Estimating Foreign-Born Emigration Using Longitudinal Tax and Social Security Data

Thomas B. Foster, Anthony Knapp, and Erik Vickstrom

U.S. Census Bureau
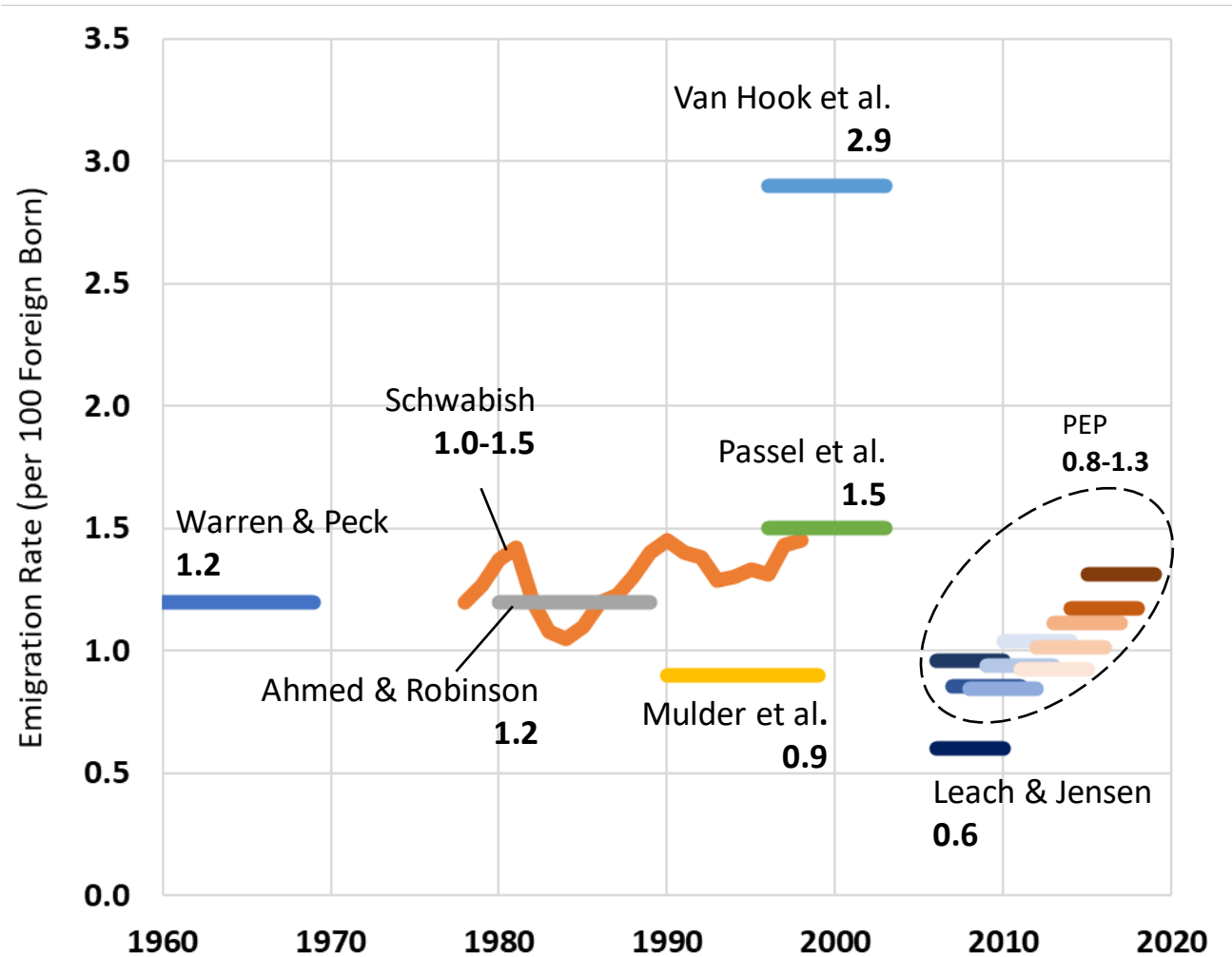
October 26, 2022

Background

- Estimating emigration is challenging because of limited data.

- Previous Sources:
  - Immigration and Naturalization Service, 1908 to 1957.
  - Alien Address Report Program, 1952 to 1981.

- Contemporary Sources:
  - U.S. Customs and Border Protection (CBP) Form I-94 Arrival/Departures.
  - Social Security Administration (SSA) files.

Background

- Lacking data directly detecting emigration, "residual" methods aim to estimate emigration indirectly.

- Warren and Peck (1980) developed a Census-Census residual method to derive foreign-born emigration from population change measured between two decennial censuses.

- Ahmed and Robinson (1994) and Mulder et al. (2003) updated the rates after each census.

- Leach and Jensen (2013) replaced census data with the American Community Survey (ACS).

Selected Foreign-Born Emigration Rates from Prior Research

Generally, 1 to 1.5 out of every 100 foreign born emigrate per year over the long run.

Source: Simulation of emigration rates using PEP methodology on 1-year ACS PUMS with adjustments.
Note: Figure derived from comparative analyses from Schwabish (2011) and Leach and Jensen (2013).

