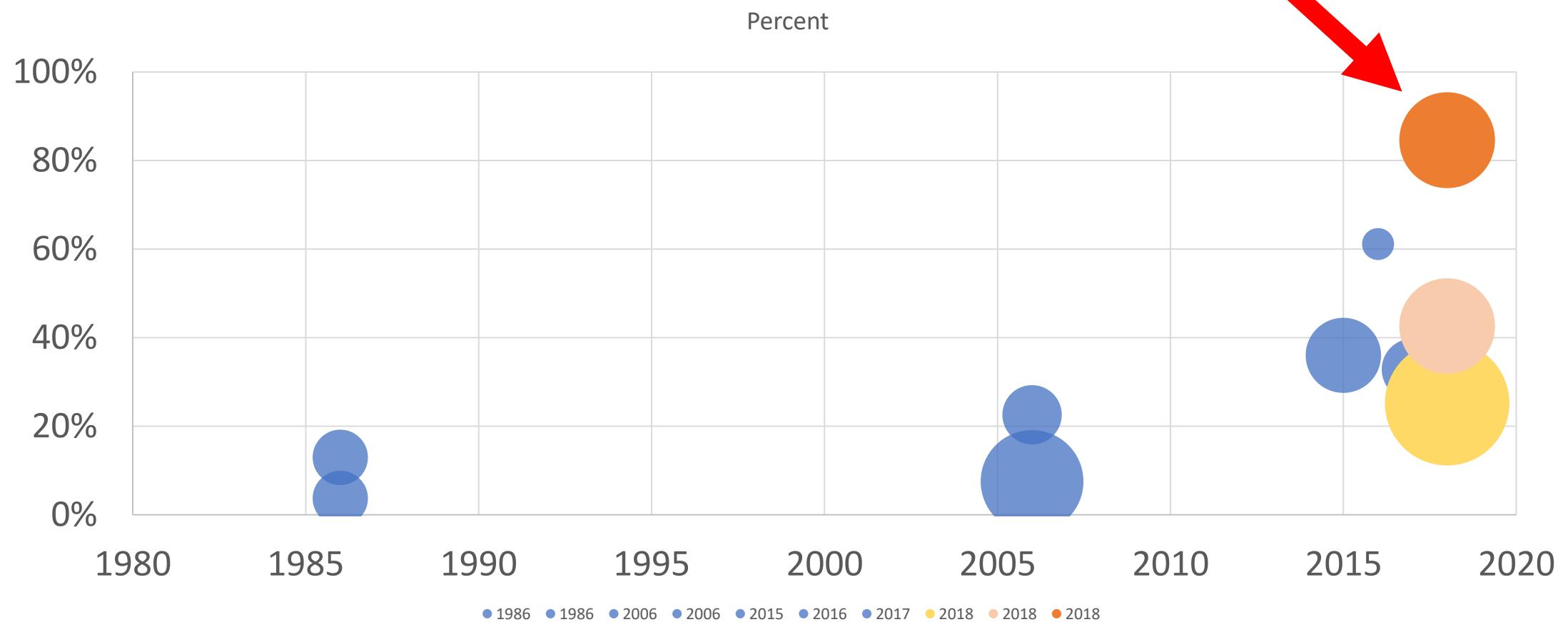




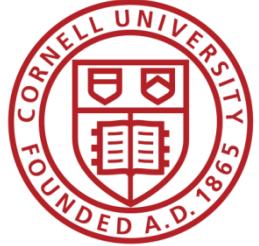
# Current efforts at the AEA



# Remember this?



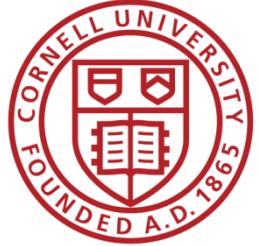
Method



# Undergraduate Summer Research Assistants

2014 - 2018





# Undergraduate Summer Research Assistants

2014 - 2018



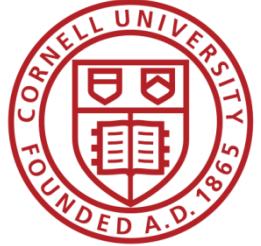
Economics



Comp Sci.

OpRe





# Undergraduate Summer Research Assistants

2014 - 2018



Economics

Comp Sci.

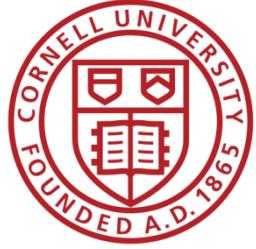
OpRe

Reproduction

VCS

Cloud Computing

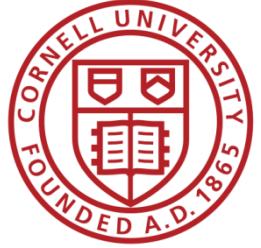




# Method

A computer monitor displays a Google Form titled "Entry\_Questionnaire\_Draft". The form has a purple header bar. The first question is "NetID \* (only if Cornell student, enter email otherwise)" with a text input field labeled "Your answer". The second question is "DOI \*" with a text input field labeled "Your answer". The third question is "TypeOfArticle \*" with a dropdown menu labeled "Choose". At the bottom of the form are "NEXT" and "PREVIOUS" buttons, and a progress bar indicating "Page 1 of 15".

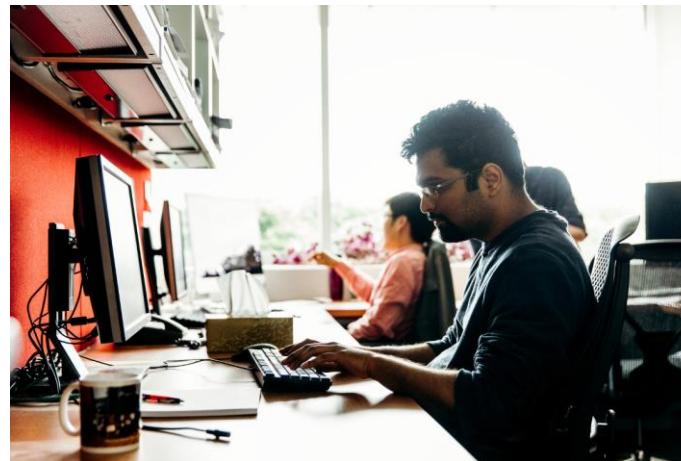
≡ Assessment Form



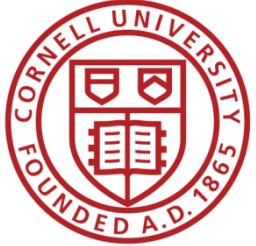
# Method

A computer monitor displays a Google Form titled "Entry\_Questionnaire\_Draft". The form includes fields for "NetID" (marked as required), "DOI" (with a note about including the full DOI if it's a Cornell student), and "TypeOfArticle" (with a note about including empirical work, simulations, or experimental work). The form has a "NEXT" button and a note at the bottom about never sending passwords through Google Forms.

❖ Assessment Form



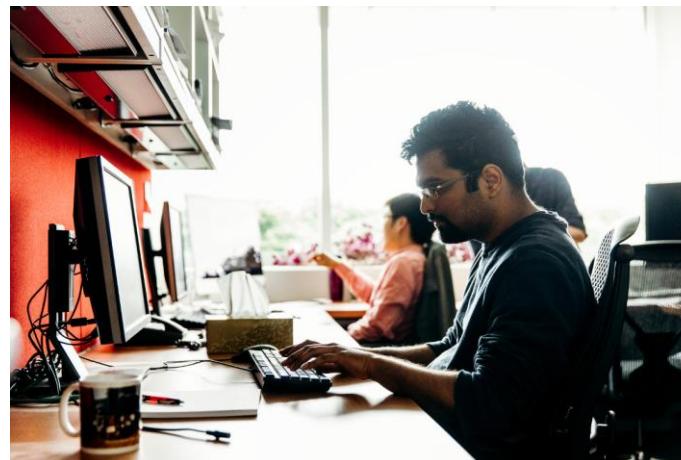
❖ Reproduction Exercise



# Method

A computer monitor displays a Google Form titled "Entry\_Questionnaire\_Draft". The form includes fields for "NetID" (marked as required), "DOI" (marked as required), and "TypeOfArticle" (multiple choice options: "Empirical work", "Simulations", or "Experimental work"). A "NEXT" button is visible at the bottom.

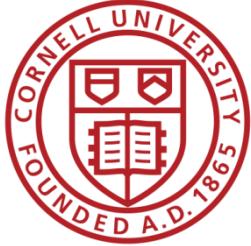
❖ Assessment Form



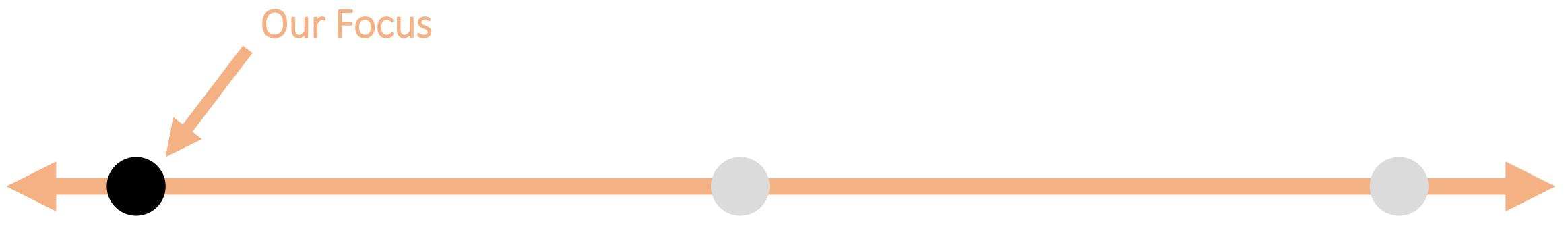
❖ Reproduction Exercise

A tablet device displays a Google Form titled "Exit\_Questionnaire\_Draft". The form includes fields for "NetID or email" (marked as required), "DOI" (marked as required), and "Replication\_Success" (radio button options: "Yes", "No", or "Partial"). A "NEXT" button is visible at the bottom.

❖ Exit Questionnaire



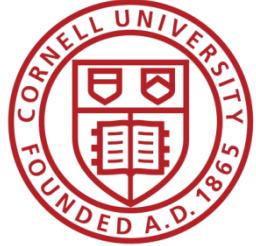
## Replication continuum



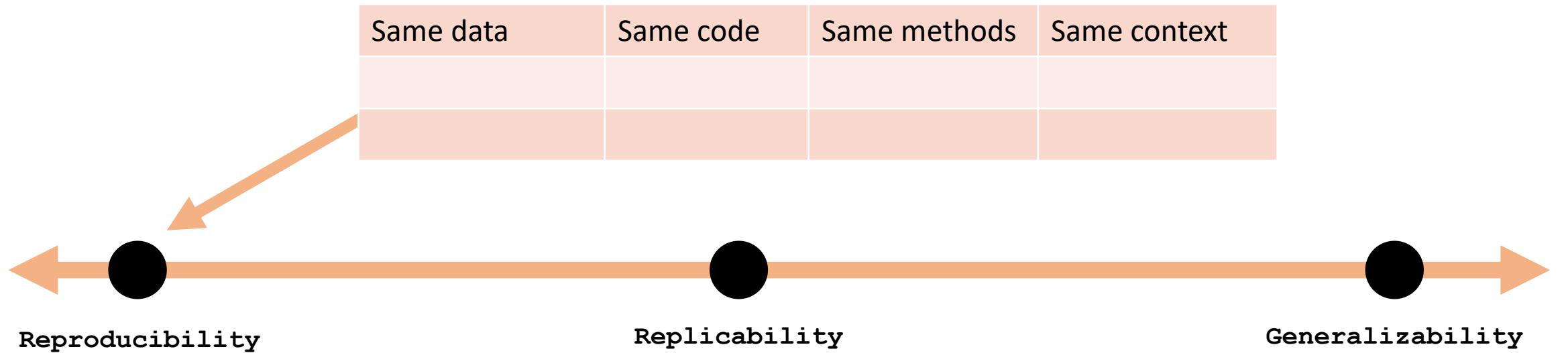
- Narrow Replication (Pesaran 2003)
- Pure Replication (Hamermesh 2007)
- Verification (Clemens 2015)

- Wide Replication (Pesaran 2003)
- Statistical Replication (Hamermesh 2007)
- Reproduction/Reanalysis (Clemens 2015)

- Wider Replication (Pesaran 2003)
- Scientific Replication (Hamermesh 2007)
- Reanalysis/Robustness (Clemens 2015)



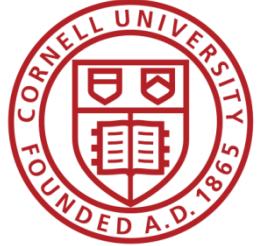
## Replication continuum



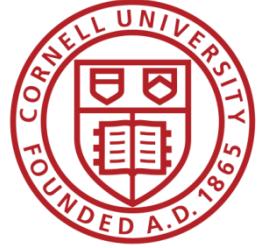
- Narrow Replication (Pesaran 2003)
- Pure Replication (Hamermesh 2007)
- Verification (Clemens 2015)

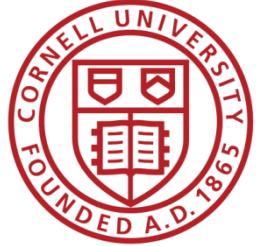
- Wide Replication (Pesaran 2003)
- Statistical Replication (Hamermesh 2007)
- Reproduction/Reanalysis (Clemens 2015)

- Wider Replication (Pesaran 2003)
- Scientific Replication (Hamermesh 2007)
- Reanalysis/Robustness (Clemens 2015)



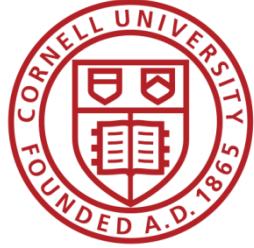
© Nick photography





# From Post- to Pre-Publication Verification

- Cornell Replication Lab has been verifying published articles
- Now switching to manuscripts in the submission workflow
- For now pilot
  - *Authors have submitted prior to announcement of new data policy*



# Email during pilot phase

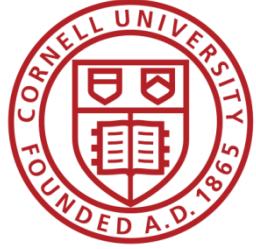
Dear [AUTHOR]:

I am writing about your paper [MC ID]. I am ready to give you a "conditional accept" on the paper.

