

# How Does Data Access Shape Science? Evidence from the Impact of U.S. Census's Research Data Centers on Economics Research

Matteo Tranchero  
University of California, Berkeley

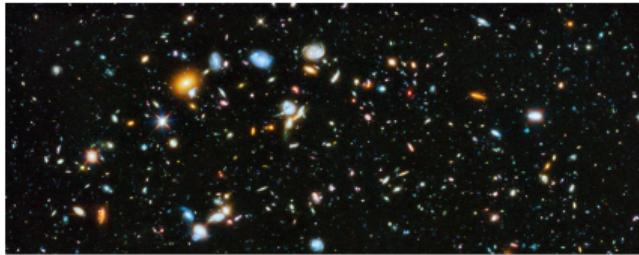
(with Abhishek Nagaraj)

DRUID

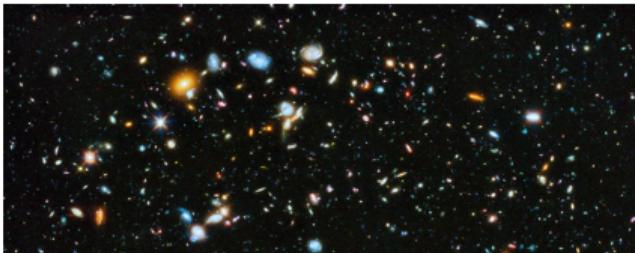
12 June 2023

Supported by the Alfred P. Sloan Foundation

# The James Webb Telescope and Data Deluge



# The James Webb Telescope and Data Deluge



**The New York Times**

## Get ready for an outpouring of Webb telescope science.

The pictures and other data released Tuesday were selected from a list of possible sources by a small team of imaging experts and public outreach specialists for their ability to show off the Webb telescope's range and power, and to knock our socks off.

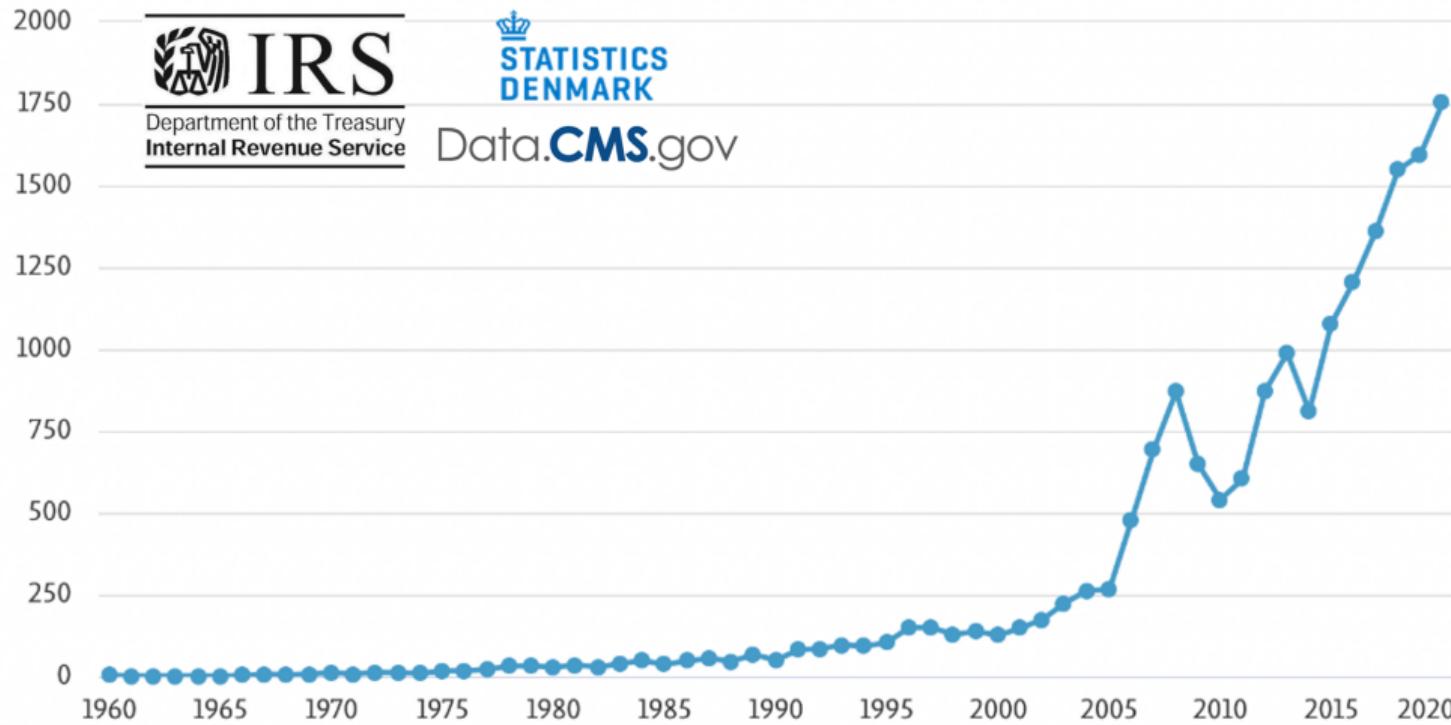
But other major scientific research is already underway, and in the months ahead, the results of a group of programs that NASA calls [Early Release Science](#) will be pouring into science journals and other repositories as fast as the astronomers can write them. Some of the results will be available on Thursday, including the data gathered while the telescope was being set up in space and images of distant galaxies that are even deeper than the one President Biden showed yesterday.

# Administrative Data and Social Science Research



Data.**CMS**.gov

# Administrative Data and Economics Research



# Balancing Privacy Costs vs Scientific Benefits



- ▶ Findings based on administrative data can help further science
- ▶ Better science can, in principle, shape better economic policy
- ▶ These benefits need to be traded off against privacy concerns of making admin data more available
- ▶ Yet most current discussions have no way to account for these potential benefits vis-á-vis costs

## Research Question

- ▶ How does access to large-scale administrative data affect the quality and direction of research?

## Research Question

- ▶ How does access to large-scale administrative data affect the quality and direction of research?
- ▶ Are there spillover effects on non-adopters of administrative data, and if so, why?
- ▶ Do administrative data lead to more policy-relevant empirical research?

# Federal Statistical Research Data Center (FSRDC) Network



- ▶ Network of secure facilities providing access to confidential administrative data

# Federal Statistical Research Data Center (FSRDC) Network



- ▶ Network of secure facilities providing access to confidential administrative data
- ▶ Gradual expansion across the country since 1994 to over 30 centers

# Federal Statistical Research Data Center (FSRDC) Network



- ▶ Network of secure facilities providing access to confidential administrative data
- ▶ Gradual expansion across the country since 1994 to over 30 centers
- ▶ Combined with bibliographic data can be used to estimate the impact of access to data

## Preview: Four Key Findings



### **Increase Diffusion**

Use of Census data  
more than doubles  
→ also +53%  
cites to past  
Census papers!

## Preview: Four Key Findings



### Increase Diffusion

Use of Census data more than doubles  
→ also +53% cites to past Census papers!

### Positive Effects

Empiricists produce about 24% more publications/year in top journals

# Preview: Four Key Findings



## Increase Diffusion

Use of Census data more than doubles  
→ also +53% cites to past Census papers!

## Positive Effects

Empiricists produce about 24% more publications/year in top journals

## Spillovers

Spillovers on non-users via exploration of new topics and use of better research designs

# Preview: Four Key Findings



## Increase Diffusion

Use of Census data more than doubles  
→ also +53% cites to past Census papers!

## Positive Effects

Empiricists produce about 24% more publications/year in top journals

## Spillovers

Spillovers on non-users via exploration of new topics and use of better research designs

## Policy Impact

Empirical research becomes more likely to be cited by policy documents

# Contributions



## **Admin. Data**

First systematic investigation of the effects of administrative data on academic research (Atrostic, 2007; Card et al., 2010; CES, 2017; Cole et al., 2020)

# Contributions



## **Admin. Data**

First systematic investigation of the effects of administrative data on academic research (Atrostic, 2007; Card et al., 2010; CES, 2017; Cole et al., 2020)

## **Inequality in Science**

Data access as a driver of differences in productivity (Feenberg et al., 2017; Sarsons et al., 2021; Heckman and Moktan, 2020; Card et al., 2020, 2022)

# Contributions



## Admin. Data

First systematic investigation of the effects of administrative data on academic research (Atrostic, 2007; Card et al., 2010; CES, 2017; Cole et al., 2020)



## Inequality in Science

Data access as a driver of differences in productivity (Feenberg et al., 2017; Sarsons et al., 2021; Heckman and Moktan, 2020; Card et al., 2020, 2022)



## Research Inputs

Role of data in shaping quality and policy relevance of science (Furman and Stern, 2011; Murray et al., 2016; Teodoridis, 2017; Williams, 2013; Nagaraj et al., 2020)

# Overview

Introduction

Setting

Data

Research Design

Results

Discussion

## U.S. Government Administrative Data

- ▶ Administrative data: any information not originally collected for research purposes

# U.S. Government Administrative Data

- ▶ Administrative data: any information not originally collected for research purposes
- ▶ Center for Economic Studies at the U.S. Census Bureau
- ▶ Established in 1982 to provide access to manufacturing data
- ▶ Enabled examination of firm-level heterogeneity beyond representative firm models

Gross job creation, gross job destruction, and employment reallocation

[SJ Davis, J Haltiwanger](#) - The Quarterly Journal of Economics, 1992 - academic.oup.com

This study measures the heterogeneity of establishment-level employment changes in the US manufacturing sector over the 1972 to 1986 period. We measure this heterogeneity in ...

[☆ Save](#) [99 Cite](#) [Cited by 2362](#) [Related articles](#) [All 18 versions](#) [»»](#)

The dynamics of productivity in the telecommunications equipment industry

[S Olley, A Pakes](#) - 1992 - nber.org

Technological change and deregulation have caused a major restructuring of the telecommunications equipment industry over the last two decades. We estimate the

[☆ Save](#) [99 Cite](#) [Cited by 8050](#) [Related articles](#) [All 26 versions](#) [»»](#)

The empirical analysis is based on an extremely rich plant-level panel constructed from data collected by the U.S. Bureau of the Census. It is clear from the data that during

## Broaden Access: Research Data Center (FSRDC) Network



Chicago RDC



FEDERAL RESERVE BANK *of* KANSAS CITY



**TRDC**  
Triangle Research Data Center  
*a multi-institutional consortium*

**FEDERAL RESERVE BANK**  
PHILADELPHIA



## Access Restrictions

1. Projects are reviewed in advance

## Access Restrictions

1. Projects are reviewed in advance
2. Background checks + Special Sworn Status

## Access Restrictions

1. Projects are reviewed in advance
2. Background checks + Special Sworn Status
3. Analyses can only be done in the FSRDC

## Access Restrictions

1. Projects are reviewed in advance
2. Background checks + Special Sworn Status
3. Analyses can only be done in the FSRDC
4. All results undergo disclosure review

## Access Restrictions

1. Projects are reviewed in advance
2. Background checks + Special Sworn Status
3. Analyses can only be done in the FSRDC
4. All results undergo disclosure review



*NBER Cambridge FSRDC  
(courtesy of Jeff Furman)*

## Variation in Data Access

Census and the NSF (which funded FSRDCs) was explicitly trying to balance researchers' demand with equitable geographical coverage across the United States.

## Variation in Data Access

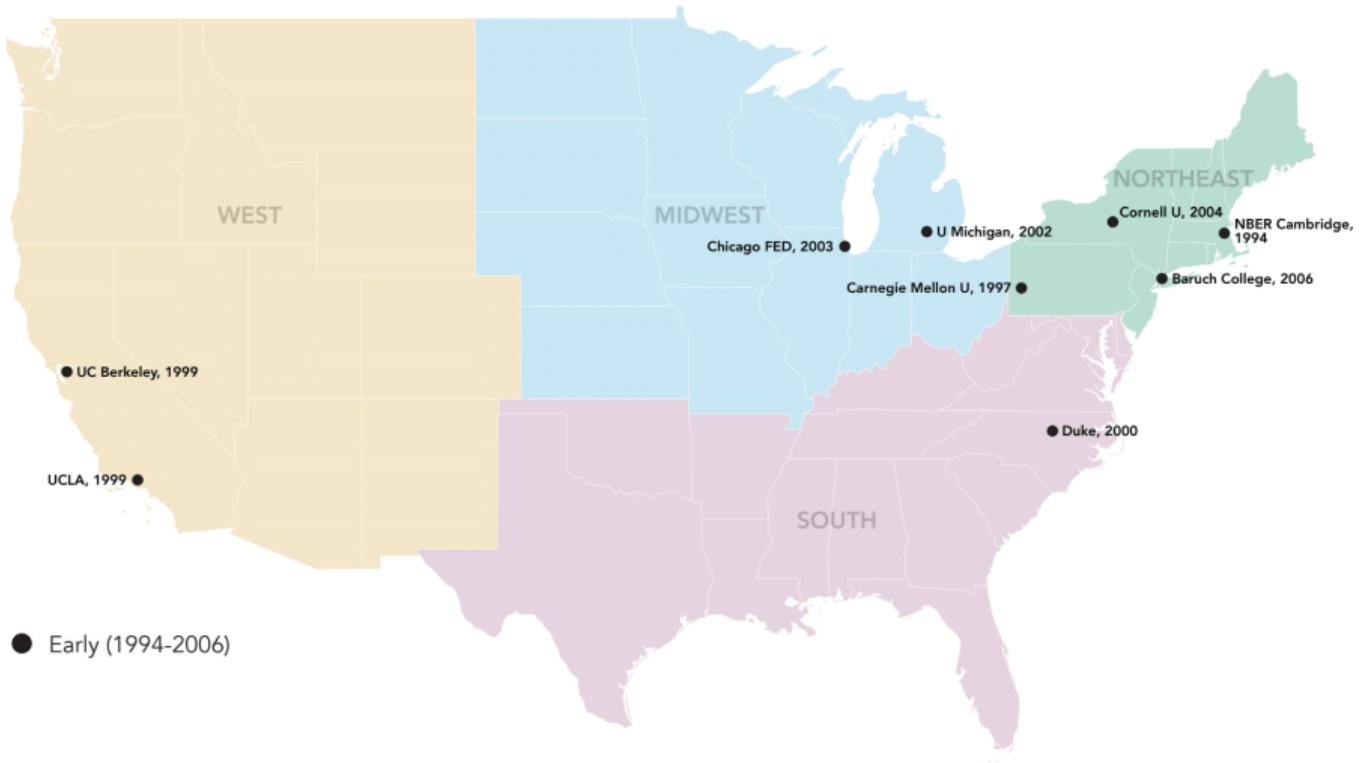
Census and the NSF (which funded FSRDCs) was explicitly trying to balance researchers' demand with equitable geographical coverage across the United States.



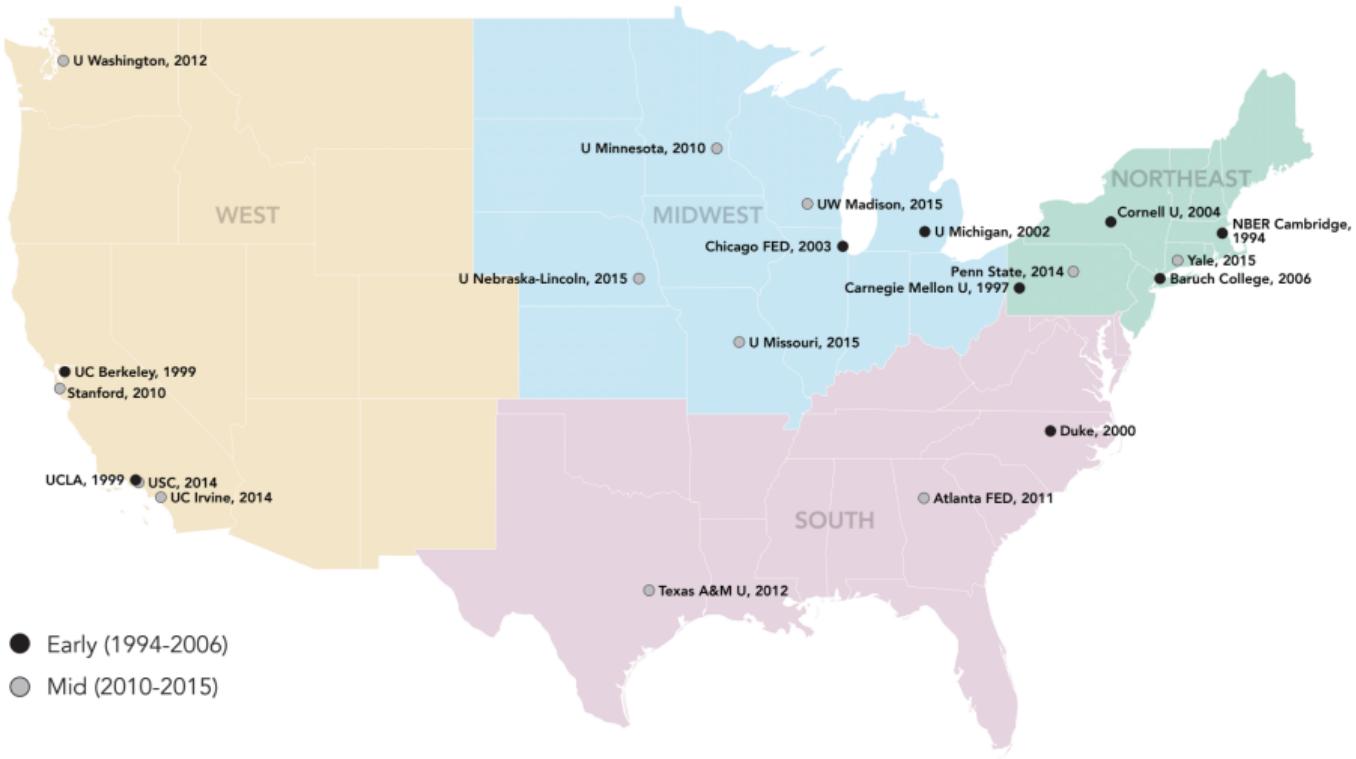
“Many institutions were interested in opening an FSRDC, but the NSF was interested in kind of parity across the US so that researchers in one part of the country had the same access as researchers in another part of the country did.”