United States® Census Bureau

Motivation

- Population Estimates Program (PEP) uses ACS data to estimate foreign-born emigration.
  - Estimates derived from ACS inputs are lagged, not year-specific, and can be unstable.
  - Coverage and non-coverage errors create uncertainty in the estimates.
- Ongoing work to leverage existing administrative data at the Census Bureau to evaluate ACS-based rates.
  - Better leverage the longitudinal nature of tax data, implementing and evaluating a "signs-of-life" approach.
  - Increase coverage of foreign-born non-filers with IRS 1099 or "information returns".

Data

- IRS 1040 federal tax returns and IRS 1099 "information returns" (W2s, 1098s, 1099s)
  - Tax Years 2009 through 2020
- Numident file (2022)
- Link these individual records using Protected Identification Keys (PIKs)
- Comparison data based on 2020 PEP methodology: [1]
  - Simulated emigration rates from ACS-ACS residual method
  - Characteristics come from ACS proxy universe method

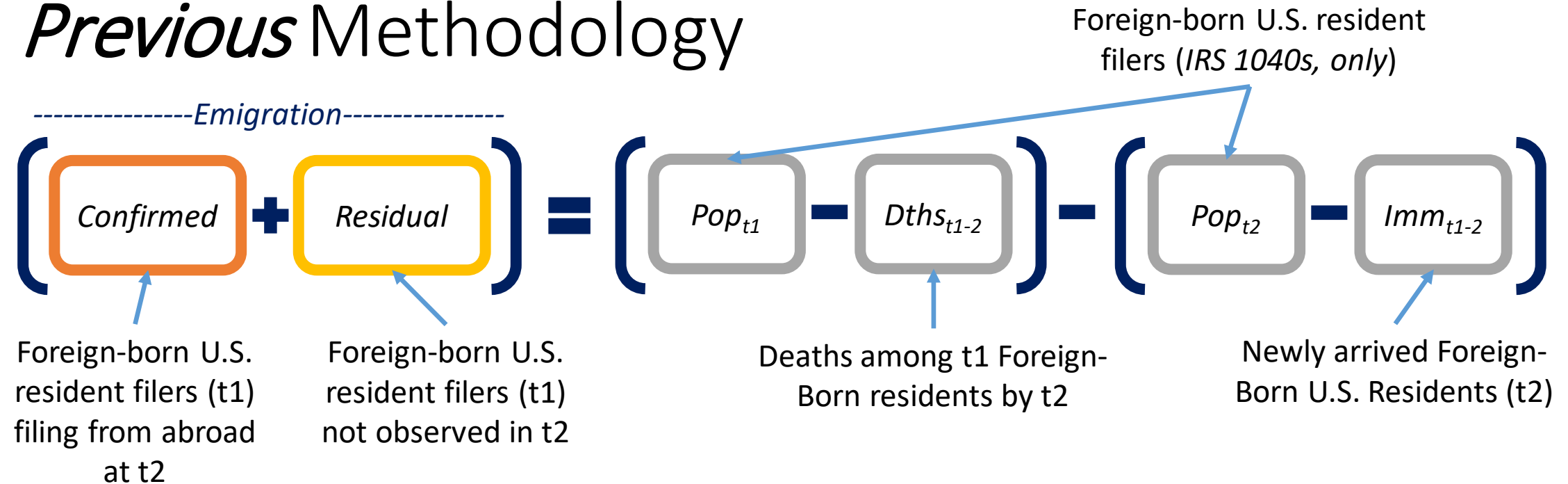| Data Set	Variables	Universe
IRS 1040s	• Geography • Tax Year • ITIN Flag	• Filers in U.S. *or abroad*
IRS 1099s	• Geography • Tax Year	• Filers and Non-Filers in U.S.
Numident	• Date of Birth • Date of Death • Country of Birth • Year Social Security Number (SSN) Issued • Foreign-Born Flag	• Individuals ever issued an SSN

[1] PEP methodology: https://www.census.gov/programs-surveys/popest/technical-documentation/methodology.html

Coverage Improvements & ITIN Issues

Coverage	Previously (1040s Only)	Currently (1040s and 1099s)
Best	Permanent immigrants, naturalized citizens, non-citizen workers, refugees	Permanent immigrants, naturalized citizens, non-citizen workers, refugees
Limited	Non-citizen students, non-citizen spouses and children	Non-citizen students, non-citizen spouses and children, ***non-filers, inconsistent filers***
Worst	Undocumented/unauthorized migrations, asylum seekers	Undocumented/unauthorized migrations, asylum seekers
None	***Non-filers, inconsistent filers***, Non-SSNs (ITINs)	Non-SSNs (ITINs)

- While we do observe ITINs in the IRS 1040s and 1099s (15% of the resident foreign-born population, on average), unresolved issues stemming from PIK assignment remain.
  - Cannot distinguish an *ITIN $\rightarrow$ residual emigrant* from an *ITIN $\rightarrow$ SSN non-emigrant*.

