Business Dynamics and Workforce Characteristics: Integrating Demographic Data with Business Administrative Data

NBER Summer Institute

July 2017

Kristin McCue, U.S. Census Bureau

Martha Stinson, U.S. Census Bureau

Cristina Tello-Trillo, U.S. Census Bureau

Disclaimer

Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau.

All results have been reviewed to ensure that no confidential information is disclosed.

Benefits of adding workforce information to firm data

- Use worker flows to improve quality of longitudinal firm links
  - Essential for accurately calculating number of firm births and deaths
- Answer research questions such as:
  - How does the workforce differ across types of firms?
    - What types of workers are employed at growing firms, shrinking firms, or start-ups?
  - How are worker flows related to firm dynamics?
    - As firms grow/become more productive, what happens to the earnings distribution? Does the workforce become more homogenous or more diverse?
  - Have these relationships changed over time?
- Workforce information can augment published statistics on firm dynamics

What we gain from W-2 tax returns

- Consistent national time series for 10 years (2005-2015) with potential to extend back to 1978.
- Same business frame as Census administrative business data
  - Same business entities file quarterly tax returns with IRS and annual W-2 forms with SSA
- Additional earnings fields allow creation of total earnings measure
- Drawbacks:
  - data are annual not quarterly
  - Some types of work don't require W-2s (i.e. gig economy)

United States™
Census Bureau

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

Data Sources

Business Register (BR)
- Annual list of all establishments
- Includes employment, payroll, industry, and firm organization

Longitudinal editing →

Longitudinal Business Database (LBD)
- Accurately link establishments over time

link by firm tax identifier (EIN) ↓

5% of firms have no matching W-2 records

Internal Revenue Service Form W-2 File
- Universe of workers in a year
- Includes uncapped annual earnings
- Matches workers to firms

link by person identifier (PIK) ↓
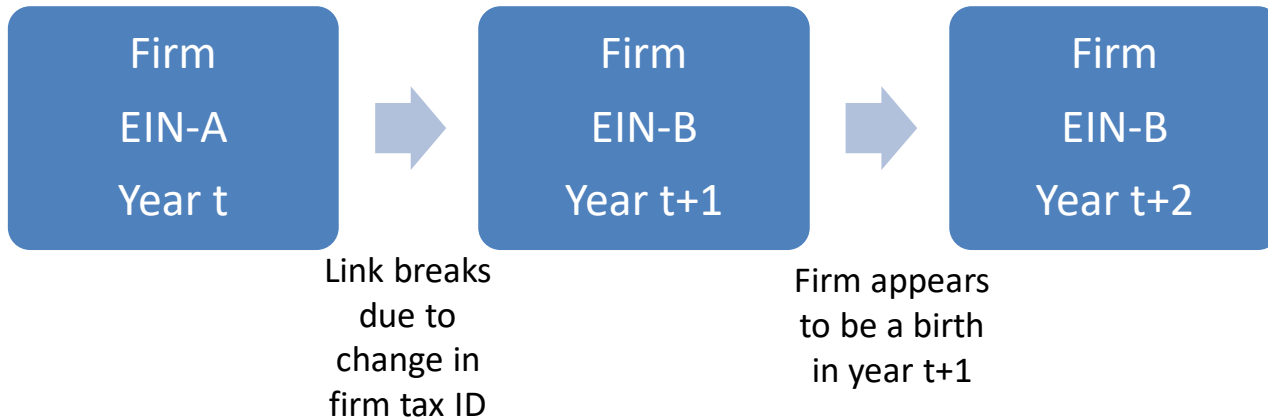
3% of people missing at least one trait

