

Integrated Database for International Migration: Foreign-born Immigration to the United States

Peter Borsella

International Migration Branch

UNECE Online Meetings on Migration Statistics

27-28 October, 2021

Integrated Database for International Migration: IDIM

- Estimates of foreign-born immigration to the United States using American Community Survey (ACS) data are robust, but with some caveats
 - Increased variance at lower levels of geography and smaller demographic classifications
 - Time delay between data collection and processing
- Linking administrative and survey data can overcome these weaknesses
 - Administrative data are comprehensive for their population universe, which should allow for more accurate estimation of small classifications
 - Administrative data are available much sooner than survey data (3-4 months vs 1-2 years)

Numident (Social Security Administration)

- Social Security Number (SSN), unique identifiers for every native-born citizen, naturalized foreign-born citizen, and foreign-born non-citizen authorized to work
 - Does not include unauthorized migrants
- Includes demographic data (age, sex, race) and citizenship data (non-citizen/naturalized, country and city of birth), but no geographic data for current residence
- Can use entry into the Numident combined with citizenship status as a proxy for foreign-born immigration; however, significant over/under-estimation of certain populations
 - Includes all applications for SSNs, ignoring if an immigration event in fact occurred
 - Does not include non-working dependents of non-citizens authorized to work
- Need to identify short-term migrants, exclude non-migrants, include non-working dependents of immigrants, and match/assign geographic data

Tax Filings

(Internal Revenue Service 1040 Form)

- SSNs for primary tax filer as well as dependents
 - Individual Taxpayer Identification Number (ITIN) are used for non-citizens not eligible for a SSN
 - Does not include workers who fall under income threshold for filing, those who fail to file taxes, or those who are not claimed as dependents
- Theoretically can use ITIN in combination with geolocation as a proxy for foreign-born immigration; however, ITIN data has proven unreliable
 - For primary filers, ITINs are not representative of the immigrant population
 - For dependents, changes in tax rules in 2017 mean ITINs are no longer assigned in many cases
- Includes geolocation data for where taxes were filed from, but no demographic data
- Needs some way to identify the foreign born and match/assign demographic data

IDIM (cont.)

- Linking the Numident to tax filings using SSN solves some of the issues with using these data sources individually for estimating foreign-born immigration

Source	Issue	Solution
Numident	Excluding non-migrants	Can confirm entry into the Numident is a migration event by checking for “signs of life” in tax filings
Numident	Identify short-term migrants	Can estimate length of stay using consecutive years of tax filings
Numident	Including non-working dependents of migrants	Can create observations for non-working dependents using tax filing dependent claims
IRS 1040 Form	Identifying foreign-born migrants	Can identify foreign-born tax filings by linking to Numident database, with new filings as a proxy for entry into the United States

- Demographic data primarily sourced from Numident, geographic data primarily sourced from tax filings, with ACS as a secondary source for both demographic and geographic data