Previous Methodology

Foreign-born U.S. resident filers (*IRS 1040s, only*)

--------------Emigration--------------

$$\Big(\text{Confirmed} + \text{Residual}\Big) = \Big(\text{Pop}_{t1} - \text{Dths}_{t1-2}\Big) - \Big(\text{Pop}_{t2} - \text{Imm}_{t1-2}\Big)$$

Foreign-born U.S. resident filers (t1) filing from abroad at t2

Foreign-born U.S. resident filers (t1) not observed in t2

Deaths among t1 Foreign-Born residents by t2

Newly arrived Foreign-Born U.S. Residents (t2)
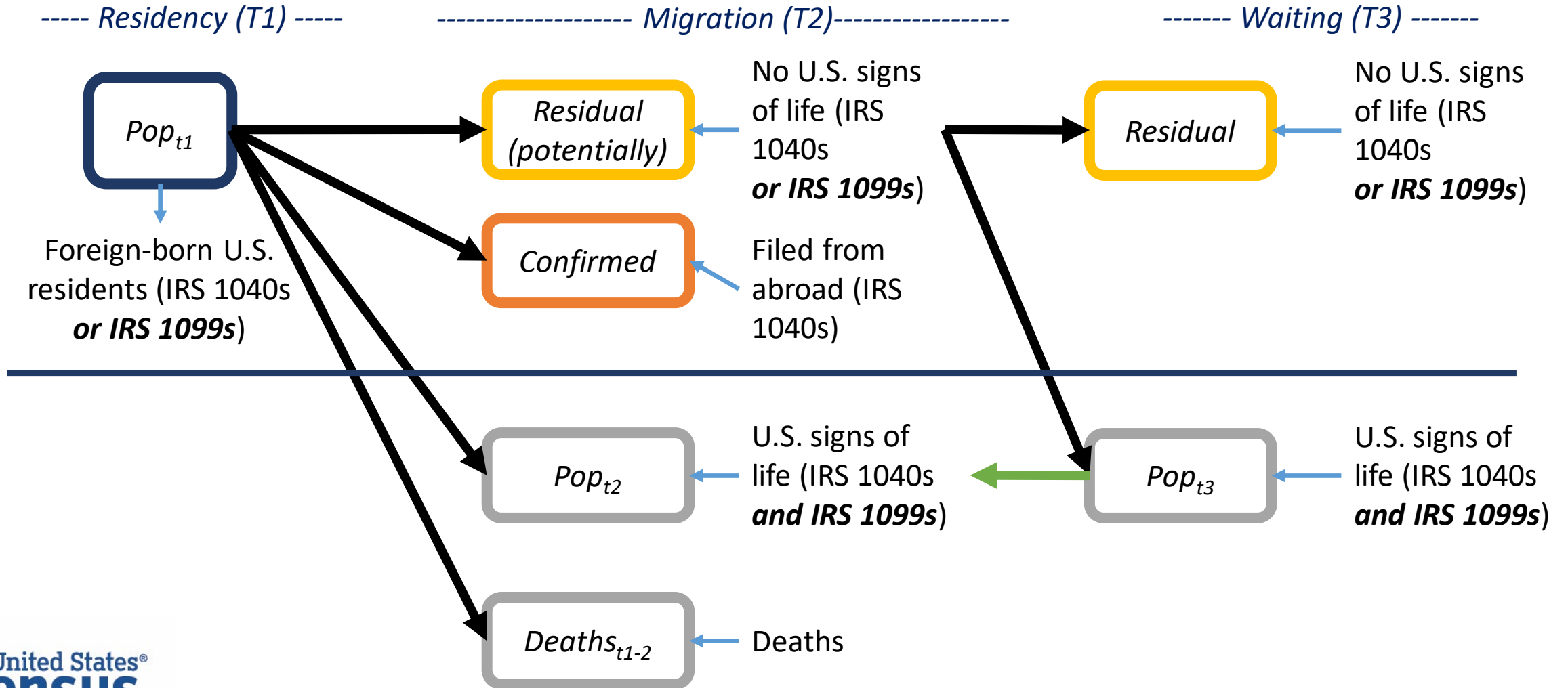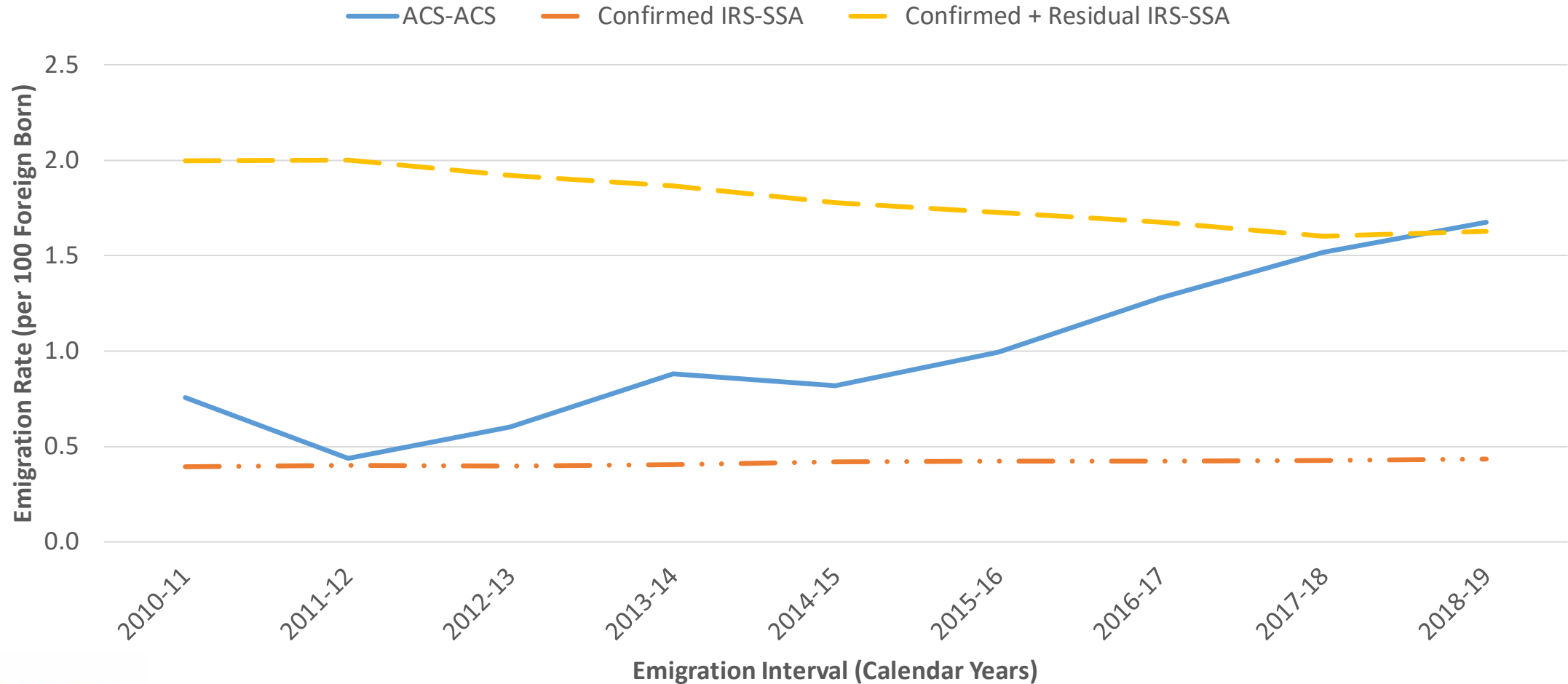
- Stable confirmed emigrant component, but erratic residual.