The condition is simply to work with the AEA's data co-editor, Lars Vilhuber (copied on this email), to prepare the data and code for publication. The data is an important contribution of the paper. He is copied on this email.

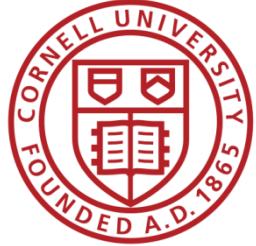
Thank you for submitting your work.

I am looking forward to seeing the paper in print.



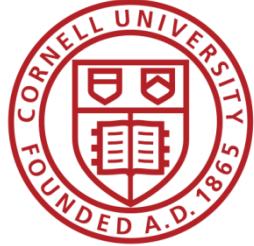
# Future efforts

AEA, Social Sciences, elsewhere



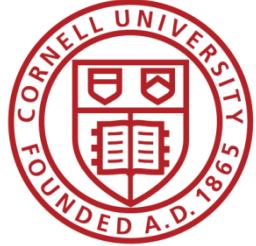
# Evolving Journal and Data Infrastructure

- More self-deposit repositories in the social sciences
  - Dataverse
  - Figshare
  - openICPSR
  - Zenodo
  - Qualitative Data Repository (QDR)
  - Others...



# More and easier data deposit

- More integration into **Journal Workflows**
  - AJPS + Dataverse
  - QJE, Restat (both Harvard...) + Dataverse
- Data deposit occurring during submission

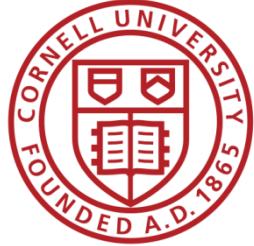


## At the AEA...

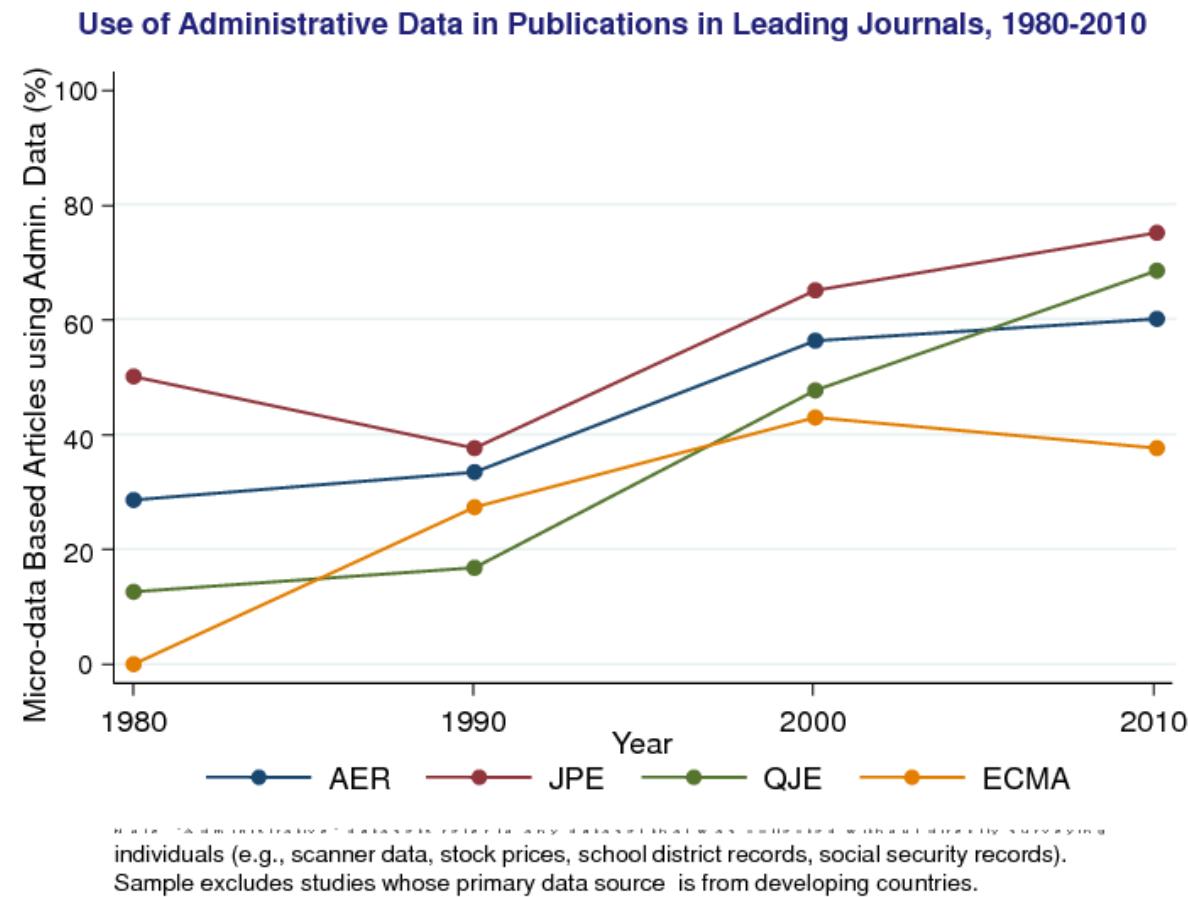
- We will create a journal-specific repository for AEA journals
  - *Open for deposit 2019 Q1*
- We will migrate **all historical supplements** to the archive
  - *Explode the ZIP files: greater transparency*

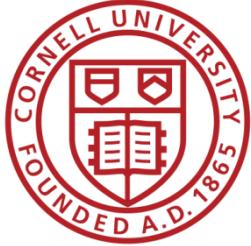
And then we'll kill it

Let me explain...



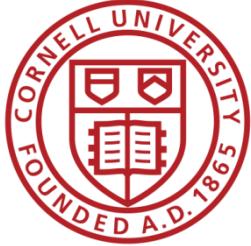
# Remember this figure?





# Assessing accessibility

- Chetty's assessment was entirely manual
  - Download the article
  - Read what kind of data is used
- And yet does not provide information such as
  - Who can access?
  - How long can you access?
  - Characteristics of the data
- ... and not for want of trying, but because that information is not required and not provided



# From earlier examples

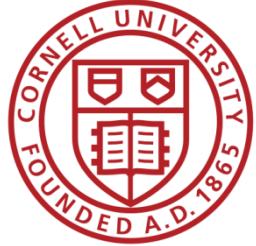
- **Costly access but good intentions:**

“researchers could submit programs that [...] research assistants would run. Alternatively, researchers wishing to work directly with the data could come and work on the Institute's premises.”

Autor/Houseman doi:10.1257/app.2.3.96

- **Uncertain access**

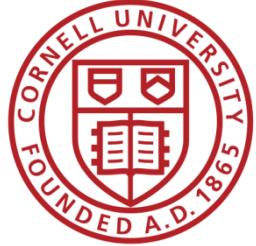
“Data [...] is proprietary and owned by the Alachua County, Florida School District. The corresponding author [...] holds the deidentified dataset [...] and will provide copies to authors who receive written permission from the Alachua County Public Schools.”



# Current Data Availability Policies are Broken

- If the Data is  
not open-access,

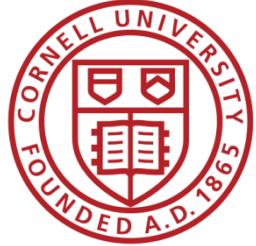
**no systematic information is  
collected  
("exemption")**



# Current Data Availability Policies are Broken

- If the Data is  
not open-access,

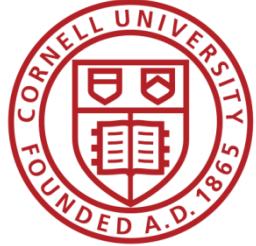
even if the *Data Owner*  
has already deposited it



# Current Data Availability Policies are Broken

- If the Data **is open-access**,  
you are asked to deposit the data in the  
“journal repository”

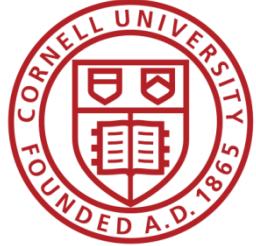
**even if the *Data Owner*  
has already deposited it**



# Current Data Availability Policies are Broken

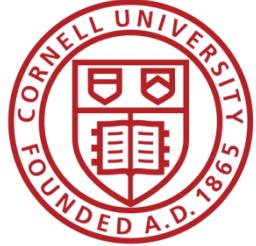
- If the Data **is open-access**,  
you are asked to deposit the data in the  
“journal repository”

**even if** you have followed *best practices* and  
**have already deposited it**



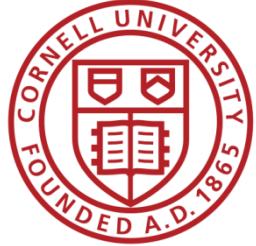
Illustration

If you used the ACS,  
should you deposit it  
*again*?



Illustration

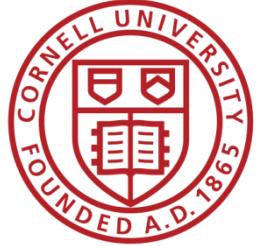
If you used the LEHD,  
has the Census Bureau  
not *already archived it?*



Illustration

If you used the LEHD,

can others  
not **access it**?

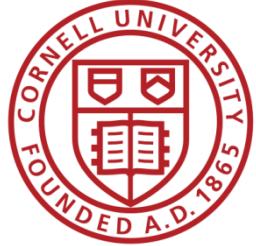


Illustration

If you used files at  
the National Archives,

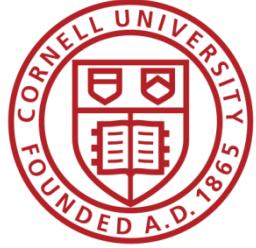
are they not

*already archived?*



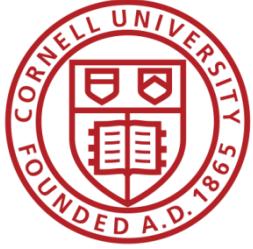
# Best practices

- Deposit and archive early
  - If you collect data, archive it  
*(possibly privately)*
  - If you finish the manuscript, deposit the analysis files  
*(possibly privately)*



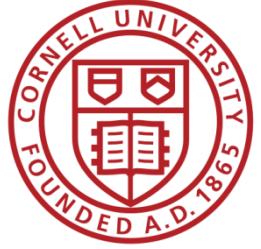
# Best practices

- Follow robust coding
  - Ensure that code reliably produces results  
*(possibly automated)*
  - Before you finish the manuscript, run all analysis code again  
*(if not too onerous)*

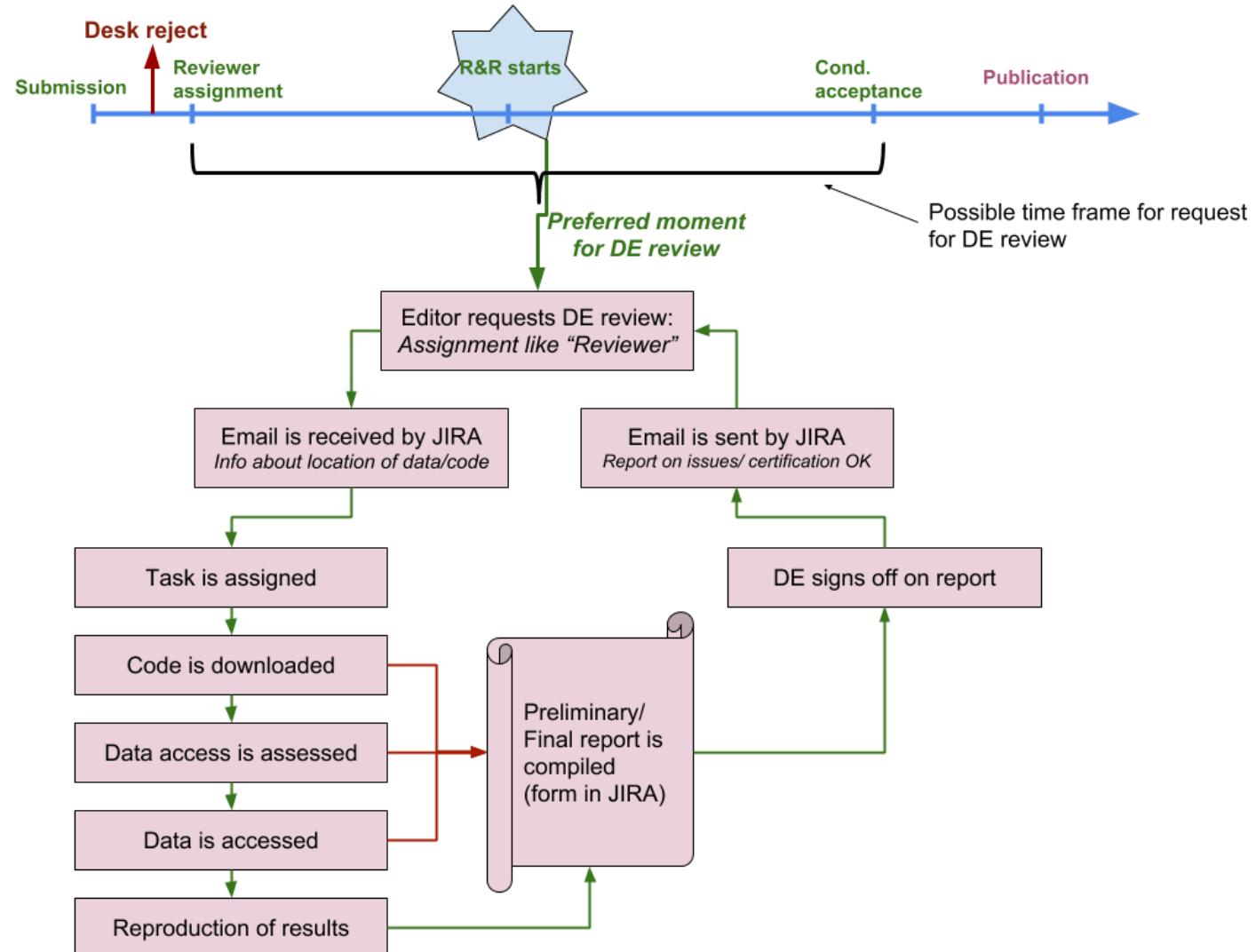
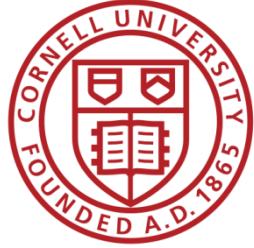


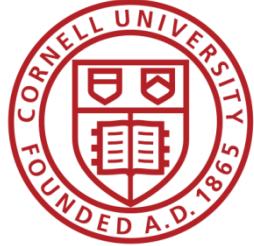
# The goal at the AEA

- Ensure **reproducibility** of computational code
- Encourage **best practices**
- Not require double-work
  - If you have deposited your work (*in a reliable location*), leave it there
  - If you have had your code verified (*by a reliable third party*), we will accept that



Ensuring scalable reproducibility





# New AEA Data and Code Availability Policy

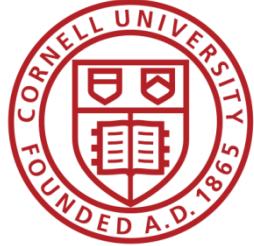
Dear [AUTHOR]:

I am writing about your paper [MC ID]. I am ready to give you a  
"conditional accept" on the paper.

