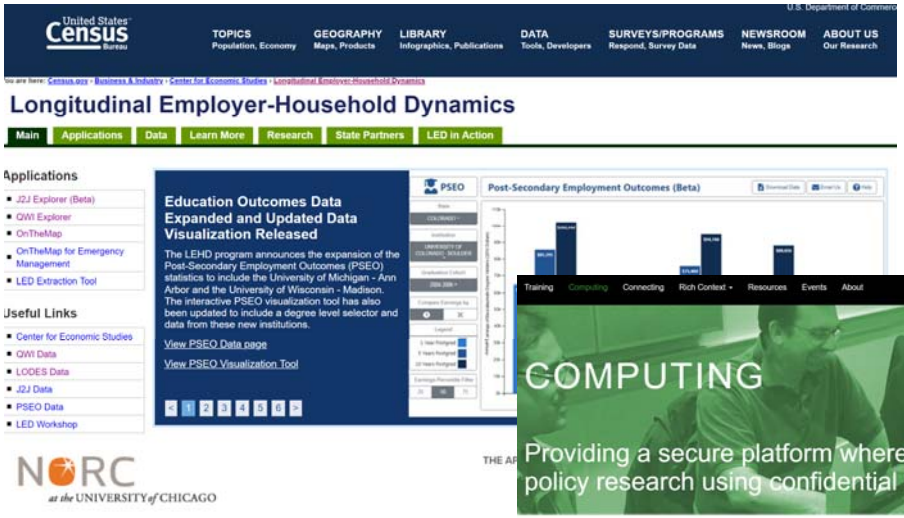
Implications

Julia Lane

NYU

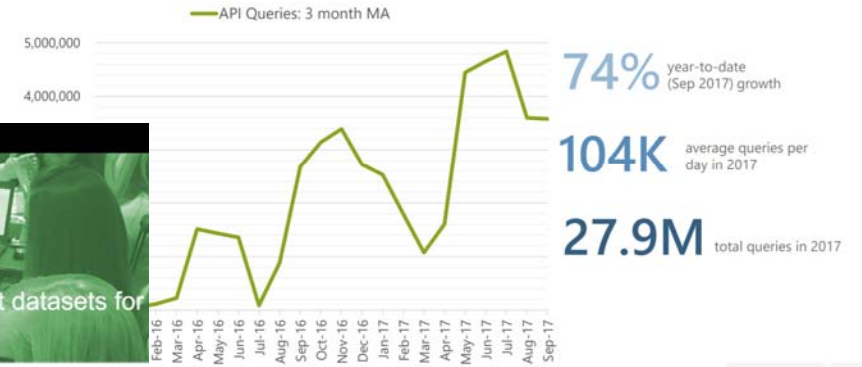


Home > Our Work > Services and Solutions > Data Enclave

Data Enclave

NORC is a recognized innovator in the field of secure data management, analysis, and dissemination. The NORC Data Enclave®—NORC's constantly evolving data storage and management solution—is an integral part of the **Advanced Data Solutions Center**. The Enclave's high performance computing environment includes cutting-edge statistical, analytical, visualization, and reporting tools.

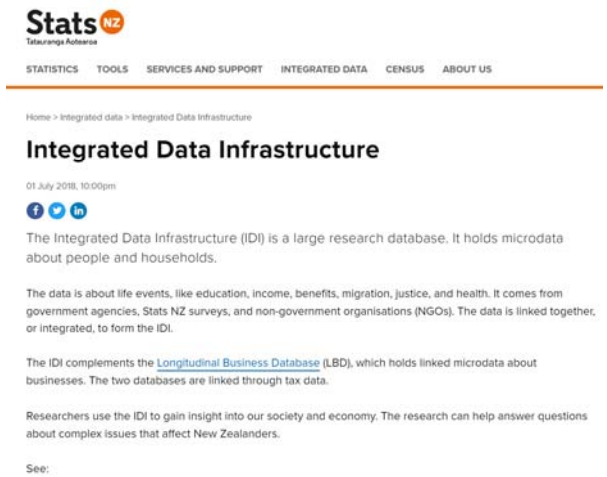
Since 2006, state and federal agencies, research institutes, foundations, and universities have used the Enclave to securely house and provide remote access to confidential data. Enclave-based research informs a wide spectrum of public and private sector decision-making, as well as journal articles, books, position papers, conference presentations, dissertations, etc. At any given time, the Enclave supports over 1000 researchers via contracts and grants with a wide variety of government, academic, nonprofit, and commercial clients.