- Residual component includes non-emigrants who simply do not file in t2 (retirees, low earners, etc.)

- Requires accurate estimation of population stock and immigration component in t2.
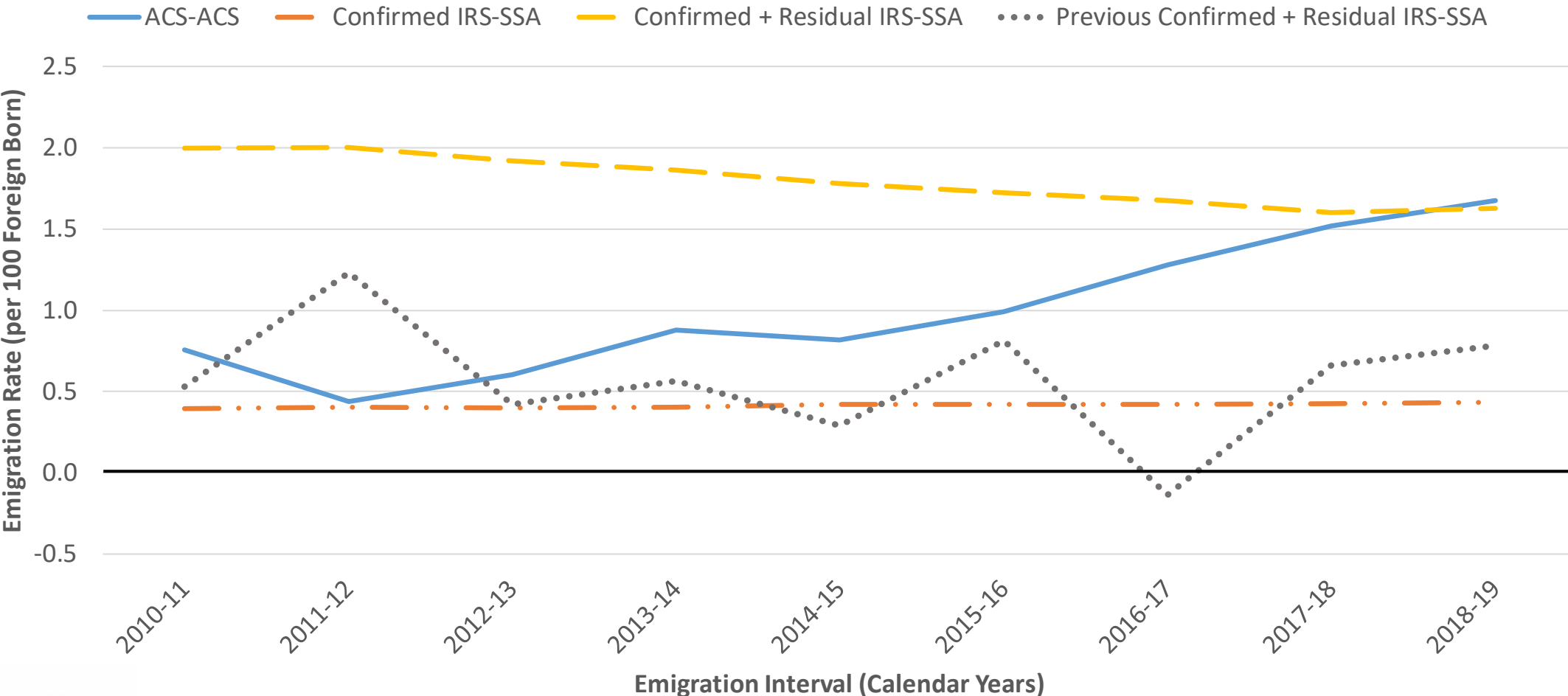
United States® Census Bureau

