Global Migration and Demography

Residual estimates of unauthorized immigrants for the US and states:
Methods, data and assumptions

Jeffrey S. Passel

Senior Demographer

Pew Research Center

Pew Research Center is a nonpartisan <u>fact tank</u> that informs the public about the issues, attitudes and trends shaping America and the world.

We conduct public opinion polling, demographic research, content analysis and other data-driven social science research. **We do not take policy positions.**

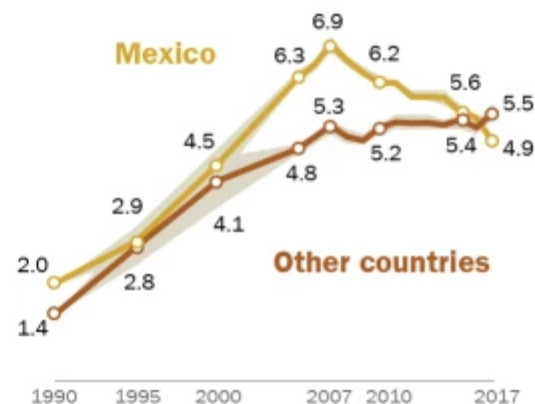We are a subsidiary of The Pew Charitable Trusts, our primary funder.

Pew Research Center

www.pewhispanic.org

Pew Research Center *Hispanic Trends*

HISPANIC | JUN 12, 2019

Mexicans Decline to Less Than Half the U.S. Unauthorized Immigrant Population

For the first time, Mexican unauthorized immigrants no longer account for the majority of the 10.5 million living in the U.S. illegally.

- Interactive: Unauthorized immigrant population trends
- Video: How we estimate the U.S. unauthorized immigrant population

U.S. unauthorized immigrants by origin, in millions

Pew Research Center

Today's presentation

- **Measurement methods**
  - Description of ***Residual Method***
  - Who is **"authorized"** vs. **"unauthorized"**
  - Microdata **status assignments** and state estimates

- **Data and assumptions**
  - Survey data, microdata and variables needed
  - Key assumptions
  - Alternative estimates

- **Sample results**
  - Potential impact on apportionment

But, first...

Unauthorized immigrants grew 500K/year to reach 12+ million in '07; Sudden reversal, DROP of 500K/year; Slower declines to '17 as population drops to '04 (!) level

Unauthorized immigrants (millions)

12.2

11.3

11.1

11.0

10.7

8.6

5.7

3.5

10.5

Shaded area represents 90% confidence interval

1990   1992   1994   1996   1998   2000   2002   2004   2006   2008   2010   2012   2014   2016

Pew Research Center

Outline of Pew methodology

- **Estimate unauthorized in survey (A)**
  - **Residual** = survey "immigrants" *minus* lawful (LPRs)
  - Demographic estimate of LPRs from admin records

- ~~Adjust for omissions~~ (not needed for Census 2020)
  - Based on several external data sources

- **Assign legal status in microdata (B)**
  - Create status variable for individuals, families, h'holds
  - Deterministic for legal temps, refugees, naturalized
  - Probabilistic "multiple imputations" for lawful/unauthorized

- **Final estimates**
  - "Analytic" or "direct" residual estimate — US & 6 states (A)
  - "Tabulated" estimate — remaining 44 states and DC (B)
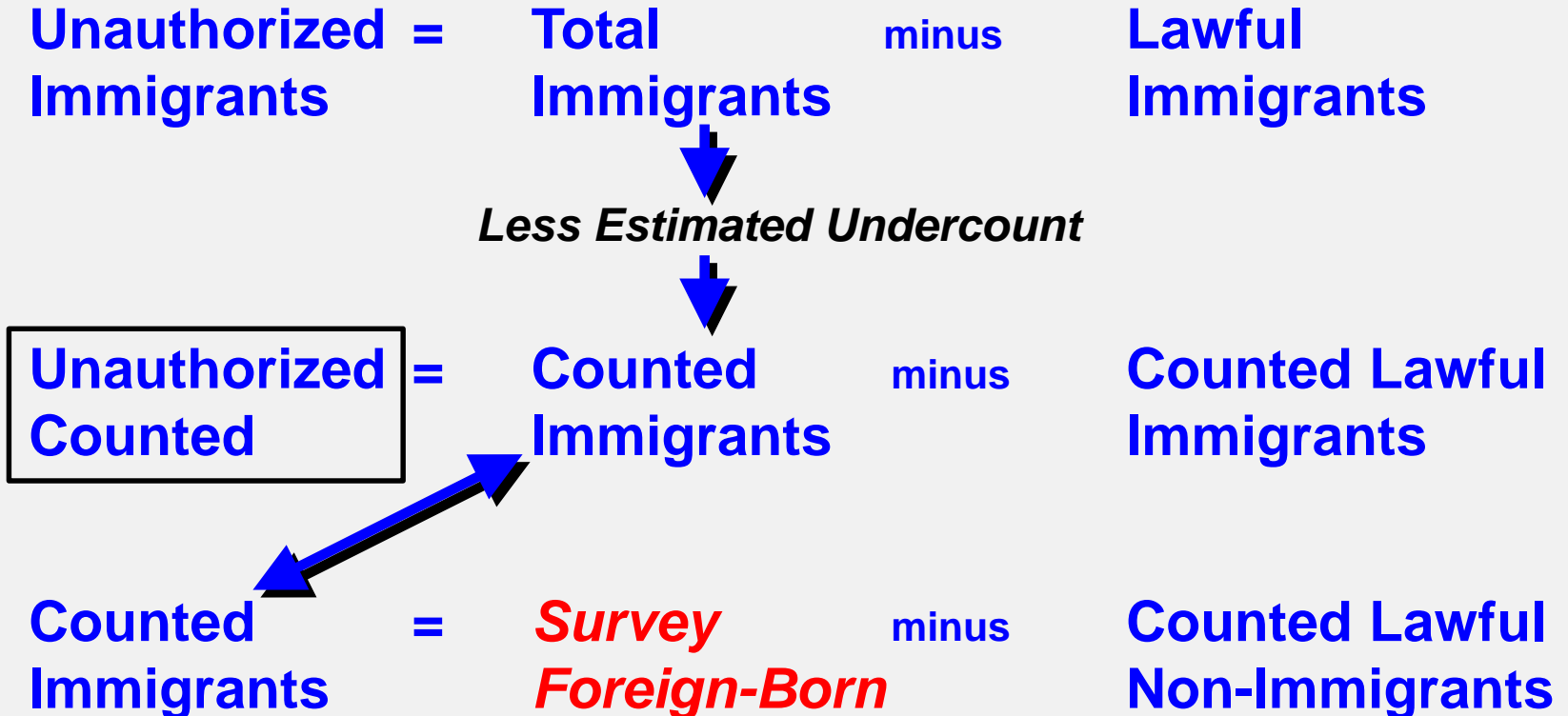
Estimating the population