— Former FSRDC administrator (25<sup>th</sup> November 2019)

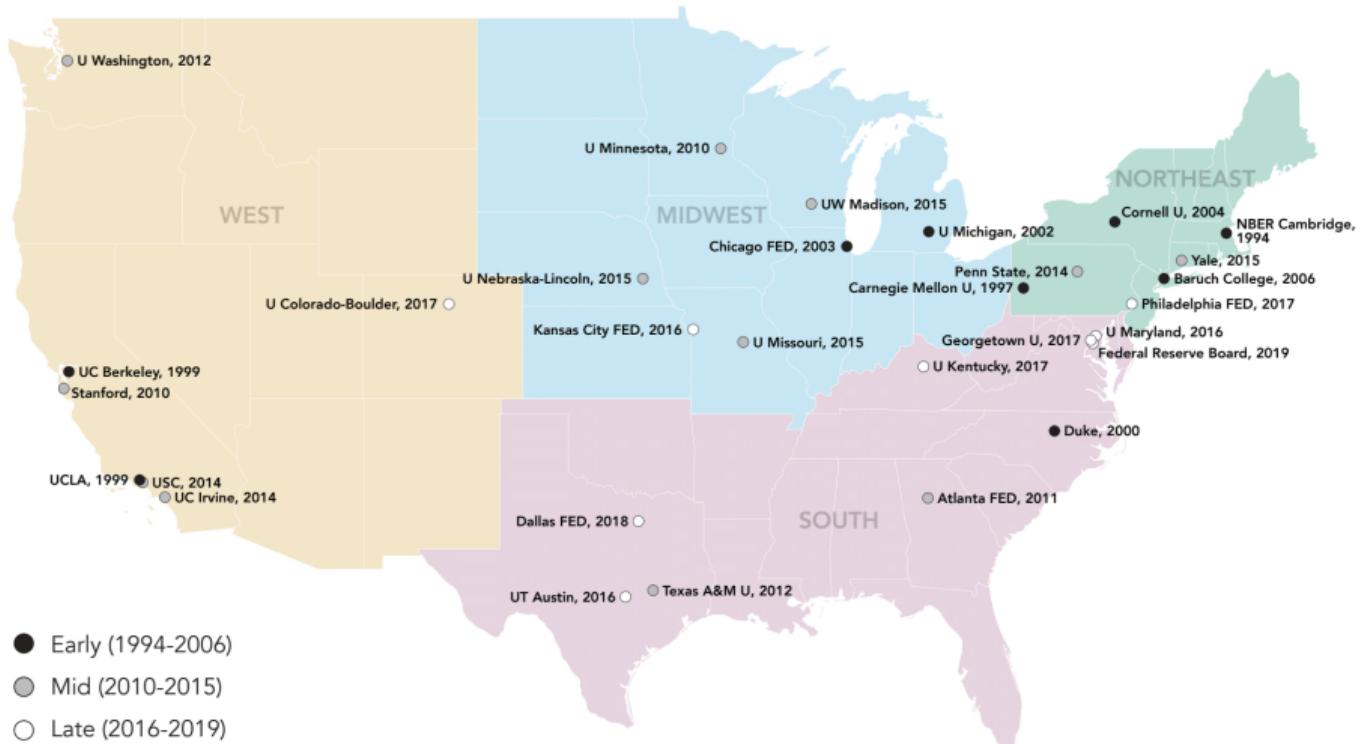
# Expansion of FSRDCs: 29 Centers in 25 Years



# Expansion of FSRDCs: 29 Centers in 25 Years



# Expansion of FSRDCs: 29 Centers in 25 Years



# Overview

Introduction

Setting

Data

Research Design

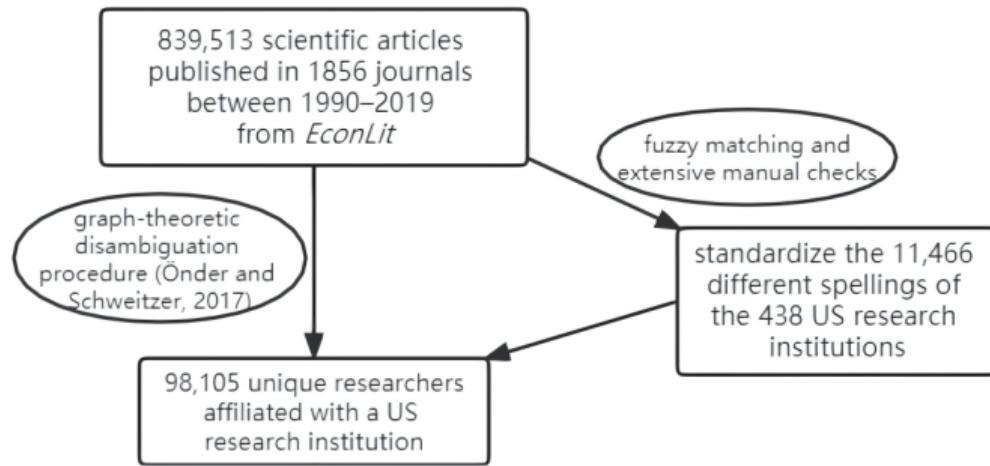
Results

Discussion

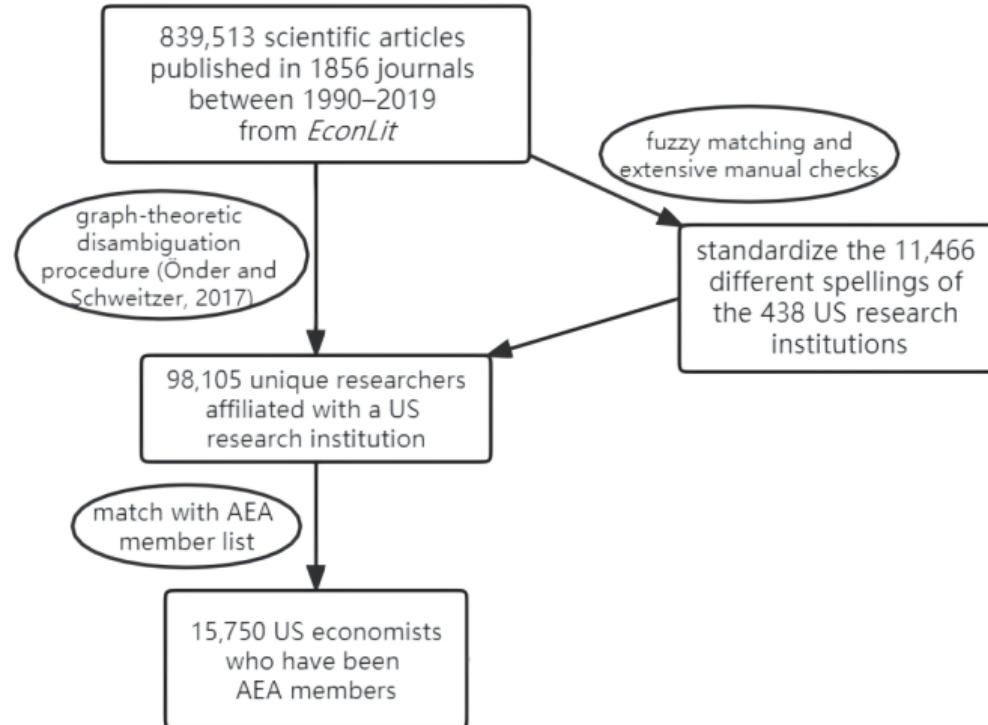
# Universe of Publishing Economists

839,513 scientific articles  
published in 1856 journals  
between 1990–2019  
from *EconLit*

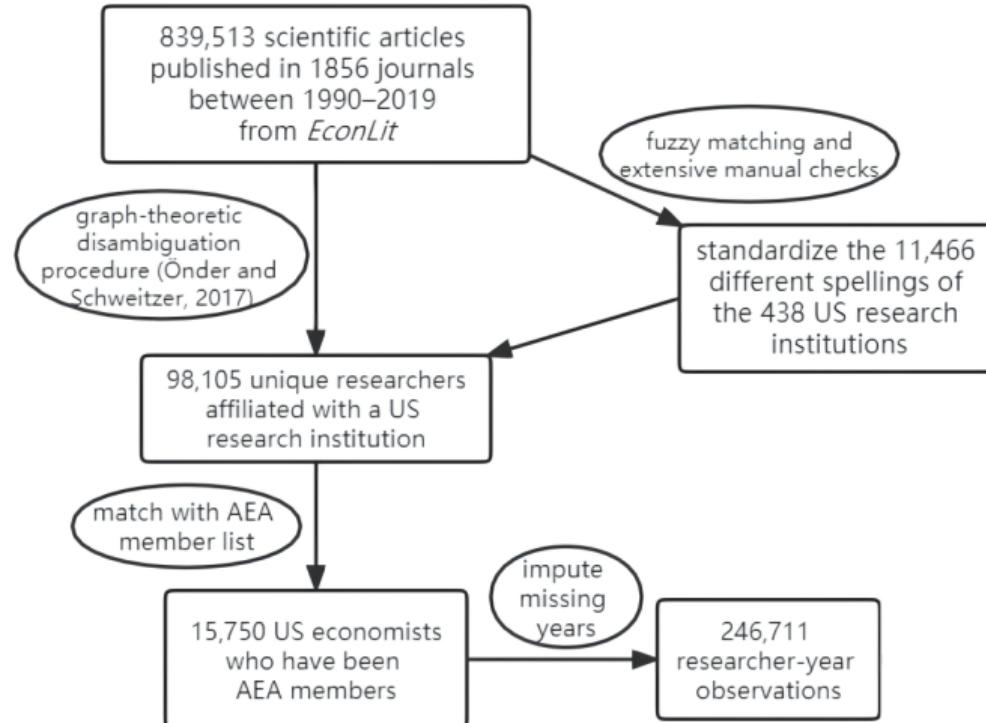
# Universe of Publishing Economists



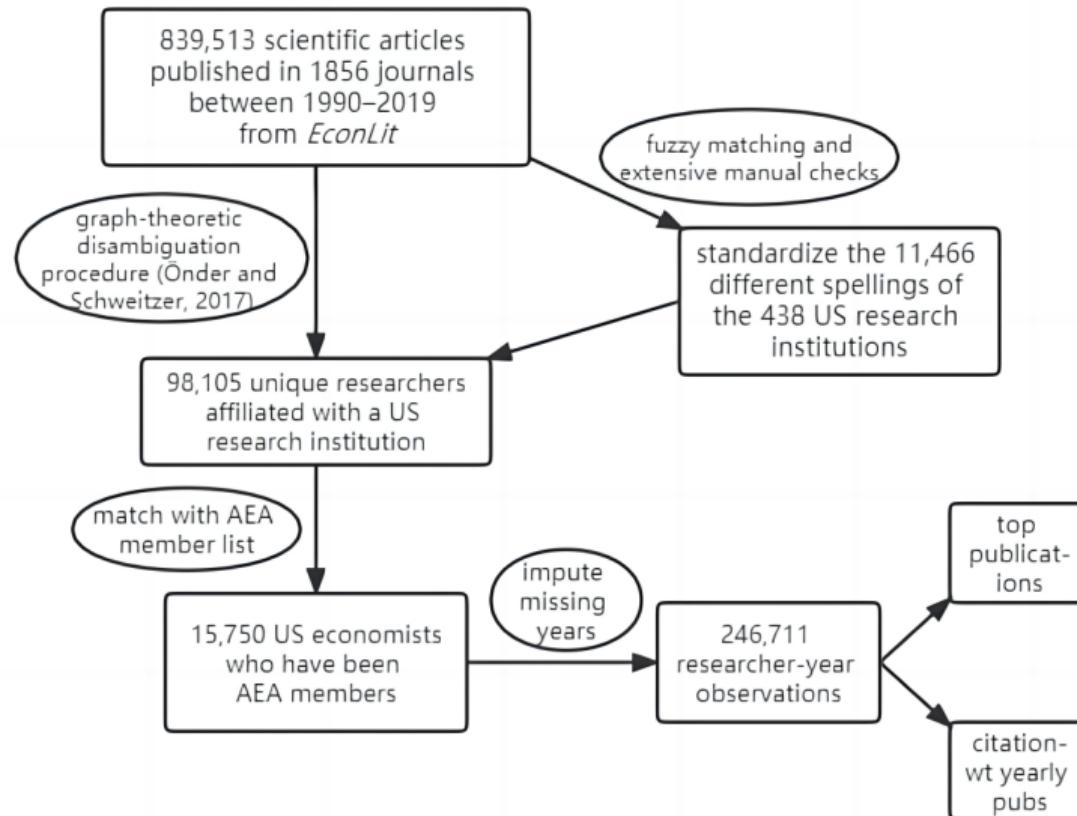
# Universe of Publishing Economists



# Universe of Publishing Economists



# Universe of Publishing Economists



## Adoption of U.S. Census Microdata

- ▶ We manually assembled list of all articles that directly employ restricted-access microdata accessible only in an FSRDC
- ▶ In the absence of an official bibliographic record, we used a few strategies:
  - ▶ Keyword searches among papers' acknowledgement
  - ▶ Papers submitted to CES working papers series
  - ▶ CES yearly reports
  - ▶ FOIA requests of all approved FSRDC projects
- ▶ We also record all articles *citing* the papers written using FSRDC data

► Descriptive Statistics

► FSRDC Papers

► Dataset-level Evidence

► Dataset Lifecycle

# Measuring Orientation: Empirics vs Theory

Classify a paper as either empirical or theoretical based on Angrist et al. (2020) (those with greater than 50% empirical are empiricists)

*Econometric Theory*, 13, 1997, 467–505. Printed in the United States of America.

## GAUSSIAN ESTIMATION OF MIXED-ORDER CONTINUOUS-TIME DYNAMIC MODELS WITH UNOBSERVABLE STOCHASTIC TRENDS FROM MIXED STOCK AND FLOW DATA

A.R. BERGSTROM  
*University of Essex*

This paper develops an algorithm for the exact Gaussian estimation of a mixed-order continuous-time dynamic model, with unobservable stochastic trends, from a sample of mixed stock and flow data. Its application yields exact maximum likelihood estimates when the innovations are Brownian motion and either the model is closed or the exogenous variables are polynomials in time of degree not exceeding two, and it can be expected to yield very good estimates under much more general circumstances. The paper includes detailed formulae for the implementation of the algorithm, when the model comprises a mixture of first- and second-order differential equations and both the endogenous and exogenous variables are a mixture of stocks and flows.

$$P(\text{Theory}) = 83.6\%$$

► Measuring Validations

## DO ENERGY EFFICIENCY INVESTMENTS DELIVER? EVIDENCE FROM THE WEATHERIZATION ASSISTANCE PROGRAM\*

MEREDITH FOWLIE  
MICHAEL GREENSTONE  
CATHERINE WOLFRAM

A growing number of policies and programs aim to increase investment in energy efficiency, because conventional wisdom suggests that people fail to take up these investments even though they have positive private returns and generate environmental benefits. Many explanations for this energy efficiency gap have been put forward, but there has been surprisingly little field testing of whether the conventional wisdom is correct. This article reports on the results of an experimental evaluation of the nation's largest residential energy efficiency program—the Weatherization Assistance Program—conducted on a sample of approximately 30,000 households in Michigan. The findings suggest that the upfront investment costs are about twice the actual energy savings. Furthermore, the model-projected savings are more than three times the actual savings. Although this might be attributed to the "rebound" effect—when demand for energy end uses increases as a result of greater efficiency—the article fails to find evidence of significantly higher indoor temperatures at weatherized homes. Even when accounting for the broader societal benefits derived from emissions reductions, the costs still substantially outweigh the benefits; the average rate of return is approximately –7.8% annually. *JEL Codes:* Q4, Q48, Q5.