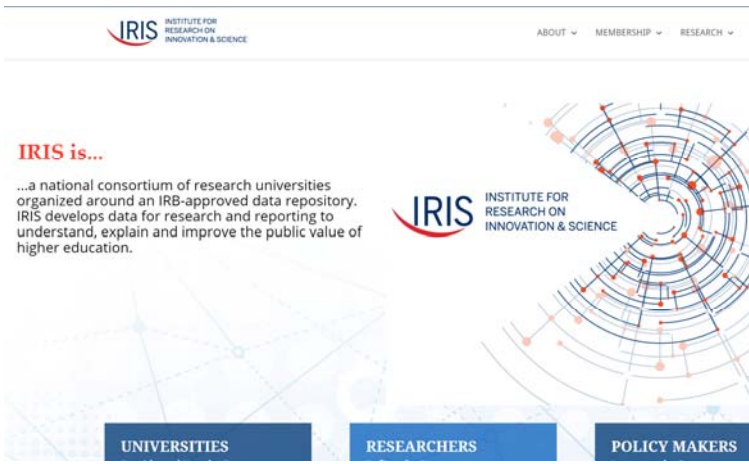


ADRF

The **Administrative Data Research Facility** (ADRF) provides a secure platform to host confidential micro-data. The ADRF is designed to promote collaboration, facilitate documentation, and provide information about use to data stewards. See our whitepapers for a more thorough description of ADRF components.

The ADRF has hosted almost 50 confidential government datasets from 12 different agencies at all levels of government. The ADRF has received Authorization to Operate from the Census Bureau, has achieved a FedRAMP Moderate approval and is listed on the FedRAMP Marketplace.





Overview

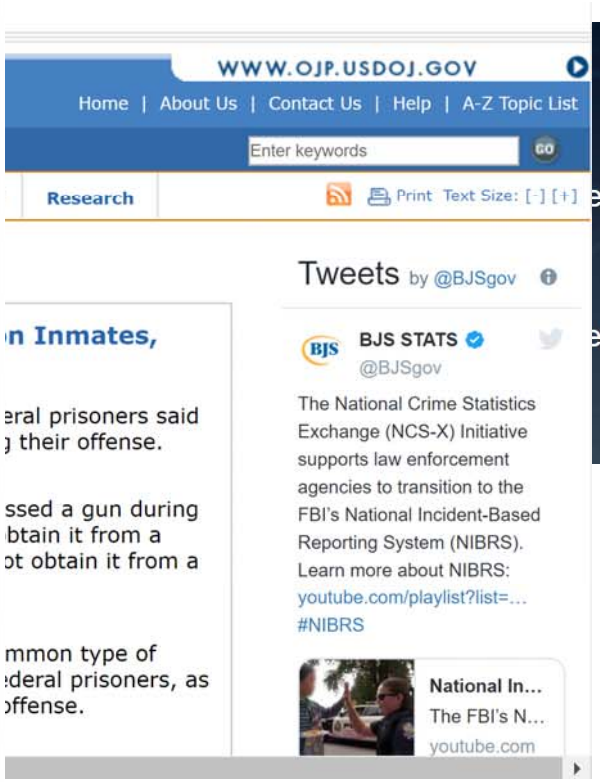
- Three issues for statistical agencies
  - Remember core mission
  - Be realistic
  - Build ecosystem

- Three areas for future investment
  - Utility measures
  - Access
  - Output protection