Residual estimates of unauthorized immigrants

Unauthorized immigrants =

Total immigrants (Survey)

minus

Lawful immigrants (Estimate)

- *Widely used:*
  - DHS-OIS '05–; Warren '80–'00; Van Hook et al.
  - Passel (et al.) '80–present
  - Binational Study '96
  - Warren & Warren '13 (variant)
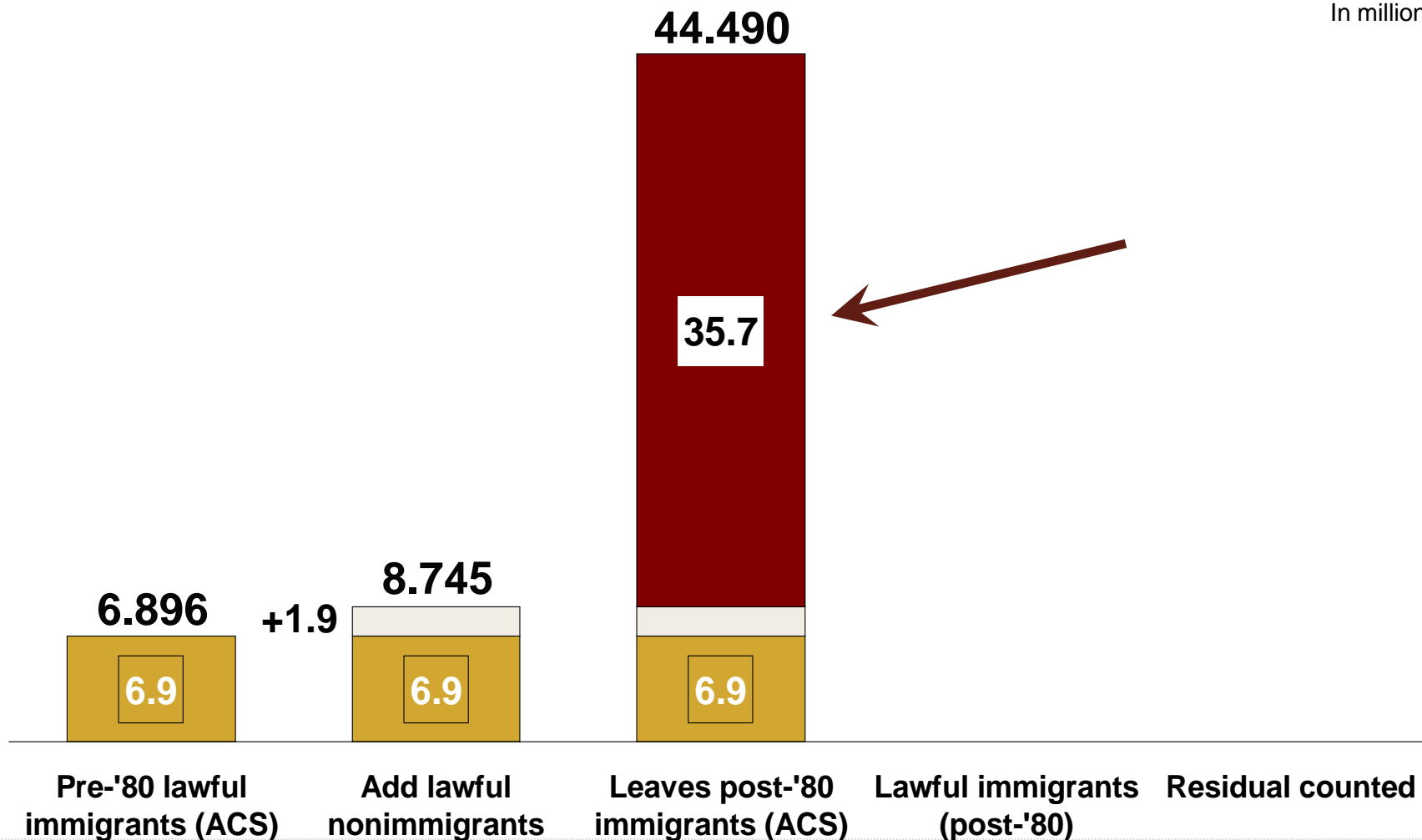
Equations for Pew estimate of unauthorized immigrants

Unauthorized Immigrants = **Total Immigrants** minus **Lawful Immigrants**

Less Estimated Undercount

Unauthorized Counted = **Counted Immigrants** minus **Counted Lawful Immigrants**

Counted Immigrants = *Survey Foreign-Born* minus **Counted Lawful Non-Immigrants**

Note: All populations are for post-1980/1982 entrants.

