

ICPSR



Journal Data Workflows with ICPSR

Jared Lyle

“Wishlists and Workflows: Integrating Research Transparency into Editorial and Publishing Processes” Data-PASS Pre-APSA Workshop
Washington, D.C.

28 August 2019



- DATA & PUBLICATIONS
- EVERYTHING ELSE

Search Data and Publications

Search

Log In/Create Account

ICPSR

Sharing data to advance science

News



Jun 22, 2017

An ICPSR team gives viewers a spin through our new Data Manager, and a peek at what's next



Jun 19, 2017

ICPSR celebrates launch of new data deposit system!



Jun 16, 2017

Father's Day stats from US Census Bureau

More Headlines

Jun 12, 2017

New Releases through 2017-06-11

May 31, 2017

Due June 19: Call for abstracts for the Jamaica Statistics Symposium 2017, being put on by The Jamaica Statistical Society

May 12, 2017

"Who was the most important person in raising you?" Explore data from the Flint Adolescent Study

[view all news](#)

<http://www.icpsr.umich.edu>

ICPSR



Philip Converse, Warren Miller, and Angus Campbell

- Established 1962
- Originally 21 Members, now consortium of 776 world-wide
- Originally Political Science, now all social and behavioral sciences

ICPSR

- Current holdings
 - 11,000+ studies, many curated
 - 1500+ are *restricted studies*, almost always to protect confidentiality
 - Bibliography of Data-related Literature with 84,000 citations
- Approximately 60,000 active MyData (“shopping cart”) accounts
- Thematic collections of data about addiction and HIV, aging, arts and culture, child care and early education, criminal justice, demography, health and medical care, and minorities

Curated Collections

The 1915 Iowa State Census Project (ICPSR 28501)


Version Date: Dec 14, 2010 [Cite this study](#) | [Share this page](#)

Principal Investigator(s):

[Claudia Goldin](#), Harvard University, and National Bureau of Economic Research; [Lawrence Katz](#), Harvard University, and National Bureau of Economic Research

<https://doi.org/10.3886/ICPSR28501.v1>

Version V1

Download 

 Analyze Online (0)

At A Glance

Data & Documentation

Variables

Data-related Publications

Export Metadata

▼ Project Description



Summary

The 1915 Iowa State Census is a unique document. It was the first census in the United States to include information on education and income prior to the United States Federal Census of 1940. It contains considerable detail on other aspects of individuals and households, e.g., religion, wealth and years in the United States and Iowa. The Iowa State Census of 1915 was a complete sample of the residents of the state and the returns were written by census takers (assessors) on index cards. These cards were kept in the Iowa State Archives in Des Moines and were microfilmed in 1986 by the Genealogical Society of Salt Lake City. The census cards were sorted by county, although large cities (those having more than 25,000 residents) were grouped separately. Within each county or large city, records were alphabetized by last name and within last name by first name. This data set includes individual-level records for three of the largest Iowa cities (Des Moines, Dubuque, and Davenport; the Sioux City films were unreadable) and for ten counties that did not contain a large city. (Additional details on sample selection are available in the documentation). Variables include name, age, place of residence, earnings, education, birthplace, religion, marital status, race, occupation, military service, among others. Data on familial ties between records are also included.

Citation

Goldin, Claudia, and Katz, Lawrence. The 1915 Iowa State Census Project. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2010-12-14. <https://doi.org/10.3886/ICPSR28501.v1>

Export Citation:

- [RIS](#)  (generic format for RefWorks, EndNote, etc.)
- [EndNote](#) 

1,938

Downloads *
[Usage Report](#)

* past three years

12

[Data-related Publications](#)

Notes

- The public-use data files in this collection are available for access by the general public. Access does not require affiliation with an ICPSR member institution.
- The citation of this study may have changed due to the new version control system that has been implemented.



This study was originally processed, archived, and disseminated by [Data Sharing for Demographic Research \(DSDR\)](#), a project funded by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD).

