# Statistics of Income IRS Migration Data

October 26, 2022

Kevin Pierce

- Background on IRS Migration Data

- Data Anomalies

- Comparison between IRS Migration Data and Census Bureau's Population Estimates Program

- Future Plans

- Demonstration

# Outline

# Background: What are the Migration Data?

Migration data show the movement of individuals, via the address listed on Form 1040, over a two-year period

As an example…

Calendar Years

Migration Data ⟹ 2018 - 2019

⬇ ⬇

Tax Years: 2017 2018

The data are available at the State or county level:

- Number of inflows – residents moving in
- Number of outflows – residents moving out

# Background: What are the Migration Data?

State level
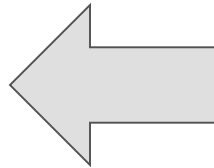
(1) State-to-State Inflow

(2) State-to-State Outflow

County level

(1) County-to-County Inflow

(2) County-to-County Outflow

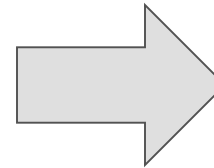| Origin from Alabama (State Code) | Destination into | | | Number of returns | Number of exemptions | Adjusted gross income (AGI) |
|---|---|---|---|---|---|---|
| | State Code | State | State Name | (1) | (2) | (3) |
| 01 | 96 | AL | AL Total Migration US and Foreign | 51,971 | 107,304 | 2,109,108 |
| 01 | 97 | AL | AL Total Migration US | 50,940 | 105,006 | 2,059,642 |
| 01 | 98 | AL | AL Total Migration Foreign | 1,031 | 2,298 | 49,465 |
| 01 | 01 | AL | AL Non-migrants | 1,584,665 | 3,603,439 | 87,222,478 |

# Background: History of the Migration Data

| Previous | | | | New | |
|---|---|---|---|---|---|
| Methodology | | | | Methodology | |

| **1990s** | **2000s** | **2010-2011** | **2011-2012** | **...** | **2018-2019** |
|---|---|---|---|---|---|

- Based on returns filed up until late September
- Matched on Primary Taxpayer only

- Based on returns for the full calendar year
- Matched on Primary, Secondary, and Dependents

What didn't change?

