A Blended-data Approach to Business Demographics Statistics: Nonemployer Statistics by Demographics (NES-D)

Adela Luque (presenter)
Renuka Bhaskar
Michela Dillon
Julia Manzella
James Noon
Kevin Rinz
Victoria Udalova
(U.S. Census Bureau)

APDU
July 9, 2019

Disclaimer: The views expressed in this presentation are those of the author(s) and not the U.S. Census Bureau. DRB approval number CBDRB-FY19-356.
Business Survey Consolidation

In response to declining response rates, increasing operational costs, stakeholders needs....

- **SBO nonemployers**
  - Demographic characteristics for businesses and business owners by gender, ethnicity, race, and veteran status

- **SBO employers**
  - ASE (employers)
  - BRDIS-M (employers)

- **Nonemployer Statistics by Demographics (NES-D)**

- **Annual Business Survey (ABS)**
  - Only employers

- **Business Demographics**
Nonemployer Demographics – Past & Present

Previously:
Collected as Part of SBO

- Survey (2012) – sample size approx. 800,000 nonemployer firms & 940,000 employer firms
- Every five years
- Three-year dissemination lag from reference year
- High respondent burden
- Low response rate => high imputation rate & costs

Currently:
Blended Data Approach

- Not a survey - Full nonemployer universe
- Leverages AR & Census data to assign demographics to business owners
- Annual – higher frequency
- Aims to disseminate closer to reference year than the SBO – more timely
- No additional respondent burden
- Reduces imputation & costs
- Aims for first (beta) release in 2020 for 2017 reference year

Meets stakeholder needs: more frequent, timely, high quality data
What are Nonemployer Businesses?

2015 Nonemployer Universe: 24,331,403
Businesses w/ no paid employment, subject to federal income taxes, & w/ annual receipts of $1,000 or more

- **Sole Proprietorships** (1 owner):
  - 21,023,170 (86%)

- **Partnerships** (>1 owner):
  - 1,803,587 (7%)

- **S-corporations** (>1 owner):
  - 1,124,020 (5%)

- **C-corporations** (>1 owner)
  - 380,626 (2%)

Source: 2015 Nonemployer Statistics (NES), U.S. Census Bureau.
NES-D Content

Tabulations of:

- Number of businesses
- Business Receipts
- Number of business owners

By:

- Demographics of business owner
  - Sex, race, Hispanic origin, veteran status (SBO core demographics)
  - Age, place of birth
- Legal Form of Organization
- Geography
- NAICS industry
- Receipt-size class
- Interactions
## Data

<table>
<thead>
<tr>
<th>Variable</th>
<th>Data Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonemployer business identification, revenues, industry, geographic location</td>
<td>Business Register</td>
</tr>
<tr>
<td>Business owner identification</td>
<td>Tax Form 1040 for sole proprietorships &amp; Schedule K-1 for partnerships &amp; S-corporations</td>
</tr>
<tr>
<td>Race (core demographic in SBO)</td>
<td>Decennial &amp; American Community Survey (ACS) (Census Numident POB)</td>
</tr>
<tr>
<td>Hispanic origin (core demographic in SBO)</td>
<td>Decennial &amp; ACS (Census Numident POB)</td>
</tr>
<tr>
<td>Gender (core demographic in SBO)</td>
<td>Census Numident (Decennial, ACS)</td>
</tr>
<tr>
<td>Veteran-status (core demographic in SBO)</td>
<td>Department of Veterans Affairs (VA) AR data</td>
</tr>
<tr>
<td>Age</td>
<td>Census Numident (Decennial, ACS)</td>
</tr>
<tr>
<td>Place of birth</td>
<td>Census Numident</td>
</tr>
</tbody>
</table>
Methodology

• How are demographics linked to business owners?
  • Unique person identifiers called Protected Identification Keys (PIKs) facilitate linkage of records across data sources

• Source of PIKs
  • Sole proprietors – Tax form 1040
  • Partnerships & S-corporations – Schedule K1 (Forms 1065 & 1022S)