Figure 1. Foreign-born Immigration Totals

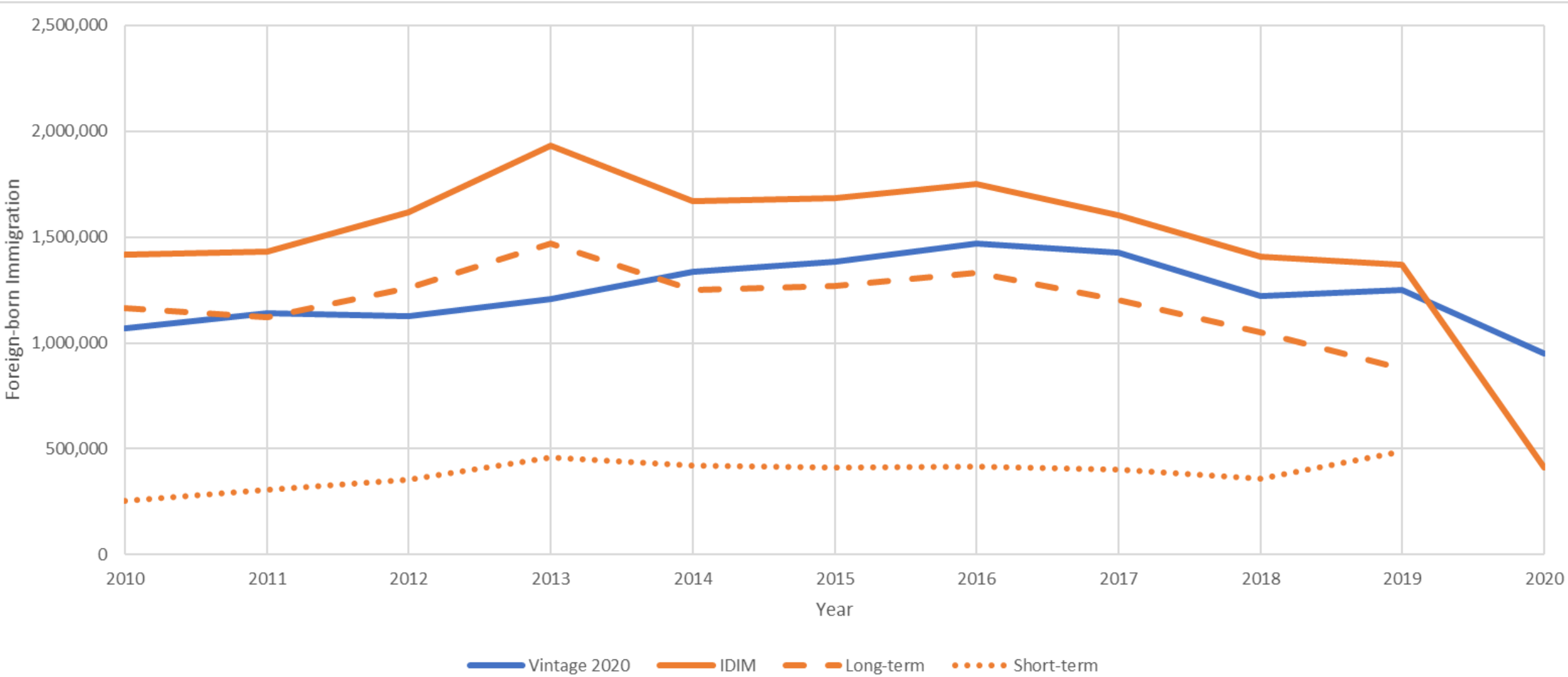
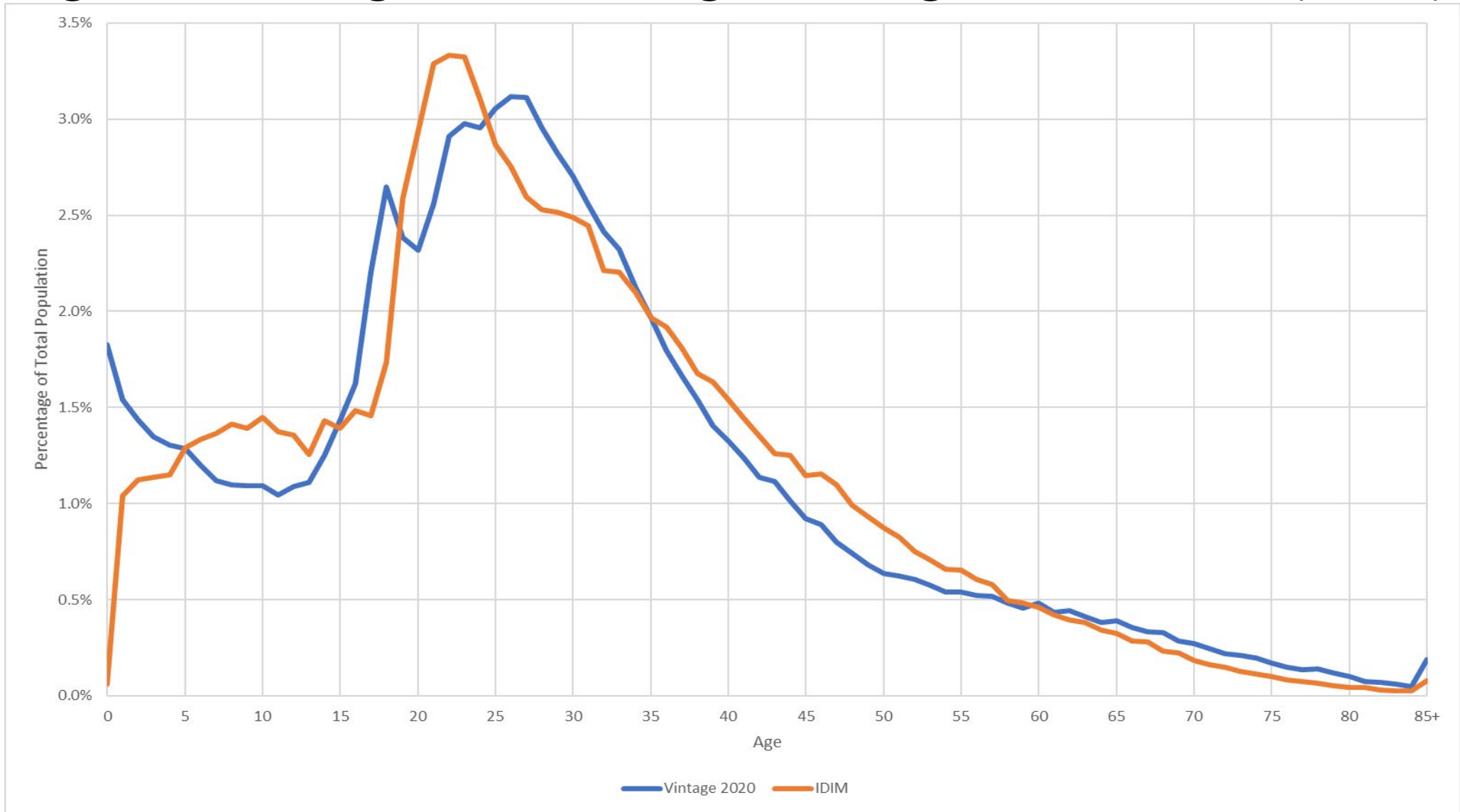


Figure 2. Foreign-born Immigration Age Distribution (2019)



Sources: Social Security Administration, Internal Revenue Service, US Census Bureau Vintage 2020 Estimates

Conclusion

- Preliminary results are promising
 - IDIM estimates of national foreign-born immigration are similar in both levels and trends over time to estimates from the Census Bureau Population Estimates Program (PEP)
 - Age/sex distributions are reasonable, with some noted exceptions
 - IDIM characteristics come from single years of data, whereas PEP estimates come from pooled 3-year ACS data
- Some caveats
 - Clear underestimation of young children < 5 years old
 - Missing student migrant flows
 - Missing unauthorized/irregular flows
- Next steps
 - Estimate race/Hispanic origin distributions
 - Evaluate sub-national geographies (state level looks good so far)
 - Develop foreign-born emigration method
 - Link more data sources

Contact Info

Peter Borsella

U.S. Census Bureau

Population Division, International Migration Branch

c.peter.borsella@census.gov