The condition is simply to work with the AEA's data co-editor, Lars Vilhuber  
(copied on this email), to prepare the data and code for publication. The data  
is an important contribution of the paper.

Thank you for submitting your work.

I am looking forward to seeing the paper in print.



# New AEA Data and Code Availability Policy

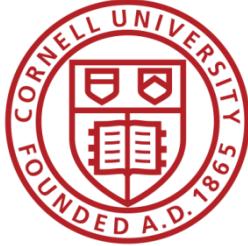
Dear [AUTHOR]:

I am writing about your paper [MC ID]. I have received two referee reports, which you will find attached. We would like you to **revise and resubmit** your paper. [...]

As you revise the paper, please also work with the AEA's data co-editor, Lars Vilhuber (copied on this email), to document and verify the reproducibility of the data and code. The data is an important contribution of the paper.

Thank you for submitting your work.

I am looking forward to seeing the revised paper.



## Policy A.2

### Data Availability Policy

It is the policy of the American Economic Association to publish papers only if the data used in the analysis are clearly and precisely documented and are readily available to any researcher for purposes of replication.

Authors of accepted papers that contain empirical work, simulations, or experimental work must provide, prior to publication,

the data, programs, and other details of the computations sufficient to permit replication.

These will be posted on the AEA Data and

## Policy B

### Data and Code Availability Policy

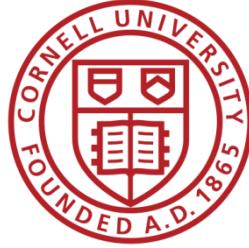
**It is the policy of the American Economic Association to publish papers only if the data and code used in the analysis are clearly and precisely documented; access to the data and code is clearly and precisely documented, and is non-exclusive to the authors.**

Authors of accepted papers that contain empirical work, simulations, or experimental work must provide,

**prior to acceptance,**

information about the data, programs, and other details of the computations sufficient to permit replication, as well as information about access to data and programs.

Data and programs should be archived in



prior to publication,

the data, programs, and other details of the computations sufficient to permit replication.

These will be posted on the AEA Data and Code Repository at *TBD*.

**prior to acceptance,**

information about the data, programs, and other details of the computations sufficient to permit replication, as well as information about access to data and programs.

**Data and programs should be archived in community-recognized or general repositories**, including the AEA Data and Code Repository at *TBD*.

The Editor should be notified at the time of submission if the data used in a paper are proprietary or if, for some other reason, the requirements above cannot be met.

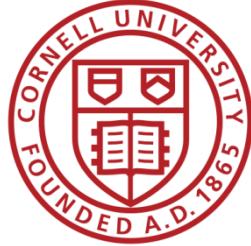
The AEA Data Editor will assess compliance with this policy.

Authors will provide access to editors and reviewers, if requested, to both data and programs prior to acceptance.

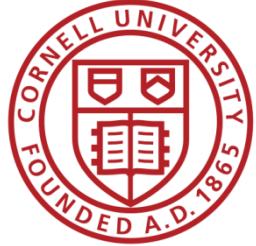
The Editor should be notified at the time of submission if access to the data used in a paper is restricted or limited, or if, for some other reason, the requirements above cannot be met.

The AEA Data Editor will assess compliance with this policy,

and will verify the accuracy of the information.



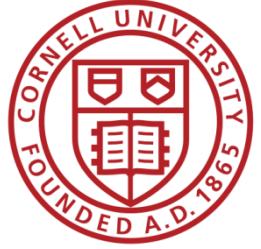
	<p><b>repositories</b>, including the AEA Data and Code Repository at <i>TBD</i>.</p>
	<p>Authors will provide access to editors and reviewers, if requested, to both data and programs prior to acceptance.</p>
The Editor should be notified at the time of submission if the data used in a paper are proprietary or if, for some other reason, the requirements above cannot be met.	The Editor should be notified at the time of submission if access to the data used in a paper is restricted or limited, or if, for some other reason, the requirements above cannot be met.
The AEA Data Editor will assess compliance with this policy.	<p><b>The AEA Data Editor will assess compliance with this policy,</b> <b>and will verify the accuracy of the information prior to acceptance by the Editor.</b></p>
	A statement describing compliance with this policy will be posted alongside the article.



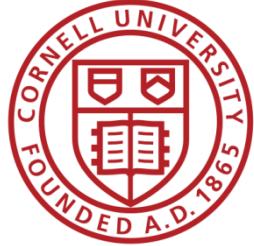
# Goal: earlier provision of data and code

“Arguably, the replication files should be submitted when the paper is **initially submitted** so that they will be available for inspection by the **referees**, if the referees so desire.”

(McCullough et al, 2006)

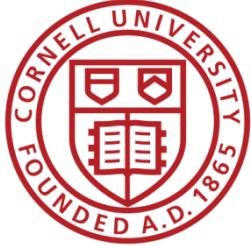


# Scalability and sustainability



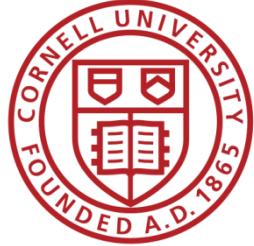
# Scalability of Verification

- 500+ articles per year at AEA (*not all have code or data*)
- Financing:
  - \$100 submission fee minus \$100 per referee report (x2)
  - Undergraduate student (~\$15/h) spends 10-20 hours on most verifications (*would be far less if automated*)  
= \$200 extra
  - Most journals are run at a loss (*but maybe Elsevier has the funds...*)



# Scalability of Verification

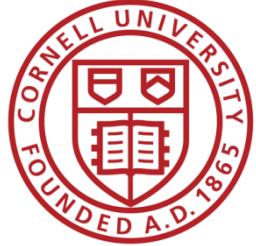
- Third party verification services already exist!
  - For research institutes and groupings (J-PAL, NBER, etc.)
  - At universities (Cornell, etc.)
  - Because authors already do it
- Even at restricted-access centers
  - Explicitly: France
  - Implicitly: German IAB data, other restricted-access data (remote processing at NCHS, Statistics Canada)
  - Interest in other institutions (including Census Bureau)



# Scalability of Verification

## Third party verification

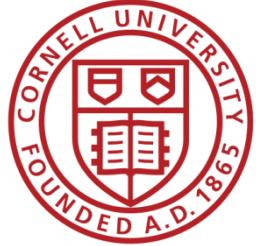
- How to fund
  - similar to open-access submission fees?
  - Requirement by funders?
- How to certify third-party certifiers?
  - How reliable is their service? What do they check?
  - Who certifies? (*Association of Data Editors?*)
- These are **open questions**
- They need to be solved for persistence of the effort



# Verifying Data and Code Deposits

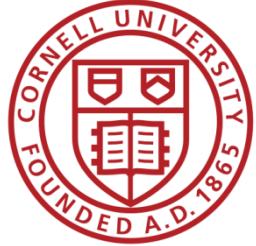
Why do journals like  
affiliated repositories  
(or website deposits)?

- They can ensure **longevity/ persistence**
- They can ensure **access**
- They can ensure **availability**



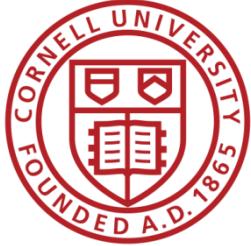
# Verifying Data and Code Deposits

- Not every data repository is created equal
  - Github, Dropbox, etc. are not data or code repositories
  - Is the institutional repository at the University of Southern Venezuela a reliable repository?
  - Is the institutional repository at Cornell University a reliable repository?
  - Is the institutional repository at Harvard University (Dataverse!) a reliable repository?
  - Are the National Archives a reliable repository?



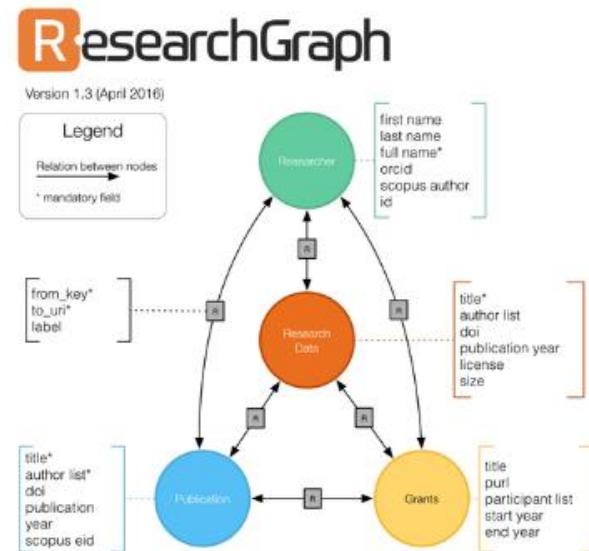
# Verifying Data and Code Deposits

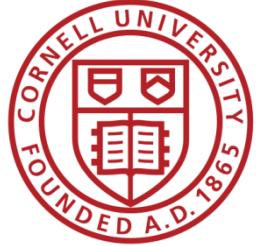
- Not every data repository is created equal
  - The Second Bank of Third City credit card data is not a data/code repository
  - Is the School Board of Third City a reliable repository?
  - Is the JPMC Institute a reliable repository?
  - Is the US Census Bureau a reliable repository?
  - **Are any restricted-access repositories a reliable repository?**



# Linking objects (data, programs), articles and researchers

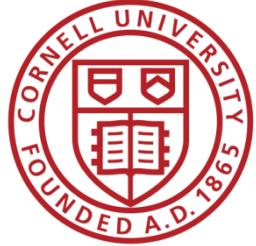
- Reminder: currently a very manual process
  - Only really works if the article cites the data
- Infrastructure starting to emerge
  - If article cites data (DOI!)
  - If archive and/or journal leverages infrastructure





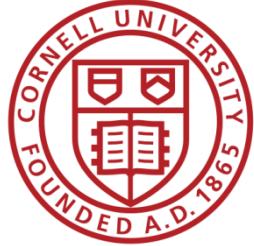
# AEA: improve transparency of all data deposits

- Developed a schema to collect relevant information
  - Relies as much as possible on lookups of existing data
  - Collects only as much information as necessary
  - Replaces existing documentation efforts
  - Can be done early in research workflow

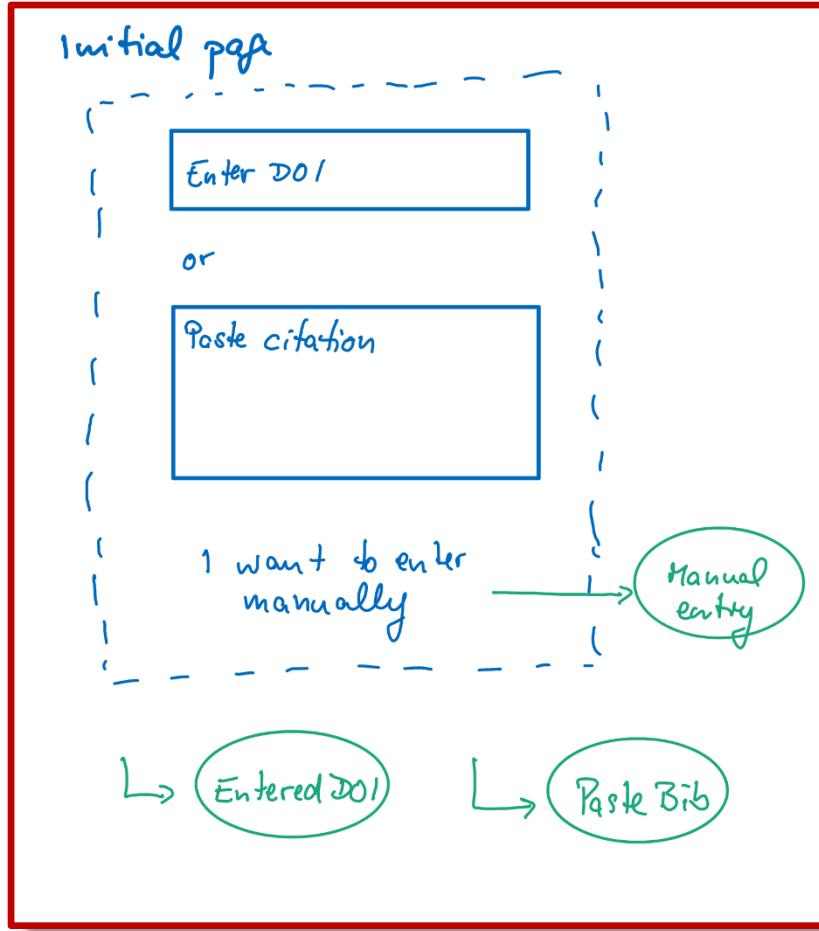


# AEA: improve transparency of all data deposits

- Developed a schema to collect relevant information
  - **Uniformly applied** to public-use and restricted-access data
  - Can be added at low-cost to existing journal websites
  - Can be data-mined



# Richer metadata, more transparency



American Economic Association [return to AEweb](#)

## Predation, Protection, and Productivity: A Firm-Level Perspective.

Abstract References Online appendix Supplementary materials Notes

### Supplementary materials

- Code and Data  
*Besley, Timothy, and Hannes Mueller. 2018. "Replication data for: Predation, Protection, and Productivity: A Firm-Level Perspective." American Economic Journal: Macroeconomics, 10 (2): 184-221. DOI: 10.1257/mac.20160120.data*  
cite!
  - Data is freely accessible at 10.1257/mac.20160120.data under CC BY-NC 4.0
  - Code verified under AEA guidelines 2.0
- Data  
*Statistics Norway. 2015. "Firm-level statistics 1975-2013 [dataset]" Norwegian Data Archive [curator], v2. DOI: 10.7654/nda.7643A::34*  
cite!
  - Data restricted-access (has residency requirement, has citizenship requirement), accessible at Norwegian Data Archive in Oslo, Norway, under Norwegian Data Access license.
  - ?Code could not be verified due to access restrictions.