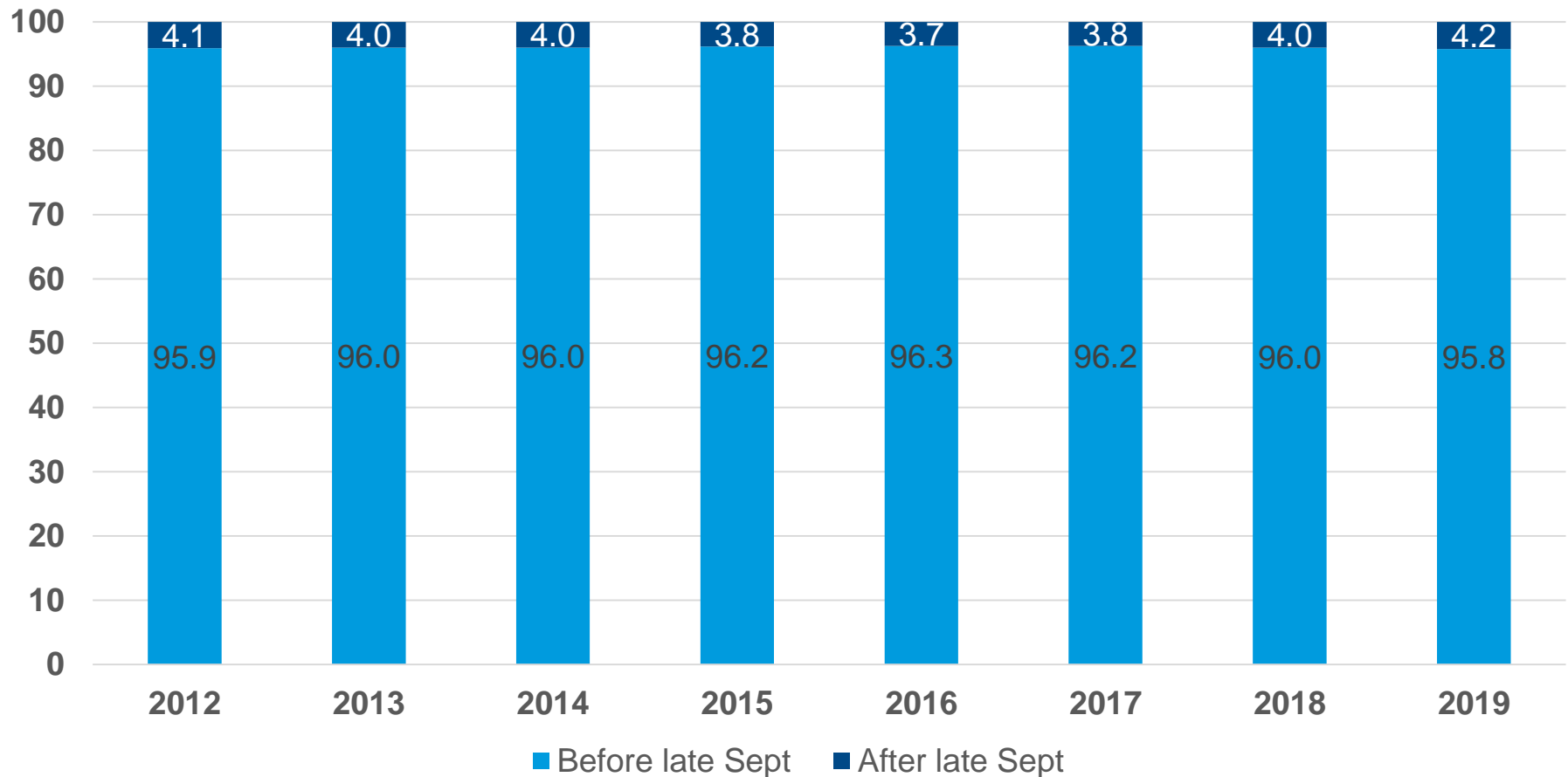Maintained consistency between final tables, layout, codes, and disclosure