$$P(\text{Empirical}) = 96.95\%$$

# Overview

Introduction

Setting

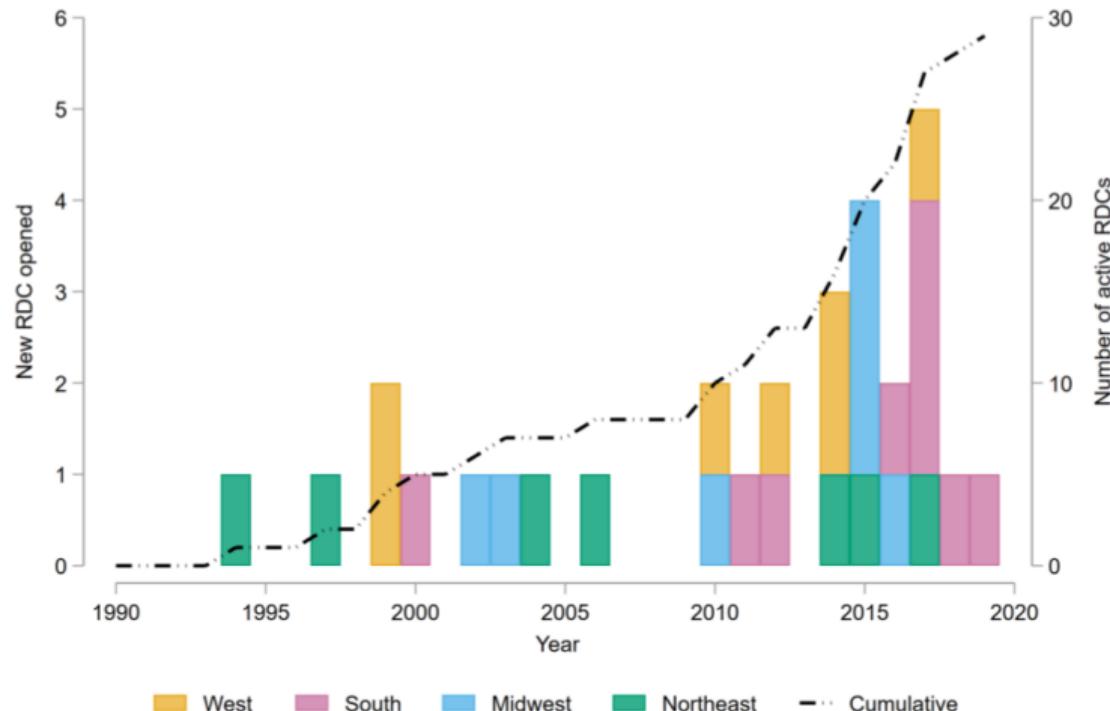
Data

Research Design

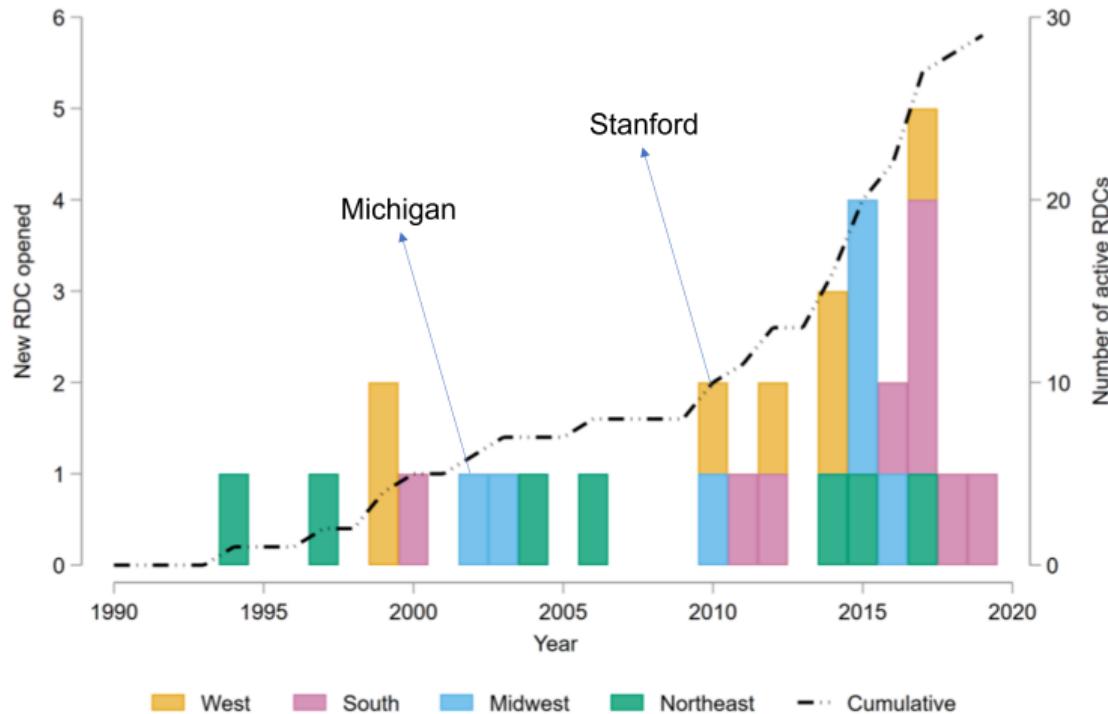
Results

Discussion

# Expansion of FSRDCs



# Expansion of FSRDCs

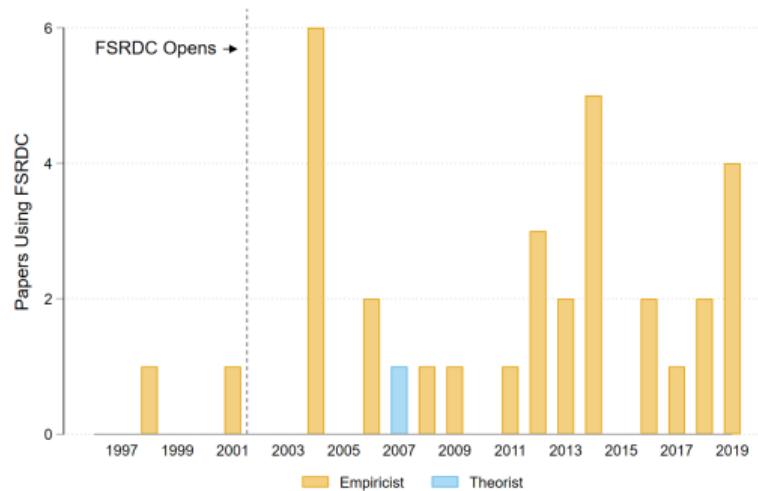


▶ Number of Researchers Gaining FSRDC Access over Time

▶ Rank of Treated Institutions

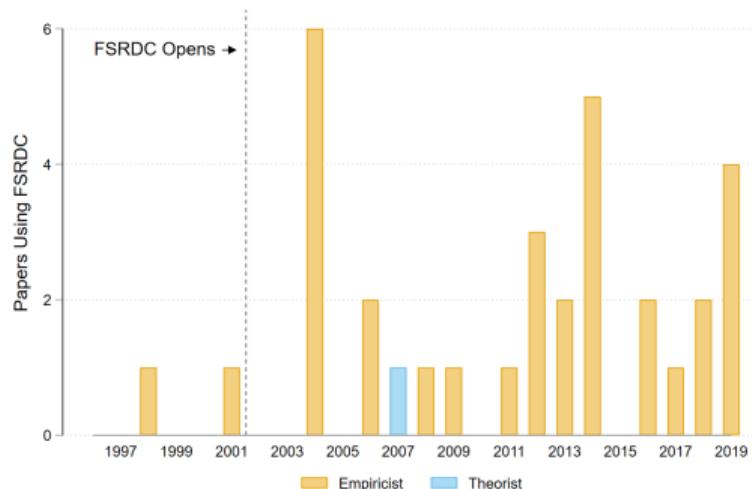
# Publications Based on FSRDC Data

*University of Michigan*

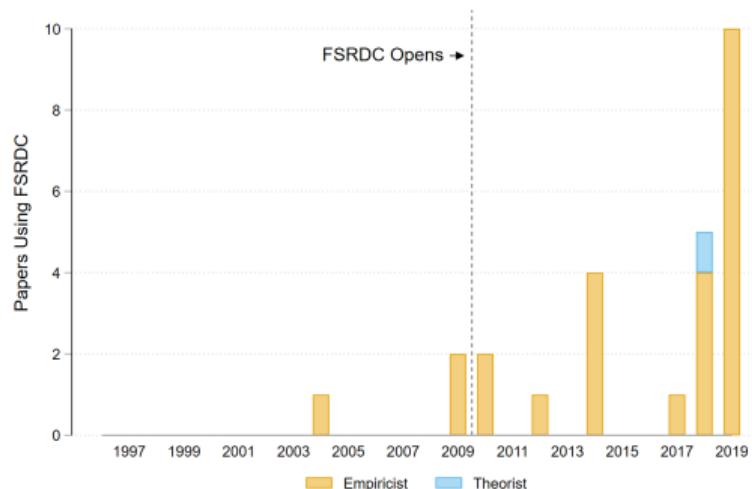


# Publications Based on FSRDC Data

*University of Michigan*



*Stanford University*



## Specification

We exploit variation in “exposure” to FSRDCs by methodological orientation, i.e., distinguishing between theoretical and applied economists within the same university

## Specification

We exploit variation in “exposure” to FSRDCs by methodological orientation, i.e., distinguishing between theoretical and applied economists within the same university

$$y_{i,j,t} = \alpha + \beta_1 PostFSRDC_{j,t} + \beta_2 PostFSRDC_{j,t} \times Empiricist_i + \mu_i + \delta_t \times w_j + \epsilon_{i,j,t} \quad (1)$$

- ▶  $y_{i,j,t}$ : various measures of scientific output

## Specification

We exploit variation in “exposure” to FSRDCs by methodological orientation, i.e., distinguishing between theoretical and applied economists within the same university

$$y_{i,j,t} = \alpha + \beta_1 PostFSRDC_{j,t} + \beta_2 PostFSRDC_{j,t} \times Empiricist_i + \mu_i + \delta_t \times w_j + \epsilon_{i,j,t} \quad (1)$$

- ▶  $y_{i,j,t}$ : various measures of scientific output
- ▶  $PostFSRDC_{j,t}$ : a dummy that equals one after an FSRDC facility is open in time  $t$  in the same city of her university  $j$

## Specification

We exploit variation in “exposure” to FSRDCs by methodological orientation, i.e., distinguishing between theoretical and applied economists within the same university

$$y_{i,j,t} = \alpha + \beta_1 PostFSRDC_{j,t} + \beta_2 PostFSRDC_{j,t} \times Empiricist_i + \mu_i + \delta_t \times w_j + \epsilon_{i,j,t} \quad (1)$$

- ▶  $y_{i,j,t}$ : various measures of scientific output
- ▶  $PostFSRDC_{j,t}$ : a dummy that equals one after an FSRDC facility is open in time  $t$  in the same city of her university  $j$
- ▶  $Empiricist_i$ : a dummy that equals one if the author is categorized as an empiricist

## Specification

We exploit variation in “exposure” to FSRDCs by methodological orientation, i.e., distinguishing between theoretical and applied economists within the same university

$$y_{i,j,t} = \alpha + \beta_1 PostFSRDC_{j,t} + \beta_2 PostFSRDC_{j,t} \times Empiricist_i + \mu_i + \delta_t \times w_j + \epsilon_{i,j,t} \quad (1)$$

- ▶  $y_{i,j,t}$ : various measures of scientific output
- ▶  $PostFSRDC_{j,t}$ : a dummy that equals one after an FSRDC facility is open in time  $t$  in the same city of her university  $j$
- ▶  $Empiricist_i$ : a dummy that equals one if the author is categorized as an empiricist
- ▶  $\mu_i$ : individual researcher fixed effects

## Specification

We exploit variation in “exposure” to FSRDCs by methodological orientation, i.e., distinguishing between theoretical and applied economists within the same university

$$y_{i,j,t} = \alpha + \beta_1 PostFSRDC_{j,t} + \beta_2 PostFSRDC_{j,t} \times Empiricist_i + \mu_i + \delta_t \times w_j + \epsilon_{i,j,t} \quad (1)$$

- ▶  $y_{i,j,t}$ : various measures of scientific output
- ▶  $PostFSRDC_{j,t}$ : a dummy that equals one after an FSRDC facility is open in time  $t$  in the same city of her university  $j$
- ▶  $Empiricist_i$ : a dummy that equals one if the author is categorized as an empiricist
- ▶  $\mu_i$ : individual researcher fixed effects
- ▶  $\delta_t \times w_j$ : Additional FE $s$  at the university-by-year level

# Overview

Introduction

Setting

Data

Research Design

Results

Discussion

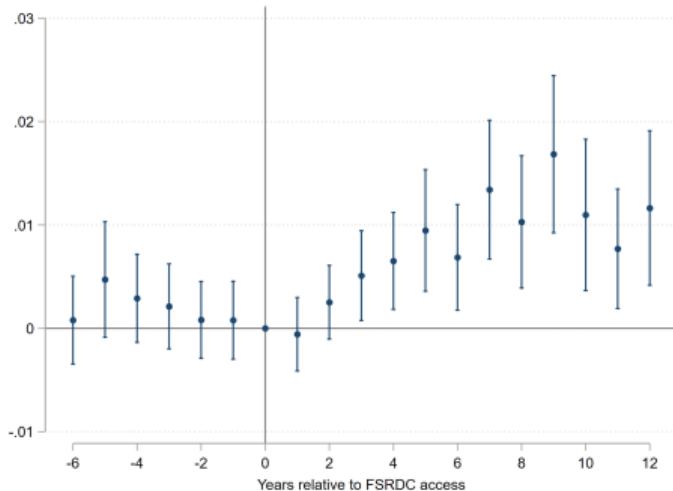
## I. Does Local Data Access Boost Use of Administrative Data?

# Event Study (First Stage)

► Table

► Other Measures of FSRDC Use

## *Papers Using FSRDC*



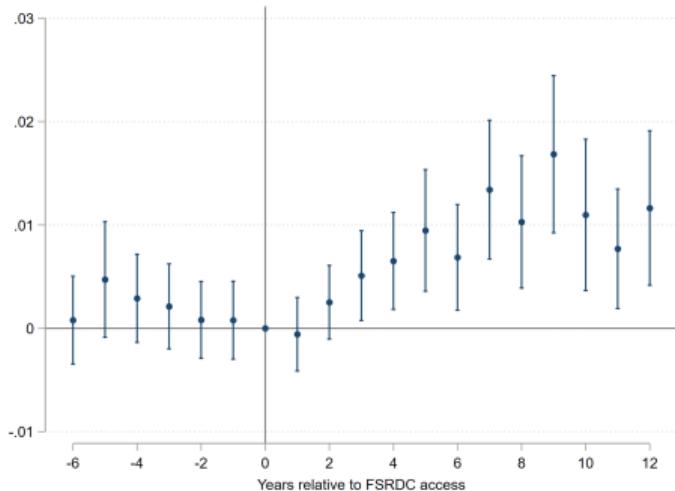
DID estimate: +131% over the mean

# Event Study (First Stage)

► Table

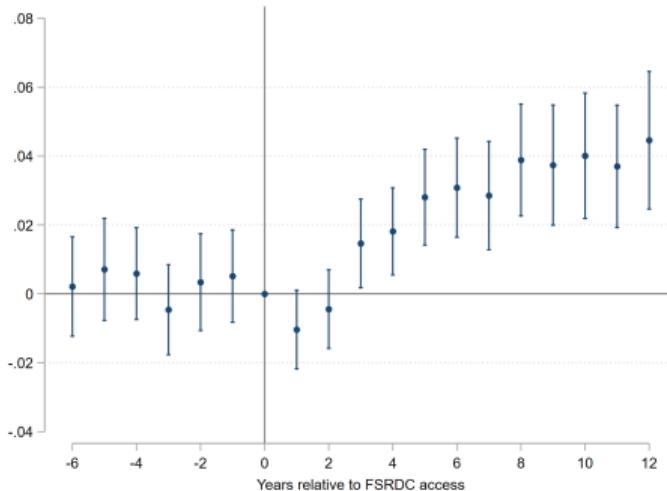
► Other Measures of FSRDC Use

*Papers Using FSRDC*



DID estimate: +131% over the mean

*Papers Citing FSRDC*

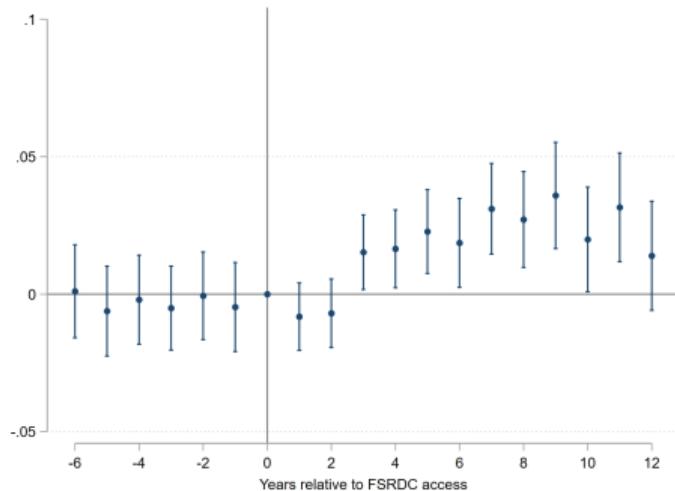


DID estimate: +53% over the mean

## II. How Does Local Data Access Affect Productivity?

# Event Study on Top Tail of Research Quality

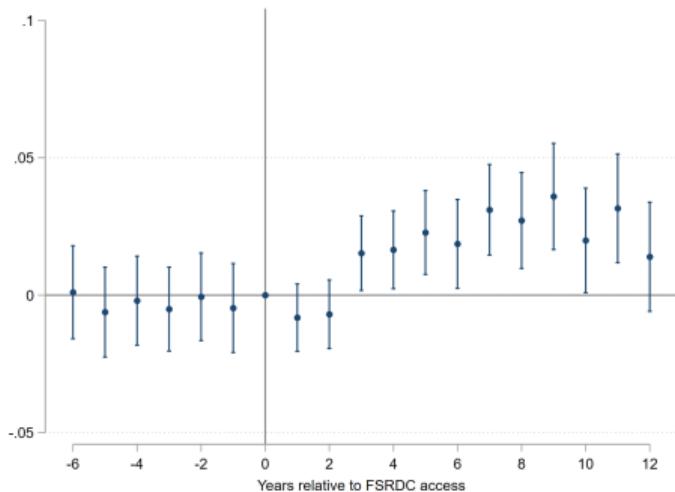
*Top Five Publications*



DID estimate: +49% over the mean

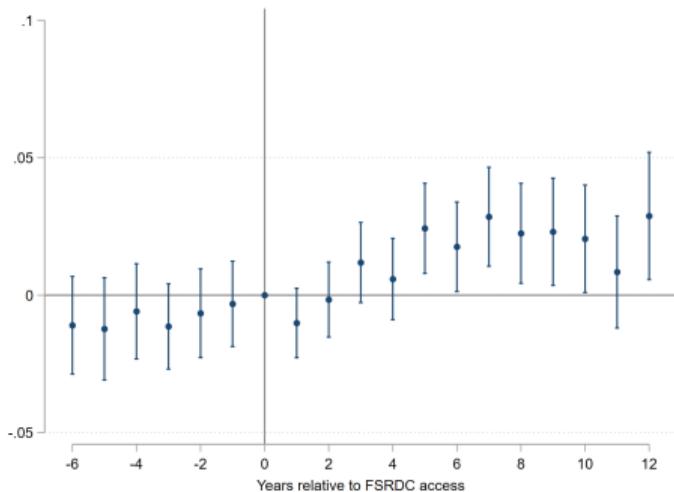
# Event Study on Top Tail of Research Quality