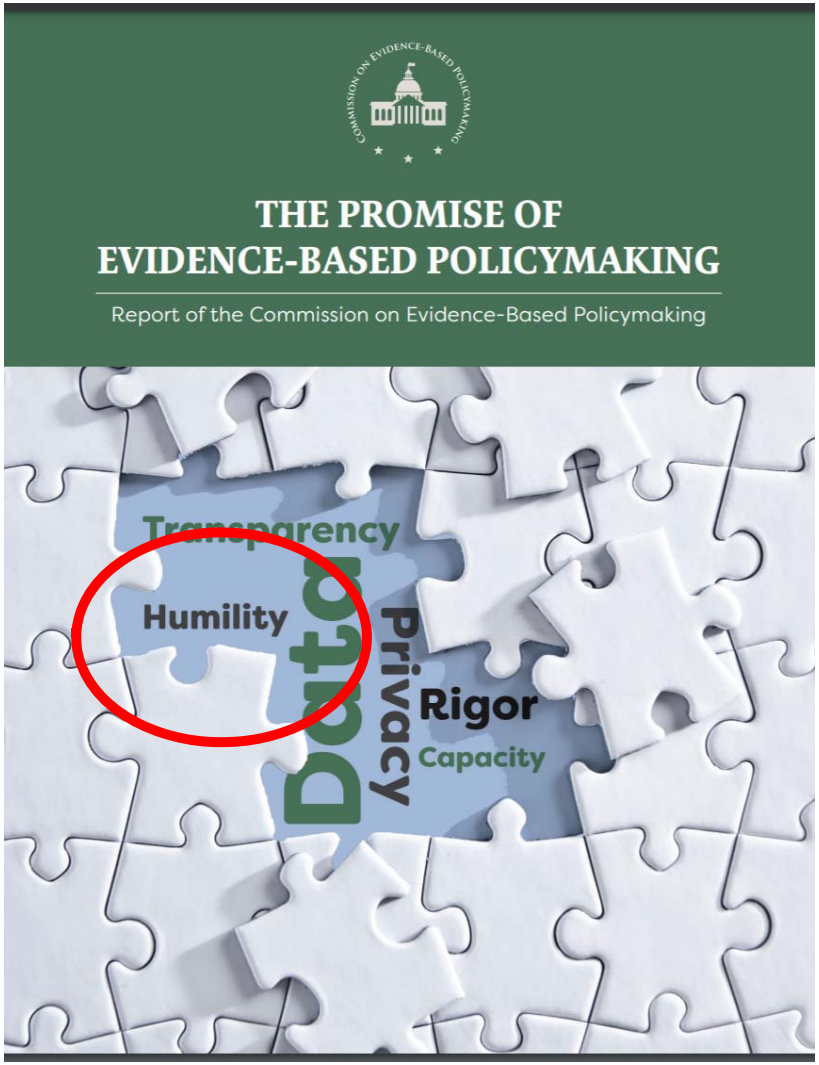
Overview

- Three issues for statistical agencies
  - Remember core mission
  - Be realistic
  - Build ecosystem
- Three areas for future investment
  - Utility measures
  - Access
  - Output protection

