

# The People Behind the Enterprises: The Use of Administrative Records in Business Owner Statistics

FedCASIC – April 22, 2026

Vitaliy Novik<sup>1</sup> (presenter)

Adela Luque<sup>1</sup>, Valeska Araujo<sup>1,2</sup>, John Earle<sup>1,2</sup>, Lorenz Eckerdt<sup>1,4</sup>, Nick LaBerge<sup>1</sup>, James Noon<sup>1</sup>, Jared Wold<sup>1,2</sup>, and Sammy Young<sup>1,3</sup>

*1 – U.S. Census Bureau*

*3 – Arizona State University*

*2 – George Mason University*

*4 – Stony Brook University*

# Disclaimer

*Any opinions and conclusions expressed herein are those of the author(s) and do not reflect the views of the U.S. Census Bureau. The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data (Project No. P-7504866, Disclosure Review Board (DRB) approval number: CBDRB-FY25-ESMD010-012).*

# Outline

**Background**

**Methodology**

**Results**

- Microdata example
- Coverage
- QA - Comparing adrec & survey based demographics

**Illustration:** Where do exiting employer owners go (by age of owner)?

**Challenges**

**Looking forward**

# Concept & Background

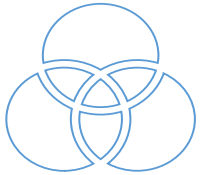
- Use administrative records (AdRec) & census data to assign demo & other characteristics to universe\* of employer business owners & their businesses
  - ✓ Produce adrec-based business demographic/other characteristics microdata & estimates
    - E.g., What share of businesses are female-owned? Has the average age of startups owners decreased over time? What share of employers started as non-employers & does that vary by owner characteristics?
  - ✓ Employer demographics currently estimated via survey (Annual Business Survey)
  - ✓ Adrec-based demographics already accomplished for non-employer firms (Nonemployer Statistics by Demographics, [NES-D](#))
  - ✓ Current work: Can it be done for employers...? Yes!

# Preview of Benefits & Uses



## **New adrec-based microdata & estimates:**

- Owner- and business-level microdata with demo/other characteristics for universe of employer firms in U.S. (annual series linkable over time)
- Business demographic & other characteristics estimates
- New paths to entrepreneurship research (e.g., owners background/history)



## **Data/Program/Survey improvements, such as:**

- *Unburdening* surveys (Annual Business Survey)
- *Enhancing* business data, sampling frames & programs with demo/other characteristics
- More granularity and flexibility



## **Reduce costs**



## **Facilitate Linkages:**

- Providing owner-to-business links
- Ability to link to external & internal business and/or individual data
  - Linking to other data by person or business identifier (e.g., PIK, EIN)

# In-Scope Legal Forms

	EINs (thousands)	Employment (millions)	Firms (thousands)
Total	6,120	117.4	5,974

	Column %		
Partnerships	13%	14%	13%
Sole Proprietorships	13%	4%	13%
S Corporations	58%	31%	59%
C Corporations	16%	50%	15%

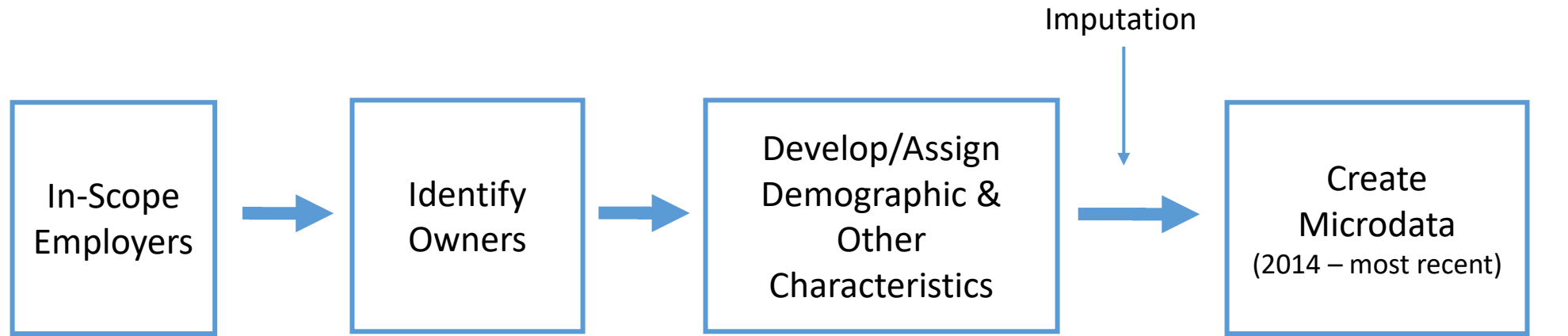
} **Able to assign demo/char.\***

Sources: 2022 CBP and 2022 CBPBR.