Pew Research Center

Residual estimate using 2017 American Community Survey

In millions

44.490

35.7

8.745

+1.9

6.896

6.9 | 6.9 | 6.9

Pre-'80 lawful immigrants (ACS) | **Add lawful nonimmigrants** | **Leaves post-'80 immigrants (ACS)** | **Lawful immigrants (post-'80)** | **Residual counted**

Residual estimate of unauthorized immigrants

Total unauthorized immigrants **=** **Counted unauthorized immigrants** **plus** *Missed unauthorized immigrants*

Some assumptions:

- **a. *Estimated undercount***
  - **% Undercount for <u>lawful immigrants</u> – based on Census coverage studies – *age-sex-race* * 1.75 for recent arrivals**
  - **% Undercount for <u>unauthorized immigrants</u> –**
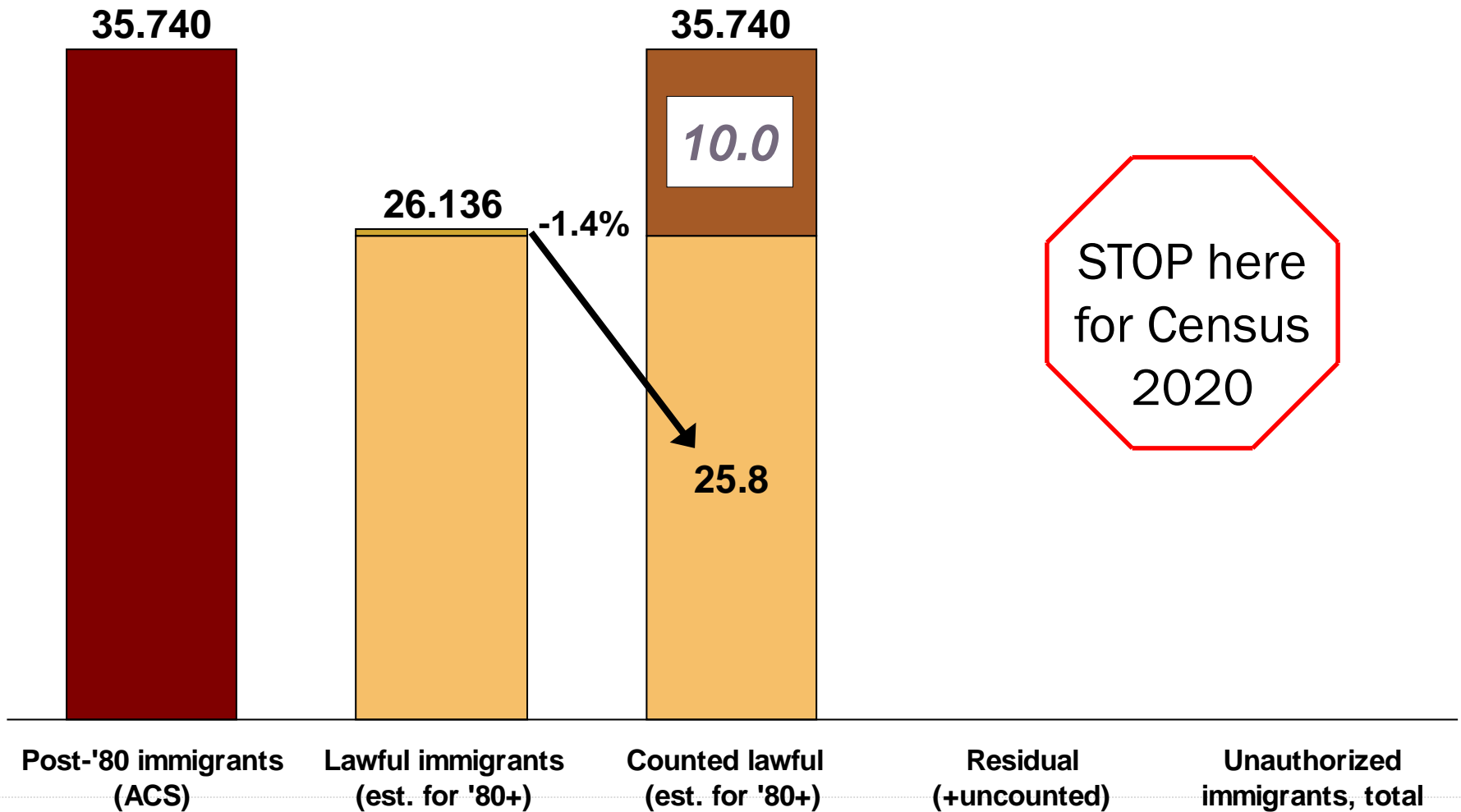 - **1.83 * specific rates for lawful immigrants**
- **b*. Estimated undercount (2)***
  - **New work (Van Hook et al. 2014) shows improvements in ACS/CPS coverage for Mexicans from '90s to late '00s**
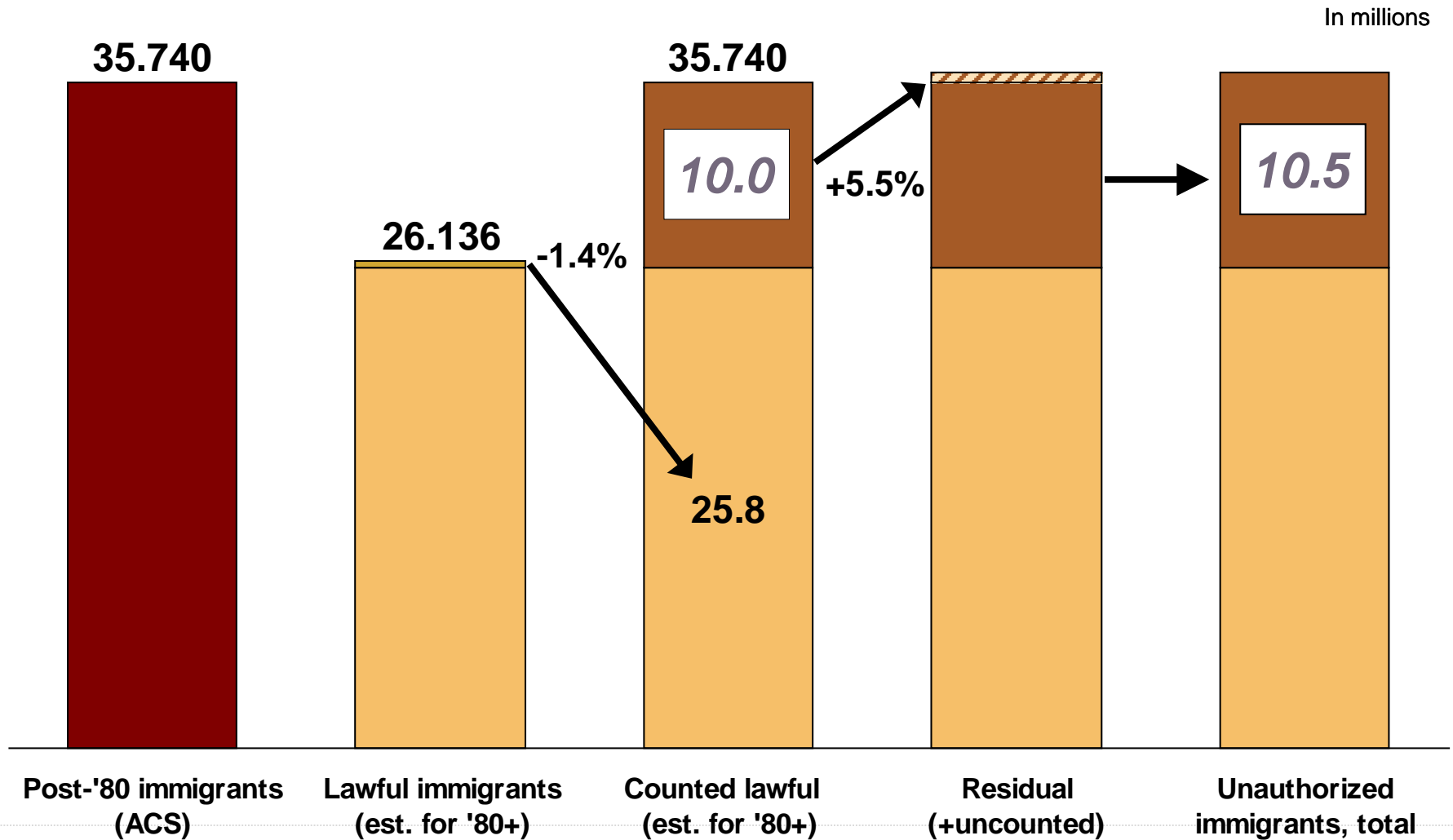- **c. *Internal migration* – ACS rates for F-B for lawful pop**