Social Security PIK Database and 2000 Census of U.S. Population
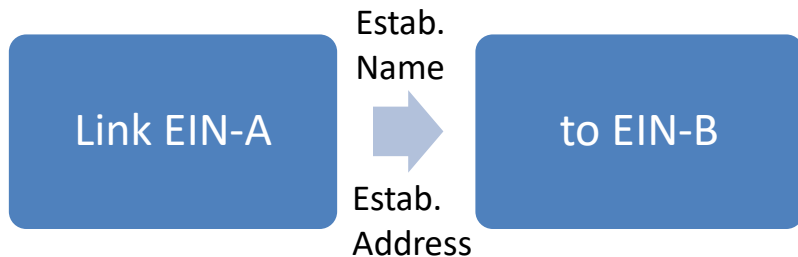- Includes gender, date of birth, race, ethnicity, and education

Data Improvement

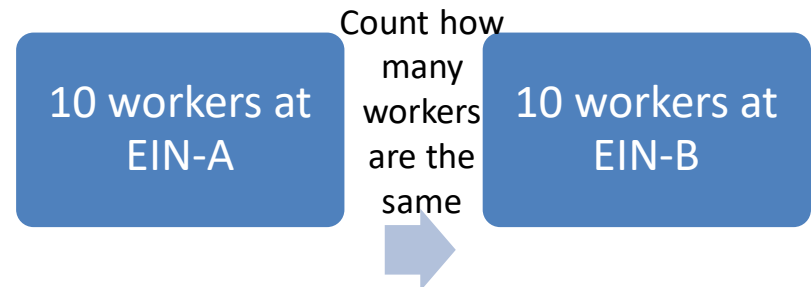# Tracking single-establishment firms over time

Inherent Problem: Re-organization

| Firm
EIN-A
Year t | → | Firm
EIN-B
Year t+1 | → | Firm
