



**INSTITUTE FOR SOCIAL RESEARCH**  
UNIVERSITY OF MICHIGAN

# CenHRS: Census-Enhanced Health and Retirement Study

NBER Summer Institute 2021 Aging Workshop

July 23, 2021

# Disclaimer

Any opinions and conclusions expressed herein are those of the authors and do not represent the views of the Federal Reserve Bank of Boston, the Board of Governors of the Federal Reserve System, or the U.S. Census Bureau.

All results have been reviewed to ensure that no confidential information is disclosed (CBDRB-FY21-CED006-0019).

The CenHRS is supported by the Alfred P. Sloan Foundation through the CenHRS project at the University of Michigan with additional support from the Michigan Node of the NSF-Census Research Network (NCRN) under NSF SES 1131500. The Health and Retirement Study is supported by NIA and SSA.

# CenHRS project

## Research Team

- Co-PIs: Maggie Levenstein, John Abowd, Kristin McCue, Matthew Shapiro, David Weir
- Michigan, Cornell, and Census faculty, staff, and graduate students
- Co-authors of linkage paper: John Abowd, Joelle Abramowitz, Margaret Levenstein, Kristin McCue, Dhiren Patki, Trivellore Raghunathan, Ann Rodgers, Matthew D. Shapiro, Nada Wasi, Dawn Zinsser

# What is CenHRS?

*CenHRS integrates Health and Retirement Study (HRS) and the U.S. Census Bureau's Business Register (BR)*

- Link to HRS respondents' employers and co-workers
- Labor supply from survey; labor demand from administrative data
- What leads/enables/requires that people work longer as they age?

# Why is CenHRS important?

## **Creating infrastructure to study older people at work**

- HRS respondents' households in context
  - Labor demand and supply
  - Health and quality of life
- Effects of firm dynamics on workers
  - Technology, imports, public policy
  - Effects on employment, wages, health and well-being
- Peer and cohort effects
  - Wage and age distribution
  - Retirement norms

## CenHRS: Two data products

### Crosswalk between HRS respondents' job and Census BR

- Access within FSRDC
- Requires approved Title 13/Title 26 project
- Crosswalk allows linkage to Census business data
- Available this year; apply now

### Enhanced HRS

- Employer and co-worker characteristics
- Aim: Disclosable statistics
- Aim: Access under HRS restricted-data protocols

*Research within FSRDC to create relevant, disclosable variables to enhance HRS with employer information*

# What is HRS?

## **Longitudinal study of households with at least one person over the age of 50**

- Panel of ~25,000 people surveyed every two years
- Started in 1992
  - New cohorts added every six years
  - Measures biology and health, psychosocial well being, income and wealth, consumption, employment
  - Links to a number of other data resources including Social Security and private pension plan records

# What is the BR?

## **Annual data on essentially every employer in country**

- About 10 million establishments per year
- Begins in 1975
- Each record includes:
  - Number of employees, wage bill, NAICS, detailed geography
  - Sales, other variables depending on year and type of business
  - Employer Identification Number (EIN)
  - Mapping of establishment to firm
  - Links to all Census data on businesses



# The Business Register links to many other data sources

- **Economic Censuses and Surveys**

- Output and products
- Inputs and costs
- Technology
- Firm organization

- **Longitudinal Employer Household Dynamics (LEHD)** provides information on characteristics of coworkers

- Earnings
- Demographics
- Tenure
- Turnover

- **Business Administrative Data**

- Import and export transactions
- Patents
- Pensions and benefits
  - Form 5500
  - Medical Expenditure Panel Survey - Insurance Component

- **Other Sources**

- Compustat and Execucomp
- Toxic Release Inventory
- LexisNexis Directory of Corporate Affiliations

# Creating the CenHRS

Challenge: Linking data from different frames (households and businesses) where distribution of employment is very skewed