# Why change the methodology?

- Greater flexibility

- Increased population

- Inclusion of higher income tax filers

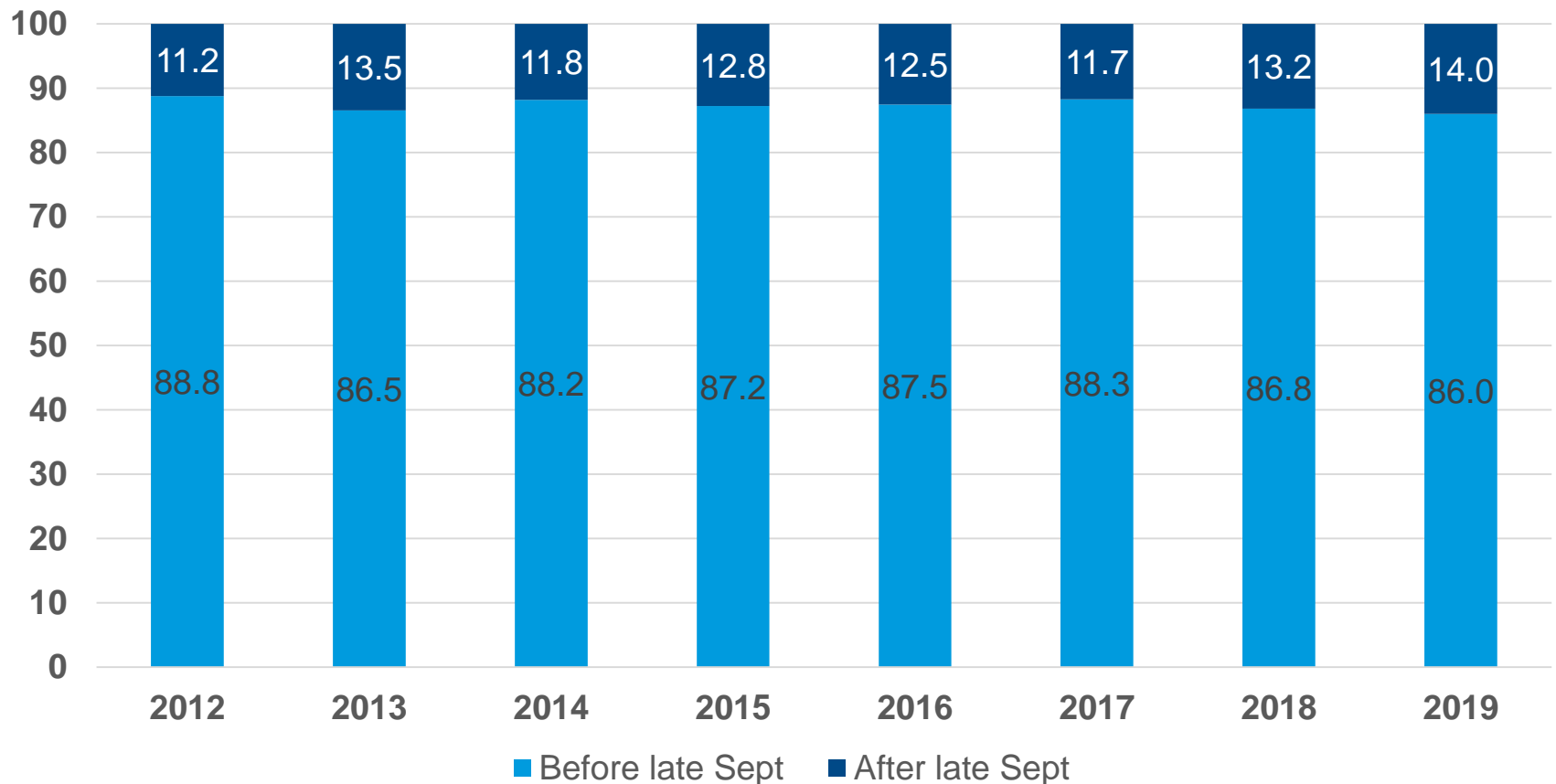# Including Tax Returns after late September, Increases Population Size by 4%

**Percent of total**

Percentage of Returns, Calendar Years 2012-2019

| Year | After late Sept | Before late Sept |
|------|-----------------|------------------|
| 2012 | 4.1 | 95.9 |
| 2013 | 4.0 | 96.0 |
| 2014 | 4.0 | 96.0 |
| 2015 | 3.8 | 96.2 |
| 2016 | 3.7 | 96.3 |
| 2017 | 3.8 | 96.2 |
| 2018 | 4.0 | 96.0 |
| 2019 | 4.2 | 95.8 |

■ Before late Sept ■ After late Sept

# Tax Returns after late September account for 4% of Returns, but over 10% of Total Income

**Percent of total**

Percentage of AGI, Calendar Years 2012-2019



| Year | After late Sept | Before late Sept |
|------|-----------------|------------------|
| 2012 | 11.2 | 88.8 |
| 2013 | 13.5 | 86.5 |
| 2014 | 11.8 | 88.2 |
| 2015 | 12.8 | 87.2 |
| 2016 | 12.5 | 87.5 |
| 2017 | 11.7 | 88.3 |
| 2018 | 13.2 | 86.8 |
| 2019 | 14.0 | 86.0 |

■ Before late Sept   ■ After late Sept

# Tax Returns after late September have Higher Average AGI

**Dollars**

Average AGI, Calendar Years 2012-2019



| Year | Before late Sept | After late Sept |
|------|------------------|-----------------|
| 2012 | $52,945 | $157,658 |
| 2013 | $56,067 | $210,134 |
| 2014 | $56,323 | $181,724 |
| 2015 | $59,148 | $217,402 |
| 2016 | $61,091 | $225,809 |
| 2017 | $61,931 | $211,035 |
| 2018 | $64,642 | $234,344 |
| 2019 | $67,601 | $250,223 |