EIN-B
Year t+2 |

Link breaks due to change in firm tax ID

Firm appears to be a birth in year t+1

LBD Solutions

| Link EIN-A | → | to EIN-B |

Estab. Name

Estab. Address

Method 1: conservative matching
Method 2: broader matching

Potential of Worker Flows

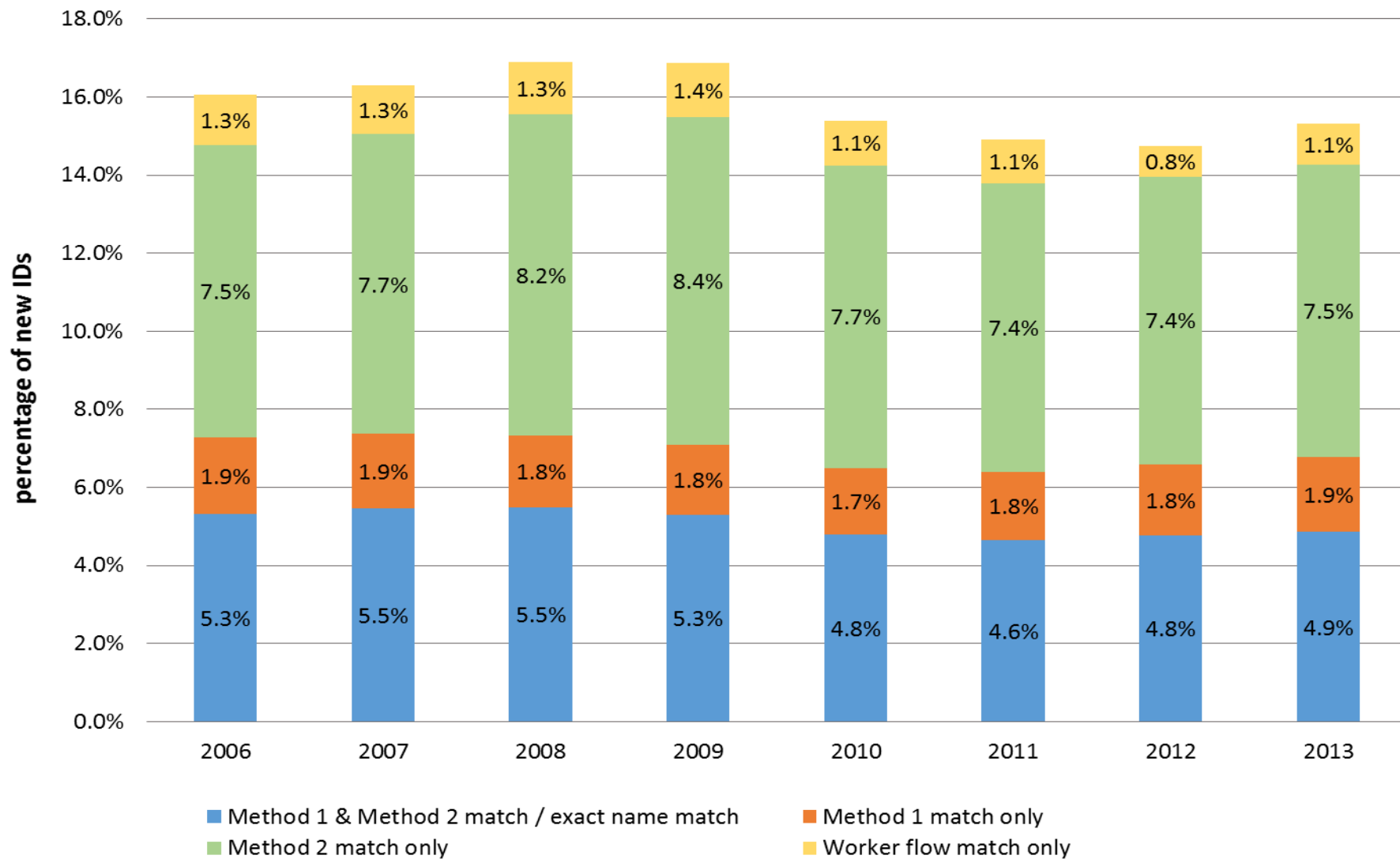| 10 workers at EIN-A | → | 10 workers at EIN-B |