Restricted-Use Holdings

- Over 1,500 restricted-use datasets available for analysis
- Approximately 20 datasets accessible only on site in physical enclave
- Over 2,000 active restricted-use data agreements
 - Across 500+ institutions
 - Representing 150+ studies
 - Approximately 50 new agreements executed per month

Methods of Accessing Restricted-Use Data

- ICPSR offers three methods of restricted-use data access:
 - Physical Enclave
 - Secure Dissemination
 - Virtual Data Enclave (VDE)

Principal Investigator(s): Rosay, Andre, University of Alaska-Anchorage; Henry, Tara, Forensic Nurse Services, Inc.

Summary:

This project examined the characteristics of sexual assault victimizations in Alaska, as observed and recorded by sexual assault nurse examiners in Anchorage, Kodiak, Bethel, Soldotna, Nome, Fairbanks, Homer, and Kotzebue. The sample utilized for this study included all sexual assault nurse examinations conducted in Anchorage from 1996 to 2004, in Bethel and Fairbanks in 2005 and 2006, and in Homer, Kodiak, Kotzebue, Nome, and Soldotna in 2005. A total of 1,699 examinations were collected. More ... ([more info](#))

Access Notes

- One or more files in this study are not available for download due to special restrictions ; consult the [restrictions note](#) to learn more. You can [apply online for access to the data](#). A login is required to apply for access.

Dataset(s)

Dataset

Documentation: [Codebook.pdf](#) [Questionnaire.pdf](#) [User guide.pdf](#)

Download: No downloadable data files available.

Bibliography of Data-related Literature

Data citation allows us to answer:

- Who uses the data?
- How are they used?
- With what impact?

Examples of a poor data citation

Poorly described and cited data

differently. Using survey data from 272 staff who worked at a Midwestern maximum-security state prison, this study examined the association between the occupational stressors of perceived dangerousness of the job, role conflict, role ambiguity, repetitiveness, and work-on-family conflict with the three forms of organizational commitment. The effects of the occupational stressors varied for each form of organizational commitment. Specifically, work-on-family conflict had a significant positive association with continuance commitment. Role conflict and repetitiveness had negative associations with moral commitment. Finally, all five stressors had significant negative associations with affective commitment.

Excessive human search effort, extensive collection knowledge

Examples of a good data citation

Citing data with a DOI

Monto, M. A. (1999a). *Clients of street prostitutes, in Portland, Oregon, San Francisco and Santa Clara, California, and Las Vegas, Nevada, 1996-1999* (ICPSR version). Portland, OR: University of Portland [Producer], Ann Arbor, MI: Inter-university Consortium for Political and Social Research [Distributor], 2000. doi:10.3886/ICPSR02859.v1

Smith, T. W., Marsden, P. V., & Hout, M. *General Social Survey, 1972-2010* [Cumulative File] (ICPSR31521-v1). Storrs: Roper Center for Public Opinion Research, University of Connecticut/Ann Arbor, MI: Inter-university Consortium for Political and Social Research [Distributors], 2011-08-01. doi:10.3886/ICPSR31521.v1

Minimal human search effort

openICPSR


OPEN ICPSR Find Data Share Data openICPSR Repositories ▾

GO Sign In







Find Data / Intergenerational Education Mobility Trends by Race and Gender in the United States


Intergenerational Education Mobility Trends by Race and Gender in the United States

Principal Investigator(s):  Joseph Ferrare, University of Washington Bothell

Version:  V1

AERA

Name 	File Type 	Size 	Last Modified 
 InterGenEdMobility_AERA_Open.sps	text/x-spss-syntax	26.7 KB	08/27/2019 09:48:AM
 Read_Me.docx	application/vnd.openxmlformats-officedocument.wordprocessingml.document	13.8 KB	08/27/2019 09:48:AM

 DOWNLOAD THIS PROJECT

Project Citation:

Ferrare, Joseph. Intergenerational Education Mobility Trends by Race and Gender in the United States. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2019-08-27. <https://doi.org/10.3886/E111586V1>

Persistent URL: <http://doi.org/10.3886/E111586V1>

Usage Metrics

Overall Project Metrics

5

Views

1

Downloads

1

Publications

[Download Detailed Metrics](#)

Published Versions

[V1 \[2019-08-27\]](#)

Export Metadata

[Dublin Core](#)

[DDI 2.5](#)

Project Description

Summary:

Researchers have examined racial and gender patterns of intergenerational education mobility, but less attention has been given to the ways that race and gender interact to further shape these relationships. Based on data from the General Social Survey, this study examined the trajectories of education mobility among Blacks and Whites by gender over the past century. Ordinary least squares and logistic regression models revealed three noteworthy patterns. First, Black men and women have closed substantial gaps with their White counterparts in intergenerational education mobility. At relatively low levels of parental education, these gains have been experienced equally among Black men and women. However, Black men are most disadvantaged at the highest levels of parental education relative to Black women and Whites in general. Finally, the advantages in education mobility experienced by White men in the early and midpart of the 20th century have largely eroded. White women, in contrast, have made steady gains in education mobility across a variety of parental education levels.

Intergenerational Education Mobility Trends by Race and Gender in the United States

Principal Investigator(s):  Joseph Ferrare, University of Washington Bothell

Version:  V1

Name 	File Type 	Size 	Last Modified 
 InterGenEdMobility_AERA_Open.sps	text/x-spss-syntax	26.7 KB	08/27/2019 01:48:PM

Generate Citation for this File

To view the citation for the overall project, see <http://doi.org/10.3886/E111586V1>.

Intergenerational Education Mobility Trends by Race and Gender in the United States

Principal Investigator(s):  Joseph Ferrare, University of Washington Bothell

Version:  V1

Name 	File Type 	Size 	Last Modified 
 InterGenEdMobility_AERA_Open.sps	text/x-spss-syntax	26.7 KB	08/27/2019 01:48:PM

File Citation:

Ferrare, Joseph. Intergenerational Education Mobility Trends by Race and Gender in the United States: InterGenEdMobility_AERA_Open.sps. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2019-08-27. <https://doi.org/10.3886/E111586V1-13185>

Persistent URL: <http://doi.org/10.3886/E111586V1-13185>

To view the citation for the overall project, see <http://doi.org/10.3886/E111586V1>.

Related Publications

The following publications are supplemented by the data in this project.

- Bloom, Nicholas, Raffaella Sadun, and John Van Reenen. "Americans Do IT Better: US Multinationals and the Productivity Miracle." *American Economic Review* 102, no. 1 (February 2012): 167–201. <https://doi.org/10.1257/aer.102.1.167>.



American Educational Research Association Data Repository

Depositing Data with openICPSR to fulfill the American Educational Research Association's Archiving and Replication Requirements

Sharing data and procedures at the publication stage enhances the value of publication for authors and users. In keeping with the American Educational Research Association's proposed data sharing and archiving policies, *AERA Open* requires authors to deposit data and other study materials with the openICPSR repository. This allows other researchers to replicate analyses in a published article without having to request any additional information from the author, while at the same time enhances the credit afforded to authors for data and related materials they produced as part of a research work.

The openICPSR repository automatically generates a citation when the data are "published" in openICPSR. Authors are required to include this citation pointing to the data in the reference section of the final version of the article sent to the *AERA Open* editors. At the same time, all publications in any AERA (and, we hope other) journals using or referring to these data will be required to include a citation to them in their own bibliographies. This enables authors to be credited, through citations, for the data that they create.

Deposits should include all data, annotated program code, command files, and documentation that are needed to replicate the findings from an AERA Open article. For guidelines on options for different kinds of data, see Archiving of Article-related Data. For general instructions for depositing data, see [Instructions for Depositing Data](#). For a discussion of best practice in preparing data for sharing, please refer to ICPSR's [Guide to Social Science Data Preparation and Archiving, 5th Edition](#).

[Start Your Deposit](#)

ICPSR's Role

ICPSR does not approve or alter datasets deposited in openICPSR in any way. Data are distributed in the same condition and format submitted by the depositor. Data approved for public use are distributed at no cost to users. ICPSR may charge a fee for the costs of managing data that have been restricted to protect confidential information about research subjects.

Confidentiality and Copyright

Investigators submitting data to ICPSR are entirely responsible for ensuring the confidentiality of research subjects. The presence of identifiable or sensitive information must be indicated during the submission process. ICPSR does not review submissions for disclosure risk. Investigators are also required to affirm that they have the right to publish this material. ICPSR requires a license for distribution of data, but copyright remains with the author.