~ 6 million employers

~ 0.3% of firms employ 53% of all workers

Separately for records with EINs and for records with no EINs but with employer names and addresses:

1. Create ground truth (training data)
2. Train model: Use machine learning techniques to estimate posterior probability of match of HRS job with BR employer, within block
3. Multiply impute (10 implicates) to incorporate uncertainty arising from the matching process

*Paper available shortly*

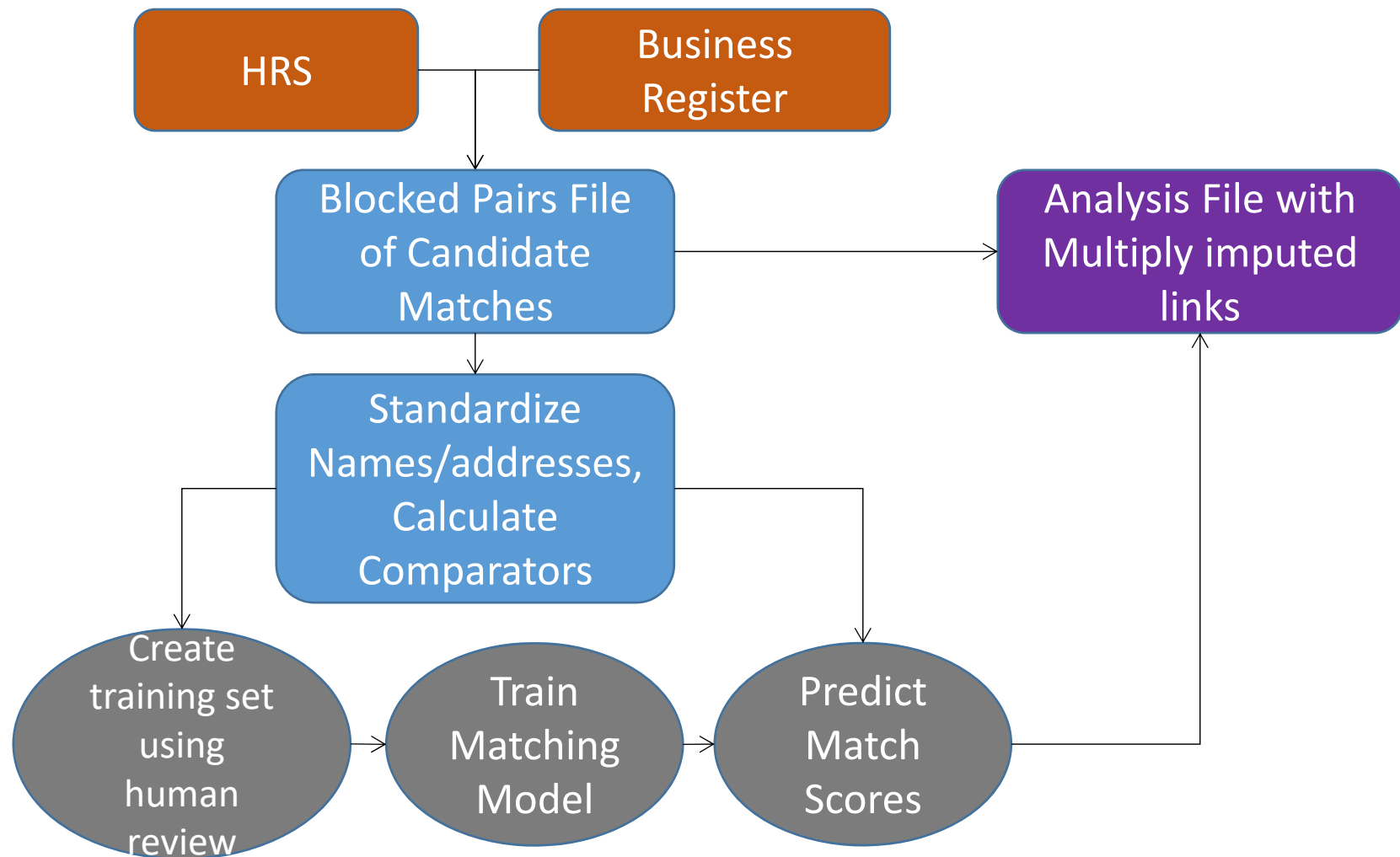
# If no EIN, finding needles in haystack