*Top Five Publications*



DID estimate: +49% over the mean

*Top 5% Cited Publications*



DID estimate: +65% over the mean

## Robustness Checks

- ▶ Timing of NSF Grants 
- ▶ Excluding NSF Grant Applicants 
- ▶ Excluding Authors in Institutions Hosting FSRDCs 
- ▶ Excluding Authors Treated After Mobility Events 
- ▶ Excluding Direct FSRDC Users 

### III. Who Benefits from Local Access to FSRDCs?

## Direct Data Use vs Spillover Effects

- ▶ Potential channels driving the productivity effect:

## Direct Data Use vs Spillover Effects

- ▶ Potential channels driving the productivity effect:
  - ▶ *directly* by using of confidential data
  - ▶ *indirectly* through spillovers to those not using the data

## Direct Data Use vs Spillover Effects

- ▶ Potential channels driving the productivity effect:
  - ▶ *directly* by using of confidential data
  - ▶ *indirectly* through spillovers to those not using the data
- ▶ Excluding direct users: coefficients 24–39% smaller but still large and significant

## Direct Data Use vs Spillover Effects

- ▶ Potential channels driving the productivity effect:
  - ▶ *directly* by using of confidential data
  - ▶ *indirectly* through spillovers to those not using the data
- ▶ Excluding direct users: coefficients 24–39% smaller but still large and significant
- ▶ Indirect channel: *exposure* to research based on administrative data

## Direct Data Use vs Spillover Effects

- ▶ Potential channels driving the productivity effect:
  - ▶ *directly* by using of confidential data
  - ▶ *indirectly* through spillovers to those not using the data
- ▶ Excluding direct users: coefficients 24–39% smaller but still large and significant
- ▶ Indirect channel: *exposure* to research based on administrative data
- ▶ Explains the large increase in citations to past FSRDC publications

## Direct Data Use vs Spillover Effects

- ▶ Potential channels driving the productivity effect:
  - ▶ *directly* by using of confidential data
  - ▶ *indirectly* through spillovers to those not using the data
- ▶ Excluding direct users: coefficients 24–39% smaller but still large and significant
- ▶ Indirect channel: *exposure* to research based on administrative data
- ▶ Explains the large increase in citations to past FSRDC publications
- ▶ We would further expect:

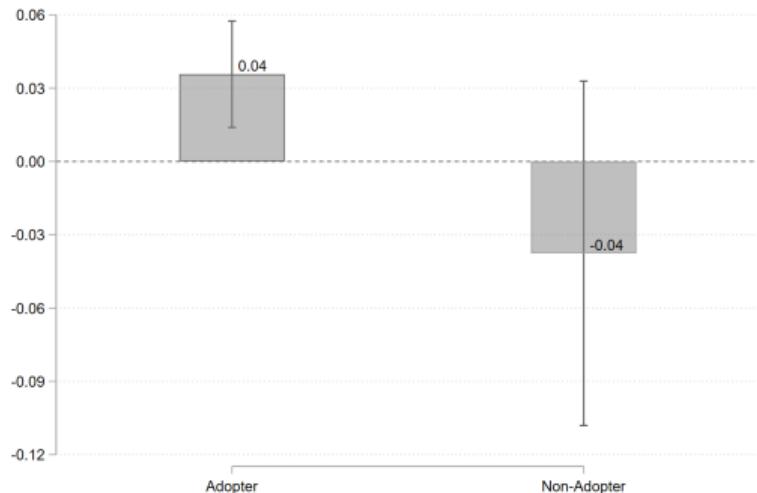
## Direct Data Use vs Spillover Effects

- ▶ Potential channels driving the productivity effect:
  - ▶ *directly* by using of confidential data
  - ▶ *indirectly* through spillovers to those not using the data
- ▶ Excluding direct users: coefficients 24–39% smaller but still large and significant
- ▶ Indirect channel: *exposure* to research based on administrative data
- ▶ Explains the large increase in citations to past FSRDC publications
- ▶ We would further expect:
  - ▶ Spillovers concentrated in departments where someone uses FSRDC
  - ▶ Spillovers larger on researchers building on FSRDC findings

## Exposure Channel (1): No Spillovers in Departments Without Adoption

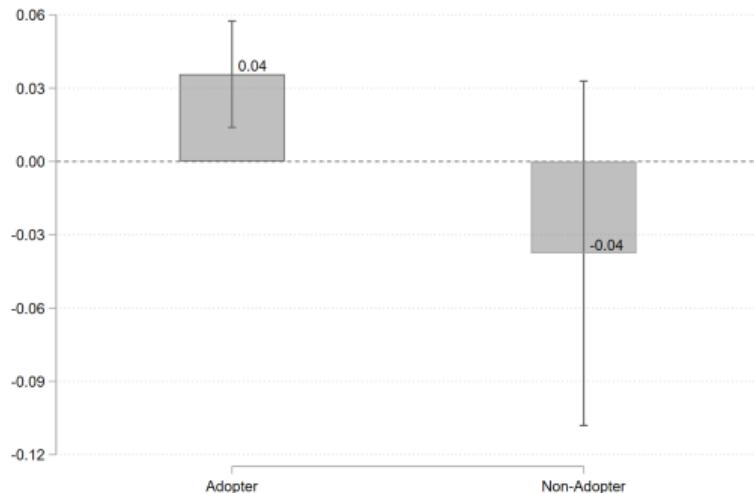
# Exposure Channel (1): No Spillovers in Departments Without Adoption

*Top Pubs*

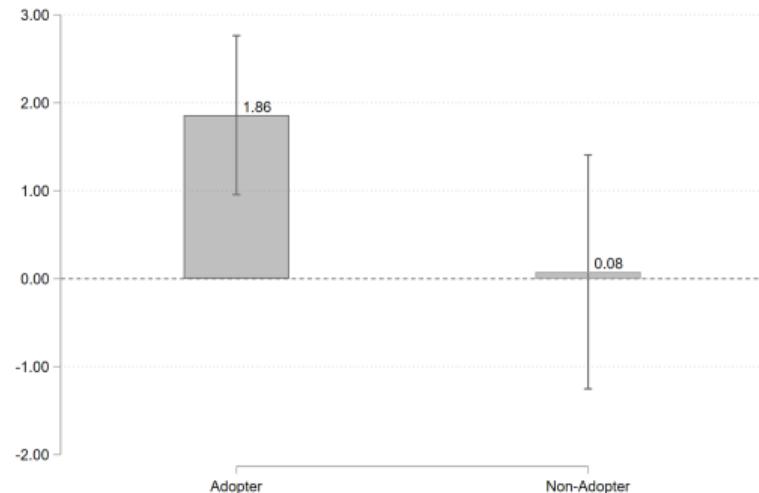


# Exposure Channel (1): No Spillovers in Departments Without Adoption

*Top Pubs*



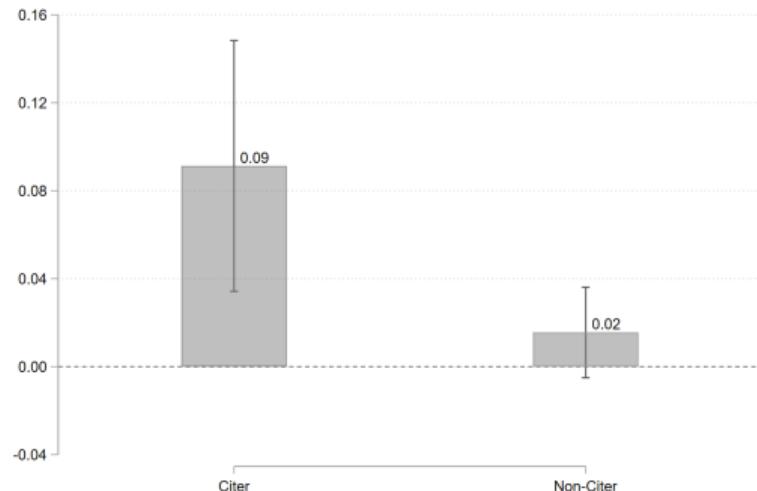
*Cite-weighted Pubs*



## Exposure Channel (2): Spillovers Concentrated on Citers

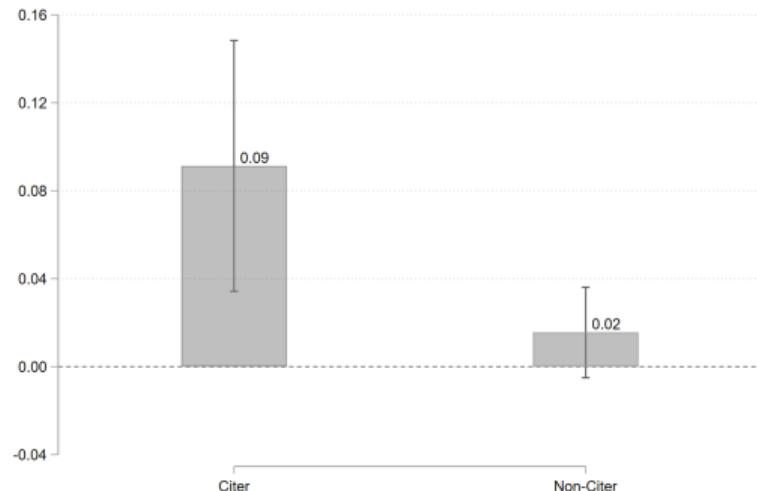
## Exposure Channel (2): Spillovers Concentrated on Citors

*Top Pubs*

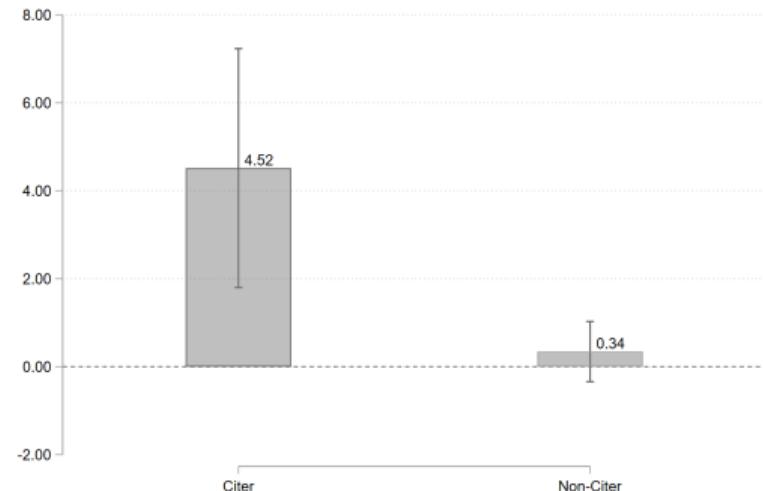


## Exposure Channel (2): Spillovers Concentrated on Citors

*Top Pubs*



*Cite-weighted Pubs*



## How Do Spillovers Operate?

# How Do Spillovers Operate?



## Exploration of New Topics

Exposure to FSRDC research encourages exploration of new domains

► Table

# How Do Spillovers Operate?



## Exploration of New Topics

Exposure to FSRDC research encourages exploration of new domains

[► Table](#)



## Design of Empirical Research

Exposure to FSRDC research inspires use of new data and empirical methods

[► Table](#)

## IV. Do Administrative Data Lead to More Policy-Relevant Research?

## Administrative Data and Policy Relevance of Research

- ▶ Expansion of access to administrative data is predicated on their potential to inform evidence-based policies (e.g., the Federal Data Strategy)

## Administrative Data and Policy Relevance of Research

- ▶ Expansion of access to administrative data is predicated on their potential to inform evidence-based policies (e.g., the Federal Data Strategy)
- ▶ Evidence remains mostly anecdotal to date (Card et al., 2010; Chetty, 2012; Einav and Levin, 2014)

# Administrative Data and Policy Relevance of Research

- ▶ Expansion of access to administrative data is predicated on their potential to inform evidence-based policies (e.g., the Federal Data Strategy)
- ▶ Evidence remains mostly anecdotal to date (Card et al., 2010; Chetty, 2012; Einav and Levin, 2014)
- ▶ We use novel data from Altmetric.com to track *citations* of economic research in policy documents



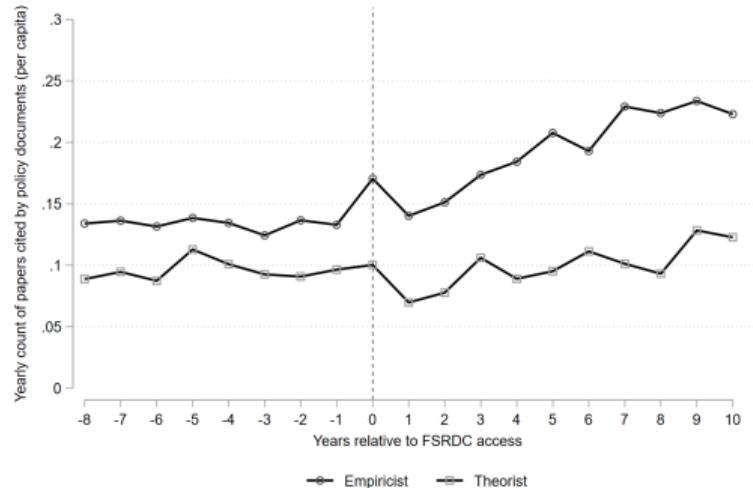
Identifying Agglomeration Spillovers:  
Evidence from Winners and Losers  
of Large Plant Openings

Michael Greenstone, Richard Hornbeck, and Enrico Moretti



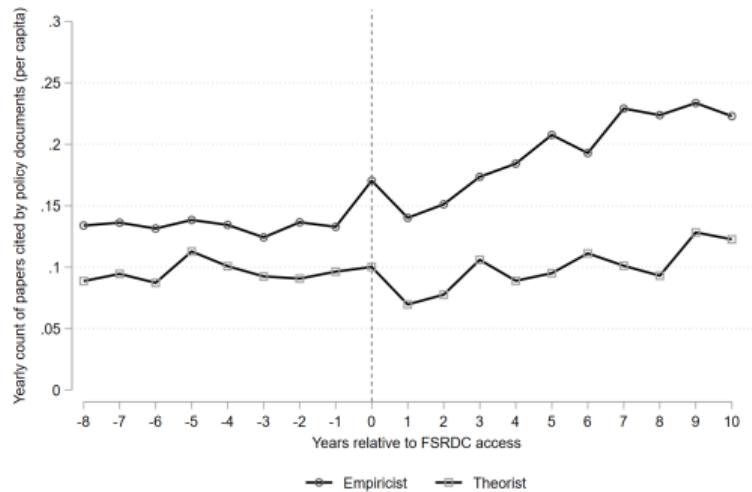
# Effect of Data Access on Policy Impact (Event Study)

*Papers cited by policy (raw means)*

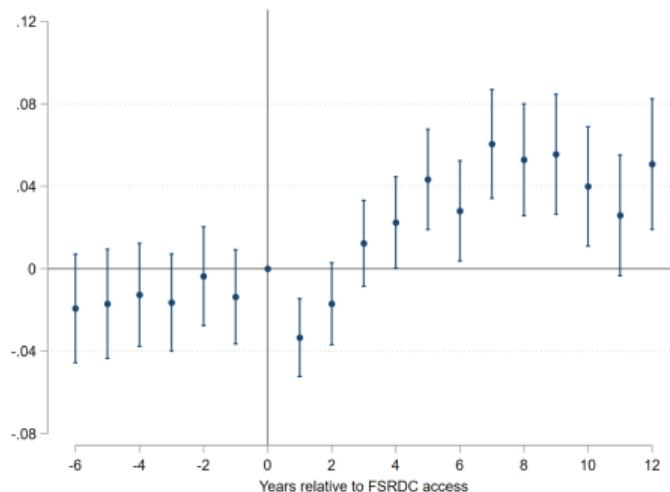


# Effect of Data Access on Policy Impact (Event Study)

*Papers cited by policy (raw means)*



*Papers cited by policy (event study)*



## Better Science Leads to Larger Policy Impact

- ▶ Access to administrative data increases the policy impact of applied work by 46-54% over the sample mean

## Better Science Leads to Larger Policy Impact

- ▶ Access to administrative data increases the policy impact of applied work by 46-54% over the sample mean
- ▶ Effect of Census data is larger among U.S.-based policy sources  Impact in U.S.