Notes: Legal forms include for-profit private firms, including publicly-traded.

Totals may not sum to 100% due to rounding.

# Process Overview



Business Register	Schedule Cs (Sole Props)	Decennial ACS
	Schedule K-1 (Partnerships/S-corps)	SSA Numident
	SS-4 (EIN appl) form	Veteran Affairs
		DOD (upcoming)
		Tax data
		Other non-tax adreces
		3 <sup>rd</sup> party

Business-Owner level  
Business level

Primary Data Sources



# Resulting Microdata – Owner & Business w/ Charact.

## Owner-Business Level File

- Includes owner identifiers (PIK\*), each owner’s demographics, etc.



Owner	Firm	EIN	Ownership Share	Sex	Vet	...
PIK1	A1	E1	60%	F	Y	
PIK2	A1	E1	40%	M	N	
PIK3	A2	E2	100%	M	N	

## Firm Level File

- Includes summarized business-level demo & other characteristics
- Follows NES-D/ABS methodology



Firm	EIN	Sex	Vet	....
A1	E1	F	Y	
A2	E2	M	N	

## Complex cases

- Some complex parent firms have under their “ownership umbrella” more than 1 distinct business (i.e., EIN), each with their own set of owners
- For these cases, we develop a methodology that “rolls-up” the characteristics of the “child” businesses into the parent level
- Lots of documentation & flags in microdata to provide flexibility to users to create their own parent-level roll-up

# Adrec Coverage

## Sole proprietorships

- Approx. 97%
- Demographic & other characteristics imputed for the rest (3%)

## S-corporations

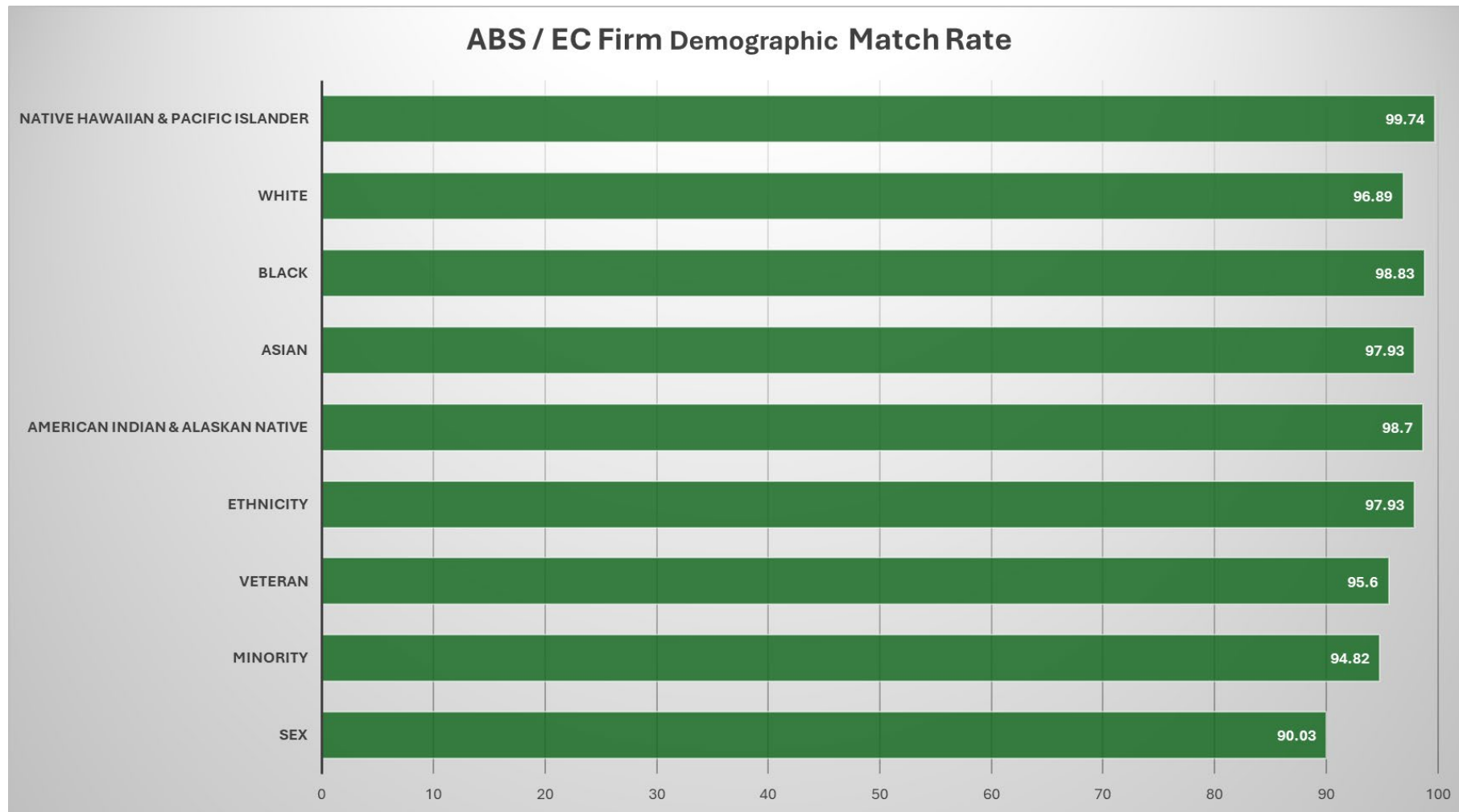
- Approx. 96%
- Characteristics imputed for the rest (4%)