About the AERA

The American Educational Research Association (AERA), a national research society, strives to advance knowledge about education, to encourage scholarly inquiry related to education, and to promote the use of research to improve education and serve the public good. AERA publishes seven highly respected, peer-reviewed journals that feature the field's leading research.

Report for AEA Data and Code Repository

Show 10 ▾ entries Print CSVSearch:

Study ID	Project Name	Status	Views	Downloads	Updated Date	Depositor Email
100236	Test	PUBLISHED	3	0	2019-06-13 13:58:48.262	harshau@umich.edu
100826	Technical Testing 2	PUBLISHED	3	0	2019-07-23 14:32:22.879	harshau@umich.edu
107872	Testing	PUBLISHED	1	0	2019-04-19 08:45:21.466	harshau@umich.edu
107968	TestModule	PUBLISHED	2	0	2019-07-23 14:30:34.034	bingshe@umich.edu
107969	TestModule	PUBLISHED	2	0	2019-06-11 12:58:25.169	bingshe@umich.edu
107972	Testaea	PUBLISHED	0	0	2019-03-18 11:02:45.557	bingshe@umich.edu
107987	AEA Test project #1	Deposit In Progress	0	0	2019-03-20 13:53:15.781	kilsang@umich.edu
107989	AEA project test #3 V10	PUBLISHED	12	0	2019-03-19 16:30:30.939	kilsang@umich.edu
107993	title1	PUBLISHED	2	0	2019-08-20 11:57:11.601	rarav@umich.edu
107995	title3	PUBLISHED	0	0	2019-03-14 09:30:00.173	rarav@umich.edu

Showing 1 to 10 of 2,782 entries

Previous

1

2

3

4

5

...

279

Next

Customization

Scope of Project

JEL Classification: ?

D24 Production; Cost; Capital; Capital, Total Factor, and Multifactor Productivity; Capacity

E23 Macroeconomics: Production

F23 Multinational Firms; International Business

M10 Business Administration: General

M16 International Business Administration

O30 Innovation; Research and Development; Technological Change; Intellectual Property Rights: General

Migration



AEA Data Editor

@AeaData

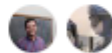
Follow



We will migrate the entire back archive of more than 3,000 data and code supplements (!) into the same infrastructure. Authors wishing to enhance the visibility of their previous data and code archives should contact me.

1:07 PM - 16 Jul 2019

1 Retweet 2 Likes



1



2

The Push to Replace Journal Supplements with Repositories

Broken links, clunky formats, and outdated platforms have both authors and publishers turning to alternative solutions.

Aug 19, 2019
DIANA KWON



Earlier this summer, [Vaughn Cooper](#), an evolutionary biologist at the University of Pittsburgh, was busy promoting a new secondary school curriculum for teaching evolution to scientists and educators. He and his colleagues had published the program in [Evolution: Education and Outreach](#) in April, and they were eager to spread the word before the start of the upcoming school year. So when Cooper received an email from a colleague who couldn't access his manuscript's supplementary files because of broken hyperlinks, he was frustrated by the news.

ABOVE: © ISTOCK.COM,
PIXELWELL

The supplementary documents contained important information, such as the experimental protocols for students that his team had tested. This was not the first time that he'd come across issues with these types of files. "I've had multiple instances from multiple publishers where the supplementary material has gone missing," he says, adding that this has occurred with both his papers and others'.

Cooper went to [Twitter](#) to vent his frustration. In response, other scientists noted that they, too, had experienced [similar problems](#). "I am afraid this is not uncommon," tweeted [Peter Murray-Rust](#), a chemist at the University of Cambridge. "Many (not all) journals generally regard supplementary data as a pain in the neck."

In addition to broken links, scientists point to other problems plaguing these files—such as their increasing length and the inaccessibility of the formats they are published in. As a result of these issues, both academics and publishers are increasingly turning to independent, online repositories as one potential solution.

“I’ve had multiple instances from multiple publishers where the supplementary material has gone missing,” he says, adding that this has occurred with both his papers and others’....

“I am afraid this is not uncommon,” tweeted [Peter Murray-Rust](#), a chemist at the University of Cambridge. “Many (not all) journals generally regard supplementary data as a pain in the neck.”