Residual estimate using 2017 ACS —
Initial estimate of "counted unauthorized"

In millions

35.740

35.740

10.0

26.136

-1.4%

STOP here
for Census
2020

25.8

Post-'80 immigrants
(ACS)

Lawful immigrants
(est. for '80+)

Counted lawful
(est. for '80+)

Residual
(+uncounted)

Unauthorized
immigrants, total

Pew Research Center ❋

Residual estimate using 2017 ACS —
Add uncounted → Total unauthorized

In millions

35.740

35.740

10.0

26.136

-1.4%

+5.5%

10.5

25.8

Post-'80 immigrants (ACS)

Lawful immigrants (est. for '80+)

Counted lawful (est. for '80+)

Residual (+uncounted)

Unauthorized immigrants, total

Pew Research Center

Who are the "Lawful immigrants"?

Components of lawful immigrants:

   a. Refugee arrivals (ORR, DHS/INS, State)

   b. Asylum approvals (DHS/INS)

   c. Cuban-Haitian entrants, Amerasians, some parolees (ORR)

   d. Other entrants, other parolees (DHS/INS)

   e. IRCA legalizations approved—SAWs & "LAWs" (INS)

   f.   DHS/INS "new arrival" green cards (except Amerasians in d.)

   g. Adjustments to LPR status, except from a.-e. above (DHS/INS)

   h. Pre-1980 (or 1982) arrivals (ACS/CPS counts)

a.-g. are combined year-by-year using allowances for *mortality, emigration & state-to-state movement* **to estimate the legal population**

Pew Research Center

Who are the "Unauthorized"?