## Partnerships

- Approx. 79%
- Characteristics imputed for the rest (21%)

➔ Overall, business-level demographics will be imputed for approx. 7% of sole props, partnerships & S-corps

# High Agreement With ABS Demographics

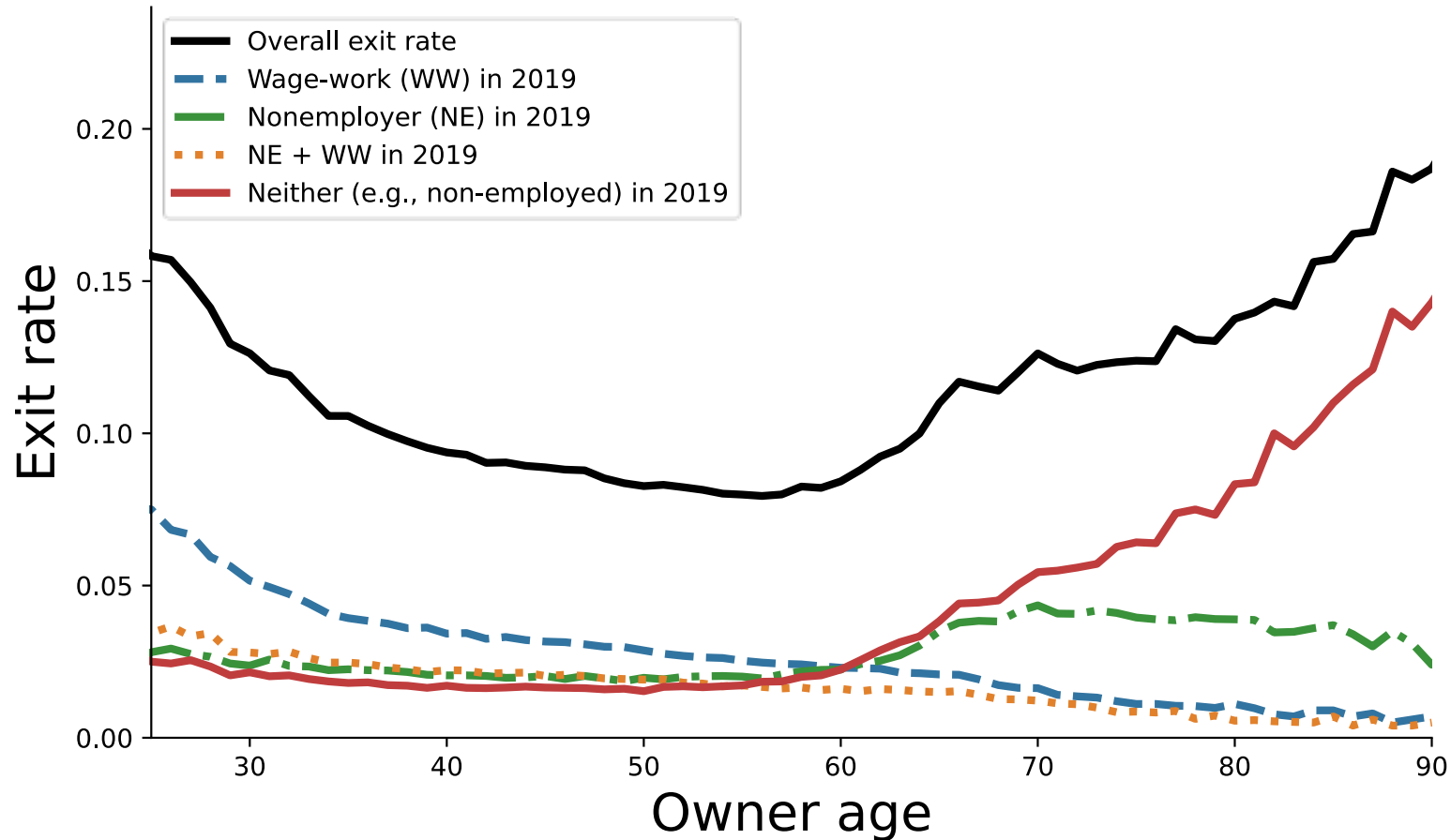


Note: Compares adres-based demographics to the same firm's reported (not imputed) demographics found in the corresponding ABS reference year. Figure shows the number of (rounded) demographic matches over the (rounded) number of Total Firm ID matches. ABS 2017-2021.

# Owner Exit Rate from Employer Ownership by Age of Owner

2018 employer exit rates, by owner age and y+1 income source

By Number of Owners

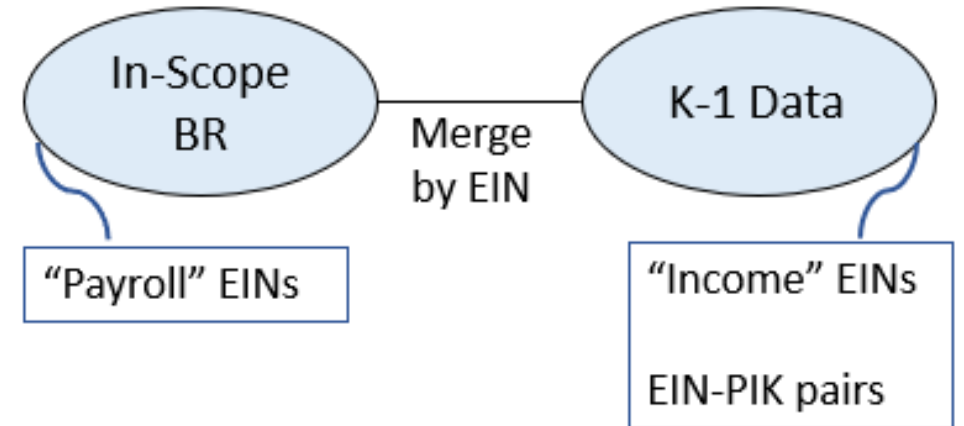


Notes: Included are the top 10 owner PIKs from each EIN. Exiters are PIKs absent from (top 10) employer ownership in 2019. Unweighted.

# Challenges

## Complex ownership structures of employer firms

- Linking issues
  - EIN is key linking variable for owner identification
  - Firms may have >1 EIN for different purposes (payroll vs. income) → Link to data with owner info is broken → Impute
- Firms owning other firms
  - Tiered entities
  - “Chase person ownership down” or un-tiering
    - New methodology => New owners database



## C-corps (SS-4 & 3<sup>rd</sup>-party data)

- Currently not able to identify owners of C-corps, but developing alternative strategies using 3rd party & SS-4 data (Equity share, voting power, board/exec roles), and imputation for smaller C-corps

# Conclusions & Looking forward

## **Many benefits (as already discussed)**

→ Reduce respondent burden, cost savings, new longitudinal microdata linkable to other programs/data, enhancement of other programs and sampling frames, provides linkages between business universe & owners universe, new research paths to entrepreneurship

**Current efforts to expand non-employers (w/ demographics) base** to include self-employed contractors (using 1099 tax data)

→ Ongoing cross-Directorate effort at Census (ESMD & CES)

**Continue developing more AdRec-based owner statistics beyond demographics**

→ Experimental NES-D (<https://www.census.gov/data/experimental-data-products/ned.html>)

# Thank you!

vitaliy.a.novik@census.gov