• Owners of U.S. nonemployer C-corporations cannot be identified with tax or AR data (only 2% of nonemployers & 4% of receipts)
  • Evaluating feasibility of imputation

• Missing demographics will be imputed
Preliminary findings on coverage

• PIK information available for approximately 98-99 percent of nonemployer firms

• Availability of gender, age, and place of birth from Numident for approximately 99 percent of owners (w/ PIKs)

• Availability of race and ethnicity from Decennial Census, ACS & Numident POB data for approximately 90-95 percent of owners (w/ PIKs)
  • Exploring use of additional AR sources to increase coverage

• Approximately 90-95 percent of owners (w/ PIKs) are not missing any demographics
  • Only 1 percent are missing 3 or more demographics
Three-year timeline

• Year 1 (2018): Assessment of PIK coverage & prevalence of missingness in AR
  • Work begins April, 2018
  • Coverage & missingness assessment
  • Preliminary firm and owner-level tabulations by LFO and gender, race, Hispanic origin, veteran status, age, place of birth - limited to 1 year of data (2015) and national level statistics
  • Report => Center for Economic Studies, Census Bureau, WP 2019-01

• Year 2 (2019): Check longitudinal consistency and imputation of missing demographics
  • Imputation of missing values & explore feasibility of C-corp imputation
  • Add additional years (2014-16) and assess longitudinal consistency of estimates
  • Explore geography and industry: tabulations by state and sector
  • Report

• Year 3 (2020): Estimates of Nonemployers Statistics by Demographics
  • Release of NES-D prototype – 2017 reference year of nonemployers

NOTE: Traditionally SBO statistics were released with a 3 year lag so no loss from transition to new program
Beyond 3-year timeline

• Annual releases starting 2020

• Explore additional characteristics such as:
  • Household characteristics: marital status, number of dependents, home/property ownership
  • Gig economy related: contract work, does nonemployer also do salaried work?
  • Transitions to employer universe
  • Trade status: importer/exporter?
  • Innovation metric: patenting activity?
  • Sources of financing
  • Education
Thank you!

Corresponding author: Adela.Luque@census.gov
Some challenges & limitations

• Owners of U.S. nonemployer C-corporations cannot be identified with tax or AR data (only 2% of nonemployers)
  • Evaluating feasibility of imputation

• Veteran status
  • Conceptual differences between what the SBO/ABS view as a veteran and VA’s definition: VA concept is narrower
  • Potential for supplementary AR source of military service data

• Equally-owned category for sole proprietorships (1 owner) possible in survey but not in AR sources

• Interagency data agreements
Preliminary Results

Race Distribution of Nonemployer Businesses by State

Race Distribution of Nonemployer Businesses by Sector

Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

13
## Some preliminary tabulations

<table>
<thead>
<tr>
<th></th>
<th>Sex</th>
<th>Place of birth</th>
<th>Veteran status</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Owner level</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>56.0%</td>
<td>Born in US</td>
<td>78.5% Veteran</td>
</tr>
<tr>
<td>Female</td>
<td>43.0%</td>
<td>Born outside US</td>
<td>20.5% Non-veteran</td>
</tr>
<tr>
<td>Missing</td>
<td>1.0%</td>
<td>Missing</td>
<td>1.0%</td>
</tr>
<tr>
<td><strong>Firm level</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>55.1%</td>
<td>Born in US</td>
<td>77.5% Veteran</td>
</tr>
<tr>
<td>Female</td>
<td>41.5%</td>
<td>Born outside US</td>
<td>21.1% Non-veteran</td>
</tr>
<tr>
<td>Equal</td>
<td>2.4%</td>
<td>Equal</td>
<td>0.4% Equal</td>
</tr>
<tr>
<td>Not assigned</td>
<td>1.1%</td>
<td>Not assigned</td>
<td>1.1% Not assigned</td>
</tr>
</tbody>
</table>