- Not "lawful" or nonimmigrants (temporary)

- Overstays (~40-45% or more) & EWIs

- "Quasi"-lawful, including:
  a. DACA Beneficiaries (~700,000 in 2017)
  b. TPS, especially Central Americans (~300,000)
  c. Asylum applicants, defensive/affirmative (600,000-900,000)
  d. NCARA & ABC beneficiaries, DED
  e. Adjustment applicants (esp. K, V Visas)

- Overlapping categories (& little data)

- Possibly *one-sixth* of unauthorized in a.-c.

Status assignments
(microdata)

Types of data and estimates

- ***Analytic*** estimates of unauthorized
  a. "Counted" in survey
  b. ***Totals by country/region birth*** (i.e., Mexico, Latin America, Asia, rest of world)
  c. ***6 States*** (CA, FL, IL, NJ, NY, TX) and balance of US
  d. Some, limited demographic information:
    * ***Age (<18/18+)*** and sex
    * Period of entry
  e. "Corrected" for undercount

- ***Status assignments*** into survey

- ***Final estimates*** may differ from ***analytic*** totals for some countries, age groups, regions, etc.

Microdata status assignments (i) — Groups & survey/census data used

- **_Lawful temporary immigrants_**
  a. Year of arrival
  b. Occupation, industry, school enrollment & income
  c. Household relationships **
  d. Program participation (e.g. food stamps)
  e. [No targets — deterministic]

- **_Refugee & asylee arrivals_**
  a. Country of birth
  b. Year of arrival
  c. [Demographic estimates (targets) control assignments]

- **_Naturalized citizens_**
  a. Reported citizenship and country/region of birth
  b. Citizenship of spouse
  c. [No targets — deterministic]

** Items in decennial census.

Microdata status assignments (ii) — Groups & survey/census data used

- ***"Definite" lawful immigrants***
  a. Naturalized citizens, refugees/asylees, legal temps (i)
  b. Entered before 1982
  c. Occupation and industry (e.g., licensed, government)
  d. Program participation (e.g. food stamps)
  e. Family members ** of lawful immigrants
  f. [No targets — deterministic]
- ***Unauthorized immigrants ("potential")***
  a. Country of birth, year of arrival
  b. Family relationships **
- ***Unauthorized immigrants (final assigned)***
  a. [Demographic residual estimates are targets]
  b. Model-based random assignments within groups
  c. Family-based edits for consistency
  d. Iterative assignments until estimates agree with targets

Status assignments *(I)*

- ## *Lawful temporary immigrants*
  a. Students (F, M visas) & scholars (J visa)
  b. "High-tech" guest workers (H-1B), intracompany transfers (L visas), nurses (H-1A) and physicians (J visas)
  c. International organizations (G visas), diplomats and embassy employees (A visas)
  d. Religious workers (R visas), athletes, artists, entertainers and "outstanding" talent (O,P visas)
  e. Au pairs and exchange visitors (J visas)
  f. Based on occupation, period of entry, relationships in household, other characteristics (e.g. programs)
  g. No targets, tends to understate DHS estimates
  h. Many more in ACS than CPS

Status assignments *(II)*

- ***Refugees/asylees*** (at entry)
  a. Country of birth
  b. Year/period of entry
  c. Demographic estimates (targets) control the assignments

- ***Naturalized citizens***
  a. In US < 6 years → edit to alien (except spouses)
  b. In US >= 6 years →
     - ***Mexico*** → potential unauthorized (revert to naturalized)
     - ***Central America*** → same as Mexico
     - ***All others countries*** → keep as naturalized
  c. No demographic targets

Status assignments *(IIa—Lawful)*

- ***Definite lawful immigrants***
  a. Naturalized citizens, refugees, legal temps (previous)
  b. Entered US before 1982/1980
  c. Occupations/employers →
     - Government workers, veterans, military
     - Law enforcement and related occupations
     - Court and legal employees
     - Licensed occupations (e.g., medical)
     - Security occupations (esp. with licensing; e.g., police, ATC)
     - Assorted other minor occupations
  d. Program participation (SSI, TANF, Medicaid*, Medicare)
  e. Family members (most) of lawful immigrants
  f. No demographic targets
- ***Others:*** *"Potential unauthorized"* →

Status assignments *(////-Unauthorized)*

- ***Targets***
  a. 6 ***States*** (CA, FL, IL, NJ, NY, TX) and balance of US
  b. Total, ***Under 18***
  c. ACS: ***Mexico, Latin America, Asia, all other*** (may collapse)
  d. CPS: *Mexico, all other*

- ***Random assignments of Potentials***
  a. Replicate households for "multiple" imputation
     1) Weights are fractionalized (10 in ACS; 100 in CPS)
     2) Random assignments separately for individuals in each replicate
     3) "Like" assignments collapsed at end
  b. Initial ***p***'s = Target (above) / potential unauthorized (previous)
     1) Parents and non-parents done separately
     2) Assignments for parents use child estimates as targets

Status assignments *(IIIa-*Unauthorized*)*

- **Final status assignments of Potentials**
  - a. Household edits for consistency of parents, children & relatives (based on date of arrival)
  - b. Adjustment of *p*'s from (Target /potential unauthorized)
    1) Initial modification based on IRCA LPS occupation groups
    2) Possibility of using other data here (e.g., SIPP, Marcelli)
    3) Adjust each state-region of birth for convergence to targets
    4) Separate adjustments for parents/non-parents by state-region
  - c. Iterate assignments and edits until targets are hit
  - d. Relatively insensitive to initial assignments (a.-b.) since targets are 80–95% of *potential unauthorized targets*