...“This data [within supplementary materials] is often deposited as a token of depositing data, rather than people actually wanting to reuse it.”

Correspondence | Published: 07 December 2005

Supplementary data need to be kept in public repositories

Carlos Santos, Judith Blake & David J. States

Nature **438**, 738 (2005) | [Download Citation](#) ↓

Sir

The reality of the genomics age is that there are many very large data sets that are most usefully saved and manipulated in electronic form. Many journals add online 'supplementary material' to articles as a service to authors wishing to publish volumes of such data that cannot be accommodated within the body of an article.

Supplementary-material collections maintained by publishers serve as archival repositories directly connected with the peer-reviewed scientific literature, often competing with or substituting for the deposition of data in public repositories.

Americans Do IT Better: US Multinationals and the Productivity Miracle

Nicholas Bloom

Raffaella Sadun

John Van Reenen

AMERICAN ECONOMIC REVIEW
VOL. 102, NO. 1, FEBRUARY 2012
(pp. 167-201)

[Download Full Text PDF](#)

Article Information

Abstract

US productivity growth accelerated after 1995 (unlike Europe's), particularly in sectors that intensively use information technologies (IT). Using two new micro panel datasets we show that US multinationals operating in Europe also experienced a "productivity miracle." US multinationals obtained higher productivity from IT than



Additional Materials

[Data Set \(1.08 MB\)](#)

[Online Appendix \(302.40 KB\)](#)

Replication data for: Americans Do IT Better: US Multinationals and the Productivity Miracle

Principal Investigator(s):  Nicholas Bloom; Raffaella Sadun; John Van Reenen

Version:  V1




AMERICAN
ECONOMIC
ASSOCIATION

Name 	File Type 	Size 	Last Modified 
 ADIB_Posted_Dec2011			08/23/2019 09:07:AM
 LICENSE.txt	text/plain	14.6 KB	08/23/2019 05:07:AM

Project Citation:

Persistent URL: <http://doi.org/>

Project Description

Summary:  US productivity growth accelerated after 1995 (unlike Europe's), particularly in sectors that intensively use information technologies (IT). Using two new micro panel datasets we show that US multinationals operating in Europe also experienced a "productivity miracle." US multinationals obtained higher productivity from IT than non-US multinationals, particularly in the same sectors responsible for the US productivity acceleration. Furthermore, establishments taken over by US multinationals (but not by non-US multinationals) increased the productivity of their IT. Combining pan-European firm-level IT data with our management practices survey, we find that the US IT related productivity advantage is primarily due to its tougher "people management" practices. (JEL D24, E23, F23, M10, M16, O30)

Scope of Project


JEL Classification:

D24 Production; Cost; Capital; Capital, Total Factor, and Multifactor Productivity; Capacity
E23 Macroeconomics: Production
F23 Multinational Firms; International Business
M10 Business Administration: General
M16 International Business Administration
O30 Innovation; Research and Development; Technological Change; Intellectual Property Rights: General

Related Publications

The following publications are supplemented by the data in this project.

- Bloom, Nicholas, Raffaella Sadun, and John Van Reenen. "Americans Do IT Better: US Multinationals and the Productivity Miracle." *American Economic Review* 102, no. 1 (February 2012): 167–201. <https://doi.org/10.1257/aer.102.1.167>.

 DOWNLOAD THIS PROJECT

Usage Metrics

Overall Project Metrics

6

Views

0

Downloads

1

Publications

[Download Detailed Metrics](#)

Published Versions

[V1 \(2019-08-23\)](#)

Export Metadata

[Dublin Core](#)

[DDI 2.5](#)

[DDI 3.1](#)

Report a Problem

Found a serious problem with the data, such as disclosure risk or copyrighted content? Let us know.

New requests

Update workflow for repository administrator to review and approve before release.



Thank you!

lyle@umich.edu

Applying for Access

- Users must apply for access
 - Describe proposed research and why restricted data is required
- All users must sign a Restricted Data Use Agreement
 - Agree to data security requirements and procedures
 - Principal Investigator generally must hold terminal degree
 - Students can be on agreements
 - Some projects provide access to researchers, policymakers, journalists without a terminal degree
 - Be affiliated with legitimate research institution
 - Receive IRB approval or exemption before gaining access to most data



AEA Data Editor

@AeaData

Follow



There is a new repository infrastructure, hosted at [@ICPSR](#), called the "AEA Data and Code Repository."

openicpsr.org/openicpsr/aea Data (where allowed) and code (always) will be uploaded to the repository and shared with the Data Editor prior to publication.