© 2017 Larc Vilhuber, Labor Dynamics Institute, Cornell University



Abstract



References



Online appendix



Supplementary  
materials



Notes

## Supplementary materials

- Code and Data

*Besley, Timothy, and Hannes Mueller. 2018. "Replication data for: Predation, Protection, and Productivity: A Firm-Level Perspective." American Economic Journal: Macroeconomics, 10 (2): 184-221. DOI: [10.1257/mac.20160120.data](https://doi.org/10.1257/mac.20160120.data)*

cite!

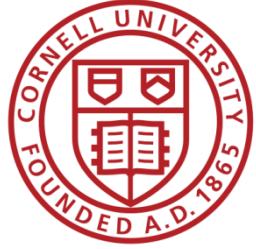
- Data is freely accessible at [10.1257/mac.20160120.data](https://doi.org/10.1257/mac.20160120.data) under CC BY-NC 4.0.
- Code verified under [AEA guidelines 2.0](#)

- Data

*Statistics Norway. 2015. "Firm-level statistics 1975-2013 [dataset]" Norwegian Data Archive [curator], v2. DOI: [10.7654/nda::7643A::34](https://doi.org/10.7654/nda::7643A::34)*

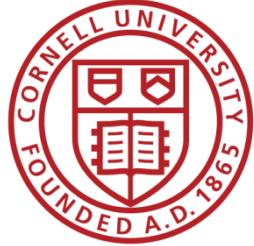
cite!

- Data restricted-access (has residency requirement, has citizenship requirement), accessible at [Norwegian Data Archive in Oslo, Norway](#), under [Norwegian Data Access license](#).
- Code could not be verified due to access restrictions.



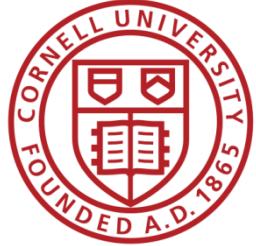
New problem: Licensing of code and data

Do you know what license  
your journal supplements  
are under?



# Really brief licensing primer

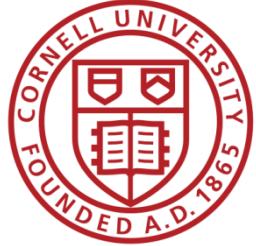
- Software licenses:
  - *Copyright*
  - *protecting yourself from being held liable*
- US: copyright is automatic
  - if you don't choose a license for your software, *no one else can use it!*
- Journals: Usually acquire copyright
  - Most likely, you also assigned the copyright to the supplement to the journal!



# Did you know?

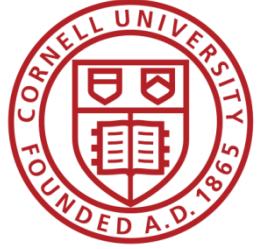
- The AEA acquired the copyright to the supplements when you assigned the article copyright
- The AEA did *not grant a default license*
- Thus, all replication and reproduction attempts are infringing on the AEA's copyright!

**We're fixing that...**



# Best practices in licensing

- Data: Creative Commons Attribution (**CC-BY**) and variants
  - Non-commercial (NC)
  - Share-alike (SA)
- Code: various Open Source licenses
  - GPL: GNU Public License (*restrictive*)
  - MIT or (**modified**) **BSD** license (*less restrictive*)

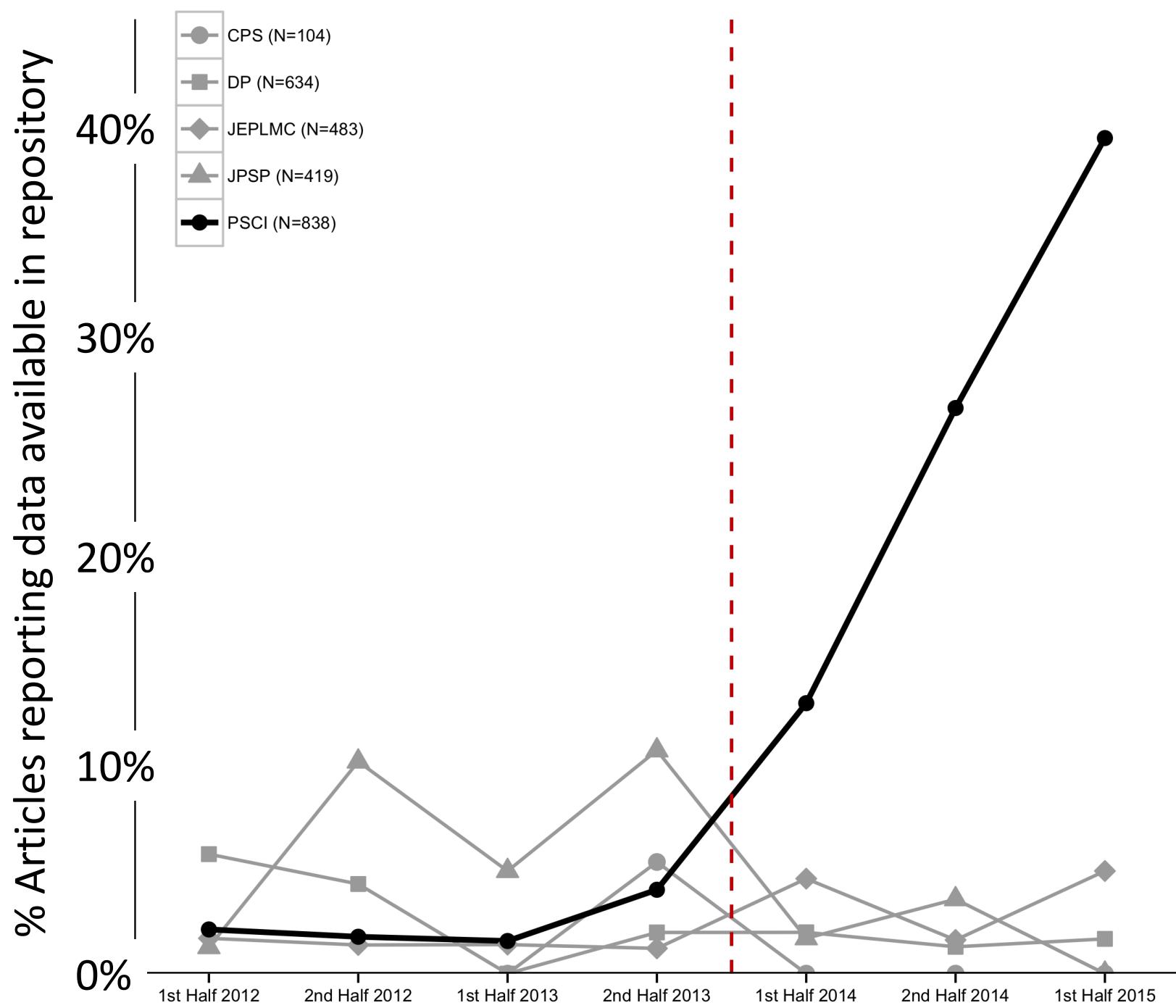


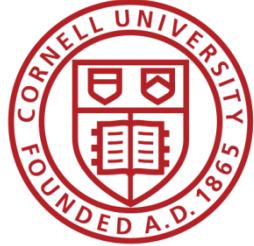
# Other initiatives

# Center for Open Science Badges



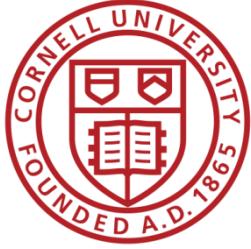
<https://cos.io/our-services/open-science-badges/>; 53 adopting journals (18 as of 2017)  
Kidwell et al., 2016





# Signals? Badges? Transparency!

- Positives:
  - Give credit for previously invisible activities
  - Can be assessed quickly
- Negatives:
  - Need to be known in order to be effective
  - May bias outcomes!
  - Require effort for verification?

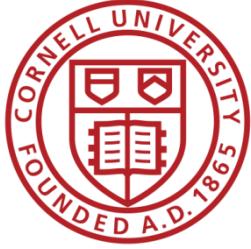


# Example: COS disclosures



**Open Data:** Authors complete two disclosure items for each Open Data badge application:

1. Provide the URL, DOI, or other **permanent** path for accessing the data in a public, **open access** repository.
  - *How do I know it is open access? **Verify!***
  - *How do I know the data is permanent? **Verify?***
  - *What do I do when **access is possible**, but **not open access**?*

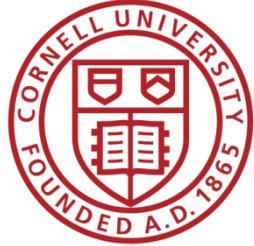


# Example: COS disclosures



**Open Data:** Authors complete two disclosure items for each Open Data badge application:

2. Is there **sufficient** information for an independent researcher to reproduce the reported results? If no, explain.
  - All researchers **claim** the information is sufficient -> our results



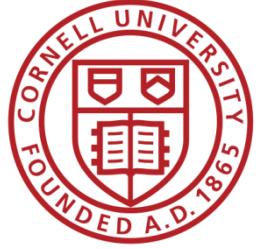
# Example: COS disclosures



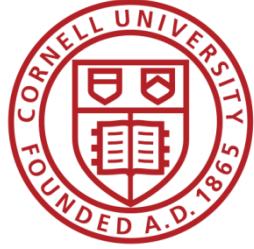
**Open Materials:** Authors complete two disclosure items for each Open Materials badge application:

1. Provide the URL, DOI, or other permanent path for accessing the materials in a public, open access repository.
2. Is there sufficient information for an independent researcher to reproduce the reported methodology? If no, explain.

**Similar concerns as before**



# Pre-registration



# Preregistration

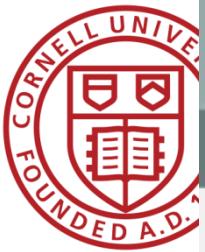
The screenshot shows the homepage of the AEA RCT Registry. At the top, there is a navigation bar with links for "About", "Registration Guidelines", "FAQ", "Advanced Search", and a search input field. Below the navigation bar, a message states "AEA RCT Registry currently lists 2268 studies with locations in 125 countries." On the right, there is a "REGISTER A TRIAL >" button. The main content area is titled "MOST RECENTLY REGISTERED TRIALS". It displays two study entries:

**Tackling sexual harassment: Evidence from India**  
LAST REGISTERED ON JANUARY 26, 2019

Goal 5 of the sustainable development goals adopted by the United Nations in 2015 aims to eliminate all forms of discrimination and violence against women in public and private spheres and to undertake reforms to give women equal rights to economic resources and access to ownership of property. Government of India has identified ending violence against women as a key national priority too. Brutal gangrape of a 23-year-old woman in 2012 in the capital of India led to an outcry against public apathy towards endemic sexual assault and harassment against women. A UN women's study showed that 92% of women surveyed in Delhi had suffered from either sexual, visual or verbal harassment. Pervasive sexual harassment can have debilitating impacts on psychological, economic and social lives of the...

**Malleability of Sustained Attention**  
LAST REGISTERED ON JANUARY 25, 2019

The economics and education literatures traditionally view human capital as an individual's stock of knowledge and skills. In this project, we posit an additional potential component: the capacity for sustained attention. In cognitive psychology, the mind's ability to direct and sustain attention is thought to underlie all activity: cognitive processes (such as solving a math problem) as well as non-cognitive activities (such as exerting self-control) (Chun et al. 2011). In this project, we examine whether the capacity for exerting sustained attention is malleable. Using a field experiment, we introduce a novel tablet-based adaptive learning platform into low-income Indian primary schools. The platform engages students in sustained practice in either mathematics or cognitive activities ...



# AEA RCT Registry

The American Economic Association's registry for randomized controlled trials

[Create Account](#) [Sign in](#)

[About](#) [Registration Guidelines](#) [FAQ](#)

[Advanced Search](#)

**SEARCH**

AEA RCT Registry currently lists 2268 studies with locations in 125 countries.

**REGISTER A TRIAL >**

## MOST RECENTLY REGISTERED TRIALS

### Tackling sexual harassment: Evidence from India

LAST REGISTERED ON JANUARY 26, 2019



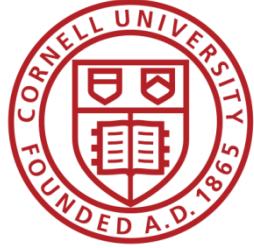
Goal 5 of the sustainable development goals adopted by the United Nations in 2015 aims to eliminate all forms of discrimination and violence against women in public and private spheres and to undertake reforms to give women equal rights to economic resources and access to ownership of property. Government of India has identified ending violence against women as a key national priority too. Brutal gangrape of a 23-year-old woman in 2012 in the capital of India led to an outcry against public apathy towards endemic sexual assault and harassment against women. A UN women's study showed that 92% of women surveyed in Delhi had suffered from either sexual, visual or verbal harassment. Pervasive sexual harassment can have debilitating impacts on psychological, economic and social lives of the...