## Better Science Leads to Larger Policy Impact

- ▶ Access to administrative data increases the policy impact of applied work by 46-54% over the sample mean
- ▶ Effect of Census data is larger among U.S.-based policy sources ▶ Impact in U.S.
- ▶ No evidence of a shift towards more explicitly policy-oriented work ▶ Policy orientation

## Better Science Leads to Larger Policy Impact

- ▶ Access to administrative data increases the policy impact of applied work by 46-54% over the sample mean
- ▶ Effect of Census data is larger among U.S.-based policy sources ▶ Impact in U.S.
- ▶ No evidence of a shift towards more explicitly policy-oriented work ▶ Policy orientation
- ▶ Suggests that increase of policy impact is due to quality of the empirical evidence and not to a crowding out of scientifically important questions by policy topics

# Overview

Introduction

Setting

Data

Research Design

Results

Discussion

# Paper in one slide

## Research Question

- ▶ How does access to data access shape the quantity, quality and policy relevance of research?

# Paper in one slide

## Research Question

- ▶ How does access to data access shape the quantity, quality and policy relevance of research?

## Data

- ▶ Unique dataset of over 15,000 publishing economists over 30 years

# Paper in one slide

## Research Question

- ▶ How does access to data access shape the quantity, quality and policy relevance of research?

## Data

- ▶ Unique dataset of over 15,000 publishing economists over 30 years

## Research Design

- ▶ Diffusion of U.S. Census Bureau's Research Data Centers (FSRDCs) across the U.S.

# Paper in one slide

## Research Question

- ▶ How does access to data access shape the quantity, quality and policy relevance of research?

## Data

- ▶ Unique dataset of over 15,000 publishing economists over 30 years

## Research Design

- ▶ Diffusion of U.S. Census Bureau's Research Data Centers (FSRDCs) across the U.S.

## Results

- ▶ **Diffusion:** Co-location with FSRDCs increases use likelihood to use and build upon census data
- ▶ **Productivity:** Treated empiricists are more likely to publish in top journals and produce more highly cited work
- ▶ **Spillovers:** Large effects on non-users by making them aware of FSRDC research, inspiring them to explore new topics and improve research design of applied work
- ▶ **Policy impact:** Applied researchers achieve larger policy impact, without crowding out scientific novelty for policy work

# Administrative Data as the Webb Telescope for Social Science?



# Thank You!

 tranchero@berkeley.edu

 www.matteotranchero.com

## Summary Statistics: Researcher Level

	N	Mean	Std. Dev.	Median	Min	Max
Ever Had FSRDC Access (0/1)	15750	0.408	0.49	0	0	1
Year of Access	6425	2007.517	6.83	2008	1994	2019
Ever Used FSRDC (0/1)	15750	0.028	0.17	0	0	1
Ever Cited FSRDC (0/1)	15750	0.234	0.42	0	0	1
Lifetime Top Publications	15750	2.278	4.41	1	0	93
Lifetime Cite-weighted Papers	15750	66.896	187.79	6	0	5536
Ever Top 5 Papers (0/1)	15750	0.236	0.42	0	0	1
Ever Top 5% Cited Papers (0/1)	15750	0.242	0.43	0	0	1
Rank of Institutions (avg)	15750	14.307	18.93	6	0	86
Empiricist (0/1)	15750	0.731	0.44	1	0	1

## Summary Statistics: Panel

	N	Mean	Std. Dev.	Median	Min	Max
Ever Had FSRDC Access (0/1)	15750	0.408	0.49	0	0	1
Year of Access	6425	2007.517	6.83	2008	1994	2019
Ever Used FSRDC (0/1)	15750	0.028	0.17	0	0	1
Ever Cited FSRDC (0/1)	15750	0.234	0.42	0	0	1
Lifetime Top Publications	15750	2.278	4.41	1	0	93
Lifetime Cite-weighted Papers	15750	66.896	187.79	6	0	5536
Ever Top 5 Papers (0/1)	15750	0.236	0.42	0	0	1
Ever Top 5% Cited Papers (0/1)	15750	0.242	0.43	0	0	1
Rank of Institutions (avg)	15750	14.307	18.93	6	0	86
Empiricist (0/1)	15750	0.731	0.44	1	0	1

## Summary Statistics: Panel

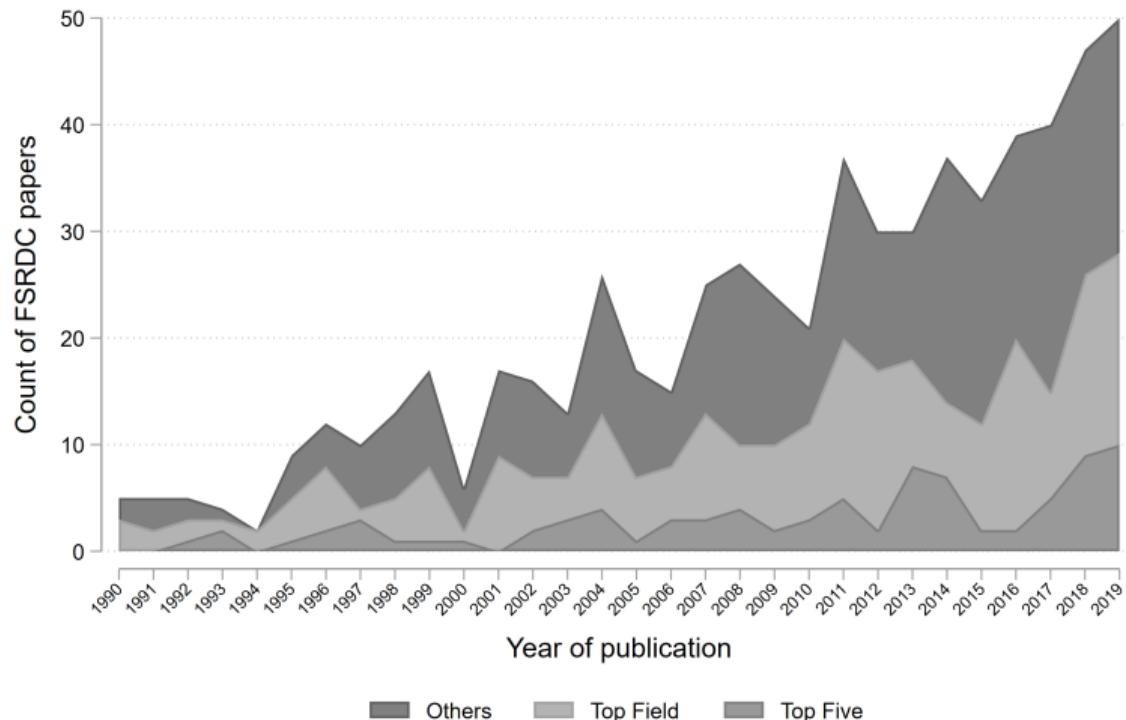
	N	Mean	Std. Dev.	Median	Min	Max
Post-FSRDC (0/1)	246711	0.258	0.44	0	0	1
Papers Using FSRDC	246711	0.003	0.06	0	0	3
Papers Citing FSRDC	246711	0.033	0.20	0	0	6
Top Publications	246711	0.145	0.43	0	0	9
Cite-weighted Papers	246711	4.271	19.79	0	0	1285
Top 5 Papers	246711	0.047	0.24	0	0	6
Top 5% Cited Papers	246711	0.044	0.23	0	0	5
New JEL Codes (0/1)	246711	0.389	0.49	0	0	1
New LDA Topics (0/1)	246711	0.297	0.46	0	0	1
Papers with FSRDC JEL (0/1)	246711	0.105	0.31	0	0	1
Papers without FSRDC JEL (0/1)	246711	0.337	0.47	0	0	1
Papers Mentioning Admin Data	246711	0.003	0.06	0	0	2
Papers Mentioning Survey Data	246711	0.005	0.08	0	0	3
Quasi-experimental Papers	246711	0.019	0.14	0	0	4
Experimental Papers	246711	0.036	0.23	0	0	10
Year	246711	2005.953	7.71	2007	1990	2019

## Summary Statistics

	N	Mean	Std. Dev.	Median	Min	Max
Post-FSRDC (0/1)	246711	0.258	0.44	0	0	1
Papers Using FSRDC	246711	0.003	0.06	0	0	3
Papers Citing FSRDC	246711	0.033	0.20	0	0	6
Top Publications	246711	0.145	0.43	0	0	9
Cite-weighted Papers	246711	4.271	19.79	0	0	1285
Top 5 Papers	246711	0.047	0.24	0	0	6
Top 5% Cited Papers	246711	0.044	0.23	0	0	5
New JEL Codes (0/1)	246711	0.389	0.49	0	0	1
New LDA Topics (0/1)	246711	0.297	0.46	0	0	1
Papers with FSRDC JEL (0/1)	246711	0.105	0.31	0	0	1
Papers without FSRDC JEL (0/1)	246711	0.337	0.47	0	0	1
Papers Mentioning Admin Data	246711	0.003	0.06	0	0	2
Papers Mentioning Survey Data	246711	0.005	0.08	0	0	3
Quasi-experimental Papers	246711	0.019	0.14	0	0	4
Experimental Papers	246711	0.036	0.23	0	0	10
Year	246711	2005.953	7.71	2007	1990	2019

► Back

# FSRDC Papers Over Time

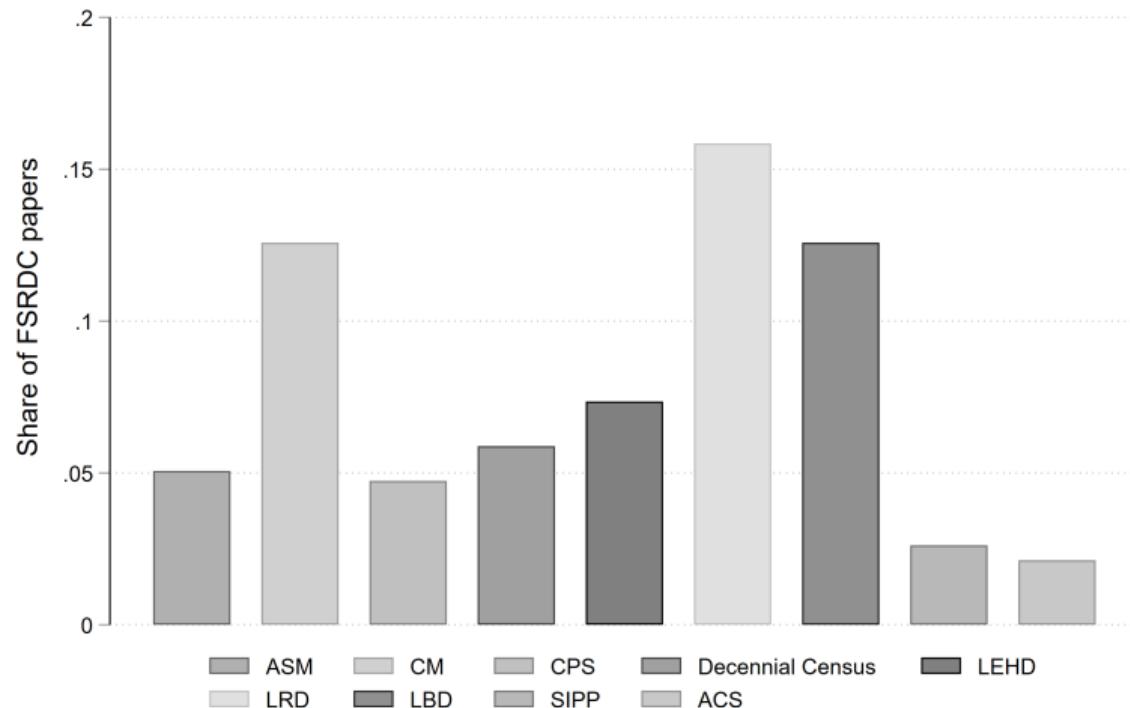


## Characteristics of FSRDC Papers

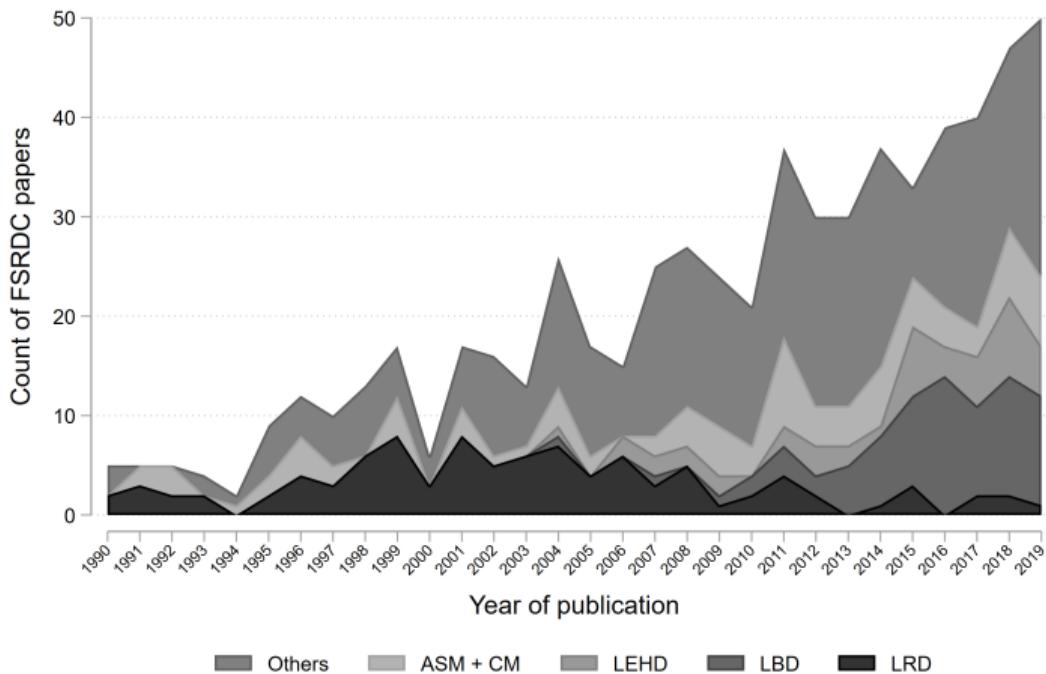
	Top Field	Top Five	Average Cites	Top 5% Cites
FSRDC Paper	0.1658*** (0.01899)	0.0954*** (0.01453)	3.5539*** (0.37347)	0.1010*** (0.01267)
Year FE	YES	YES	YES	YES
N	188,181	188,181	182,948	182,948
Mean of DV	0.1987	0.0483	3.4362	0.0568