- **Adjust weights for undercount**
  - a. From analytic estimates by state-region of birth-ages

Legal Status data

- **ACS datasets → *"Legal status"* variable**

- **Other variables created**
    a. Nuclear families ("MHUs")including parent-child linkages
    b. Family legal status (hierarchical)
    c. Household legal status (hierarchical)

- ***Uses of data***
    a. Geography — **states (except 6 "targets")**; metro areas
    b. Detailed country of birth (except Mexico "target")
    c. Family/household data by status
    d. Characteristics of legal and unauthorized populations

Selected measurement issues

- **Problematic components**, especially…
  - a. Emigration (of legal residents) since 1982
    - i. Rate declines with time in US
  - b. Mortality (of legal residents) since 1982
  - c. Backlogs for legal admission
  - d. Counts for "quasi-legal" groups, esp. asylum backlogs
- **Categorical issues**
  - a. Definition of US residence (esp. for Mexicans) & status
  - b. Gross flows in/out of immigration categories (and US)
- *Data issues*
  - a. Census/survey undercount (esp. for status groups)
    - i. Coverage improves with time in US
  - b. Accuracy of date of entry variable versus DHS/INS

Pew Research Center

ACS/survey/census issues

- **Needed data items and sample size**
  a. March CPS-ASEC
     a. Detailed income and program participation
     b. ~75K households in clustered sample (full sample)
  b. American Community Survey
     a. Adequate income and program participation
     b. 1.2M households in random sample (public-use data)
  c. Decennial census
     a. Full count but no useful data for estimation
- *Population figures*
  a. Representation in sample determines immigrant counts
  b. Census-based population controls omit nativity
  c. Inconsistencies in controls & weighting hamper trend analysis
- *Government/survey issues*
  a. More microdata, especially from DHS (OIS)
  b. Direct collection of legal status in more surveys
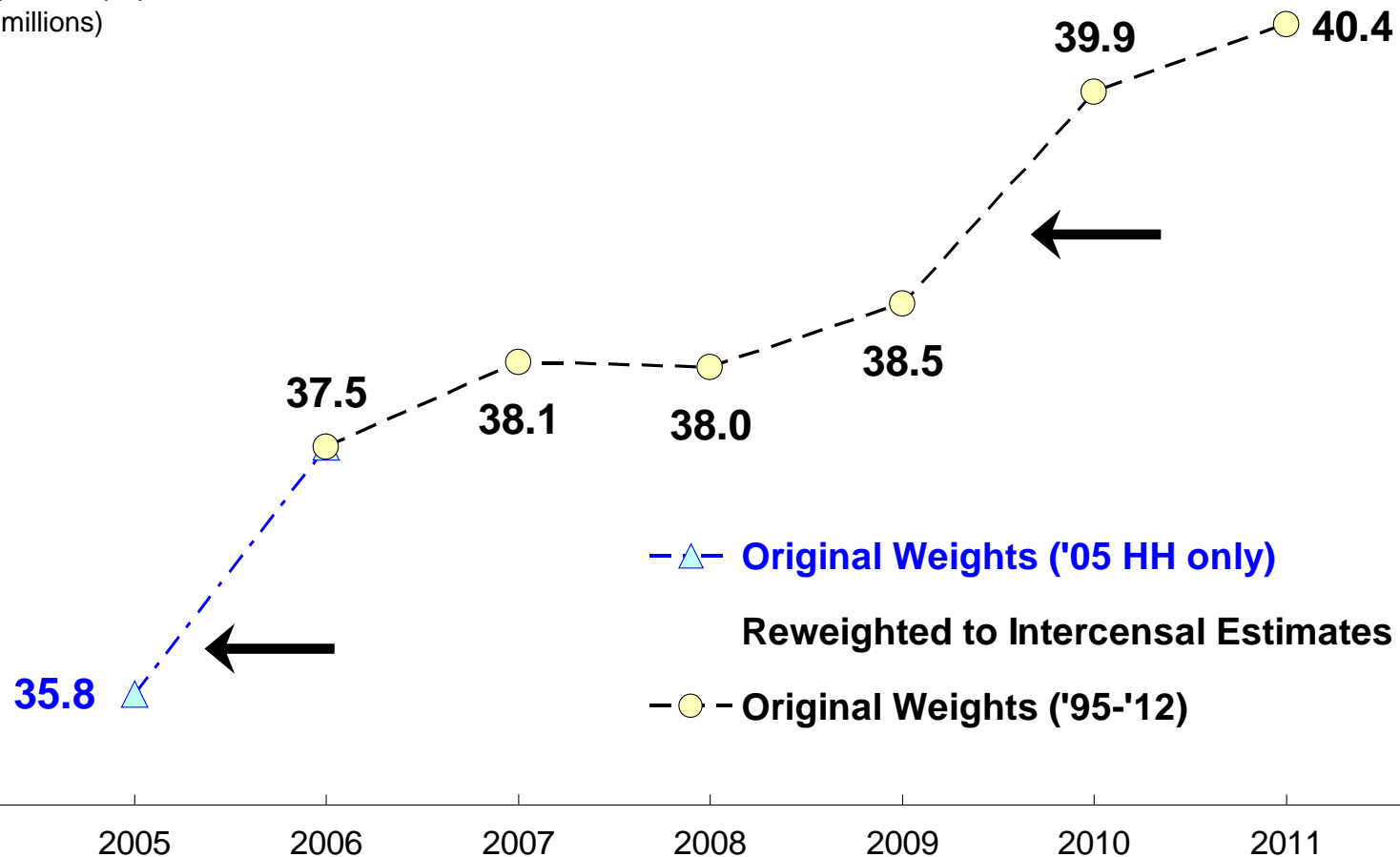
Survey data issues