Count how many workers are the same

If EIN-B
- Appears to be a birth (i.e. first year of positive payroll)
- 20%-50% of workers in common with EIN-A
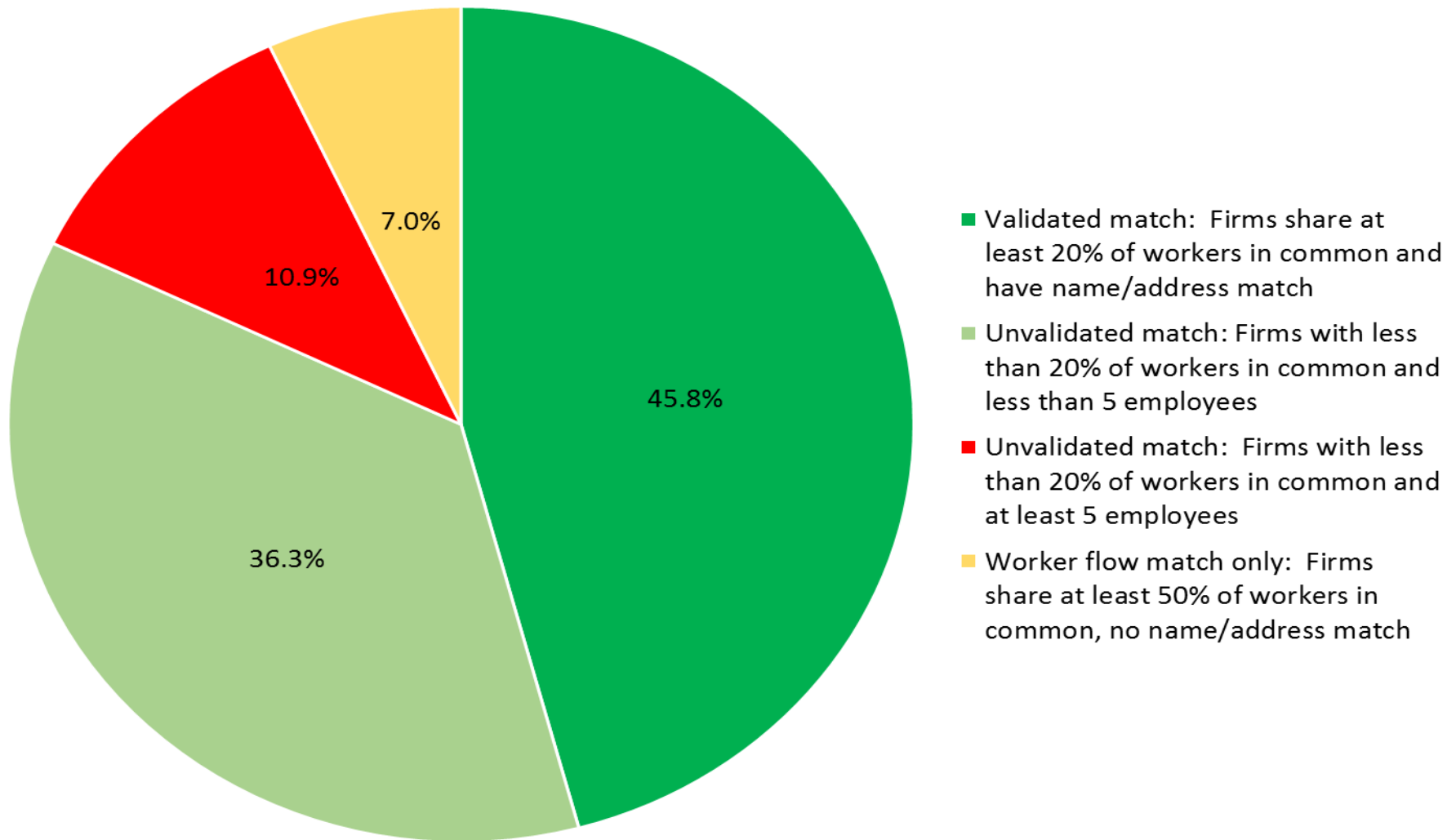- At least 5 workers