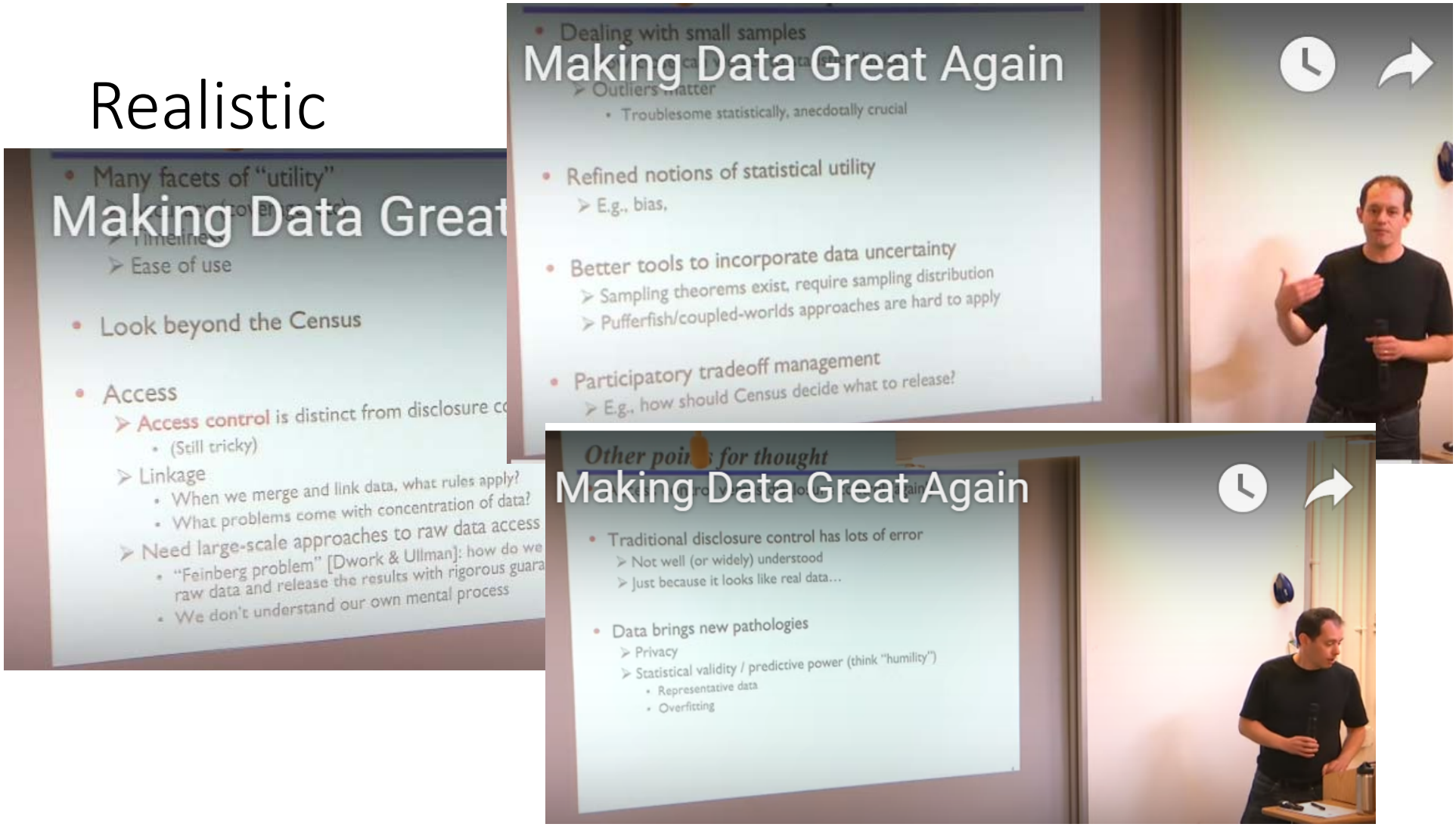
Core Mission



Be realistic



Realistic


Making Data Great
Many facets of "utility"
Ease of use
Look beyond the Census
Access
Access control is distinct from disclosure c
(Still tricky)
Linkage
When we merge and link data, what rules apply?
What problems come with concentration of data?
Need large-scale approaches to raw data access
"Feinberg problem" [Dwork & Ullman]: how do we
raw data and release the results with rigorous guara
We don't understand our own mental process
Making Data Great Again
Dealing with small samples
Outliers matter
Troublesome statistically, anecdotally crucial
Refined notions of statistical utility
E.g., bias,
Better tools to incorporate data uncertainty
Sampling theorems exist, require sampling distribution
Pufferfish/coupled-worlds approaches are hard to apply
Participatory tradeoff management
E.g., how should Census decide what to release?
Making Data Great Again
Other poin for thought
Traditional disclosure control has lots of error
Not well (or widely) understood
Just because it looks like real data…
Data brings new pathologies
Privacy
Statistical validity / predictive power (think "humility")
Representative data
Overfitting

Areas for future research

Understand the data

- Understand the structure of the core data – Census is a special case
- Consider sub samples
- Consider outliers

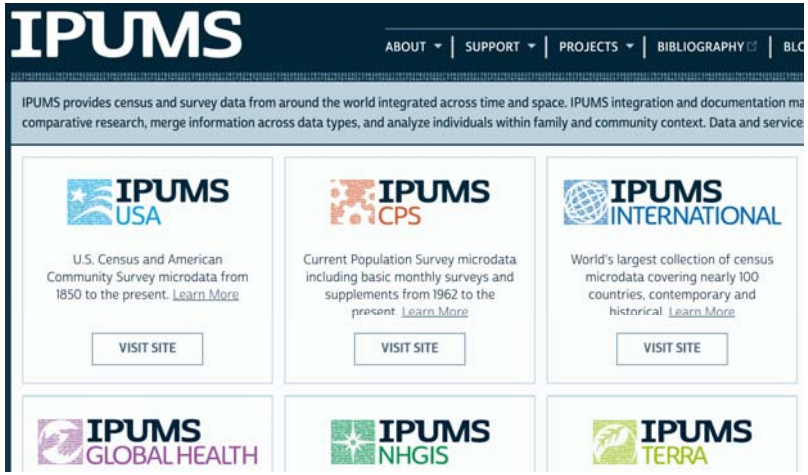
Incorporate utility

- Explicitly incorporate measures of utility in determining privacy budget
- Address more complicated variable structures in a way that preserves a sufficient amount of data utility

General comments

- Explicitly incorporate size of sample into analysis
- Don’t treat all fields equally

Build ecosystem



Overview

- Three issues for statistical agencies
  - Remember core mission
  - Be realistic
  - Build ecosystem

- Three areas for future investment
  - Utility measures
  - Access
  - Demonstrate value and move budget constraint

Context

EXECUTIVE OFFICE OF THE PRESIDENT
OFFICE OF MANAGEMENT AND BUDGET
WASHINGTON, D.C. 20503

June 4, 2019

M-19-18

MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES

FROM: Russell T. Vought
Acting Director

SUBJECT: Federal Data Strategy - A Framework for Consistency

- *Building a Culture that Values Data and Promotes Public Use (practices 1-10)*
- *Governing, Managing, and Protecting Data (practices 11-26)*
- *Promoting Efficient and Appropriate Data Use (practices 27-39)*