Current Methodology

----- *Residency (T1)* -----   ------------------ *Migration (T2)*------------------   ------ *Waiting (T3)* -------

Pop_{t1}

$Residual$ *(potentially)* → No U.S. signs of life (IRS 1040s *or IRS 1099s*)

→ Foreign-born U.S. residents (IRS 1040s *or IRS 1099s*)

$Confirmed$ → Filed from abroad (IRS 1040s)

$Residual$ → No U.S. signs of life (IRS 1040s *or IRS 1099s*)

Pop_{t2} → U.S. signs of life (IRS 1040s *and IRS 1099s*)

Pop_{t3} → U.S. signs of life (IRS 1040s *and IRS 1099s*)

$Deaths_{t1-2}$ → Deaths

United States® Census Bureau

9

Emigration Rates

Source: ACS-ACS unadjusted residual estimates; Linked IRS 1040/1099 (TYs 2009-2020) and Numident (2022). DRB Approval Numbers: CBDRB-FY2022-CES010-011 & CBDRB-FY2022-CES010-028.

# Emigration Rates (with Previous Estimates)

Emigration Rates by Years of Residence in U.S.

- Most prior research finds that emigration rates decrease as years of residence in the U.S. increase:
  - Those residing in the U.S. for 10 years or fewer emigrate at rates between 1.3 and 4.4 percent.
  - More established foreign-born residents (11+ years) emigrate at rates between 0.2 and 1.4 percent.

Prior Studies	Years of Residence in U.S.		
	0-10 Years	11+ Years	Total
Warren & Peck (1980)	4.4	-	1.2
Schwabish (2011)	2.3	-	1.3
Mulder et al. (2003)	0.3	1.4	0.9
Scopilliti (2011)	1.4	0.4	-
Leach & Jensen (2013)	1.3	0.2	0.6

Note: Rates per 100 foreign-born individuals.
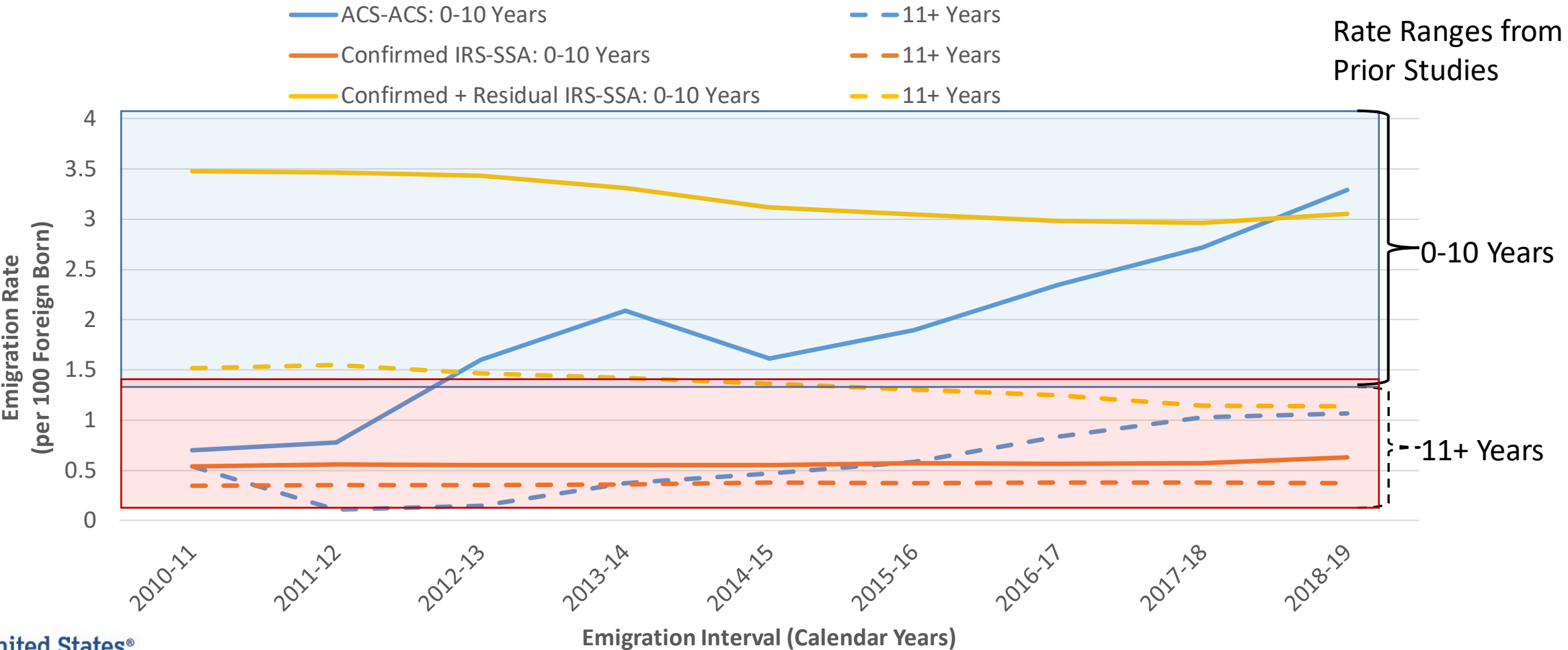
United States® Census Bureau