### Malleability of Sustained Attention

LAST REGISTERED ON JANUARY 25, 2019

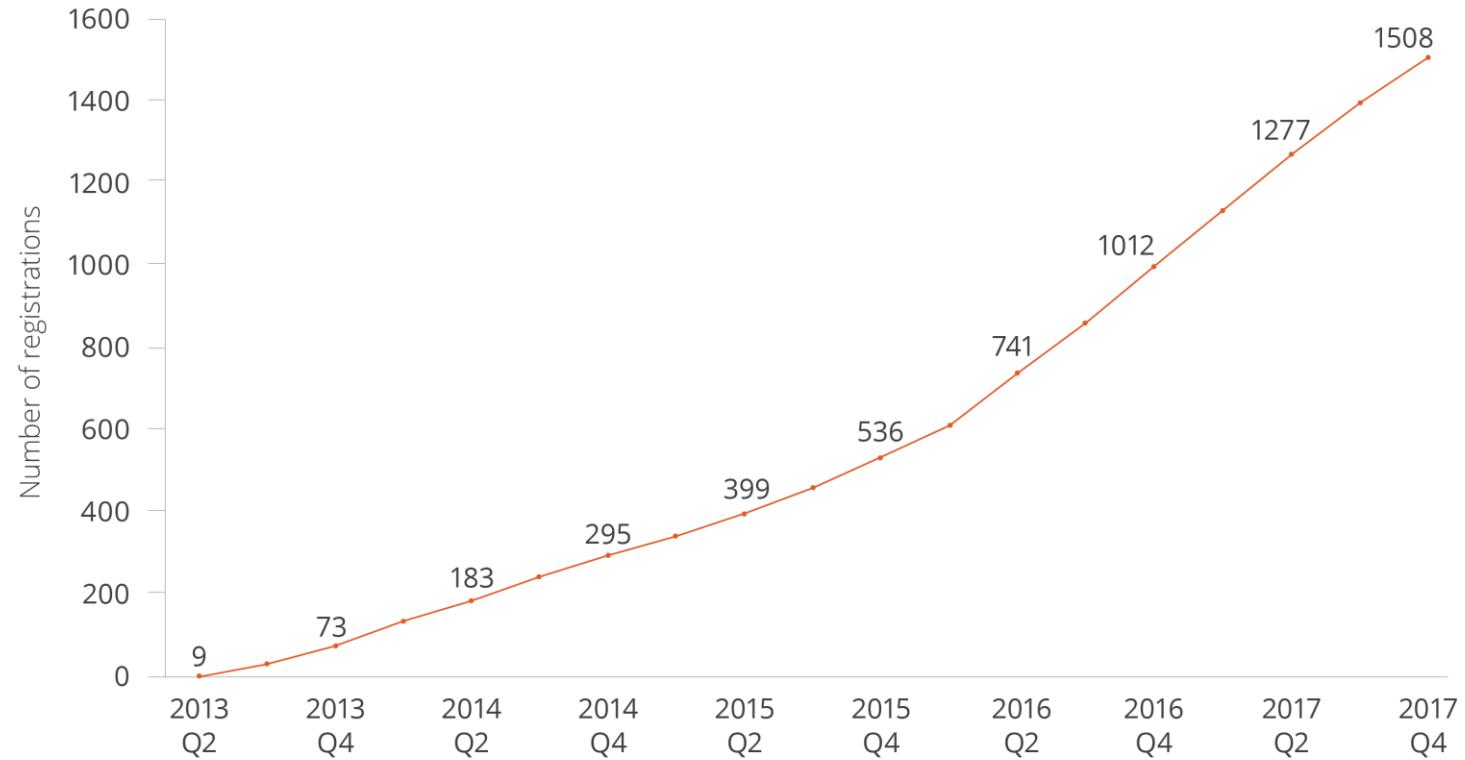


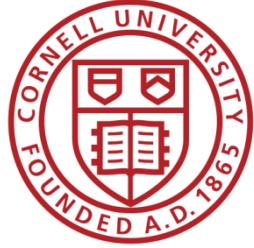
The economics and education literatures traditionally view human capital as an individual's stock of knowledge and skills. In this project, we posit an additional potential component: the capacity for sustained attention. In cognitive psychology, the mind's ability to direct and sustain attention is thought to underlie all activity: cognitive processes (such as solving a math problem) as well as non-cognitive activities (such as exerting self-control) (Chun et al. 2011). In this project, we examine whether the capacity for exerting sustained attention is malleable. Using a field experiment, we introduce a novel tablet-based adaptive learning platform into low-income Indian primary schools. The platform engages students in sustained practice in either mathematics or cognitive activities ...



# Registrations at the AEA Registry

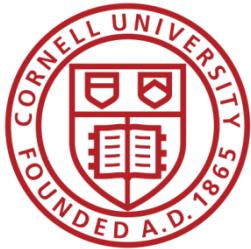
Cumulative registrations over time





# Plans at the AEA

- Integrate registrations more robustly
  - Cite through DOI
  - Link more robustly (not just citation!)
  - Highlight through badges?



# Plans at the AEA

Not just the AEA Registry

- OSF registrations
- Int'l Initiative for Impact Evaluation (3ie)'s RIDIE registry
- Evidence in Governance And Politics (EGAP)'s registry
- As-Predicted

AS PREDICTED

Create a new AsPredicted pre-registration

CREATE

See your existing AsPredicteds (e.g. approve, make public)

SEE OWN

**What's an AsPredicted?**

It is a standardized pre-registration that requires only what's necessary to separate exploratory from confirmatory analyses. You will easily generate a pre-registration document that takes less effort to evaluate than it takes to evaluate the published study itself.

About Terms of use



About Ridie ▾ FAQs for Researcher

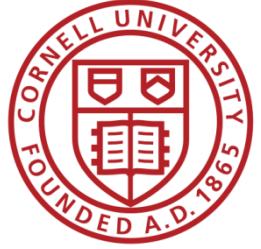
## The Registry for International Development Impact Evaluations

Welcome to the Registry for International Development Impact Evaluations (RIDIE), a project of the International Initiative for Impact Evaluation (3ie). RIDIE is a registry of impact evaluations related to development in low and middle income countries. The purpose of the registry is to enhance the transparency and quality of evaluation research as well as to provide a repository of impact evaluation studies for researchers, funders, and others.

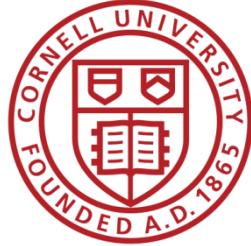
RIDIE will register any development impact evaluation that rigorously attempts to estimate the causal impacts of a program, including but not limited to randomized control trials. It is intended to be a **prospective** registry in which researchers and evaluators can record information about their evaluation designs before conducting the analysis, as well as post findings upon study completion.

ration Should I Register My Study?

- How does it work?**
- One author briefly answers 9 questions.
  - All participating authors receive an email asking for approval.
  - If everyone approves, it is saved and stays private until an author acts to make it public, or it remains private forever. ([Why?](#))
  - Authors may share anonymous .pdf with reviewers.



# Registered Reports



# Registered Reports

[Home](#)[About](#)[Events](#)[Career](#)[Opportunities](#)[Contact us](#)

Berkeley Initiative for  
Transparency in the Social Sciences

[Research](#)[Education](#)[Network](#)[Publishing](#)[Open Policy Analysis](#)[Resources](#)[Blog](#)

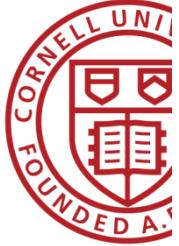
Registered Reports at  
the JDE

[About](#)[Guidelines for Authors](#)[FAQs for Authors](#)[FAQs for Reviewers](#)[Policy Updates](#)[RR Peer Review Flowchart](#)[Stage 1 Proposal Template](#)[JDE's Submissions Portal](#)

## Registered Reports at the *Journal of Development Economics*

As part of a pilot project, the *Journal of Development Economics* ([JDE](#)) now offers authors the opportunity to submit empirical research designs for review and approval *before* the results of the study are known. The *pre-results review* track is designed to award well-designed and well-executed studies regardless of whether their empirical results yield clear interpretations.

This is among the first attempts to introduce pre-results peer review in an economics journal. However, more than [156 academic journals](#) in biology, medicine, political science, psychology, and other disciplines are



FAQs for Reviewers  
Policy Updates  
RR Peer Review Flowchart  
Stage 1 Proposal Template  
JDE's Submissions Portal  
Prospective Articles  
Contact us

empirical results yield clear interpretations.

This is among the first attempts to introduce pre-results peer review in an economics journal. However, more than [156 academic journals](#) in biology, medicine, political science, psychology, and other disciplines are already implementing pre-results review, with the subsequent published articles usually called 'Registered Reports' (RR).

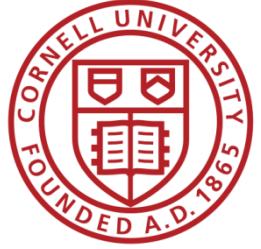
## What is Pre-Results Review?

Pre-results review splits the peer review process into two stages. In Stage 1, authors submit an introduction, methods, and analysis plan for a prospective empirical study for which data have yet to be collected (or accessed) and analyzed. This submission is evaluated based on the importance of the research question(s), the soundness of the theoretical reasoning, and the credibility and feasibility of the research design.

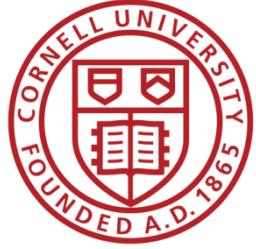


*The publishing process for pre-results review submissions at the JDE*

Positively evaluated submissions are **accepted based on pre-results review**. This constitutes a commitment by the journal to publish the full paper, regardless of the nature of the empirical results. Authors then collect and analyze their data, and submit the full paper (including results and analysis) for final review and publication (Stage 2). The final Stage 2 review provides quality assurance and ensures alignment with the research design peer reviewed in Stage 1.

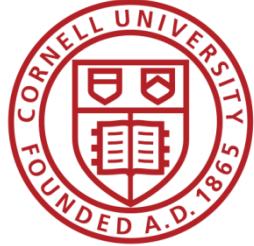


Your turn!



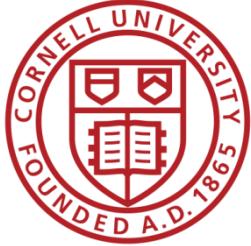
Sharing of code and data is  
coming

Are you ready?



# New skills (to learn, to teach)

- How to incorporate reproducible practices into your workflow
- How to structure your code and your data with an eye on reproducibility
- How to license your contributions!
- When to pre-register, and when not to
- Document early, and often

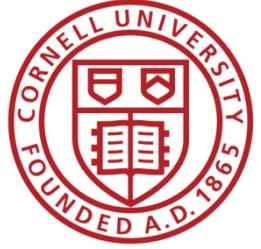


# It is good practice

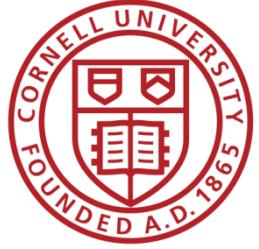
- Hamermesh (2007) already said:

"the payoff is in the influence of one's ideas: having other scholars base their work on those ideas, having students learn from them, and [...] having public policy influenced by them [...] our ideas are unlikely to be taken seriously if our empirical research is not credible"

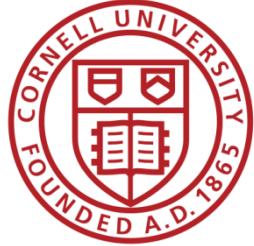




# Preservation



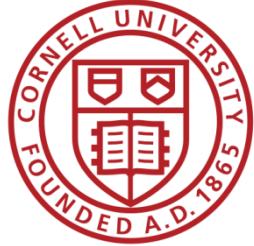
# Replicability in historic context



# Replicability in historic context

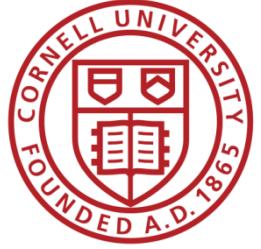
How to establish provenance of results

- Citing
  - Books
  - Tables in compendia
  - Statistical handbooks
- Math done by hand!
- Much use of public statistics

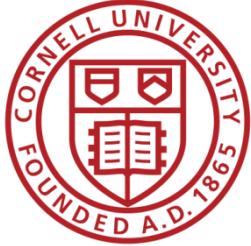


# Long drought?

- Return to this topic
  - Dewald Thursby Anderson, AER 1986 (policy at JMCB)
  - Implementation of data availability policy at AEA (2003/2005)
  - McCullough McGeary Harrison, JMCB 2006



# Definitions



# Replicability and Reproducibility

Bollen et al (2015), NSF Report

- **Reproducibility**

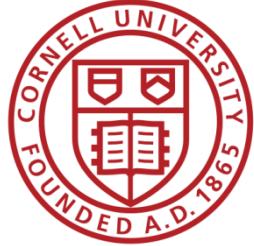
(“narrow” in J Appl Econometrics, “pure replication” Hamermesh (2007), sometimes called “replicability”)

- Same data, same code
- [Programming] aspects of a “unit test”

- **Replicability**

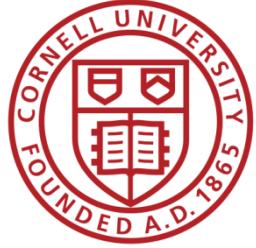
(“wide” in J Appl Econometrics, “scientific replication” Hamermesh (2007), sometimes called “reproducibility”)

- Different data and/or
- Different implementation and/or
- Different assumptions?



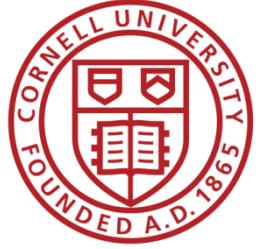
# Purpose of replication archives

- To be able to verify that the “recipe” did produce the results published (log)
- To be able to verify that the “recipe” can be used by others to produce the same results (reproducibility)
- To allow others to leverage and to build on prior work (re-use, efficiency)
- “threat of replication provides an incentive for careful empirical research” (MMH)

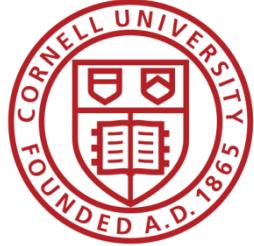


# Key ingredients

- Clear **data provenance**
- **Accessible** data
- Well-structured, comprehensible, **functioning** programs
- Available **without any additional information** or help from the authors

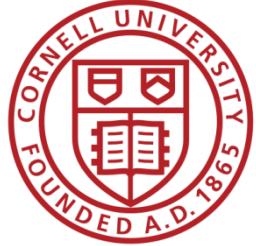


# Problems... old



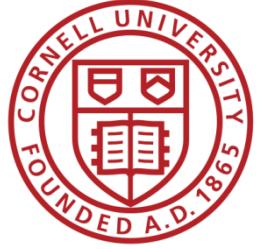
# Some key statistics

Authors	Year	Replication success	Field
Dewald Thursby Anderson	1986	$2/54 = 3.7\%$	Economics
McCullough McGahey Harrison	2006	$14/62 = 22\%$ or $14/186 = 7.5\%$	Economics

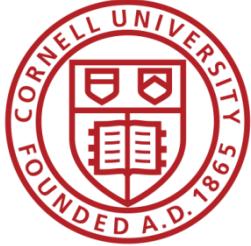


# Requesting copies of code

Authors	Year	Success (=response received)	Field
Dewald Thursby Anderson	1986	22/62= 35.5%	Economics



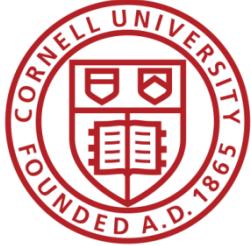
# Moving forward?