Utility (and privacy) challenges

The new world

- Datasets are not well defined entities
- Data are dirty and badly documented
- Missingness is common

Access is feature not a bug

- Linkages are error-prone and often wrong
- Outliers are where the action is
- New products are created through trial and error
- Community engagement and use is the core business model

Utility measures (examples)

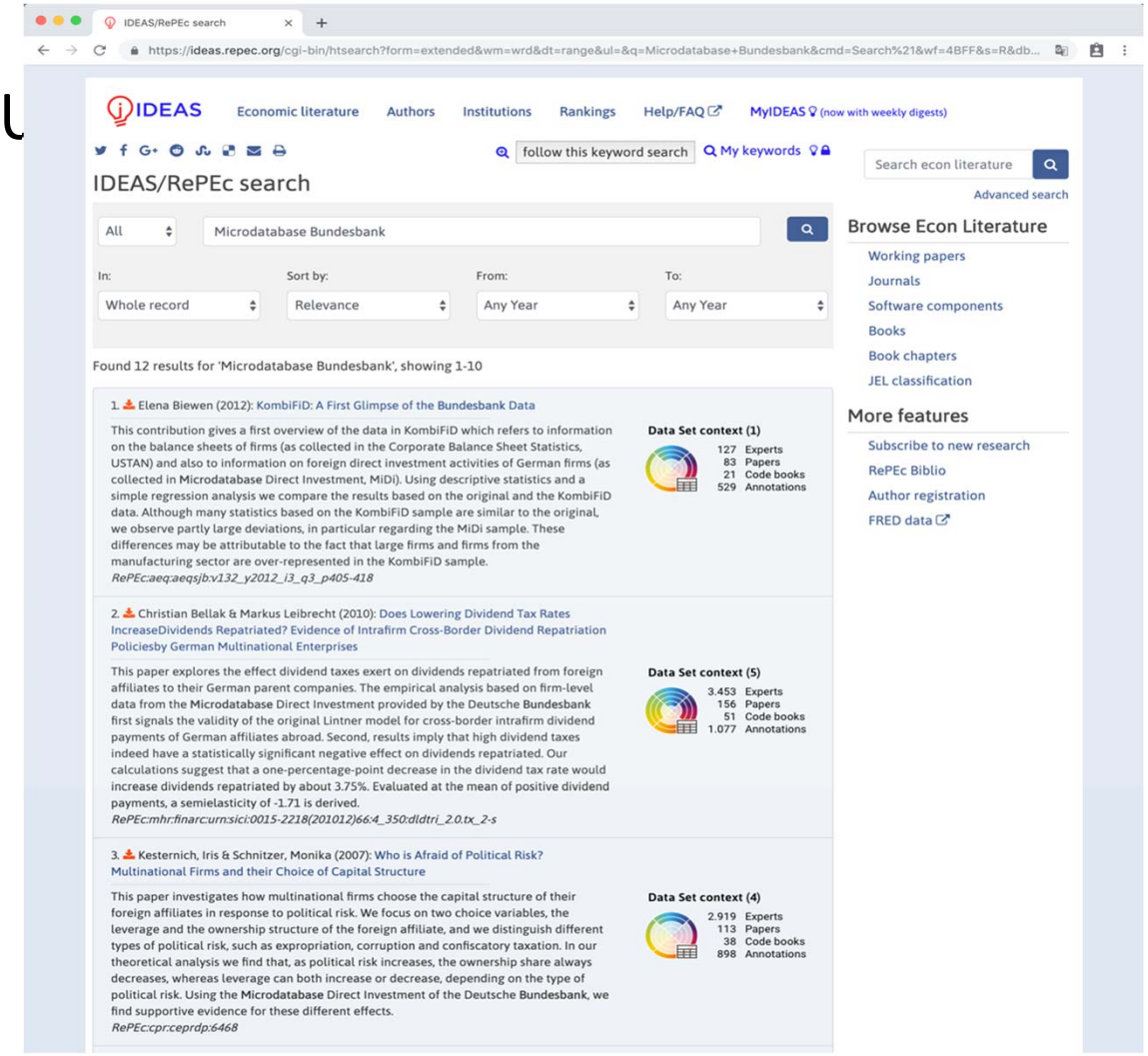
- Completeness – data rich enough
- Timeliness – data deliveries adhere to schedules
- Relevance – data satisfy user needs
- Accessibility – access to data is user friendly
- Interpretability – documentation; meta-data
- Granularity – data detailed enough
- Value - Number of ways data are put to use
- Cost-effectiveness – value for money