Emigration Rates by Years of Residence in U.S.

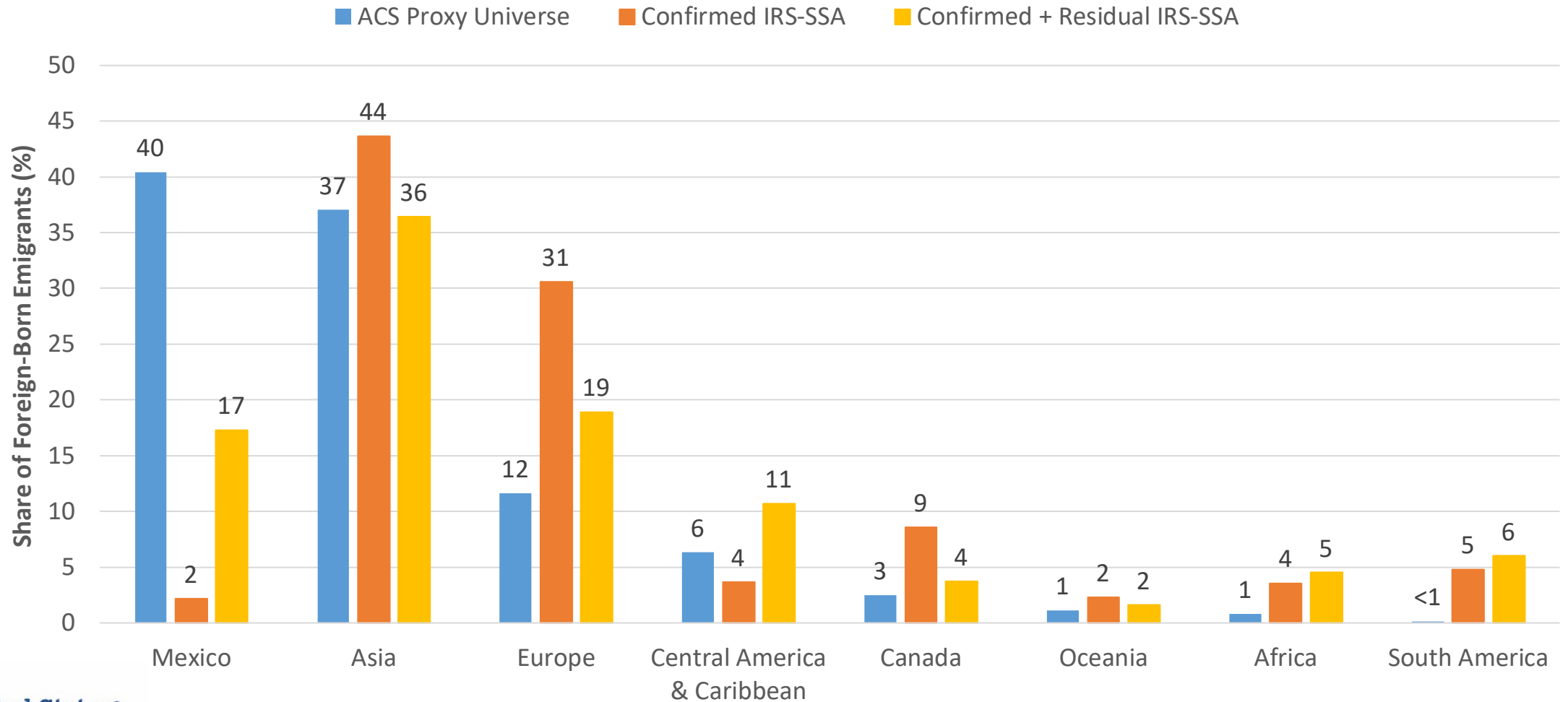Source: ACS-ACS unadjusted residual estimates; Linked IRS 1040/1099 (TYs 2009-2020) and Numident (2022). DRB Approval Numbers: CBDRB-FY2022-CES010-011 & CBDRB-FY2022-CES010-028.