► Back

# Most Frequent FSRDC Datasets



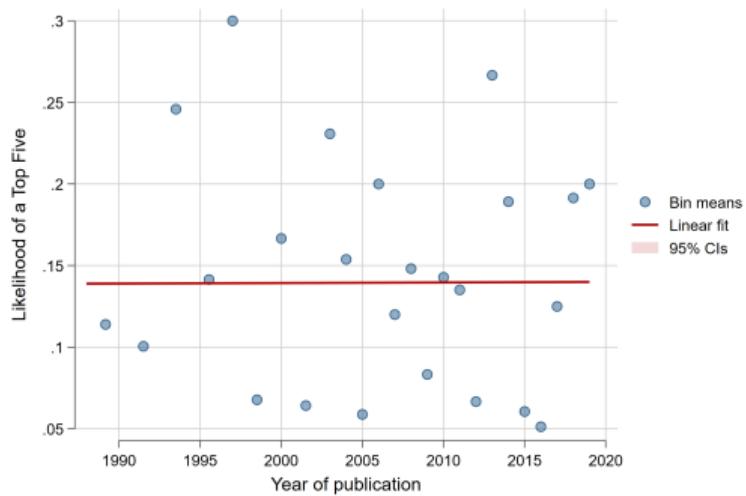
# Use of FSRDC Datasets Over Time



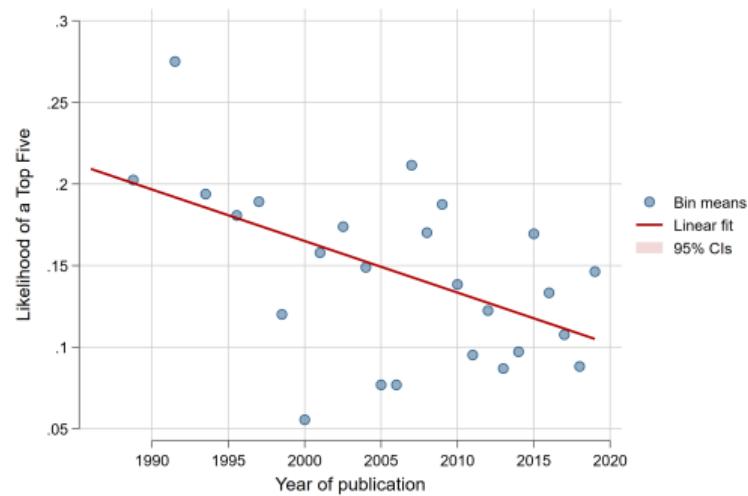
► Back

# Likelihood of Top Five Publication, FSRDC vs PSID Papers

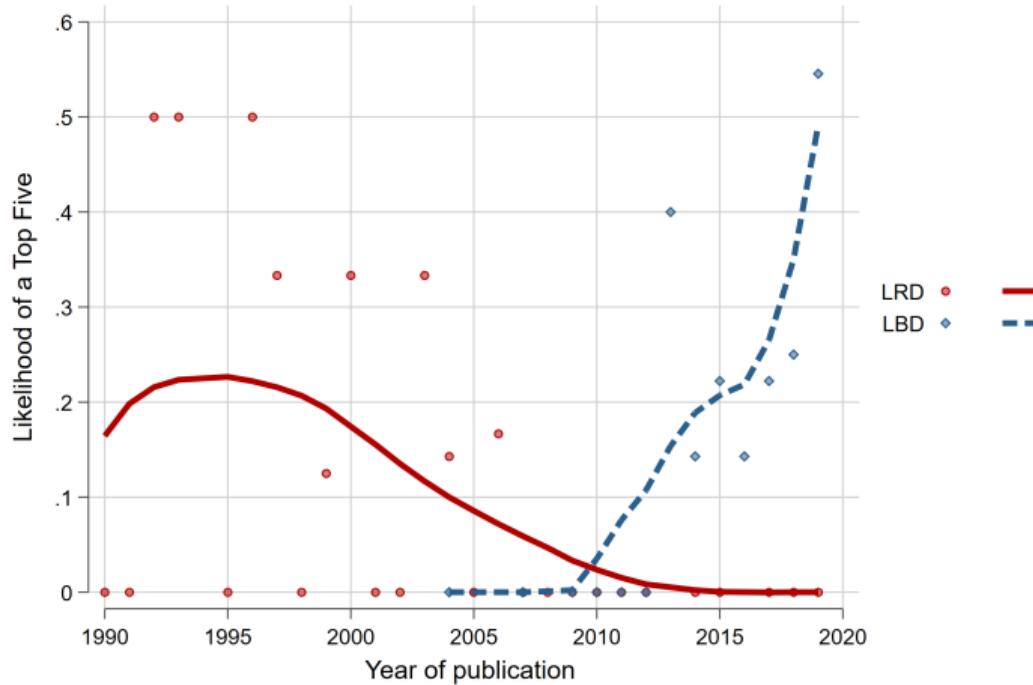
*FSRDC Papers*



*PSID Papers*



# Likelihood of Top Five Publication, LRD vs LBD Papers

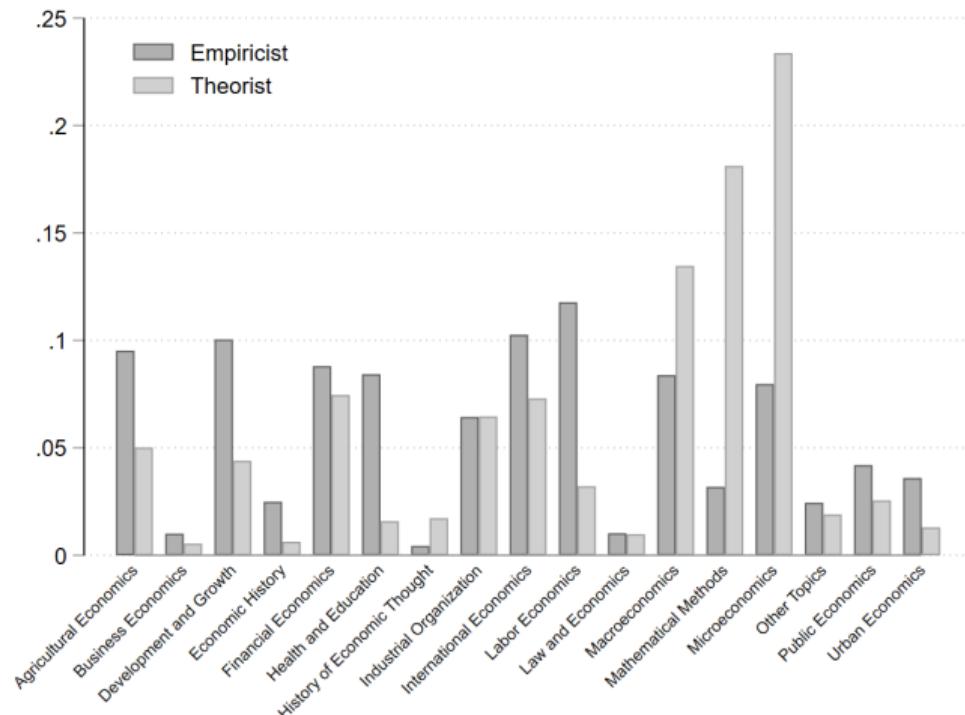


▶ Back

## Measuring Validation: Methodological Classification for the Editorial Board

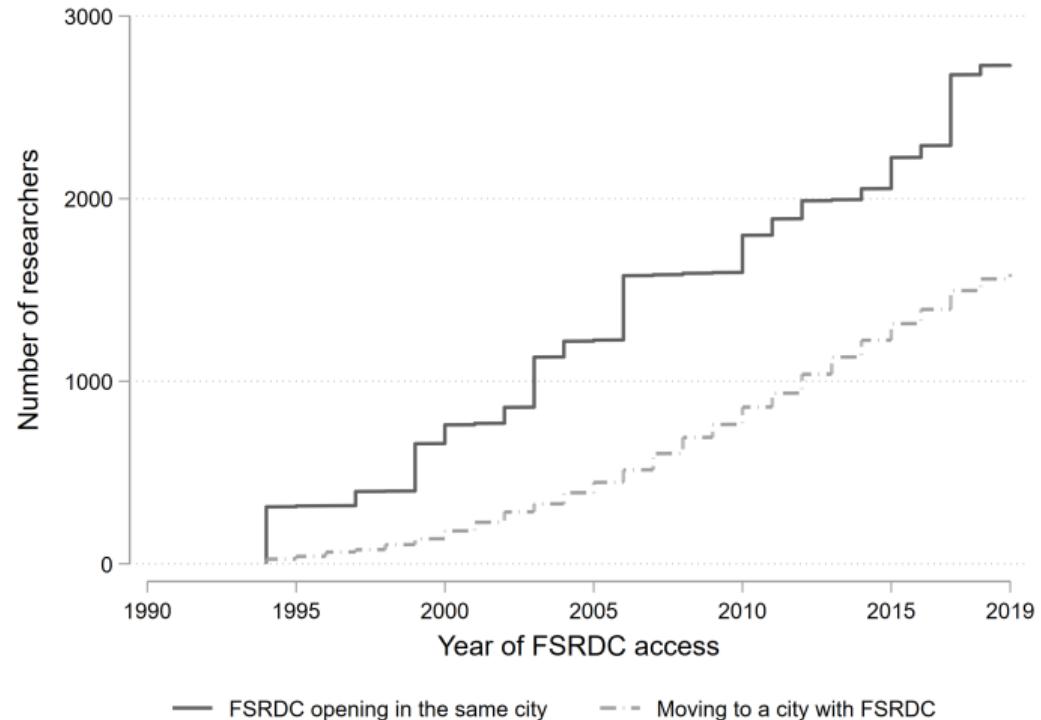
Journal	Number of Editors	Ever U.S.-affiliated	Share Empirical
Journal of Economic Theory	62	47	2.13%
AEJ: Microeconomics	21	19	31.58%
AEJ: Macroeconomics	19	18	66.67%
AEJ: Applied Economics	36	33	96.97%
AEJ: Economic Policy	37	35	97.14%

# Measuring Validation: Distribution of PhD Dissertation Fields

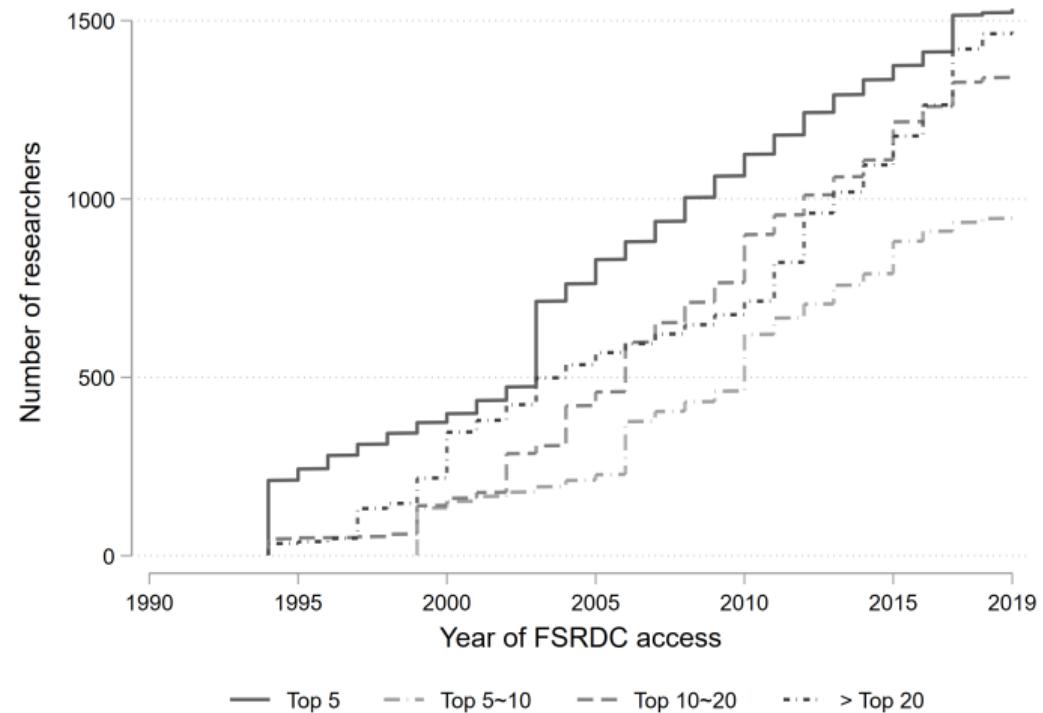


► Back

# Cumulative Number of Researchers Gaining FSRDC Access over Time



## Cumulative Number of Researchers Gaining FSRDC Access over Time



▶ Back

# Effect of FSRDC Access on Administrative Data Adoption

- ▶  $\overline{\text{Direct Usage}} = 0.003$  and  $\overline{\text{Cite FSRDC}} = 0.033$

	Papers Using FSRDC	
	(1)	(2)
Post-FSRDC	-0.000329 (0.00069)	0.00216 (0.00114)
Post-FSRDC $\times$ Empiricist	0.00392*** (0.00107)	0.00332** (0.00113)
Researcher FE	Yes	Yes
University Tier $\times$ Year FE	Yes	No
University $\times$ Year FE	No	Yes
N	246532	245556

# Effect of FSRDC Access on Administrative Data Adoption

- ▶  $\overline{\text{Direct Usage}} = 0.003$  and  $\overline{\text{Cite FSRDC}} = 0.033$

	Papers Using FSRDC	
	(1)	(2)
Post-FSRDC	-0.000329 (0.00069)	0.00216 (0.00114)
Post-FSRDC × Empiricist	0.00392*** (0.00107)	0.00332** (0.00113)
Researcher FE	Yes	Yes
University Tier × Year FE	Yes	No
University × Year FE	No	Yes
N	246532	245556

## Effect of FSRDC Access on Administrative Data Adoption

- $\overline{\text{Direct Usage}} = 0.003$  and  $\overline{\text{Cite FSRDC}} = 0.033$

	Papers Using FSRDC		Papers Citing FSRDC	
	(1)	(2)	(3)	(4)
Post-FSRDC	-0.000329 (0.00069)	0.00216 (0.00114)	-0.00884** (0.00333)	-0.00235 (0.00463)
Post-FSRDC $\times$ Empiricist	0.00392*** (0.00107)	0.00332** (0.00113)	0.0189*** (0.00425)	0.0174*** (0.00444)
Researcher FE	Yes	Yes	Yes	Yes
University Tier $\times$ Year FE	Yes	No	Yes	No
University $\times$ Year FE	No	Yes	No	Yes
N	246532	245556	246532	245556

## Effect of FSRDC Access on Administrative Data Adoption

- $\overline{\text{Direct Usage}} = 0.003$  and  $\overline{\text{Cite FSRDC}} = 0.033$

	Papers Using FSRDC		Papers Citing FSRDC	
	(1)	(2)	(3)	(4)
Post-FSRDC	-0.000329 (0.00069)	0.00216 (0.00114)	-0.00884** (0.00333)	-0.00235 (0.00463)
Post-FSRDC $\times$ Empiricist	0.00392*** (0.00107)	0.00332** (0.00113)	0.0189*** (0.00425)	0.0174*** (0.00444)
Researcher FE	Yes	Yes	Yes	Yes
University Tier $\times$ Year FE	Yes	No	Yes	No
University $\times$ Year FE	No	Yes	No	Yes
N	246532	245556	246532	245556

► Back

## Effect on FSRDC Project Approvals and CES Working Papers

	FSRDC Project Approval (0/1)	
	(1)	(2)
Post-FSRDC	-0.000996 (0.00052)	-0.00178 (0.00098)
Post-FSRDC × Empiricist	0.00457*** (0.00079)	0.00460*** (0.00080)
Researcher FE	Yes	Yes
University Tier × Year FE	Yes	No
University × Year FE	No	Yes
N	246532	245556

## Effect on FSRDC Project Approvals and CES Working Papers

	FSRDC Project Approval (0/1)	
	(1)	(2)
Post-FSRDC	-0.000996 (0.00052)	-0.00178 (0.00098)
Post-FSRDC × Empiricist	0.00457*** (0.00079)	0.00460*** (0.00080)
Researcher FE	Yes	Yes
University Tier × Year FE	Yes	No
University × Year FE	No	Yes
N	246532	245556

## Effect on FSRDC Project Approvals and CES Working Papers

	FSRDC Project Approval (0/1)		CES Working Paper (0/1)	
	(1)	(2)	(3)	(4)
Post-FSRDC	-0.000996 (0.00052)	-0.00178 (0.00098)	-0.000883 (0.00052)	-0.000448 (0.00093)
Post-FSRDC × Empiricist	0.00457*** (0.00079)	0.00460*** (0.00080)	0.00200** (0.00074)	0.00182* (0.00078)
Researcher FE	Yes	Yes	Yes	Yes
University Tier × Year FE	Yes	No	Yes	No
University × Year FE	No	Yes	No	Yes
N	246532	245556	246532	245556

## Effect on FSRDC Project Approvals and CES Working Papers

	FSRDC Project Approval (0/1)		CES Working Paper (0/1)	
	(1)	(2)	(3)	(4)
Post-FSRDC	-0.000996 (0.00052)	-0.00178 (0.00098)	-0.000883 (0.00052)	-0.000448 (0.00093)
Post-FSRDC × Empiricist	0.00457*** (0.00079)	0.00460*** (0.00080)	0.00200** (0.00074)	0.00182* (0.00078)
Researcher FE	Yes	Yes	Yes	Yes
University Tier × Year FE	Yes	No	Yes	No
University × Year FE	No	Yes	No	Yes
N	246532	245556	246532	245556

► Back

# Exposure of FSRDC Measured by the Timing of the NSF Grant

	FSRDC Use (1)	FSRDC Cite (2)	Top Pubs (3)	Cite-Pubs (4)	JEL Code (5)	LDA Topic (6)
Post-Grant	0.000894 (0.001)	-0.00810* (0.003)	-0.0143 (0.011)	-0.558 (0.362)	-0.0416*** (0.010)	-0.0512*** (0.010)
Post-Grant × Empiricist	0.00152 (0.001)	0.0138*** (0.004)	0.0296** (0.011)	1.502*** (0.411)	0.0347*** (0.010)	0.0240* (0.010)
Person FE	Yes	Yes	Yes	Yes	Yes	Yes
University × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	229756	229756	229756	229756	215048	215048

► Back

## Effect of FSRDC Access Excluding NSF Grant Applicants

	FSRDC Use (1)	FSRDC Cite (2)	Top Pubs (3)	Cite-Pubs (4)
Post-FSRDC	0.00224* (0.001)	-0.00196 (0.005)	-0.00422 (0.011)	-0.123 (0.436)
Post-FSRDC × Empiricist	0.00323** (0.001)	0.0161*** (0.004)	0.0316** (0.011)	1.664*** (0.445)
Researcher FE	Yes	Yes	Yes	Yes
University × Year FE	Yes	Yes	Yes	Yes
N	244850	244850	244850	244850

► Back

# Effect of FSRDC Access Excluding Authors in Institutions Hosting FSRDCs

	FSRDC Use (1)	FSRDC Cite (2)	Top Pubs (3)	Cite-Pubs (4)
Post-FSRDC	0.00249 (0.002)	-0.00400 (0.007)	-0.00573 (0.016)	-0.621 (0.636)
Post-FSRDC $\times$ Empiricist	0.00325 (0.002)	0.0129* (0.006)	0.0315* (0.015)	2.024** (0.654)
Researcher FE	Yes	Yes	Yes	Yes
University $\times$ Year FE	Yes	Yes	Yes	Yes
N	194110	194110	194110	194110

► Back

## Effect of FSRDC Access Excluding Authors Treated After Mobility Events

	FSRDC Use (1)	FSRDC Cite (2)	Top Pubs (3)	Cite-Pubs (4)
Post-FSRDC	0.00274 (0.001)	-0.00324 (0.005)	-0.0313* (0.015)	-0.928 (0.542)
Post-FSRDC × Empiricist	0.00347* (0.001)	0.0182*** (0.005)	0.0442*** (0.013)	2.141*** (0.560)
Researcher FE	Yes	Yes	Yes	Yes
University × Year FE	Yes	Yes	Yes	Yes
N	219941	219941	219941	219941