■ Before late Sept  ■ After late Sept
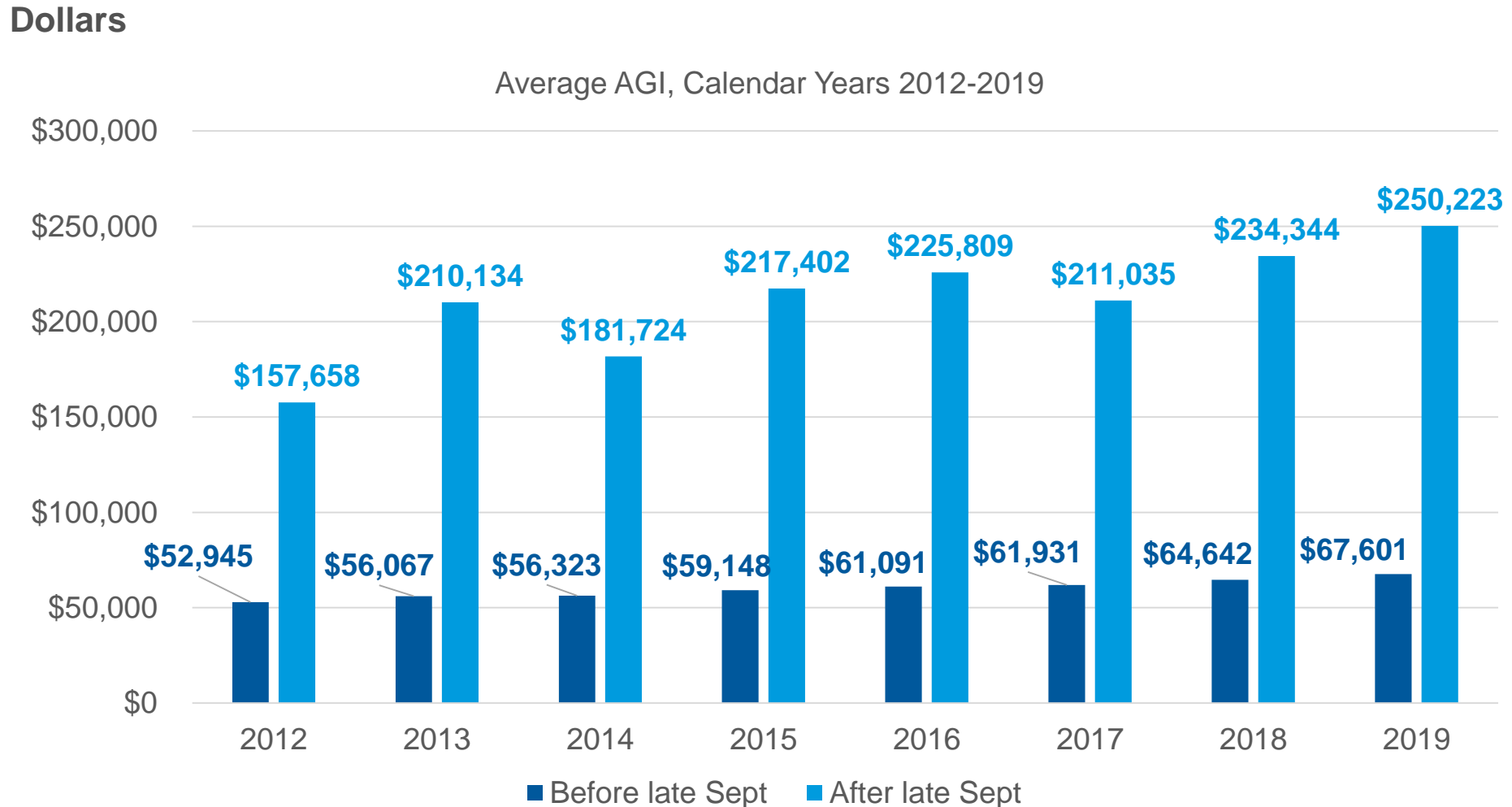
# Enhanced Merging Increases Population by 5%

| Filing position | 2011-2012 | 2013-2014 | 2014-2015 | 2016-2017 | 2017-2018 | 2018-2019 |
|---|---|---|---|---|---|---|
| **Primary-Primary** | 94.60 | 94.73 | 94.89 | 94.48 | 94.75 | 94.96 |
| **Primary-Secondary** | 0.84 | 0.85 | 1.08 | 1.07 | 1.04 | 1.02 |
| **Secondary-Primary/Secondary** | 1.67 | 1.60 | 1.83 | 1.73 | 1.69 | 1.66 |
| **Dependent-Primary/Secondary** | 2.88 | 2.82 | 2.20 | 2.72 | 2.52 | 2.36 |