American Community Survey

- *Universe:* Total population ('06→), households only '05
- *Sample size* (Random, representative, minor stratification)
  a. Full sample: ~3.5 million households (~300,000 per month), 2012+
  b. Subsampling for nonresponse → 2.1 million interviewed households
- *Key variables*
  a. State/country of birth, citizenship, year of "entry"
  b. *No* country of birth of parents
  c. Public-use sample → 1% of US population (3.2+ million cases)
- *Weighting issues*
  a. 12 merged monthly samples; 5 merged years
  b. Comparable geography, 2000-2011, 2012- PUMAs
  c. New population controls every year (!)
  d. Occasional postcensal revisions & new census create breaks in estimates for some groups
  e. Intercensal reweighting a *LARGE* impact on 2007-2009 estimates
  f. Census Bureau *RARELY* revises weights (even if needed)

Key dates for population controls

- ***Pre-2010 population estimates***
  a. Immigration revisions in Vintages 2007, 2008, 2009, 2010
- ***Census 2010-based controls***
  a. Large revisions to Asians and Hispanics
  b. Intercensal estimates revise population data for 2001-2009
  c. No revisions to weights
- ***Revisions to weights and data sets***
  a. Census Bureau almost never revises weights
  b. Pew has ***revised survey weights*** (using methods that replicate, to the extent possible, the Census Bureau's methods)
     1) ACS 2005-2009 for intercensals
     2) March CPS 1995-1999, 2001-2009 (Census revised 2000, 2010, 2011)
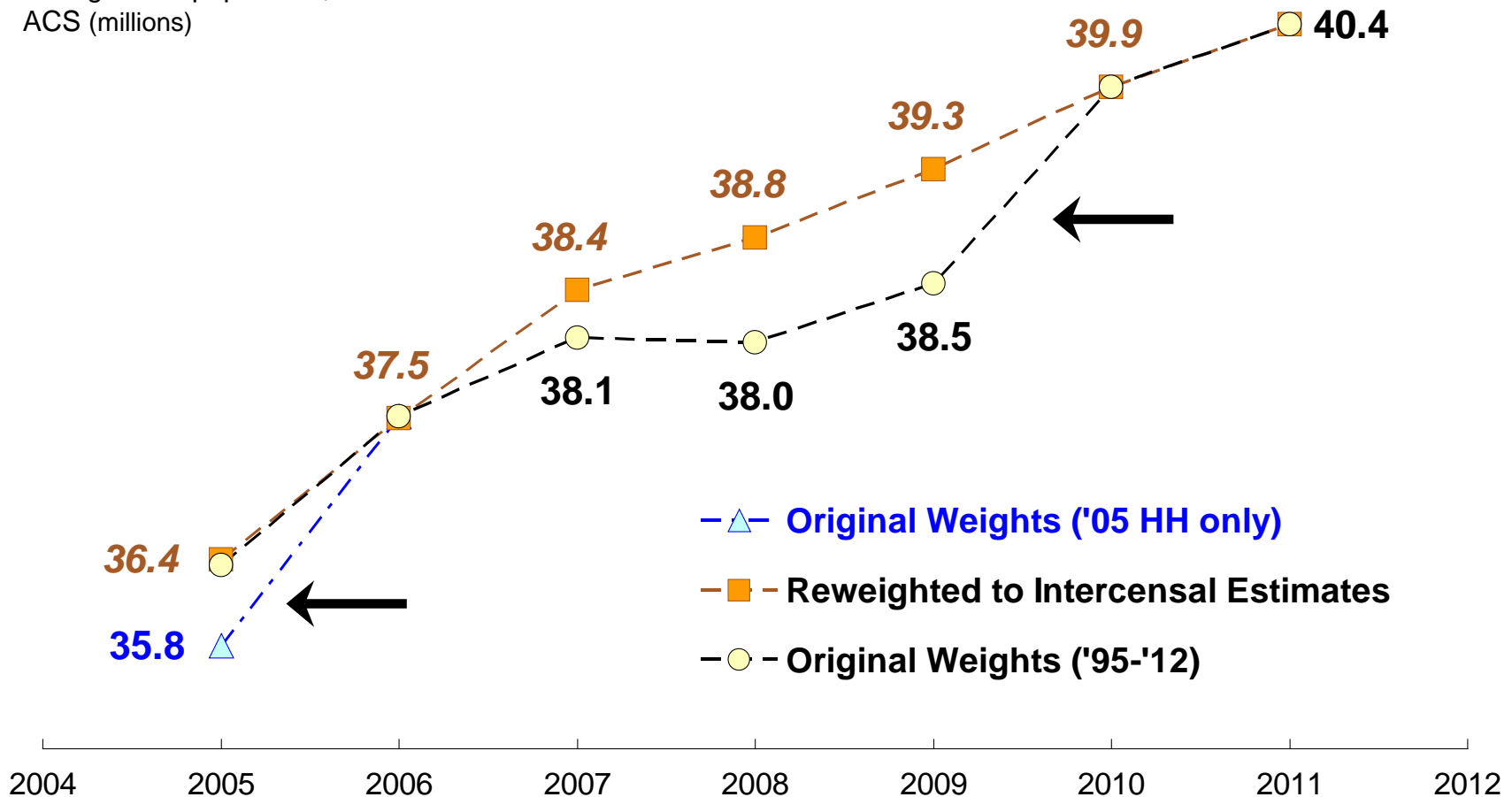 c. Pew has created new data for group quarters population in 2005 ACS to provide data for entire US population