# A new world?

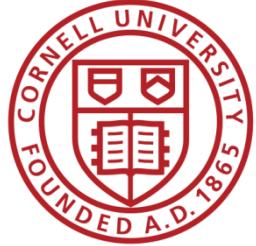
- 2003/2005 Data availability policy of the American Economic Association:

**It is the policy of the American Economic Association to publish papers only if the data used in the analysis are clearly and precisely documented and are readily available to any researcher for purposes of replication.**



# In practice

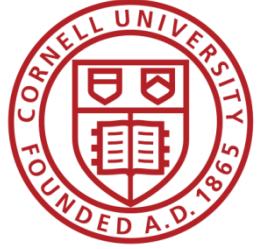
- Data and code archive provided as supplementary online materials
- Single ZIP file on journal's website
  - may contain data
  - may contain code
- Should always contain a “readme” file describing what to do in order to leverage code and data to create
  - the paper’s tables as published
  - the paper’s figures as published
  - any in-text results (less frequent!)



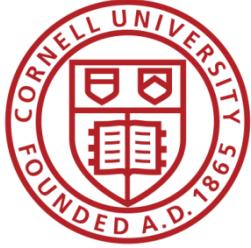
# Requesting copies of code

Authors	Year	Success (=response received)	Field
Dewald Thursby Anderson	1986	22/62= 35.5%	Economics





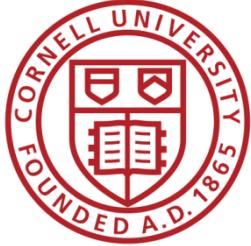
A glimpse into the future



# Ongoing elsewhere

## Journals

- have implemented **verification** of submitted code and data during the editorial process (AJPS with Odun Institute; JASA)
- **highlight** data and code access criteria (badges: OSF)  
Three hexagonal badges for OSF: "OPEN DATA" (blue), "OPEN MATERIALS" (orange), and "PRERE REGISTERED" (red).
- maintain **lists of acceptable third-party repositories** (and embed some of those within the submission workflow) (Nature, CoreTrustSeal)
- **interlink** with collaborating repositories to highlight authors' (and repositories') contributions to the data component of a scholarly work (Elsevier + ICPSR)

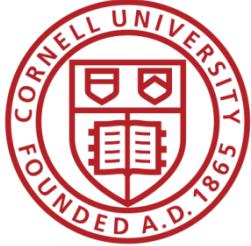


# Ongoing elsewhere

## Services

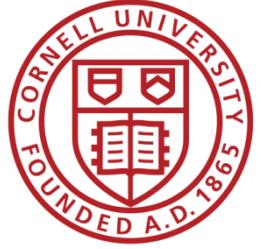
- For free or low-cost archiving  
([openICSPR](#), [Zenodo](#), [figshare](#), [dryad](#), [Dataverse](#))
- Online computational capsules  
[CodeOcean](#) (object includes OS, verification)

(note: I did not include Github, Gitlab, etc. – on purpose)

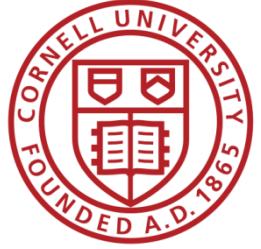


# Still early

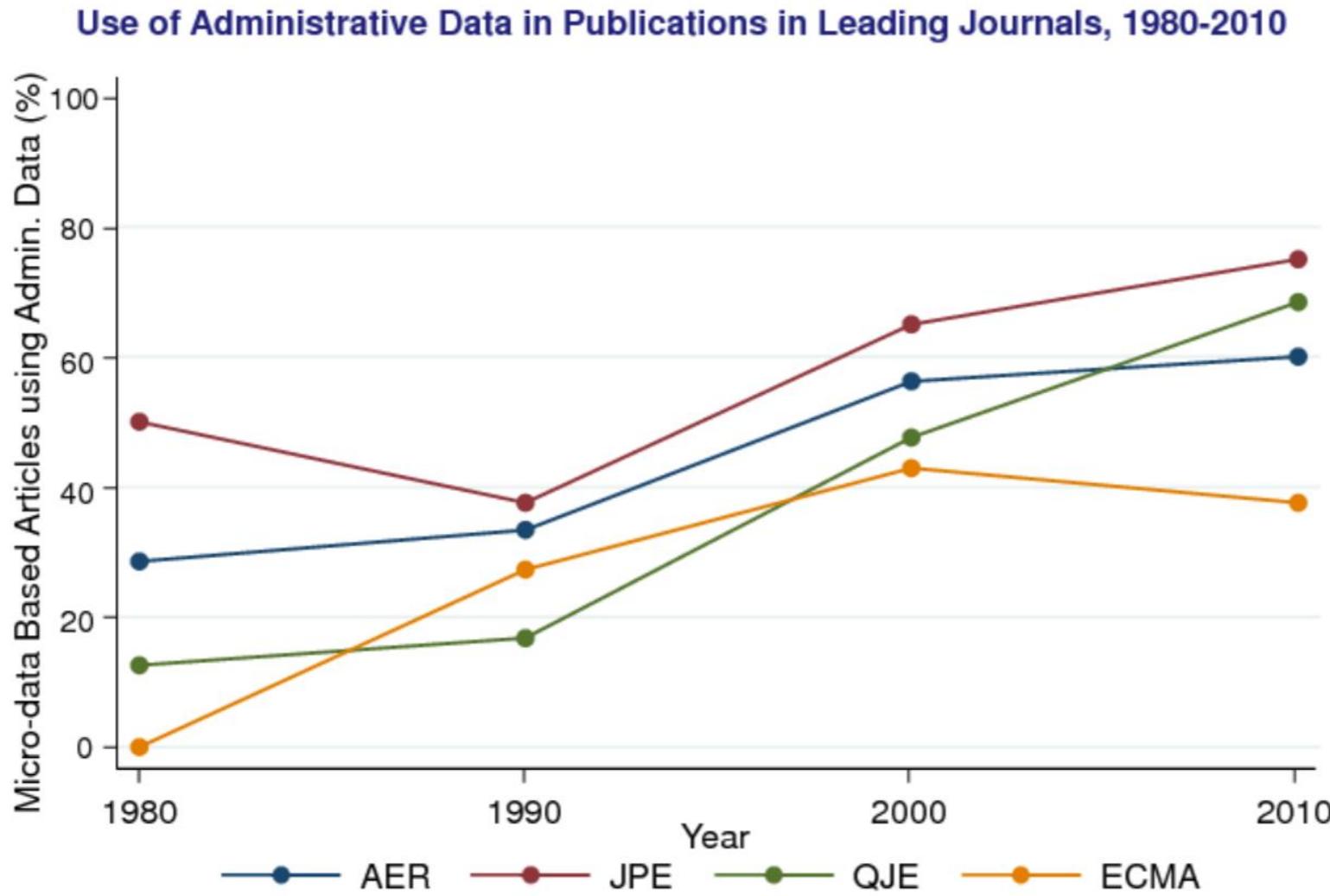
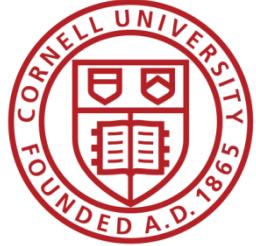
- **Badges** are visual help, but not robust yet (subject to changes, do not cover all situations)
- **Verification** can be time-consuming / costly (but see some natural sciences)
- **List of repositories** differ across journals/ providers/ central registries (OSF, Nature, re3data, Fairsharing: which is “right”? What are the criteria?)



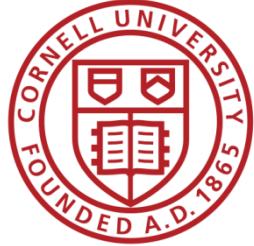
And then there was....



# Confidential data



Note: "Administrative" data refers to any dataset that was collected specifically for another purpose than the one it is used for in the study (e.g., scanner data, stock prices, school district records, social security records).  
Sample excludes studies whose primary data source is from developing countries.



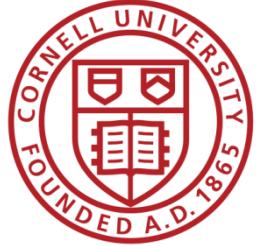
# Number of papers with “proprietary” data

Table 3: Type of Confidential Data

	Admin local	Admin National	Admin Regional	Private Commercial	Private Other	Total	
2009	0	2	1	0	1	4	/8
2010	2	8	0	4	3	17	/35
2011	2	9	4	1	0	16	/36
2012	1	10	2	0	2	15	/40
2013	2	2	1	4	2	11	/38
Total	7	31	8	9	8	63	<b>/157</b>

Table 4: Type of Access to Confidential Data

	Formal	Informal Commitment	Informal No Commitment	No Info	Total
2009	4	0	0	0	4
2010	2	3	9	3	17
2011	3	0	10	3	16
2012	1	1	0	2	15
2013	1	2	8	0	11
Total	22	6	27	8	63



# Confidentiality and Accessibility

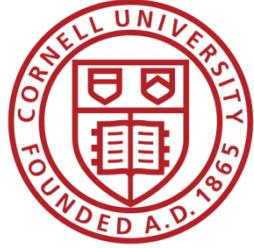


Confidential  
data



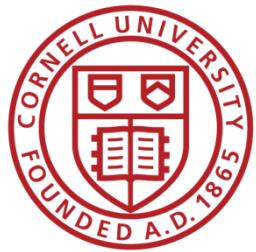
Confidentiality

Accessibility

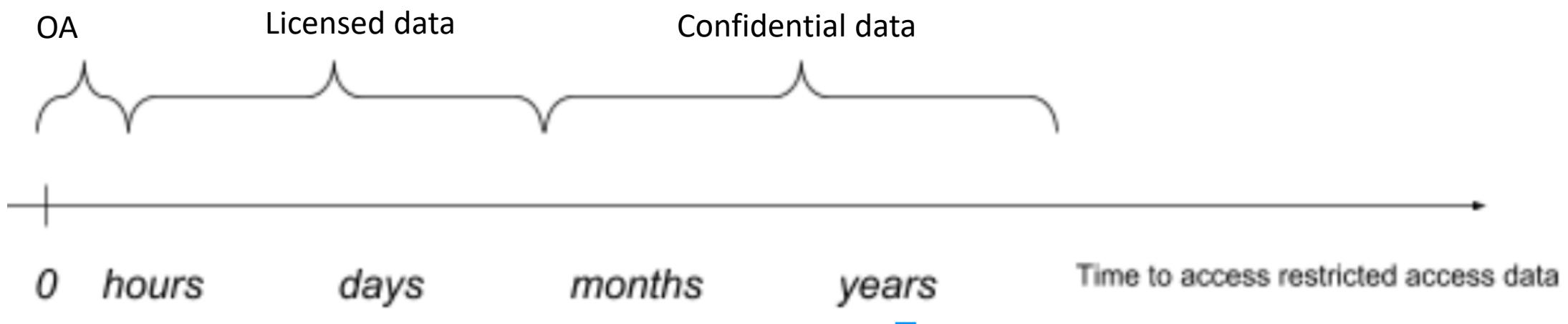


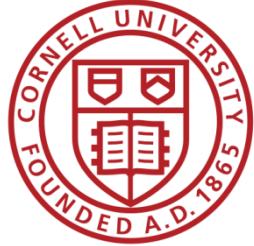
# Types of data access

- **Proprietary:** often meant for all restricted-access, more narrowly for commercial data
- **Confidential:** for any number of reasons, data cannot be anonymized without affecting the statistical validity of the data and the results
- **Licensed data:** data which is subject to conditions (use, access, redistribution)
- **Open-access data:** often used for “free” or “unencumbered” data



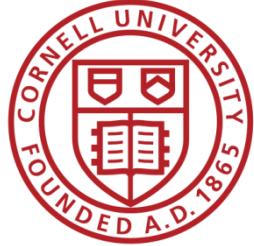
# Alternate view





# Black Box problem

- We have less insight into the “proprietary” data papers
- Less success in “help by authors” in general
- Information problem



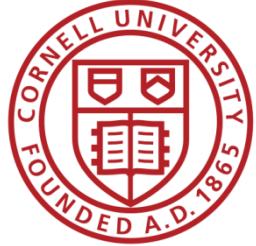
# Making data (more) accessible

National Data  
Bank (1965)

Committee on  
the Preservation  
and Use of  
Economic Data  
(1965)

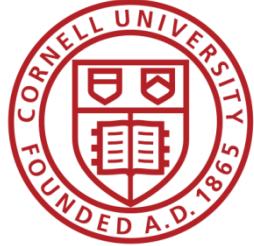
1959-  
1965

“Recommendations  
on Availability of  
Federal Statistical  
Materials to  
Nongovernmental  
Research  
Workers”  
(1959)