# Data Anomalies

- ACS, CPS, IRS Comparison

- Updated Methodology vs Previous Methodology

# Percent of Interstate Movers: ACS, CPS, IRS

# Percent of Movers: Updated Methodology vs Previous Methodology, 2008-2019
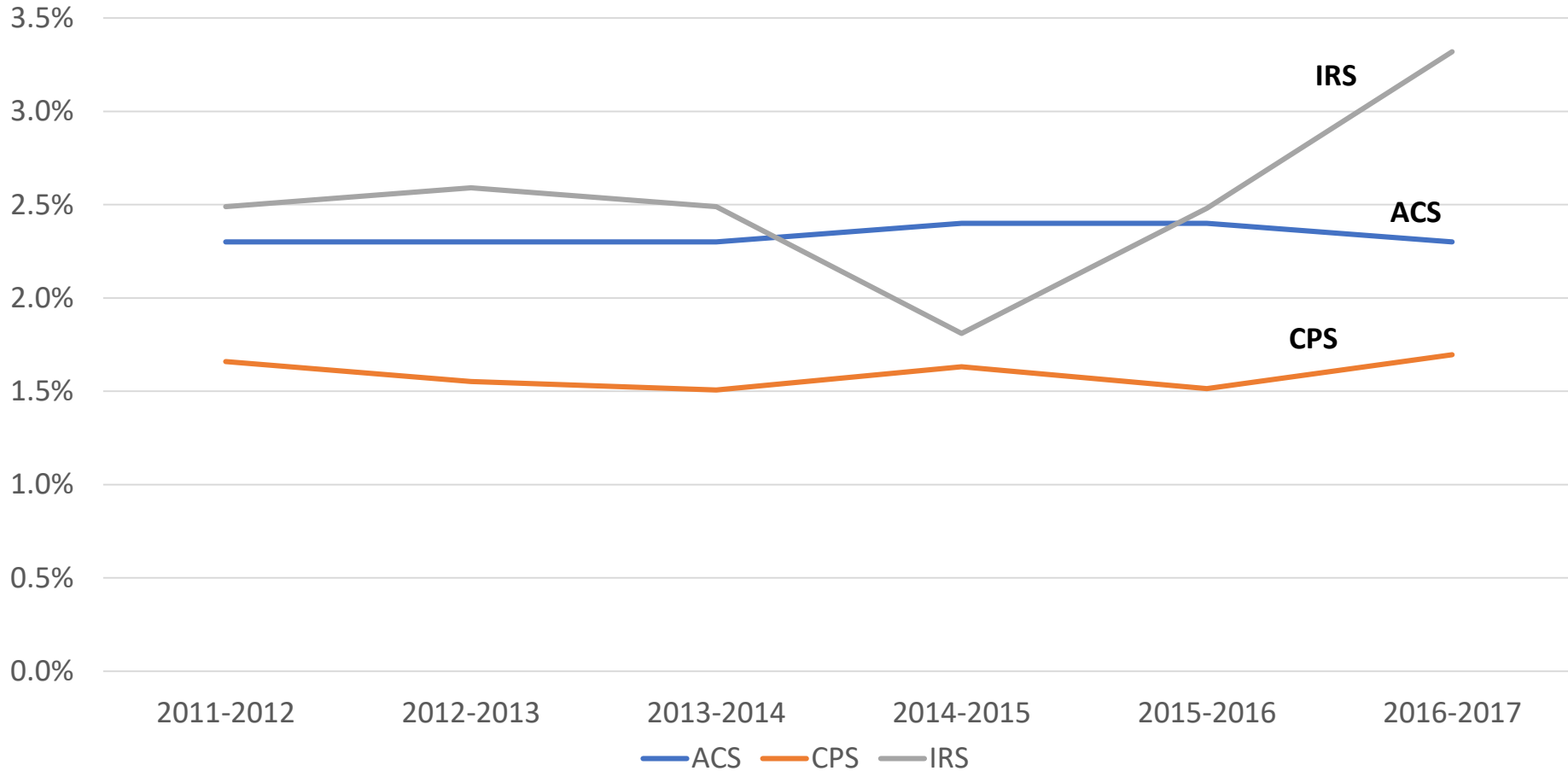
**Percent of movers**