## Effect of FSRDC Access on Research Output

- ▶  $\overline{\text{Top Pubs}} = 0.145$  and  $\overline{\text{Cite-weighted Pubs}} = 4.271$

# Effect of FSRDC Access on Research Output

- $\overline{\text{Top Pubs}} = 0.145$  and  $\overline{\text{Cite-weighted Pubs}} = 4.271$

	Top Publications	
	(1)	(2)
Post-FSRDC	-0.0139 (0.010)	-0.00712 (0.012)
Post-FSRDC $\times$ Empiricist	0.0353*** (0.011)	0.0352** (0.011)
Researcher FE	Yes	Yes
University Tier $\times$ Year FE	Yes	No
University $\times$ Year FE	No	Yes
N	246532	245556

# Effect of FSRDC Access on Research Output

- $\overline{\text{Top Pubs}} = 0.145$  and  $\overline{\text{Cite-weighted Pubs}} = 4.271$

	Top Publications	
	(1)	(2)
Post-FSRDC	-0.0139 (0.010)	-0.00712 (0.012)
Post-FSRDC $\times$ Empiricist	0.0353*** (0.011)	0.0352** (0.011)
Researcher FE	Yes	Yes
University Tier $\times$ Year FE	Yes	No
University $\times$ Year FE	No	Yes
N	246532	245556

## Effect of FSRDC Access on Research Output

- $\overline{\text{Top Pubs}} = 0.145$  and  $\overline{\text{Cite-weighted Pubs}} = 4.271$

	Top Publications		Cite-weighted Publications	
	(1)	(2)	(3)	(4)
Post-FSRDC	-0.0139 (0.010)	-0.00712 (0.012)	-0.676* (0.336)	-0.170 (0.435)
Post-FSRDC × Empiricist	0.0353*** (0.011)	0.0352** (0.011)	1.714*** (0.425)	1.737*** (0.445)
Researcher FE	Yes	Yes	Yes	Yes
University Tier × Year FE	Yes	No	Yes	No
University × Year FE	No	Yes	No	Yes
N	246532	245556	246532	245556

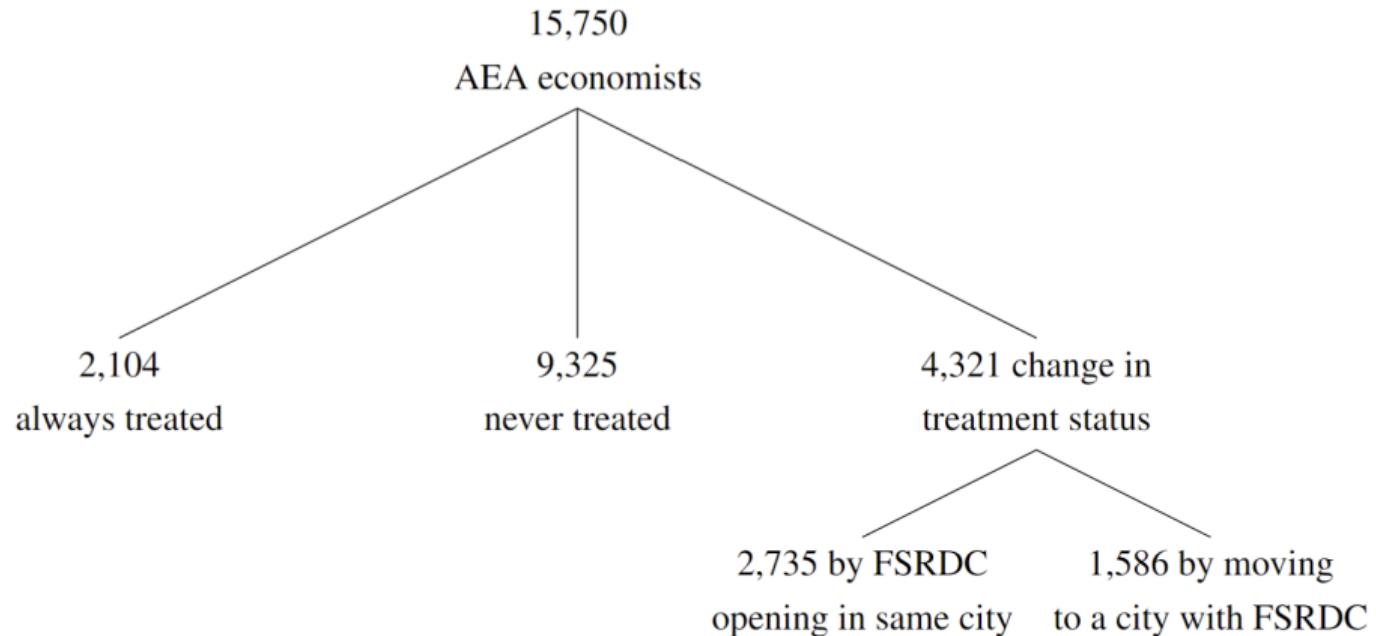
# Effect of FSRDC Access on Research Output

- $\overline{\text{Top Pubs}} = 0.145$  and  $\overline{\text{Cite-weighted Pubs}} = 4.271$

	Top Publications		Cite-weighted Publications	
	(1)	(2)	(3)	(4)
Post-FSRDC	-0.0139 (0.010)	-0.00712 (0.012)	-0.676* (0.336)	-0.170 (0.435)
Post-FSRDC $\times$ Empiricist	0.0353*** (0.011)	0.0352** (0.011)	1.714*** (0.425)	1.737*** (0.445)
Researcher FE	Yes	Yes	Yes	Yes
University Tier $\times$ Year FE	Yes	No	Yes	No
University $\times$ Year FE	No	Yes	No	Yes
N	246532	245556	246532	245556

► Back

## Variation in Exposure to FSRDCs



# Exploration of New Topics

- We fit 20 LDA topics to produce an unsupervised measure of new topics at the researcher-year level and look for new JEL codes

Topic Label	Most Salient Words				
topic 1 (innovation)	network	innovation	research	firm	social
topic 2 (institutions)	patent	r&d	election	voter	vote
topic 3 (trade)	trade	growth	export	firm	country
topic 4 (energy)	energy	electricity	policy	cost	efficiency
topic 5 (migrations)	quantile	immigrant	exit	immigration	alliance
topic 6 (corporate governance)	governance	seller	buyer	culture	girl
topic 7 (finance)	firm	bank	risk	stock	investor
topic 8 (welfare)	food	jump	maternal	nutrition	childhood
topic 9 (corporate finance)	merger	ambiguity	cultural	director	conservation
topic 10 (geography)	friction	venture	gold	building	corruption
topic 11 (econometrics)	model	forecast	method	estimator	test
topic 12 (health/education)	health	child	household	woman	income
topic 13 (natural resources)	renewable	sovereign	nuclear	screening	fine
topic 14 (macroeconomics)	rate	monetary	shock	tax	price
topic 15 (labor)	wage	worker	job	employment	labor
topic 16 (statistics)	tail	dependence	nonparametric	bootstrap	covariate
topic 17 (environmental)	emission	oil	carbon	climate	fuel
topic 18 (microeconomics)	game	equilibrium	preference	agent	information
topic 19 (accounting)	audit	dataset	reporting	auditor	transparency
topic 20 (healthcare)	patient	hospital	ecosystem	rating	student

## LDA Topics and Key Words

	New LDA Topic (0/1) (2)
Post-FSRDC	-0.00444 (0.011)
Post-FSRDC × Empiricist	0.0339*** (0.010)
Researcher FE	Yes
University × Year FE	Yes
N	229886

► Back

## Effect of FSRDC Access Excluding Direct FSRDC Users

	FSRDC Use (1)	FSRDC Cite (2)	Top Pubs (3)	Cite-Pubs (4)
Post-FSRDC	0.00274 (0.001)	-0.00324 (0.005)	-0.0313* (0.015)	-0.928 (0.542)
Post-FSRDC × Empiricist	0.00347* (0.001)	0.0182*** (0.005)	0.0442*** (0.013)	2.141*** (0.560)
Researcher FE	Yes	Yes	Yes	Yes
University × Year FE	Yes	Yes	Yes	Yes
N	219941	219941	219941	219941

► Back

## Spillover Mechanism (I): Direction of Empirical Research

- ▶  $\overline{\text{New JEL (0/1)}} = 0.389$

	New JEL Code (0/1)
	(1)
Post-FSRDC	-0.00686 (0.011)
Post-FSRDC × Empiricist	0.0475*** (0.010)
Researcher FE	Yes
University × Year FE	Yes
N	229886

## Spillover Mechanism (I): Direction of Empirical Research

- ▶  $\overline{\text{New JEL (0/1)}} = 0.389$

New JEL Code (0/1)	
	(1)
Post-FSRDC	-0.00686 (0.011)
Post-FSRDC × Empiricist	0.0475*** (0.010)
Researcher FE	Yes
University × Year FE	Yes
N	229886

## Spillover Mechanism (I): Direction of Empirical Research

- $\overline{FSRDC \ JEL \ (0/1)} = 0.105$  and  $\overline{Non-FSRDC \ JEL \ (0/1)} = 0.337$

	New JEL Code (0/1) (1)	FSRDC JELs (0/1) (3)	Non-FSRDC JELs (0/1) (4)
Post-FSRDC	-0.00686 (0.011)	-0.00332 (0.007)	-0.0125 (0.010)
Post-FSRDC $\times$ Empiricist	0.0475*** (0.010)	0.0175** (0.006)	0.0278** (0.009)
Researcher FE	Yes	Yes	Yes
University $\times$ Year FE	Yes	Yes	Yes
N	229886	245556	245556

# Spillover Mechanism (I): Direction of Empirical Research

» Back

- ▶  $\overline{FSRDC \text{ JEL} \ (0/1)} = 0.105$  and  $\overline{Non-FSRDC \text{ JEL} \ (0/1)} = 0.337$

	New JEL Code (0/1) (1)	FSRDC JELs (0/1) (3)	Non-FSRDC JELs (0/1) (4)
Post-FSRDC	-0.00686 (0.011)	-0.00332 (0.007)	-0.0125 (0.010)
Post-FSRDC × Empiricist	0.0475*** (0.010)	0.0175** (0.006)	0.0278** (0.009)
Researcher FE	Yes	Yes	Yes
University × Year FE	Yes	Yes	Yes
N	229886	245556	245556

» JEL Codes Associated to FSRDC Papers

» Alternative Def of FSRDC/Non-FSRDC JEL Codes

» Exploration of New Topics (LDA)

## Spillover Mechanism (II): Design of Empirical Research

- ▶  $\overline{\text{Admin Data}} = 0.003$  and  $\overline{\text{Survey Data}} = 0.005$

	Data	
	Admin (1)	Survey (2)
Post-FSRDC	-0.00243* (0.001)	0.00220 (0.001)
Post-FSRDC × Empiricist	0.00328*** (0.001)	0.000364 (0.001)
Researcher FE	Yes	Yes
University × Year FE	Yes	Yes
N	245556	245556

## Spillover Mechanism (II): Design of Empirical Research

- ▶  $\overline{\text{Admin Data}} = 0.003$  and  $\overline{\text{Survey Data}} = 0.005$

	Data	
	Admin (1)	Survey (2)
Post-FSRDC	-0.00243* (0.001)	0.00220 (0.001)
Post-FSRDC × Empiricist	0.00328*** (0.001)	0.000364 (0.001)
Researcher FE	Yes	Yes
University × Year FE	Yes	Yes
N	245556	245556

## Spillover Mechanism (II): Design of Empirical Research

- $\overline{\text{Quasi-Experiment}} = 0.019$  and  $\overline{\text{Experiment}} = 0.036$

	Data		Research Methods	
	Admin (1)	Survey (2)	Quasi-Exp. (3)	Experiment (4)
Post-FSRDC	-0.00243* (0.001)	0.00220 (0.001)	0.000395 (0.003)	0.00372 (0.005)
Post-FSRDC × Empiricist	0.00328*** (0.001)	0.000364 (0.001)	0.00786** (0.003)	0.000294 (0.005)
Researcher FE	Yes	Yes	Yes	Yes
University × Year FE	Yes	Yes	Yes	Yes
N	245556	245556	245556	245556

- $\overline{\text{Quasi-Experiment}} = 0.019$  and  $\overline{\text{Experiment}} = 0.036$

	Data		Research Methods	
	Admin (1)	Survey (2)	Quasi-Exp. (3)	Experiment (4)
Post-FSRDC	-0.00243* (0.001)	0.00220 (0.001)	0.000395 (0.003)	0.00372 (0.005)
Post-FSRDC × Empiricist	0.00328*** (0.001)	0.000364 (0.001)	0.00786** (0.003)	0.000294 (0.005)
Researcher FE	Yes	Yes	Yes	Yes
University × Year FE	Yes	Yes	Yes	Yes
N	245556	245556	245556	245556

## Effect of FSRDC Access on Additional Proxies of Research Design

	Big Data (1)	Data (2)	Natural Exp. (3)
Post-FSRDC	-0.000457* (0.000)	0.0204* (0.009)	-0.00139 (0.001)
Post-FSRDC × Empiricist	0.0000839 (0.000)	-0.00139 (0.008)	0.00335** (0.001)
Researcher FE	Yes	Yes	Yes
University × Year FE	Yes	Yes	Yes
N	245556	245556	245556