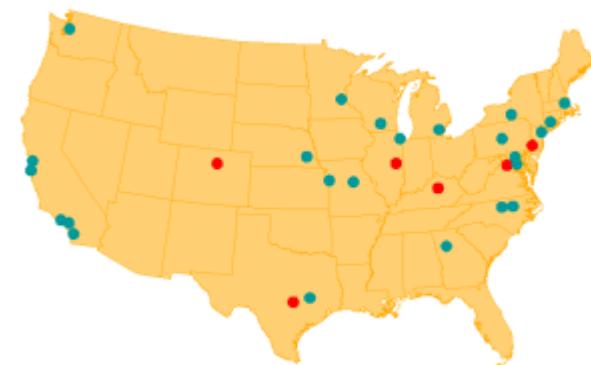
# researchers knocking on the door

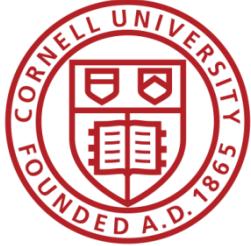




# RADEs in the US

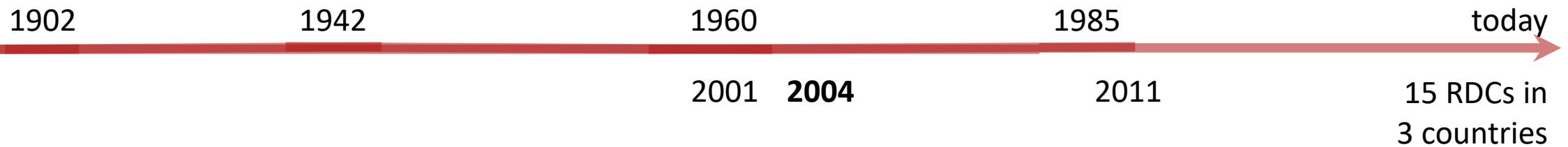
- Starting in the 1960s and 70s, increased use of public-use microdata samples and surveys
- Researcher access at Census Bureau headquarters in the 1970s
- 1990: [Computing power: 3.5 MFLOPS for \$9000]
- First RDC at Boston in the 1994
- A small number of RDCs in the 1990s
- Thin clients in the 2000s
- 2018: 31 RDCs

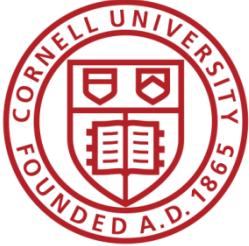




# RADEs: Germany

- Institute for Employment Research (IAB), Germany
  - *Commission to improve the informational infrastructure between the scientific community and official statistics* (KVI) recommended creation of RDCs by producers of microdata (2001)
  - RDC created in 2004 for “weakly anonymous” data
  - Scientific use files (factually anonymous data) available under licensing agreements to university data enclaves
  - 2011 RDC created at University of Michigan (with NSF funding)



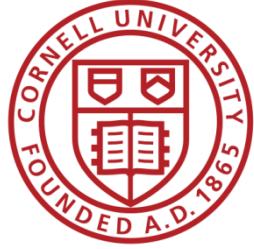


# RADEs: France

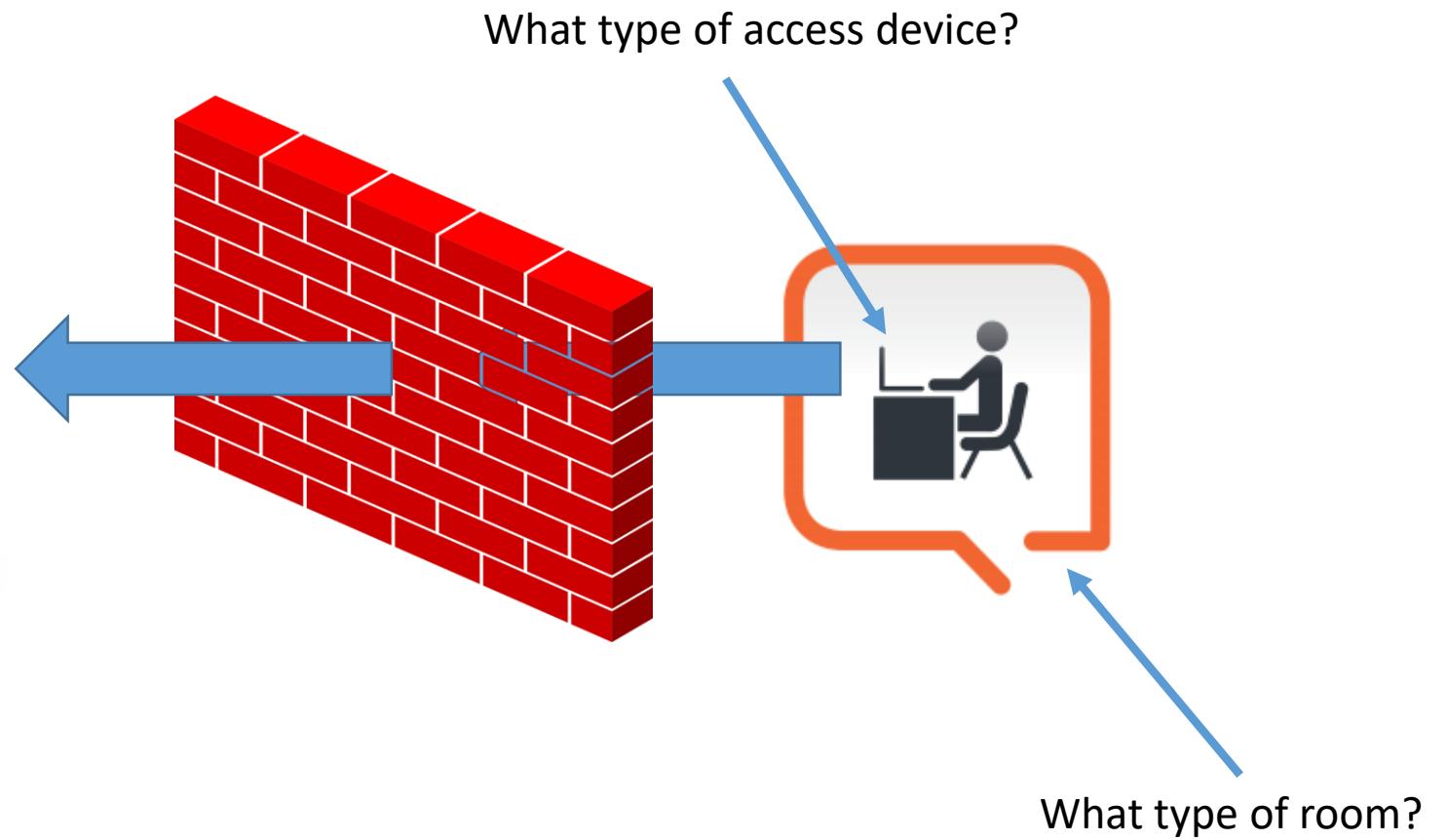
- Centre d'accès sécurisé distant (CASD, France)
  - Note: within same agency that enabled AKM (1999)
  - INSEE recommended implementing a secure center for
  - 2008 modification to Statistics Law made possible pilot
  - Pilot infrastructure becomes permanent in 2009
  - Expansion with per-project cost (invoicing) in 2012