If EIN-A appears to be a death

Declare EIN-B to match to EIN-A

Percentage of new IDs in the Census Business Register that are re-organizations of pre-existing single-unit establishments

Year	Method 1 & Method 2 match / exact name match	Method 1 match only	Method 2 match only	Worker flow match only
2006	5.3%	1.9%	7.5%	1.3%
2007	5.5%	1.9%	7.7%	1.3%
2008	5.5%	1.8%	8.2%	1.3%
2009	5.3%	1.8%	8.4%	1.4%
2010	4.8%	1.7%	7.7%	1.1%
2011	4.6%	1.8%	7.4%	1.1%
2012	4.8%	1.8%	7.4%	0.8%
2013	4.9%	1.9%	7.5%	1.1%

United States Census Bureau

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

Percentage of 2013 identified potential re-organizations confirmed by W-2 data evidence

- Validated match: Firms share at least 20% of workers in common and have name/address match
- Unvalidated match: Firms with less than 20% of workers in common and less than 5 employees
- Unvalidated match: Firms with less than 20% of workers in common and at least 5 employees
- Worker flow match only: Firms share at least 50% of workers in common, no name/address match

45.8%
36.3%
10.9%
7.0%

Research Questions
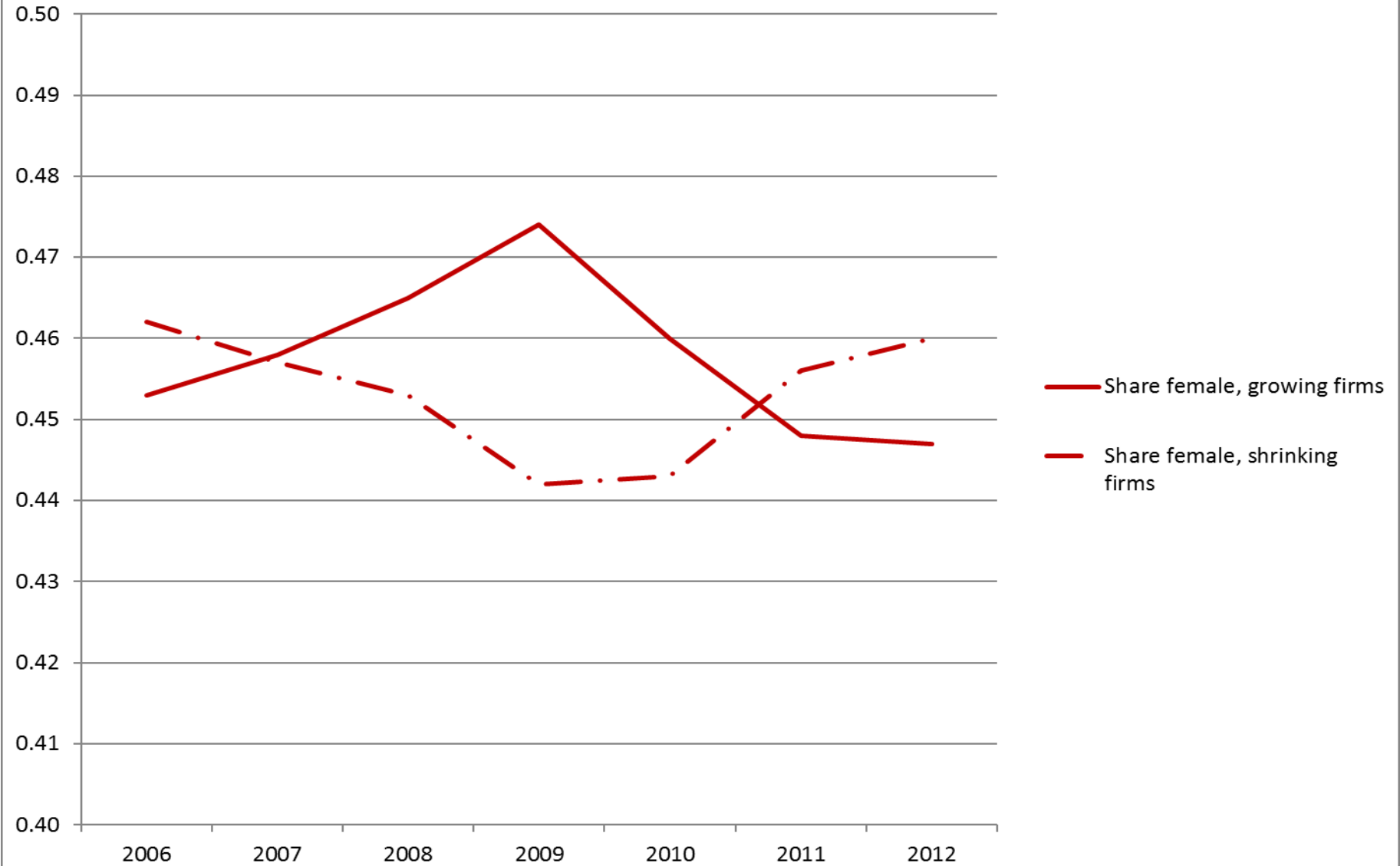# and
Useful Summary Measures