Legend: ■ Updated methodology  ■ Previous methodology
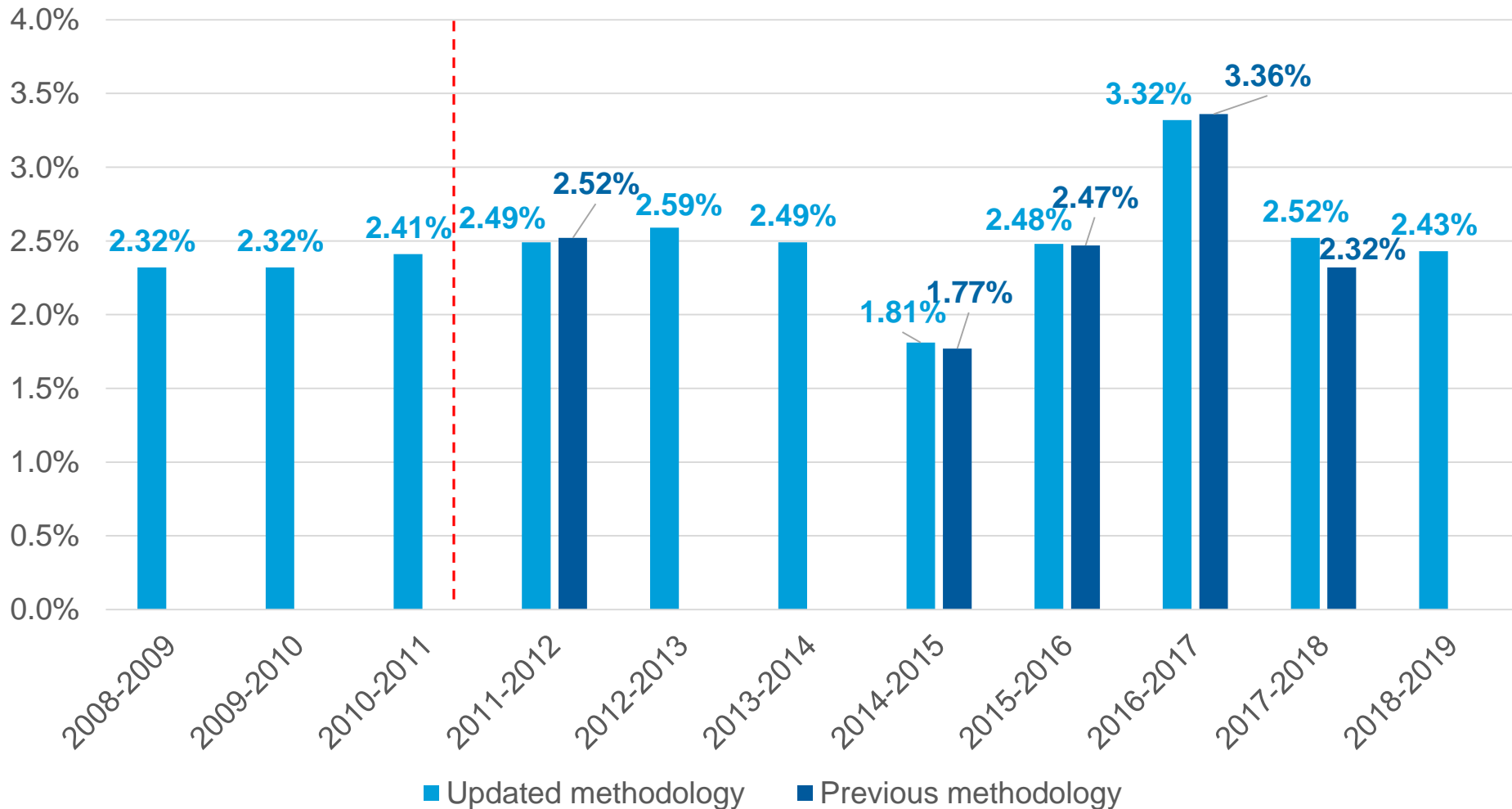
# Percent of Movers, Updated Methodology vs Previous Methodology, 2008-2019