	HRS jobs	Share of jobs	BR candidates per HRS respondent (average number of establishments)
Deterministic match, EIN available	2500	0.415	1
Probabilistic match, EIN available	1800	0.295	435
Probabilistic match, EIN not available	1800	0.291	30,050

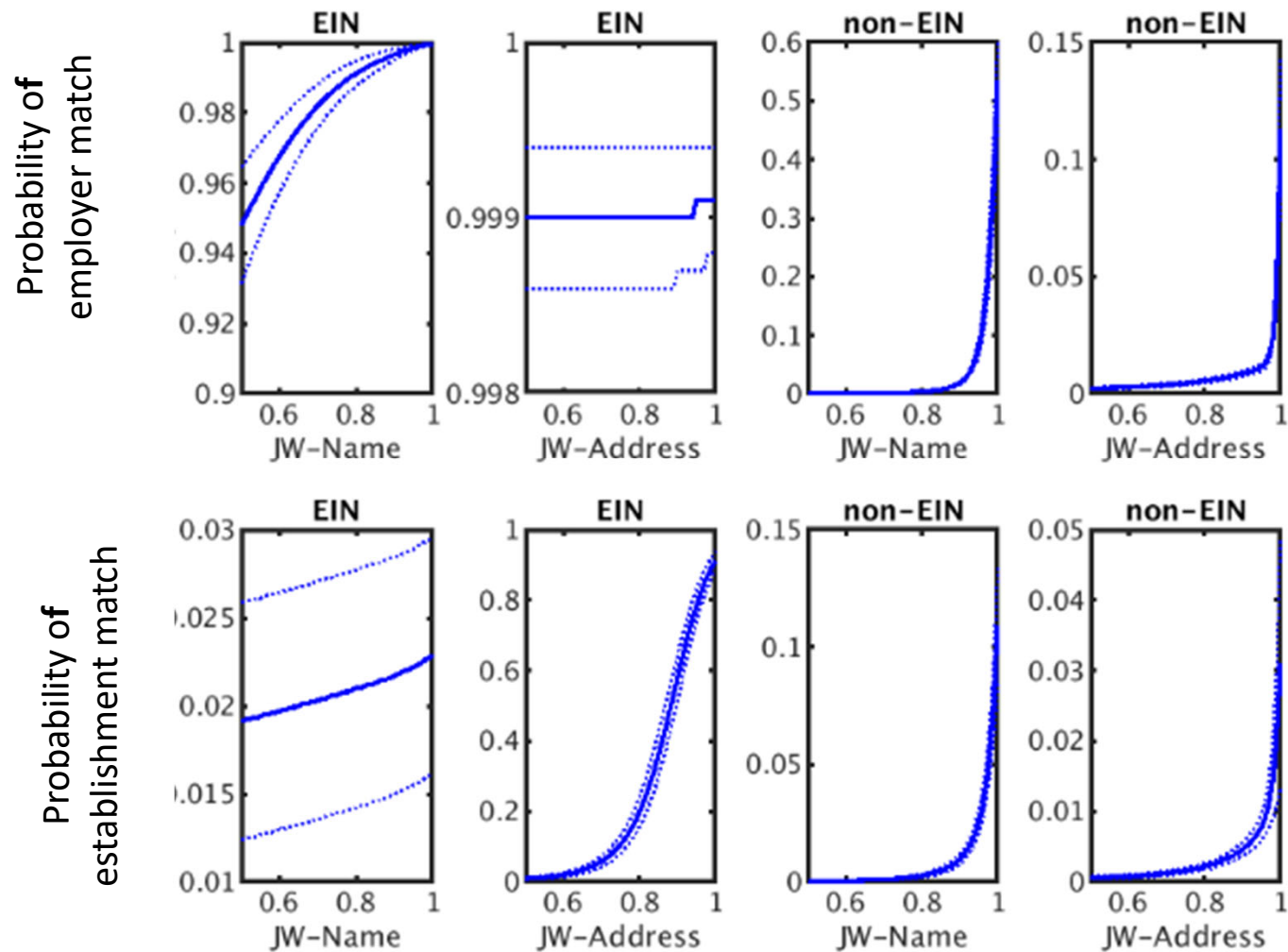
Note: 2010 wave of CenHRS

Shares do not sum to 1 because each cell independently rounded

# Linkage machine learning process



# Selected partial effects of matching models



Variables are Jaro-Winkler (JW) scores for name and address.  
 Note differences in vertical axis scales.

# Data products and access protocols

## HRS-BR crosswalk

- Restricted access in FSRDCs (available late 2021)

## HRS enhanced with employer/co-worker variables

- Restricted access under HRS protocols (under development)
- Possible variables
  - Firm characteristics
    - Categorical measures of firm age, industry, size (employment, sales), productivity, and growth
    - Ranking of firm within industry wage distribution
    - Indicators of mass layoffs and hiring freezes
    - Imports and exports
  - Co-worker characteristics
    - Distribution of age, sex, race, foreign born, tenure, and earnings
  - Location
    - Length of commute

# CenHRS restricted data approval process

- Researchers will need to request access to the crosswalk as well as any HRS and Census data they would like to link
- Any proposals requesting to use the crosswalk requires Census, IRS, CenHRS, and HRS approval
- Before beginning the application process, researchers should reach out to the CenHRS team ([CenHRSinfo@umich.edu](mailto:CenHRSinfo@umich.edu))
  - After consulting with the CenHRS team, researchers should submit an inquiry expressing interest via the ResearchDataGov portal to begin discussion with an FSRDC administrator to develop a proposal