Earnings distributions within firms

- Classify firms by size
  - Static: March 12th employment
  - Change: year-to-year difference in March 12th employment
- Use W-2 data to create earnings distribution within each firm for each year 2005-2014
  - Use LEHD quarterly work indicators to adjust for partial years
- Average distribution statistics across firms within size and growth categories
  - Difference between 90th and 10th percentiles
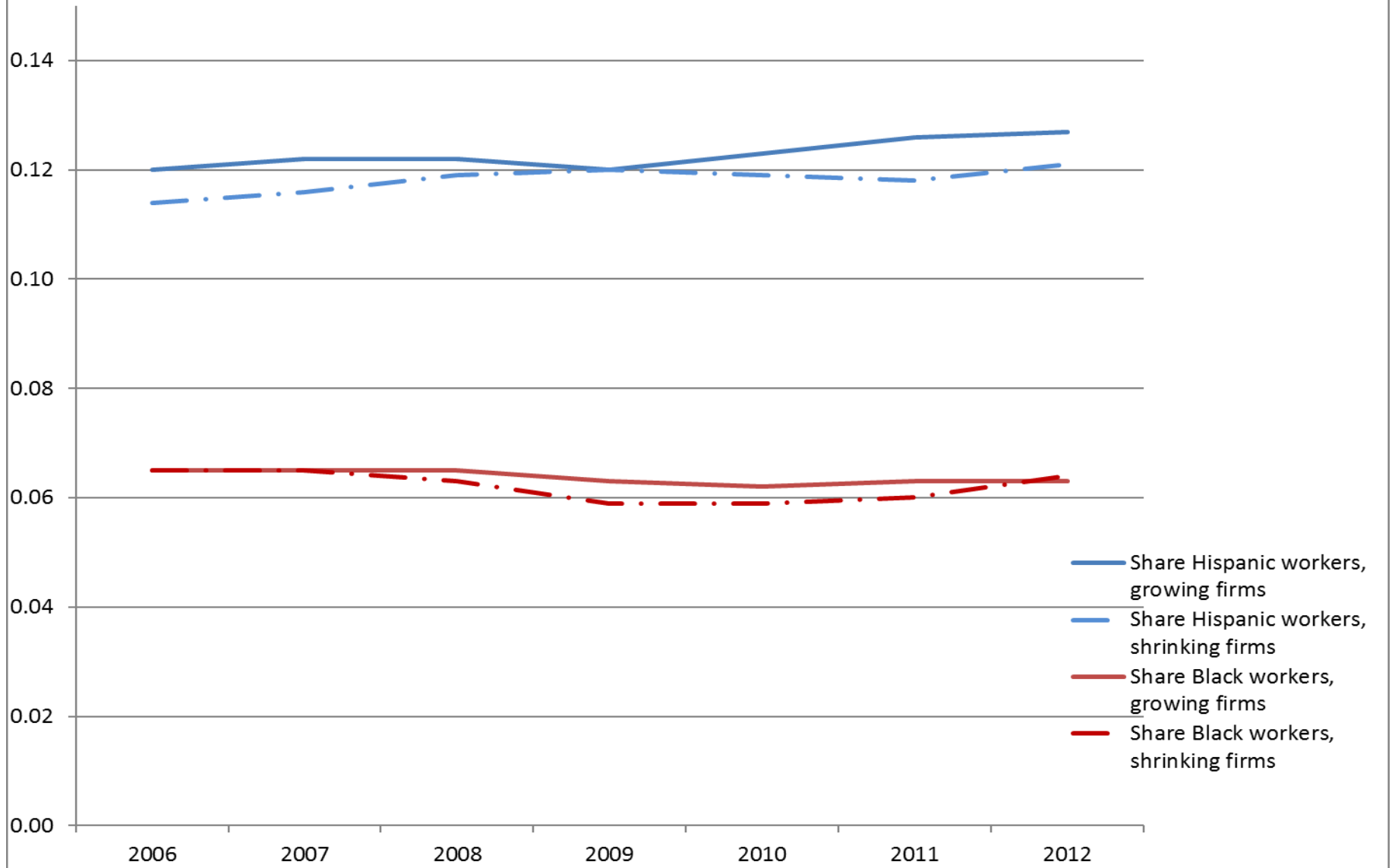 - Percentage change in this difference across adjoining years