**Percent of movers**
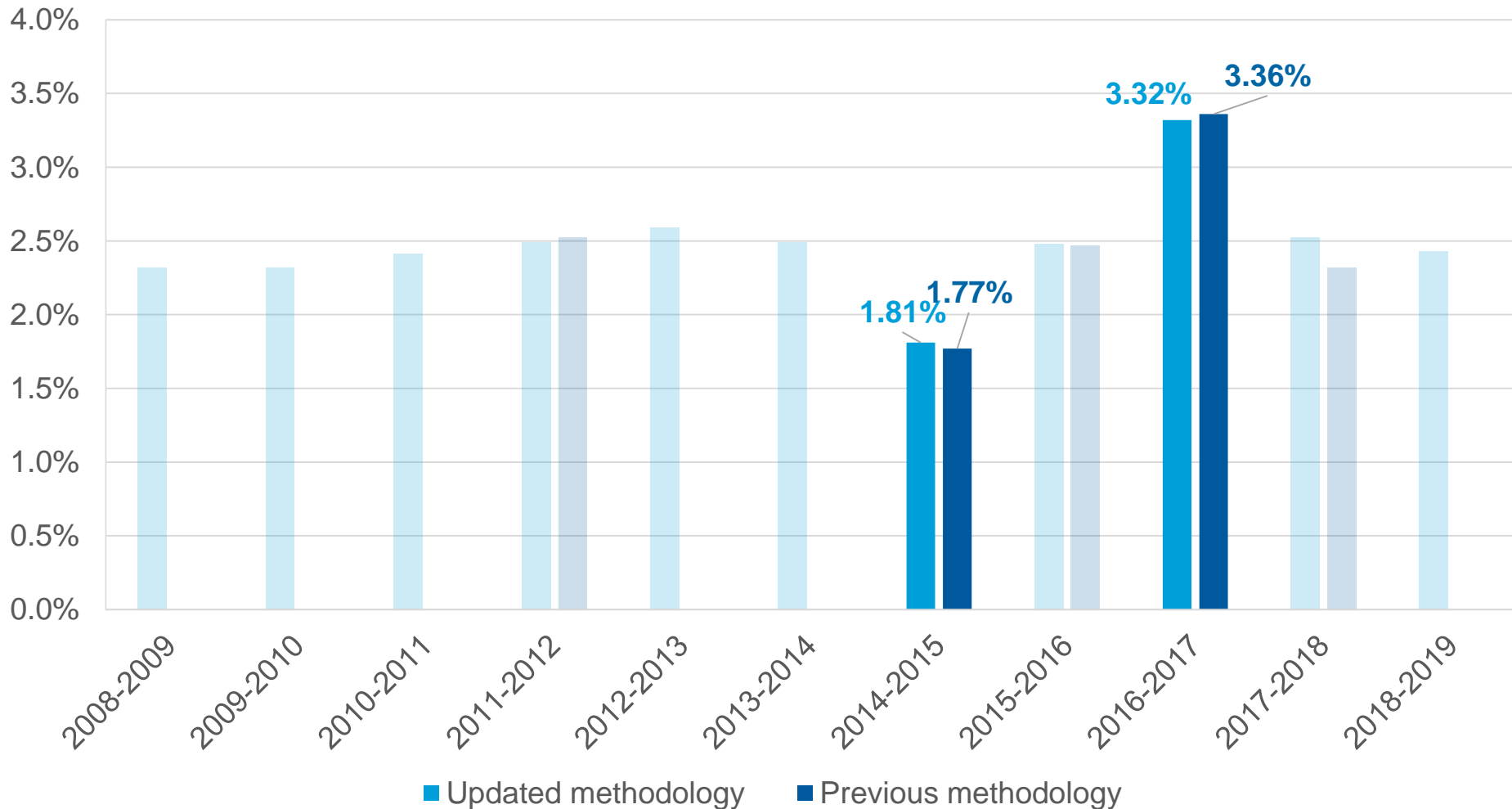
# IRS Migration and Census PEP

# IRS Migration and Census PEP Comparisons

|  | IRS Migration Data | Census PEP Migration Estimates |
|---|---|---|
| **Data used** | **IRS tax return filing population + SSA Numerical Identification File** | **IRS tax return filing population** + Medicare enrollment data + Group Quarters data + **SSA Numerical Identification File** |
| **Tax return time period covered** | Returns processed by end of December | Returns processed by late September |
| **Matching procedure** | Primary, Secondary, Dependent filers (year 1) to Primary/Secondary (year 2)* <br>• **Only matched records over 2 years are included** <br>• Dependent returns in year 1 and year 2 are excluded <br>• Nonfilers are excluded | Individuals (from IRS + Medicare data) in year 1 to individuals in year 2 <br>• **Only matched records over 2 years are included** <br>• Because of person-based method, dependent records are included <br>• Nonfilers could be represented in Medicare data |
| **Geographic coding** | **Records are geocoded based on 9-digit ZIP code** <br>• If county cannot be determined, records are excluded | **Records are geocoded based on 9-digit ZIP code** <br>• If county cannot be determined, records are coded based on address |

*Beginning with 2011-2012 IRS data

# IRS Migration and Census PEP Comparisons

| | IRS Migration Data | Census PEP Migration Estimates |
|---|---|---|
| **Data released** | Data are released and updated annually independent of previous iterations | Data are released annually and include an updated time-series beginning with the previous Census or point in time series |