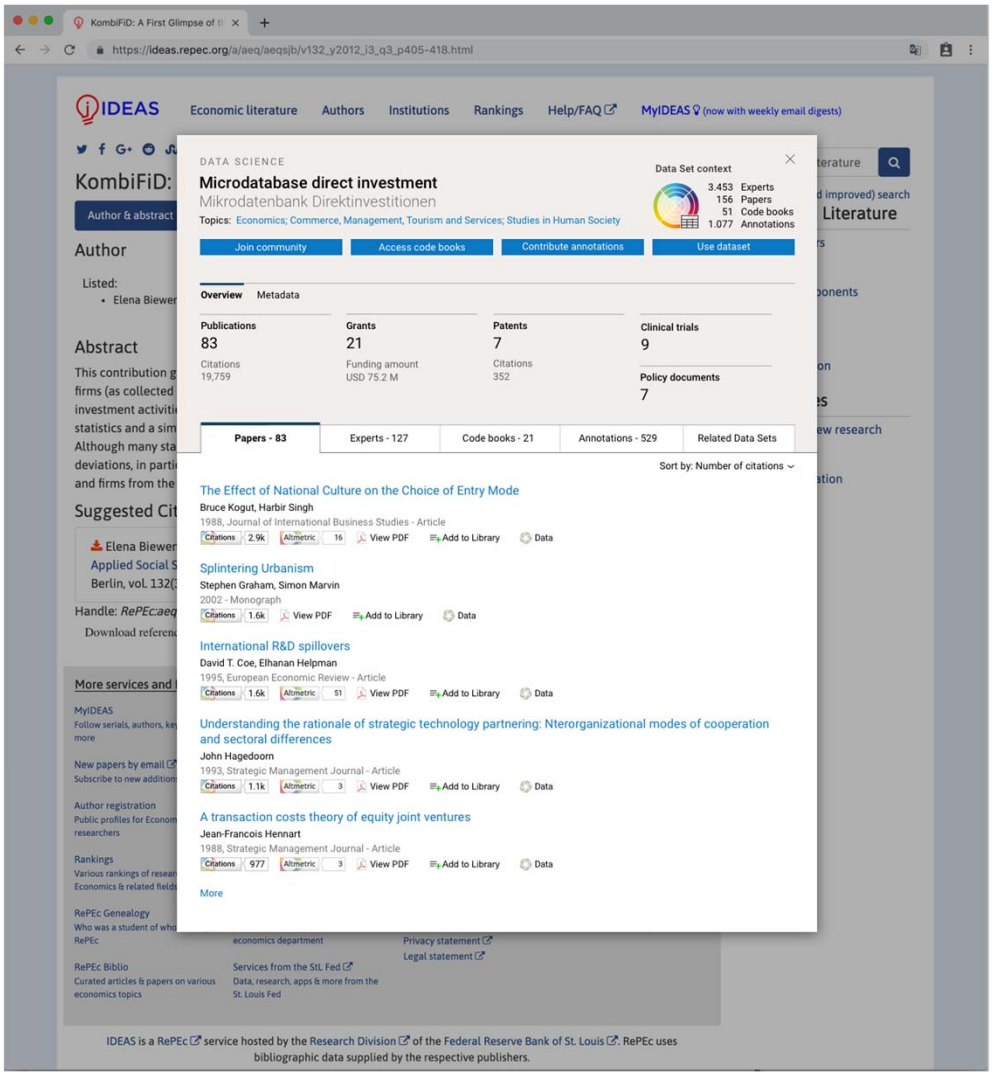
Biemer, P. (2017) Errors and Inference, Chapter 10 in BD and Social Science