1:07 PM - 16 Jul 2019

3 Retweets 4 Likes



1



3



4



Depositing Data in the AEA Data and Code Repository

The *American Economic Association Journals* require authors to deposit data and materials with a community-recognized or general repositories. The *AEA Data and Code Repository at ICPSR* serves that purpose. Please see the AEA's [Data and Code Availability Policy](#) and data citation guidance at the [Sample References](#) page for more details. **Authors are required to include a citation pointing to the deposit in the reference section of the final version of the article sent to the AEA.** The *openICPSR* repository automatically generates a citation when the data are "published."

Deposits should include all data, annotated program code, command files, and documentation that is needed to replicate the findings from the authors' submitted article.

- **Data** should be comprehensively documented (see ICPSR's [Guide to Social Science Data Preparation and Archiving, 5th Edition](#) for guidance). The **author** is responsible for removing identifying information from the data to protect [confidentiality](#). Neither the AEA nor ICPSR review submissions for disclosure risk.
- **Program** code and command files should be annotated to facilitate replication and ensure clear correspondence between code and figures, tables, and analyses in the published article.
- Authors retain ownership and copyright to the data and code. Authors are required to affirm that they have the right to publish and redistribute the material. However,
 - ICPSR requires a license for distribution of data.
 - An **open license** is required by the AEA, in order to allow others to re-use the data and code, in particular for replication. Authors can select from several license options, including CC-BY 4.0 for data and Modified BSD for software and code. If an author would like to use multiple licenses or create a customized license, she should select the "Other" license option and upload a LICENSE file alongside the data and documentation.

By depositing in the AEA Data and Code Repository, the depositors allow the AEA staff to add keywords and other metadata which are important for proper indexing in linking. Any other changes are subject to the license chosen for the materials.

[View more extensive \(unofficial\) guidance](#) [↗](#).

Start Your Deposit

Filters

Scope

ARCHIVE=openICPSR

Subject Terms

[PSID](#) 14
[J](#) 13
[Illinois](#) 12
[polling](#)
[Simon Poll](#) 12
[taxes](#) 12

[view all](#)

Geography

[United States](#) 148
[USA](#) 12
[Greece](#) 10
[Illinois](#) 10
[California](#) 7

[view all](#)

Access Level

[Public Use](#) 809
[Restricted Use](#) 42

Search Results

851 results.

[GO](#) [VIEW ALL](#)

[search tips](#)

Per page: 25

Sort by: Released/Updated

1 2 3 4 5

Study Title/Investigator

Released/Updated

1. [Did Military Service During World War I Affect the Economic Status of American Veterans?](#)
Tan, Hui Ren
[more info](#)
2. [Intergenerational Education Mobility Trends by Race and Gender in the United States](#)
Ferrare, Joseph
[more info](#)
3. [Revealed preference analysis with normal goods: Application to cost of living indices](#)
Cherchye, Laurens; Demuynck, Thomas; De Rock, Bram; Surana, Khushboo
[more info](#)
4. [Understanding gender gap further](#)
Stanfors, Maria; Burnette, Joyce
[more info](#)
5. [Google Smartphone Questionnaires - Culture](#)
Simmonds, David
[more info](#)
6. [SD3 Croatian Validation](#)
Jakšić, Krešimir; Penezić, Zvezdan
[more info](#)
7. [The Language-Content-Technology Intersection: Mrs. B Teaches Biology to English Learners](#)
Meskill, Carla; Nilsen, Jennifer; Oliveira, Alan
[more info](#)
8. [Pathways to Inequality: Between-District Segregation and Racial Disparities in School District Expenditures](#)
Sosina, Victoria; Weathers, Ericka
[more info](#)
9. [Inquiry-Based Instruction in Science and Mathematics Middle School Classrooms: Examining Its Association with Students' Attitudes by Gender and Race/Ethnicity](#)
Riegler-Crumb, Catherine
[more info](#)