**Firms 101-500 employees, 10%-50% change in employment:
Percentage real change in 90-10 diff**

	2005-2006	2013-2014

shrinking firms
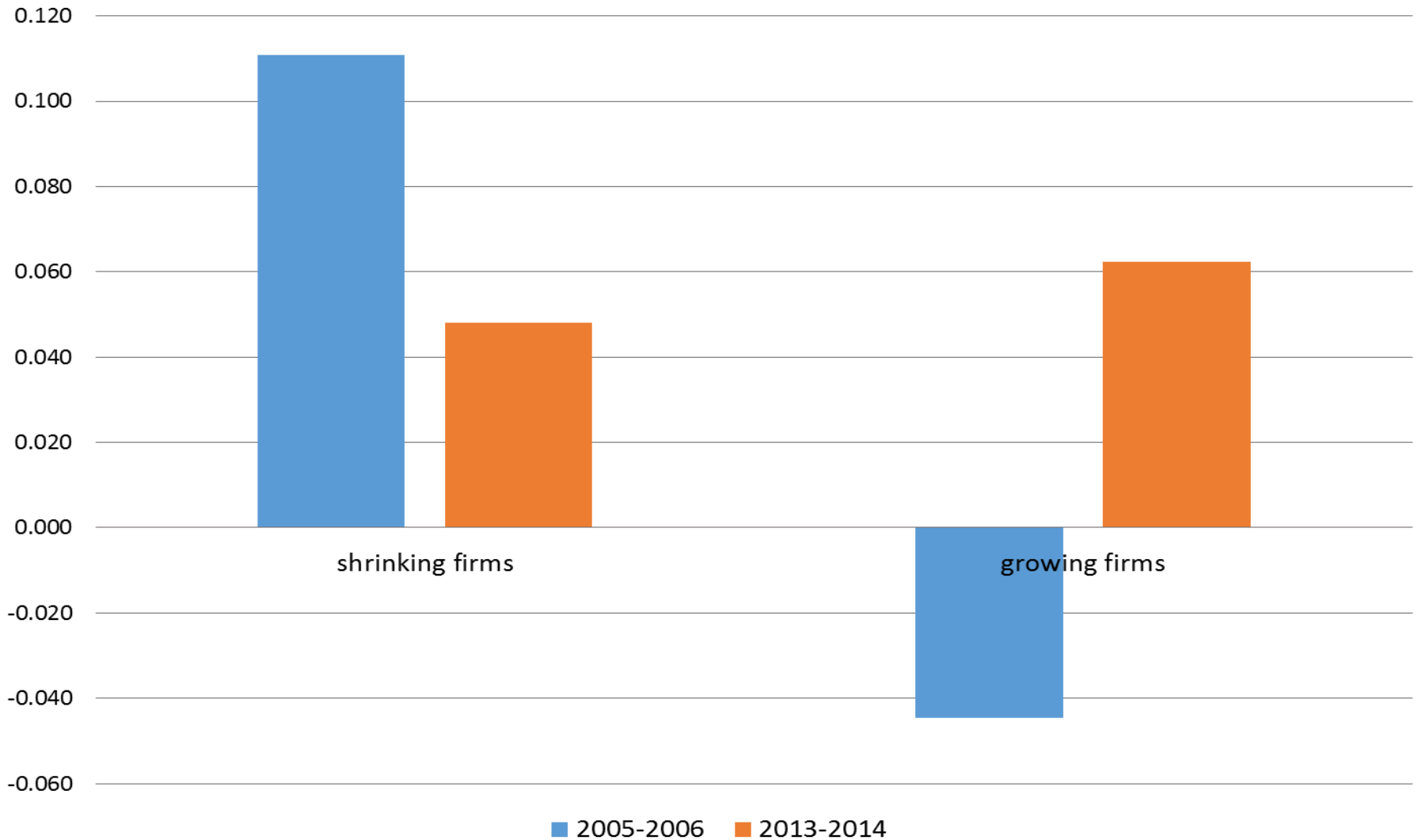
growing firms

Gender Shares at Growing and Shrinking Firms over Time

Share female, growing firms

Share female, shrinking firms

Shares of Black and Hispanic workers at growing and shrinking firms over time

Legend:
- Share Hispanic workers, growing firms
- Share Hispanic workers, shrinking firms
- Share Black workers, growing firms
- Share Black workers, shrinking firms

Thank you