Access

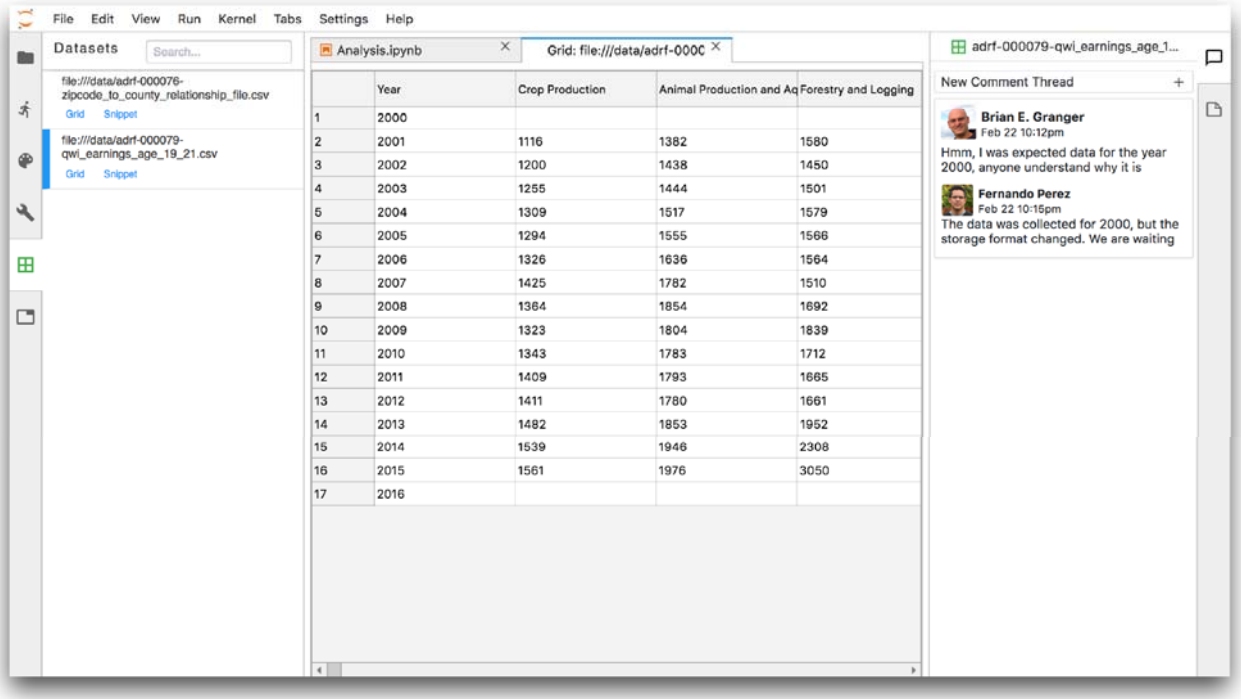
- Safe People
  - Tiers of approved and trained researchers; tiers of legal controls
- Safe Projects
  - Tiers of approved projects, consistent with agency mission and utility
- Safe Settings
  - Tiers of secure environments
- Safe Data
  - Tiers of deidentified data
- Safe Outputs
  - Disclosure reviews and export controls

Move bu


IDEAS/RePEc search
https://ideas.repec.org/cgi-bin/htsearch?form=extended&wm=wrd&dt=range&ul=&q=Microdatabase+Bundesbank&cmd=Search%21&wf=4BFF&s=R&db...
IDEAS
Economic literature
Authors
Institutions
Rankings
Help/FAQ
MyIDEAS (now with weekly digests)
follow this keyword search
My keywords
Search econ literature
Advanced search
IDEAS/RePEc search
All
Microdatabase Bundesbank
In:
Whole record
Sort by:
Relevance
From:
Any Year
To:
Any Year
Found 12 results for 'Microdatabase Bundesbank', showing 1-10
1. Elena Biewen (2012): KombiFiD: A First Glimpse of the Bundesbank Data
This contribution gives a first overview of the data in KombiFiD which refers to information on the balance sheets of firms (as collected in the Corporate Balance Sheet Statistics, USTAN) and also to information on foreign direct investment activities of German firms (as collected in Microdatabase Direct Investment, MiDi). Using descriptive statistics and a simple regression analysis we compare the results based on the original and the KombiFiD data. Although many statistics based on the KombiFiD sample are similar to the original, we observe partly large deviations, in particular regarding the MiDi sample. These differences may be attributable to the fact that large firms and firms from the manufacturing sector are over-represented in the KombiFiD sample.
RePEc:aeq:aeqsjb:v132_y2012_i3_q3_p405-418
Data Set context (1)
127 Experts
83 Papers
21 Code books
529 Annotations
2. Christian Bellak & Markus Leibrecht (2010): Does Lowering Dividend Tax Rates IncreaseDividends Repatriated? Evidence of Intrafirm Cross-Border Dividend Repatriation Policiesby German Multinational Enterprises
This paper explores the effect dividend taxes exert on dividends repatriated from foreign affiliates to their German parent companies. The empirical analysis based on firm-level data from the Microdatabase Direct Investment provided by the Deutsche Bundesbank first signals the validity of the original Lintner model for cross-border intrafirm dividend payments of German affiliates abroad. Second, results imply that high dividend taxes indeed have a statistically significant negative effect on dividends repatriated. Our calculations suggest that a one-percentage-point decrease in the dividend tax rate would increase dividends repatriated by about 3.75%. Evaluated at the mean of positive dividend payments, a semielasticity of -1.71 is derived.
RePEc:mhr:finarc:urn:sici:0015-2218(201012)66:4_350:dldtri_2.0.tx_2-s
Data Set context (5)
3.453 Experts
156 Papers
51 Code books
1.077 Annotations
3. Kesternich, Iris & Schnitzer, Monika (2007): Who is Afraid of Political Risk? Multinational Firms and their Choice of Capital Structure
This paper investigates how multinational firms choose the capital structure of their foreign affiliates in response to political risk. We focus on two choice variables, the leverage and the ownership structure of the foreign affiliate, and we distinguish different types of political risk, such as expropriation, corruption and confiscatory taxation. In our theoretical analysis we find that, as political risk increases, the ownership share always decreases, whereas leverage can both increase or decrease, depending on the type of political risk. Using the Microdatabase Direct Investment of the Deutsche Bundesbank, we find supportive evidence for these different effects.
RePEc:cpr:ceprdp:6468
Data Set context (4)
2.919 Experts
113 Papers
38 Code books
898 Annotations
Browse Econ Literature
Working papers
Journals
Software components
Books
Book chapters
JEL classification
More features
Subscribe to new research
RePEc Biblio
Author registration
FRED data