Foreign-Born Emigrants by Age: 2018-2019

Source: Internal PEP foreign-born emigrant file; Linked IRS 1040/1099 (TYs 2016-2019) and Numident (2022). DRB Approval Numbers: CBDRB-FY2022-CES010-011 & CBDRB-FY2022-CES010-028.
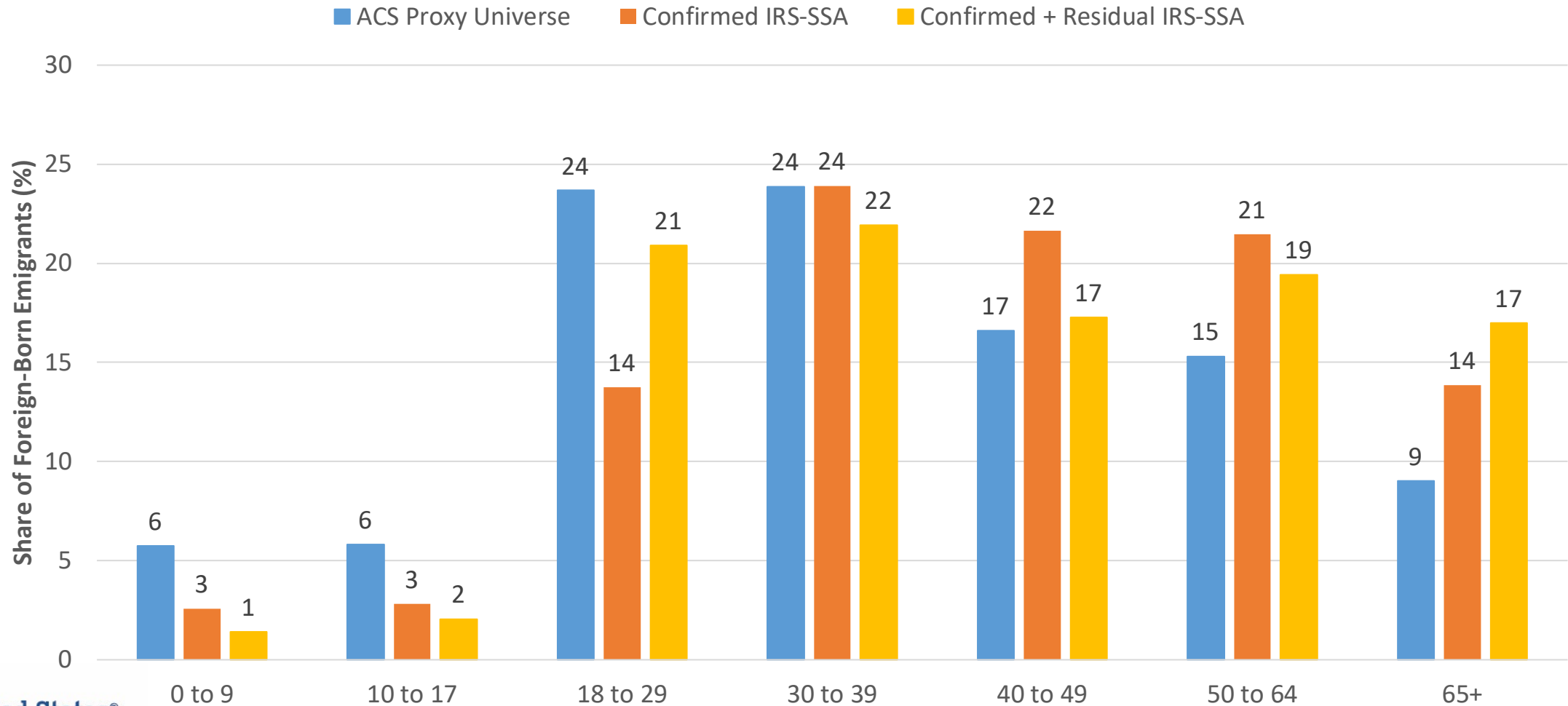
Foreign-Born Emigrants by Birthplace: 2018-2019

Legend: ■ ACS Proxy Universe ■ Confirmed IRS-SSA ■ Confirmed + Residual IRS-SSA

Y-axis: Share of Foreign-Born Emigrants (%)

Birthplace	ACS Proxy Universe	Confirmed IRS-SSA	Confirmed + Residual IRS-SSA
Mexico	40	2	17
Asia	37	44	36
Europe	12	31	19
Central America & Caribbean	6	4	11
Canada	3	9	4
Oceania	1	2	2
Africa	1	4	5
South America	<1	5	6

United States® Census Bureau

Source: Internal PEP foreign-born emigrant file; Linked IRS 1040/1099 (TYs 2016-2019) and Numident (2022). DRB Approval Numbers: CBDRB-FY2022-CES010-011 & CBDRB-FY2022-CES010-028.