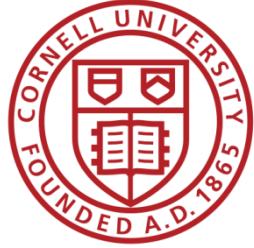


Source: presentation to the CASD Scientific Committee in 2013

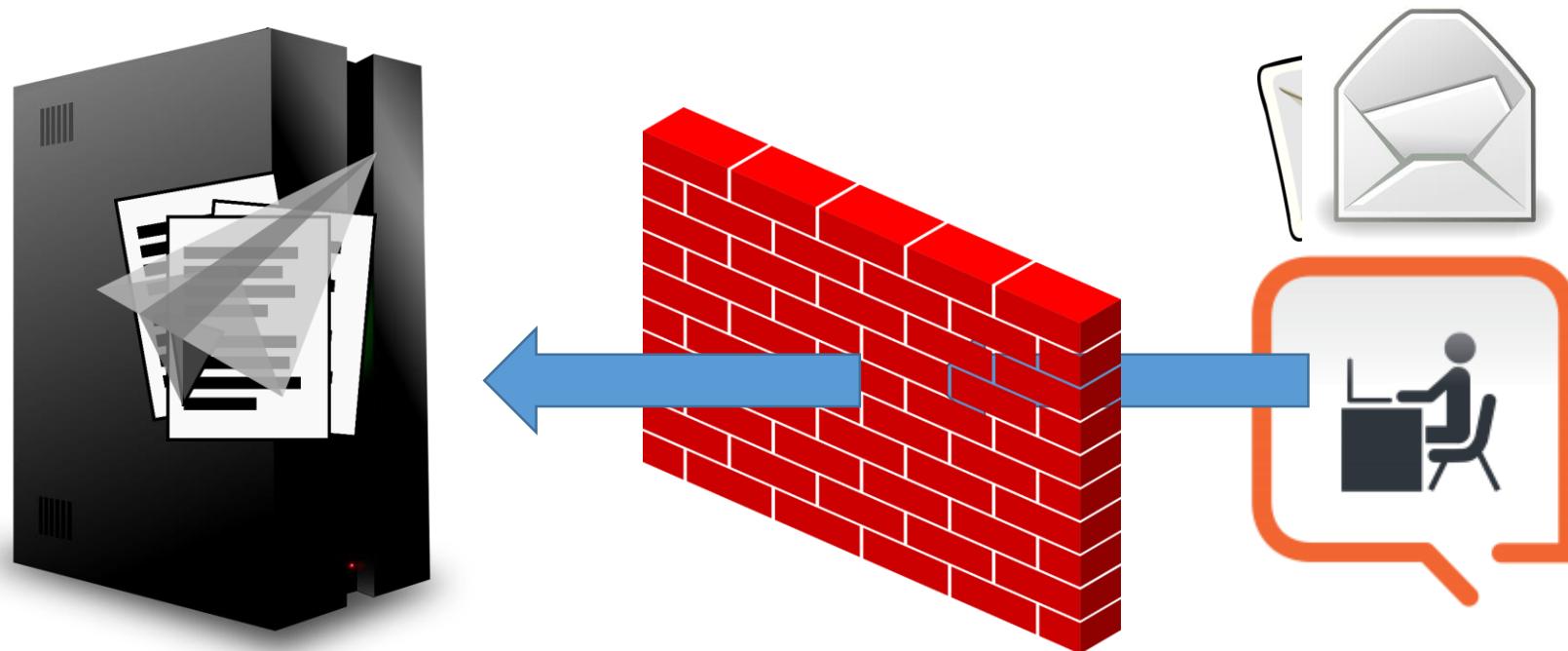


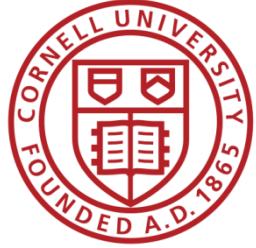
# basic paradigm



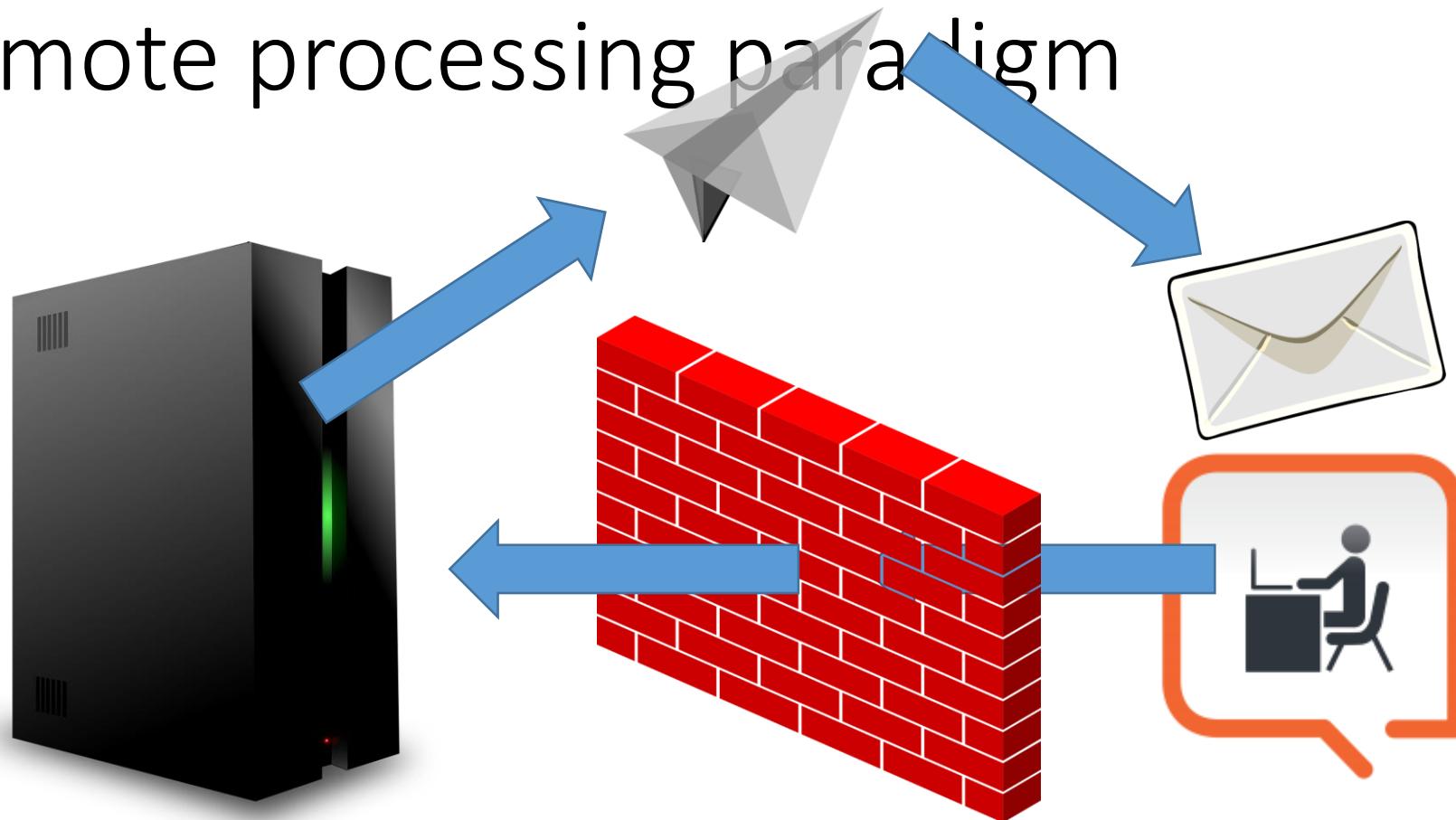


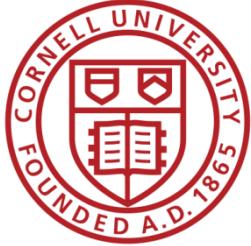
# remote processing paradigm





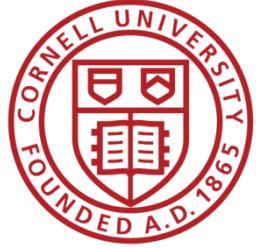
# remote processing paradigm





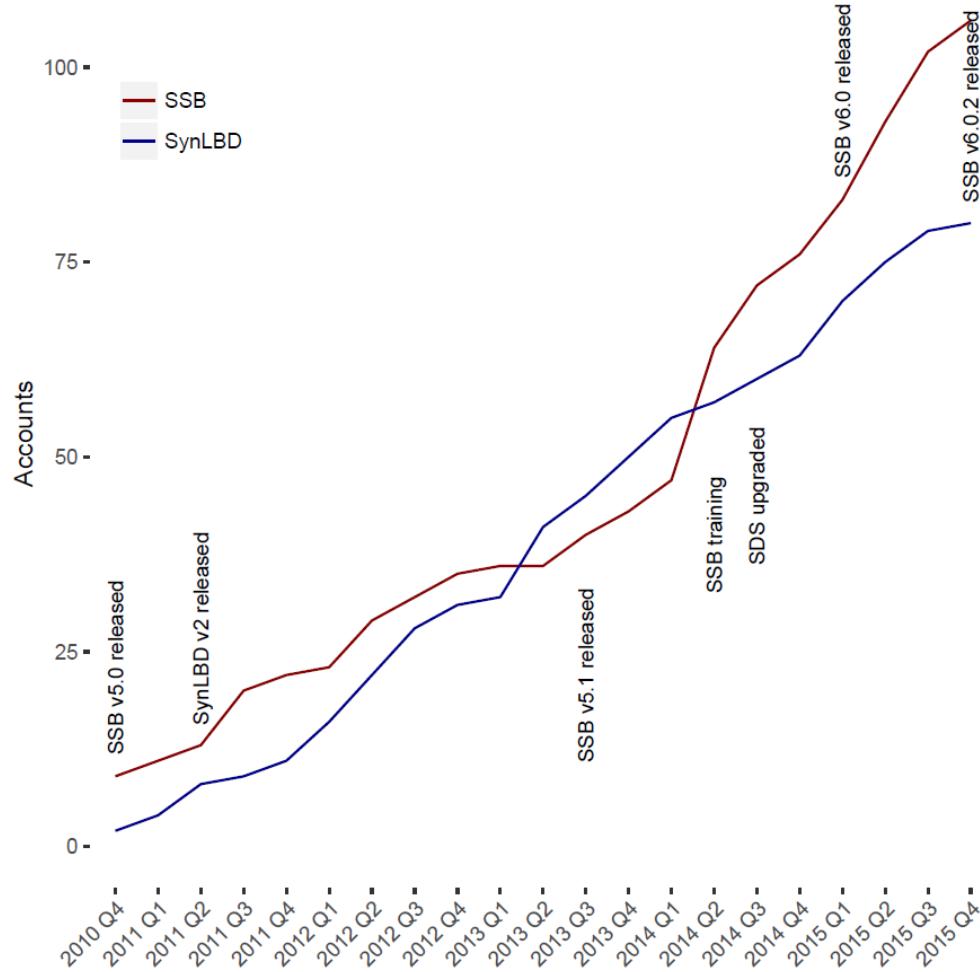
# synthetic data: validation model

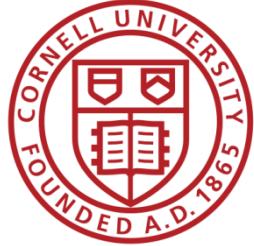
- Used for SIPP Synthetic Beta, Synthetic LBD
  - Researcher **develops model** on synthetic data
  - Assessment through **submission of programs** for “validation”
  - Researcher obtains results from **programs run on confidential data**, subject to traditional disclosure avoidance rules
  - No restrictions on types of models
- In progress since 2011
  - Approx. 200 users
  - Approx. 6-8% of users request validation
  - Some unknown fraction “self-validate” through full FSRDC project



# some results from Synthetic Data Server

6 years, 5 (versions of) synthetic datasets, over 180 users

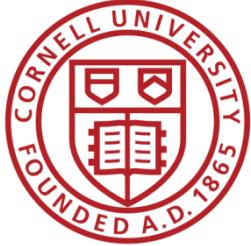




# Key insight (Lagoze and Vilhuber, 2017)

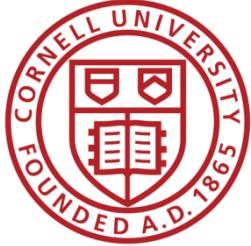
**This cycle is already a part of replicable workflow:**

- Programs are run by third party on same (some RDCs)/ similar (synthetic data)
- Accompanying documentation is needs to detail creation of results
- Results are typically logged
- (person conducting replication or reproducibility exercise)
- This serves the same function as the typical “README” from replication archives!
- Allows for assignment of DOIs!



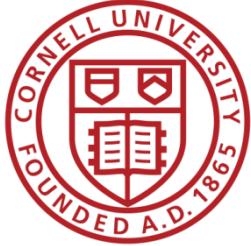
# Our recommendations

- Access is already non-exclusive, and satisfied the reproducibility criteria
  - Document the access protocol clearly and in a citable, persistent fashion. Ensure access protocol is transparent and predictable
- Provide to the researcher the information already present in administrative databased of RADEs (access rights)
  - Provenance of input data to which the researcher has access (data citation, metadata)
- Systematically release the programs used to generate the output
  - Code archives for journals
- Catalog the result files, possibly certify that they were generated from the programs
  - Provide a provable chain for the results in journals



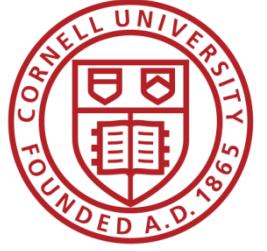
# Challenges

- Creating a public catalog of data assets with
  - persistent identifiers,
  - best-practice data citations,
  - public metadata
- Possible challenges in certifying that released results were created with the identified code
  - Used to be practice at IAB?
  - Often is a tabletop exercise (“plausibly generated by the code”)

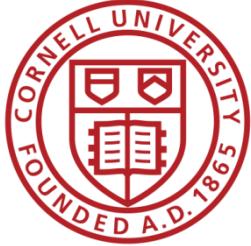


# Benefit

- Data in RADEs can become FAIR at short notice:
  - **Findable**
  - **Accessible**
  - **Interoperable**
  - **Reusable**



# Outlook

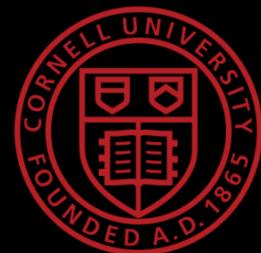


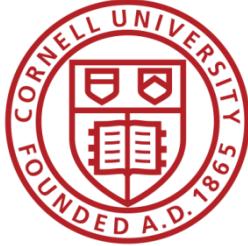
# Moving forward in economics

- Move away from “open access only” and “deposit at journal”
- Systematically require that data be in a trusted repository
  - Trusted can be controlled but not (too) discriminatory access
  - Clear documentation of access policy, retention policy
  - **Uniform treatment of open access, licensed, and confidential data**
- Systematically (but with sampling) verify reproducibility of code
  - Straightforward for OA and “easy-access” data
  - Less obvious, but not impossible for restricted-access (may be part of the process of releasing disclosable results)
- Provide incentives and guidance to researchers to incorporate replicability into scientific workflow right from the start (not an afterthought)

thank you

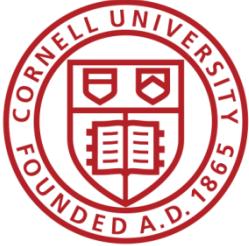
[lars.vilhuber@cornell.edu](mailto:lars.vilhuber@cornell.edu)





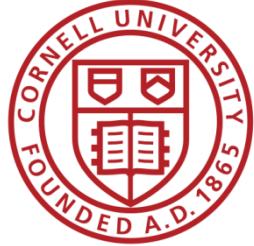
# Some References

- Anderson, Richard G, William H Greene, Bruce D McCullough, and H D Vinod. "The Role of Data and Program Code Archives in the Future of Economic Research." Working Paper. Federal Reserve Bank of St. Louis, 2005. <http://research.stlouisfed.org/wp/2005/2005-014.pdf>.
- Bender, Stefan, and Jörg Heining. "The Research-Data-Centre in Research-Data-Centre Approach: A First Step Towards Decentralised International Data Sharing." *IASSIST Quarterly / International Association for Social Science Information Service and Technology* 35, no. 3 (2011). <http://www.iassistdata.org/iq/issue/35/3>.
- Bollen, Kenneth, John T. Cacioppo, Robert M. Kaplan, Jon A. Krosnick, and James L. Olds. "Social, Behavioral, and Economic Sciences Perspectives on Robust and Reliable Science." Report of the Subcommittee on Replicability in Science Advisory Committee to the National Science Foundation Directorate for Social, Behavioral, and Economic Sciences. National Science Foundation, 2015. [https://www.nsf.gov/sbe/AC Materials/SBE Robust and Reliable Research Report.pdf](https://www.nsf.gov/sbe/AC%20Materials/SBE%20Robust%20and%20Reliable%20Research%20Report.pdf).
- Camerer, Colin F., Anna Dreber, Eskil Forsell, Teck-Hua Ho, Jürgen Huber, Magnus Johannesson, Michael Kirchler, et al. "Evaluating Replicability of Laboratory Experiments in Economics." *Science*, March 3, 2016, aaf0918. <https://doi.org/10.1126/science.aaf0918>.
- Chang, Andrew C., and Phillip Li. "A Preanalysis Plan to Replicate Sixty Economics Research Papers That Worked Half of the Time." *American Economic Review* 107, no. 5 (May 2017): 60–64. <https://doi.org/10.1257/aer.p20171034>.
- Christian, Thu-Mai, Sophia Lafferty-Hess, William Jacoby, and Thomas Carsey. "Operationalizing the Replication Standard: A Case Study of the Data Curation and Verification Workflow for Scholarly Journals," 2018. <https://doi.org/10.17605/osf.io/cfdb>.
- Clemens, Michael A. "The Meaning of Failed Replications: A Review and Proposal." *Journal of Economic Surveys* 31, no. 1 (February 1, 2017): 326–42. <https://doi.org/10.1111/joes.12139>.
- Collaboration, Open Science. "Estimating the Reproducibility of Psychological Science." *Science* 349, no. 6251 (August 2015): aac4716–aac4716. <https://doi.org/10.1126/science.aac4716>.
- Dewald, William G, Jerry G Thursby, and Richard G Anderson. "Replication in Empirical Economics: The Journal of Money, Credit and Banking Project." *The American Economic Review* 76, no. 4 (1986): 587–603.
- Frisch, Ragnar. "Editor's Note." *Econometrica: Journal of the Econometric Society* 1, no. 1 (1933): 1–4.



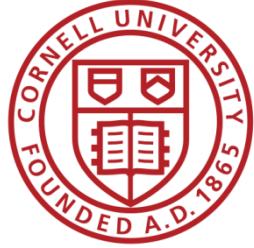
# Some References

- Hamermesh, Daniel S. "Viewpoint: Replication in Economics." *Canadian Journal of Economics/Revue Canadienne d'économique* 40, no. 3 (August 1, 2007): 715–33. <https://doi.org/10.1111/j.1365-2966.2007.00428.x>.
- Höffler, Jan H. "Replication and Economics Journal Policies." *American Economic Review* 107, no. 5 (May 2017): 52–55. <https://doi.org/10.1257/aer.p20171032>.
- Jacoby, William. "AMERICAN JOURNAL OF POLITICAL SCIENCE GUIDELINES FOR PREPARING REPLICATION FILES." *American Journal of Political Science*, 2005. <https://ajpsblogging.files.wordpress.com/2015/03/ajps-guide-for-replic-materials-1-0.pdf>.
- Kidwell, Mallory C., Ljiljana B. Lazarević, Erica Baranski, Tom E. Hardwicke, Sarah Piechowski, Lina-Sophia Falkenberg, Curtis Kennett, et al. "Badges to Acknowledge Open Practices: A Simple, Low-Cost, Effective Method for Increasing Transparency." Edited by Malcolm R Macleod. *PLOS Biology* 14, no. 5 (May 12, 2016): e1002456. <https://doi.org/10.1371/journal.pbio.1002456>.
- Kinney, Satkartar K., Jerome P. Reiter, and and Javier Miranda. "SynLBD 2.0: Improving the Synthetic Longitudinal Business Database." *Statistical Journal of the International Association for Official Statistics* 30 (2014): 129–135.
- Kinney, Satkartar K, Jerome P Reiter, Arnold P Reznek, Javier Miranda, Ron S Jarmin, and John M Abowd. "Towards Unrestricted Public Use Business Microdata: The Synthetic Longitudinal Business Database." *International Statistical Review* 79, no. 3 (December 2011): 362–384. <https://doi.org/10.1111/j.1751-5823.2011.00153.x>.
- Kohlmann, Annette. "The Research Data Centre of the Federal Employment Service in the Institute for Employment Research." *Schmollers Jahrbuch : Journal of Applied Social Science Studies / Zeitschrift Für Wirtschafts- Und Sozialwissenschaften* 125, no. 3 (2005): 437–47.
- Kraus, R. "Statistical Déjà vu: The National Data Center Proposal of 1965 and Its Descendants." *Journal of Privacy and Confidentiality*, 2013. <http://repository.cmu.edu/jpc/vol5/iss1/1/>.



# Some References

- Lagoze, Carl, and Lars Vilhuber. "Making Confidential Data Part of Reproducible Research." *Chance*, 2017 2017. <http://chance.amstat.org/2017/09/reproducible-research/>.
- McCullough, B D, Kerry Anne McGeary, and Teresa D Harrison. "Lessons from the JMBC Archive." *Journal of Money, Credit, and Banking* 38, no. 4 (June 2006): 1093–1107.
- U.S. Census Bureau. *SIPP Synthetic Beta Version 6.0.2*. Washington, DC and Ithaca, NY, USA, 2015. <http://www2.vrdc.cornell.edu/news/data/sipp-synthetic-beta-file/>.
- Vilhuber, Lars, John M. Abowd, and Jerome P. Reiter. "Synthetic Establishment Microdata Around the World." *Statistical Journal of the International Association of Official Statistics* 32, no. 1 (2016): 65–68.



# Picture references

- Tug of War 1: Johnmoore6 at English Wikipedia - CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=16357035>
- Tug of War 2: Bart Derksen at Dutch Wikipedia, CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=3218716>