KombiFiD: A First Glimpse of t
https://ideas.repec.org/a/aeq/aeqsjb/v132_y2012_j3_q3_p405-418.html
IDEAS
Economic literature
Authors
Institutions
Rankings
Help/FAQ
MyIDEAS (now with weekly email digests)
KombiFiD:
Author & abstract
Author
Listed:
Elena Biewer
Abstract
This contribution g
firms (as collected
investment activiti
statistics and a sim
Although many sta
deviations, in parti
and firms from the
Suggested Cit
Elena Biewer
Applied Social S
Berlin, vol. 132(3
Handle: RePEc:aeq
Download referenc
More services and
MyIDEAS
Follow serials, authors, ke
more
New papers by email
Subscribe to new addition
Author registration
Public profiles for Econom
researchers
Rankings
Various rankings of resear
Economics & related fields
RePEc Genealogy
Who was a student of who
RePEc
RePEc Biblio
Curated articles & papers on various
economics topics
economics department
Services from the StL Fed
Data, research, apps & more from the
St. Louis Fed
Privacy statement
Legal statement
IDEAS is a RePEc service hosted by the Research Division of the Federal Reserve Bank of St. Louis. RePEc uses
bibliographic data supplied by the respective publishers.
DATA SCIENCE
Microdatabase direct investment
Mikrodatenbank Direktinvestitionen
Topics: Economics; Commerce, Management, Tourism and Services; Studies in Human Society
Data Set context
3.453 Experts
156 Papers
51 Code books
1.077 Annotations
Join community
Access code books
Contribute annotations
Use dataset
Overview
Metadata
Publications
83
Citations
19,759
Grants
21
Funding amount
USD 75.2 M
Patents
7
Citations
352
Clinical trials
9
Policy documents
7
Papers - 83
Experts - 127
Code books - 21
Annotations - 529
Related Data Sets
Sort by: Number of citations
The Effect of National Culture on the Choice of Entry Mode
Bruce Kogut, Harbir Singh
1988, Journal of International Business Studies - Article
Citations 2.9k
Altmetric 16
View PDF
Add to Library
Data
Splintering Urbanism
Stephen Graham, Simon Marvin
2002 - Monograph
Citations 1.6k
View PDF
Add to Library
Data
International R&D spillovers
David T. Coe, Elhanan Helpman
1995, European Economic Review - Article
Citations 1.6k
Altmetric 51
View PDF
Add to Library
Data
Understanding the rationale of strategic technology partnering: Nterorganizational modes of cooperation and sectoral differences
John Hagedoorn
1993, Strategic Management Journal - Article
Citations 1.1k
Altmetric 3
View PDF
Add to Library
Data
A transaction costs theory of equity joint ventures
Jean-Francois Hennart
1988, Strategic Management Journal - Article
Citations 977
Altmetric 3
View PDF
Add to Library
Data
More

Commenting and Annotation in JupyterLab



*** Early prototype**

Summary

- Three issues for statistical agencies
  - Remember core mission
  - Be realistic
  - Build ecosystem

- Three areas for future investment
  - Utility measures
  - Access
  - Output protection

Comments welcome

- Julia.lane@nyu.edu
- https://coleridgeinitiative.org