  - HRS maintains its own portal for applying for access to restricted HRS data: researchers will need to concurrently set up a portal account and apply via the HRS portal; discussions are underway to streamline this process through the ResearchDataGov portal

Thank You!

Questions?

[CenHRSinfo@umich.edu](mailto:CenHRSinfo@umich.edu)

CenHRS website: <https://cenhrs.isr.umich.edu/>



# Census-Enhanced Health and Retirement Study Restricted Data Crosswalk

## Application Procedures

July 19, 2021

A beta version of the Census-Enhanced Health and Retirement Study (CenHRS) restricted data crosswalk will soon be made available to the research community through Federal Statistical Research Data Centers (FSRDCs). With required approvals, researchers will be able to conduct research and construct new variables using the full scope of Census economic and business data.

Before beginning the application process, researchers should reach out to the CenHRS team (CenHRSinfo@umich.edu or Joelle Abramowitz at jabramow@umich.edu).

After consulting with the CenHRS team, researchers should submit an inquiry expressing interest via the ResearchDataGov portal to begin discussion with an FSRDC administrator to develop a proposal. Interested researchers will ultimately need to submit a proposal to the Census Bureau via the ResearchDataGov portal. Proposals should request access to the CenHRS crosswalk as well as any HRS and Census data they would like to link.

Any proposals requesting to use the crosswalk will need Census, IRS, CenHRS, and HRS approval.

Proposals submitted via the ResearchDataGov portal will be forwarded to Census, IRS, and CenHRS for review. HRS maintains its own portal for applying for access to restricted HRS data: researchers will need to concurrently set up a portal account and apply via the HRS portal. Discussions are underway to streamline this process through the ResearchDataGov portal.

Once the proposal is approved by the Census Bureau, researchers will need to obtain Special Sworn Status (SSS).