Foreign-Born Emigrants by State: 2018-2019

Source: Internal PEP foreign-born emigrant file; Linked IRS 1040/1099 (TYs 2016-2019) and Numident (2022). DRB Approval Numbers: CBDRB-FY2022-CES010-011 & CBDRB-FY2022-CES010-028.
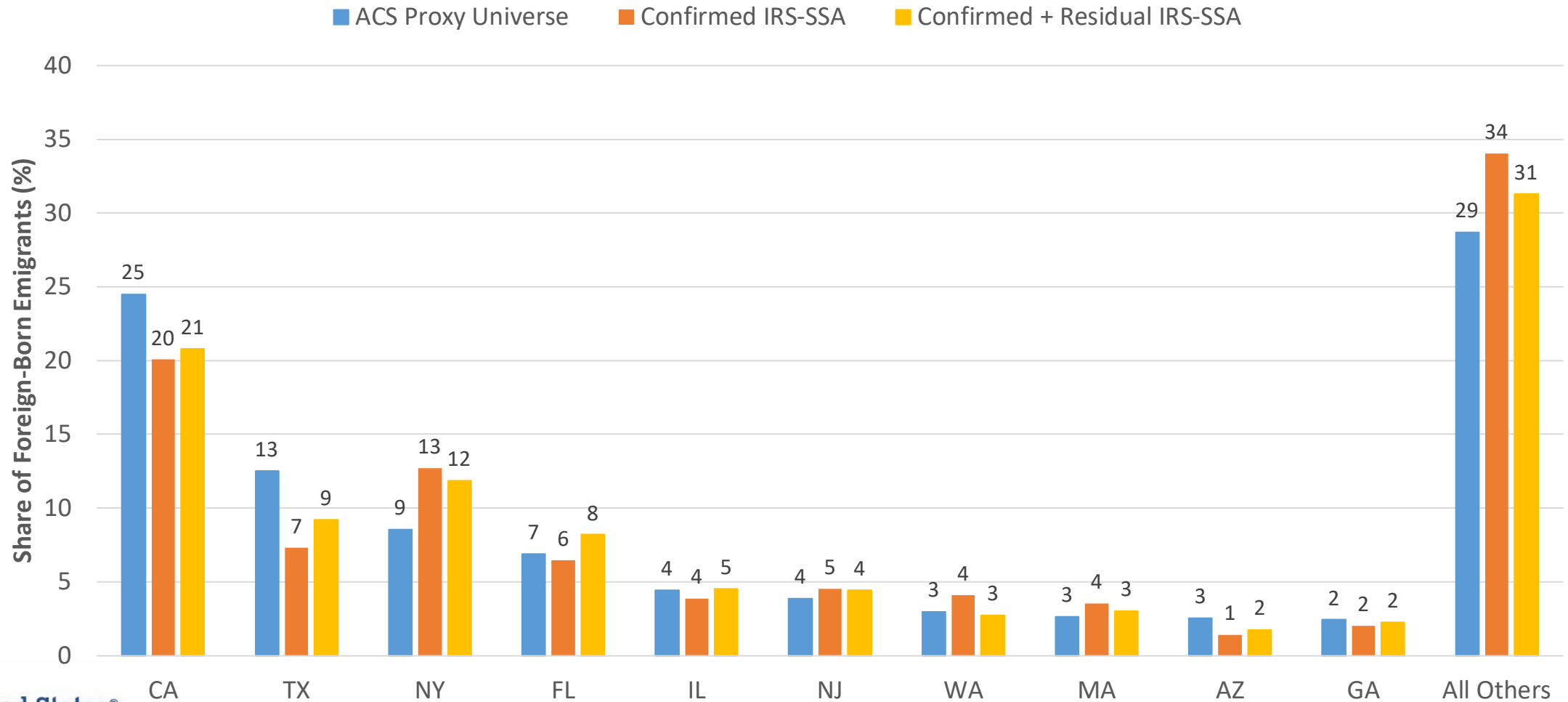
Discussion

- IRS and SSA data are not designed to measure emigration, but results show promise

- Improvements to previous method increase coverage and reduce false positives in the residual component
  - "Signs-of-Life" methods take advantage of the longitudinal nature of the data
  - Inclusion of IRS information returns capture non-filers and inconsistent filers

- However, clear signs of selection bias remain
  - IRS-SSA distribution of emigrants skews older than ACS residual estimates
  - IRS-SSA data undercount emigrants born in Mexico and leaving CA and TX

- And, declining emigration may reflect increasing coverage of tax data

United States® Census Bureau

Next Steps

- Explore methods to adjust for increasing coverage of the foreign-born population by tax data

- Potentially incorporate additional administrative data
  - Composite address or residential history files
  - Medicare beneficiary data
  - Kid link files

- Explore methods for incorporating the ITIN population into these estimates
  - Mexico-born population overrepresented among those in the tax data with ITINs?

Thank you!

Brad Foster
Center for Economic Studies
U.S. Census Bureau
thomas.b.foster@census.gov

Anthony Knapp
Population Division
U.S. Census Bureau
anthony.knapp@census.gov

Erik Vickstrom
Center for Economic Studies
U.S. Census Bureau
erik.r.vickstrom@census.gov