| **Methodology** | Direct method of calculating migration using matched population of tax returns<br>• **Identify inmigrant, outmigrant, and nonmigrant returns based on ZIP code of origin (year 1) and destination (year 2)**<br>• Microdata are aggregated into state and county flows<br>• Data below 10 returns are grouped into various "Other" categories or suppressed | Indirect method of calculating domestic migration based on matched IRS + Medicare records<br>• Data split into age groups (0-17, 18-64, and 65+)<br>• **Identify inmigrants, outmigrants, and nonmigrants based on ZIP code of origin (year 1) and destination (year 2)**<br>• Calculate net domestic migration rate (NDMR) for an area (Ins-Outs)/(Nonmigrants + Outs)<br>• Apply NDMR to household population to create net domestic migration estimate<br>• Add the domestic migration estimate for the 3 age groups to the total GQ change |

# Future Plans

# Future Plans

Future plans include the continued use of tax return information, as well as greater use of information returns (W-2, 1099-R, 1099-SSA, 1099-MISC, 1099-G)

- Include nonfilers in population
- Use of info return ZIPs + tax return ZIPs, could be less susceptible to identity theft/IRS processing changes

Under consideration…

- Moving to an individual-based matching procedure
- How should we assign income for joint tax filers?
- What types of income should be included for nonfilers?
- How far back should we apply any new methodology?
- What other items should we consider?

# IRS Migration Data Websites

IRS Statistics of Income Main Page:

Statistics | Internal Revenue Service (irs.gov)

IRS Migration Page:

SOI Tax Stats - Migration Data | Internal Revenue Service (irs.gov)

Migration Papers
- SOI Migration Data: A New Approach
- U.S. Population Migration Data: Strengths and Limitations

# Thank you.
# I welcome your feedback.
## [www.irs.gov/statistics](www.irs.gov/statistics)

Kevin Pierce

kevin.k.pierce@irs.gov