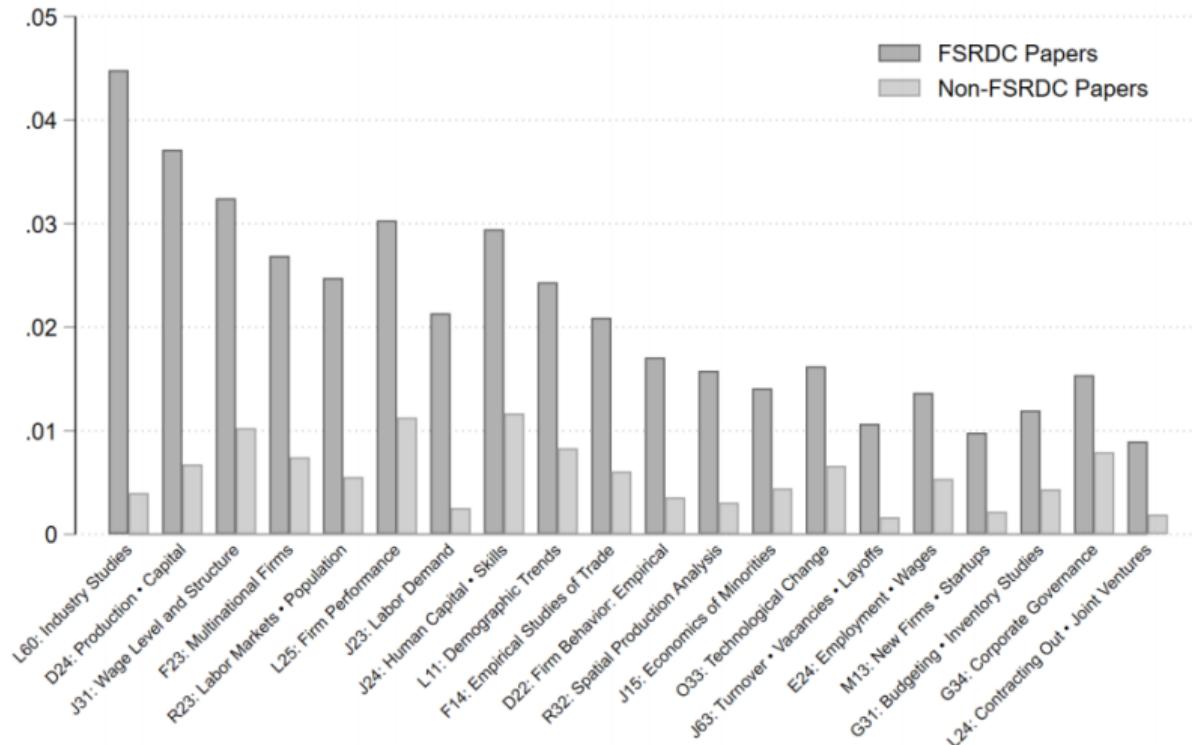
► Back

# Keywords Used to Tag Articles' Research Design and Data

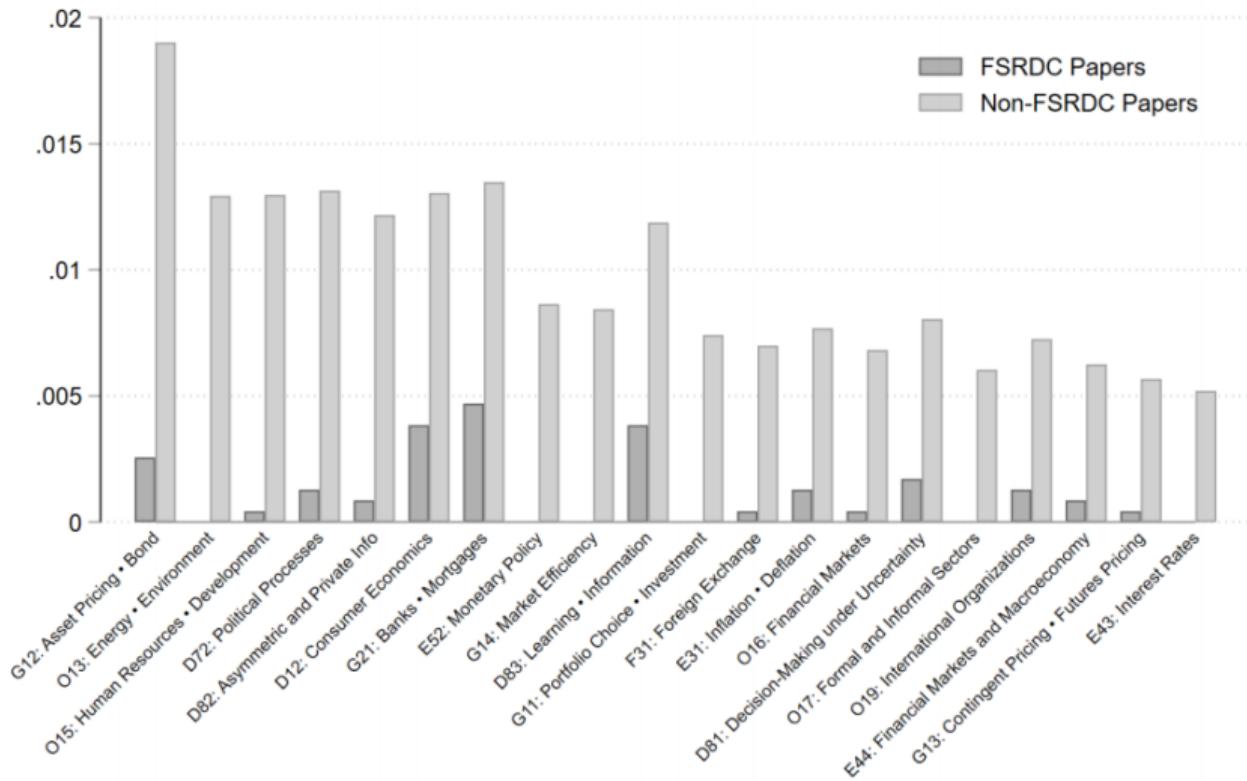
Category	Keywords (from Currie et al., 2020)
Administrative data	"administrative data" "admin data" "administrative-data" "admin-data" "administrative record" "admin record" "administrative regist" "admin regist" "registry data" "register data" "confidential data" "restricted data"
Survey data	"survey data" "phone survey" "survey administered" "household survey"
Quasi-experimental methods	"difference in diff" "differenceindiff" "differences in diff" "differencesindiff" "d-in-d" "diff in diff" "diffindiff" "event stud" "eventstud" "staggered adoption" "regression discontinuit" "regressondiscontinuit" "regression kink" "regressionkink" "rd resign" "rdresign" "rd estimat" "rdestimat" "rd model" "rdmodel" "rd regression" "rdregression" "rd coefficient" "rdcoefficient" "rk design" "rkdesign" "rdd" "rkd" "instrumental variable" "instrumentalvariable" "two stage least squares" "twostage least squares" "2sls" "tsls" "valid instrument" "exogenous instrument" "iv estimat" "ivestimat" "iv regression" "ivregression" "iv strateg" "ivstrateg" "we instrument" "i instrument" "exclusion restriction" "paper instruments" "weak first stage" "simulated instrument"
Experimental methods	"randomized controlled trial" "randomized field experiment" "randomized controlled experiment" "randomised controlled trial" "randomised control trial" "randomised field experiment" "randomised controlled experiment" "social experiment" "rct" "laboratory experiment" "lab experiment" "public good game" "public goods game" "ztree" "orsee" "showup fee" "laboratory participant" "lab participant" "pre analysis plan" "preanalysis plan" "pre registered" "preregistered" "preregistration" "dictator game" "ultimatum game" "trust game"

► Back

## JELs most representative of FSRDC research



## JELs least representative of FSRDC research



## Alternative Definitions of FSRDC and Non-FSRDC JEL Codes (Dummy)

	Definition 1		Definition 2		Definition 3	
	FSRDC (1)	Non-FSRDC (2)	FSRDC (3)	Non-FSRDC (4)	FSRDC (5)	Non-FSRDC (6)
Post-FSRDC	-0.00332 (0.007)	-0.0125 (0.010)	-0.0107 (0.006)	-0.0105 (0.010)	-0.00817 (0.008)	-0.0127 (0.010)
Post-FSRDC × Empiricist	0.0175** (0.006)	0.0278** (0.009)	0.0278*** (0.006)	0.0263** (0.009)	0.0355*** (0.007)	0.0294** (0.009)
Dependent Variable Mean	0.098	0.320	0.099	0.320	0.160	0.311
Researcher FE	Yes	Yes	Yes	Yes	Yes	Yes
University × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	245556	245556	245556	245556	245556	245556

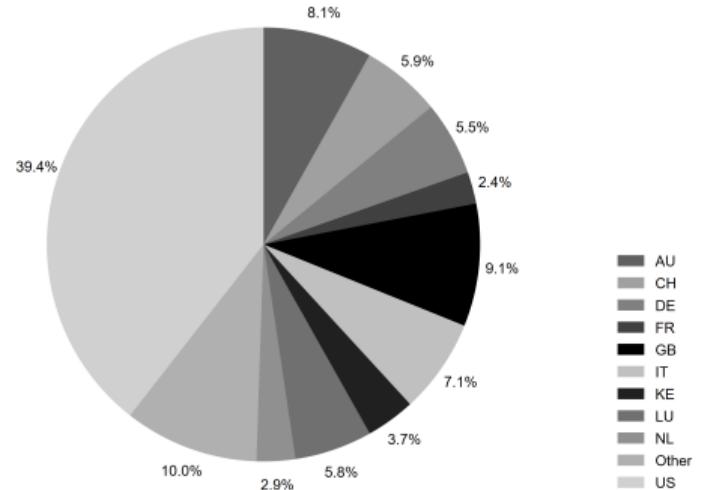
## Alternative Definitions of FSRDC and Non-FSRDC JEL Codes (Count)

	Definition 1		Definition 2		Definition 3	
	FSRDC (1)	Non-FSRDC (2)	FSRDC (3)	Non-FSRDC (4)	FSRDC (5)	Non-FSRDC (6)
Post-FSRDC	-0.00332 (0.007)	-0.0125 (0.010)	-0.0107 (0.006)	-0.0105 (0.010)	-0.00817 (0.008)	-0.0127 (0.010)
Post-FSRDC × Empiricist	0.0175** (0.006)	0.0278** (0.009)	0.0278*** (0.006)	0.0263** (0.009)	0.0355*** (0.007)	0.0294** (0.009)
Dependent Variable Mean	0.098	0.320	0.099	0.320	0.160	0.311
Researcher FE	Yes	Yes	Yes	Yes	Yes	Yes
University × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	245556	245556	245556	245556	245556	245556

► Back

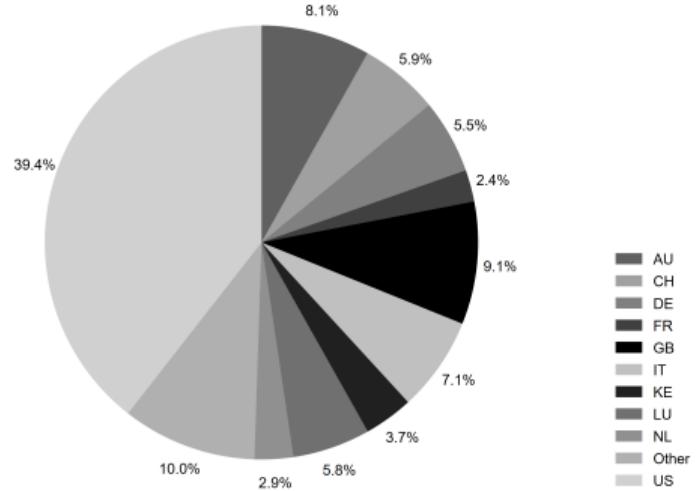
# Origin of Policy Citations to Economic Scholarship

*Country of Policy Source*

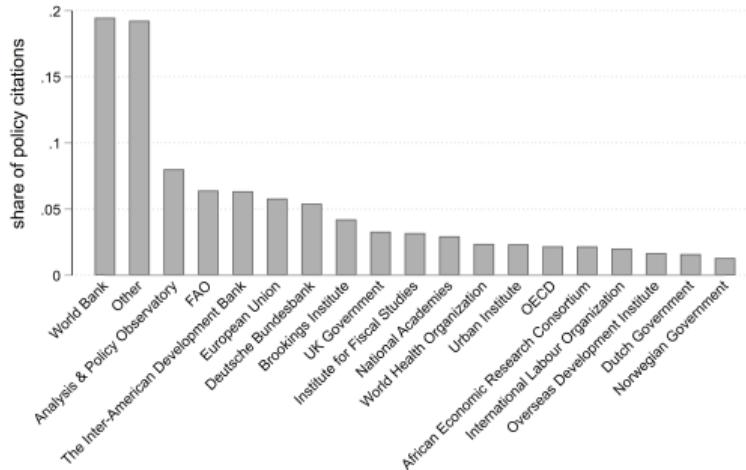


# Origin of Policy Citations to Economic Scholarship

*Country of Policy Source*

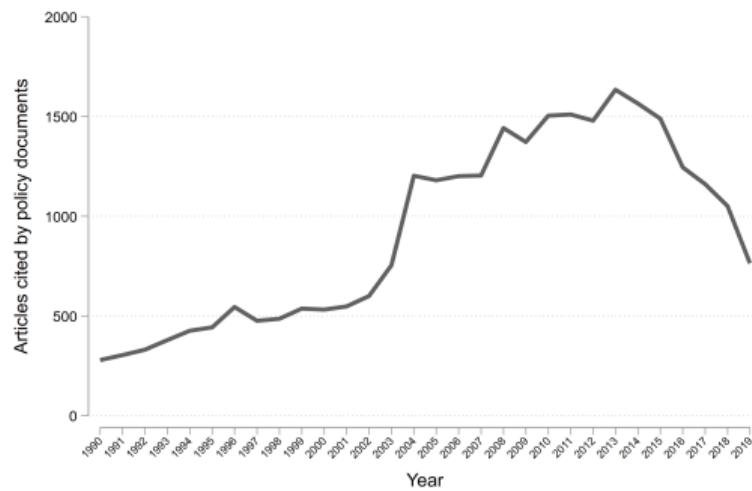


*Main Policy Sources*



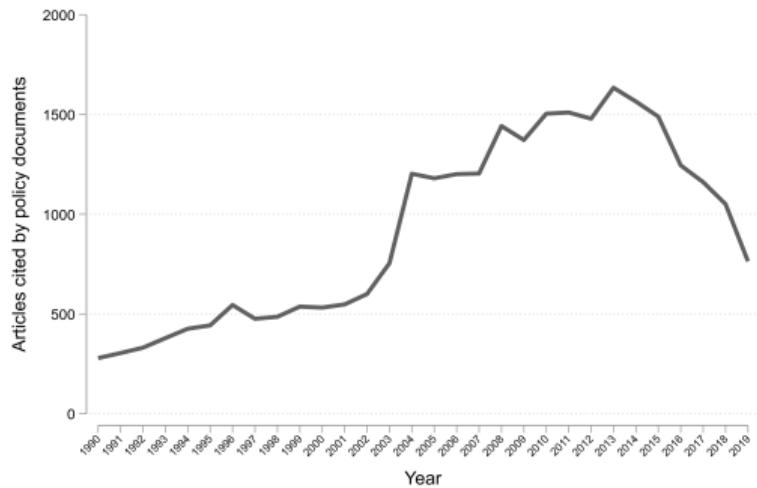
# Time Distribution of Policy Citations

*Year of Articles Cited*

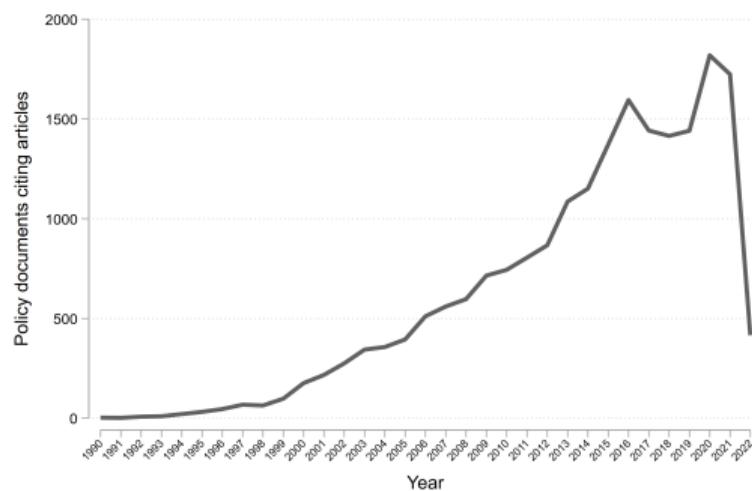


# Time Distribution of Policy Citations

*Year of Articles Cited*



*Year of Policy Documents*



► Back

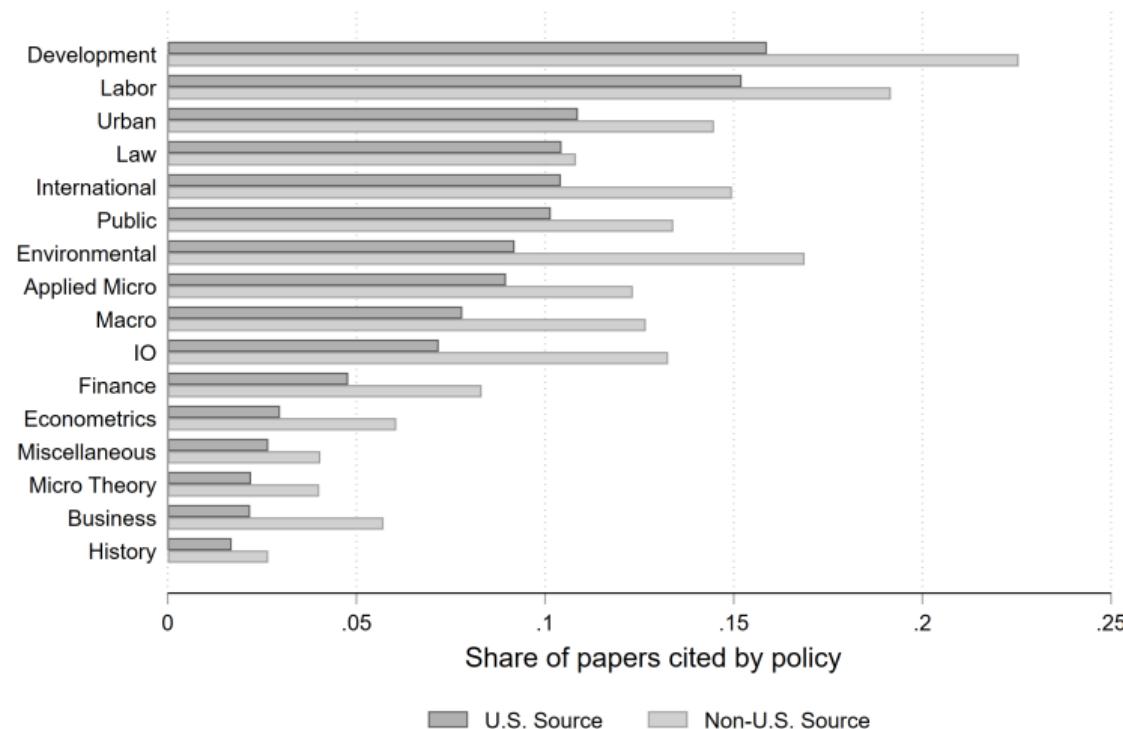
# Origin of Policy Citations to Economic Scholarship

Table: Statistical Association Between Use of Census Data and Policy Impact

	US cites (1)	Non-US cites (2)	US cites (3)	Non-US cites (4)	US cites (5)	Non-US cites (6)
Empirical (0/1)	0.1627*** (0.00939)	0.1541*** (0.01275)				
FSRDC use (0/1)			0.8086*** (0.13732)	0.3240** (0.09898)		
FSRDC cite (0/1)					0.1835*** (0.03096)	0.1871*** (0.03798)
Year FE	YES	YES	YES	YES	YES	YES
Journal FE	YES	YES	YES	YES	YES	YES
N	188,181	188,181	188,181	188,181	188,181	188,181
Mean of DV	0.1743	0.2676	0.1743	0.2676	0.1743	0.2676

Note: This table presents estimates from OLS models evaluating the average increase in policy citations for paper that are empirical, use FSRDC data, or cite other FSRDC papers, respectively. \*, \*\*, \*\*\* denote significance at 5%, 1% and 0.1% level respectively. Standard errors are clustered by year of publication.

# Share of Articles Receiving Citations from Policy Documents by Field



# Effect of FSRDC Access on the Policy Impact of Economic Research

	Count Papers Cited by Policy		
	All (4)	US Only (5)	Non-US Only (6)
Post-FSRDC	-0.0187* (0.008)	-0.0157** (0.006)	-0.0112 (0.007)
Post-FSRDC $\times$ Empiricist	0.0486*** (0.008)	0.0392*** (0.006)	0.0319*** (0.006)
Researcher FE	Yes	Yes	Yes
University $\times$ Year FE	Yes	Yes	Yes
N	245556	245556	245556

► Back

# Effect of FSRDC Access on the Policy Orientation of Empirical Researchers

	Mentions of Policy (1)	Policy JELs (all) (2)	Policy JELs (subset) (3)
Post-FSRDC	-0.00974 (0.008)	-0.000619 (0.003)	0.00379** (0.001)
Post-FSRDC × Empiricist	0.0118 (0.007)	0.00245 (0.002)	-0.00166 (0.001)
Researcher FE	Yes	Yes	Yes
University × Year FE	Yes	Yes	Yes
N	245556	245556	245556

► Back

# Average Rank of Treated Institutions