Pew Research Center

ACS foreign-born trend is *erratic*— especially '05-'06 & '09-'10

Foreign-born population, ACS (millions)

Source: Pew Research Center estimates

Consistent ACS weights matter for '05, '07-'09 and year-to-year

Foreign-born population,
ACS (millions)

39.9   40.4

39.3

38.8

38.4   38.5

37.5

38.1   38.0

36.4

35.8

— △ — **Original Weights ('05 HH only)**

— ■ — **Reweighted to Intercensal Estimates**

— ○ — **Original Weights ('95-'12)**

2004   2005   2006   2007   2008   2009   2010   2011   2012

Pew Research Center

Source: Pew Research Center estimates

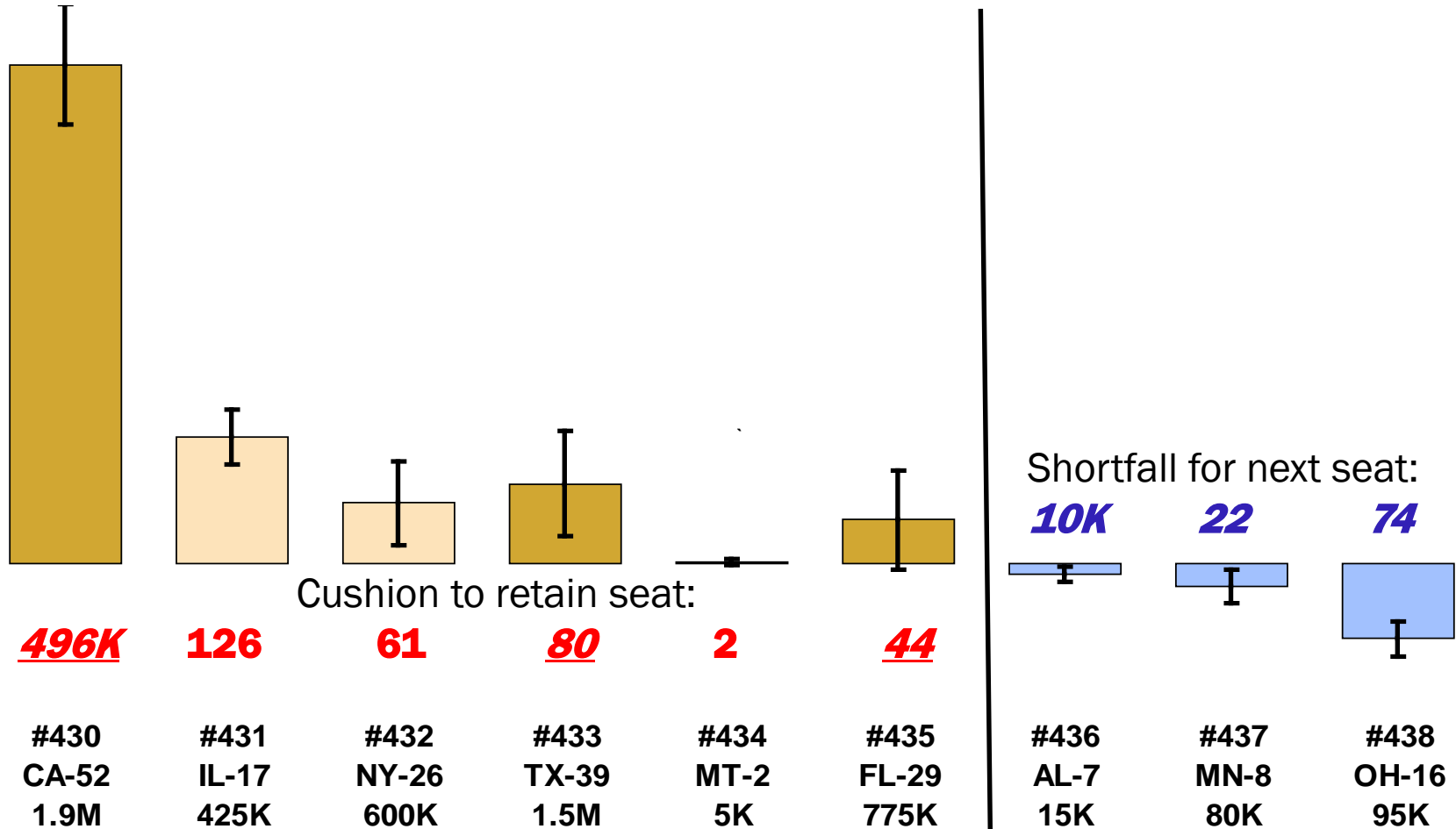# Potential impacts

Illustrative apportionment

- **Method of Equal Proportions**
  a. Each state gets one seat
  b. Priority value for seat $N = P \div$ sqrt$[N*(N-1)]$
  c. Seats assigned sequentially
  d. Assignments stop at seat 435

- *Populations for April 1, 2020*
  a. Census Bureau Vintage 2019 population estimates
  b. Extrapolated from July 1, 2019 to April 1, 2020

- *Unauthorized immigrants*
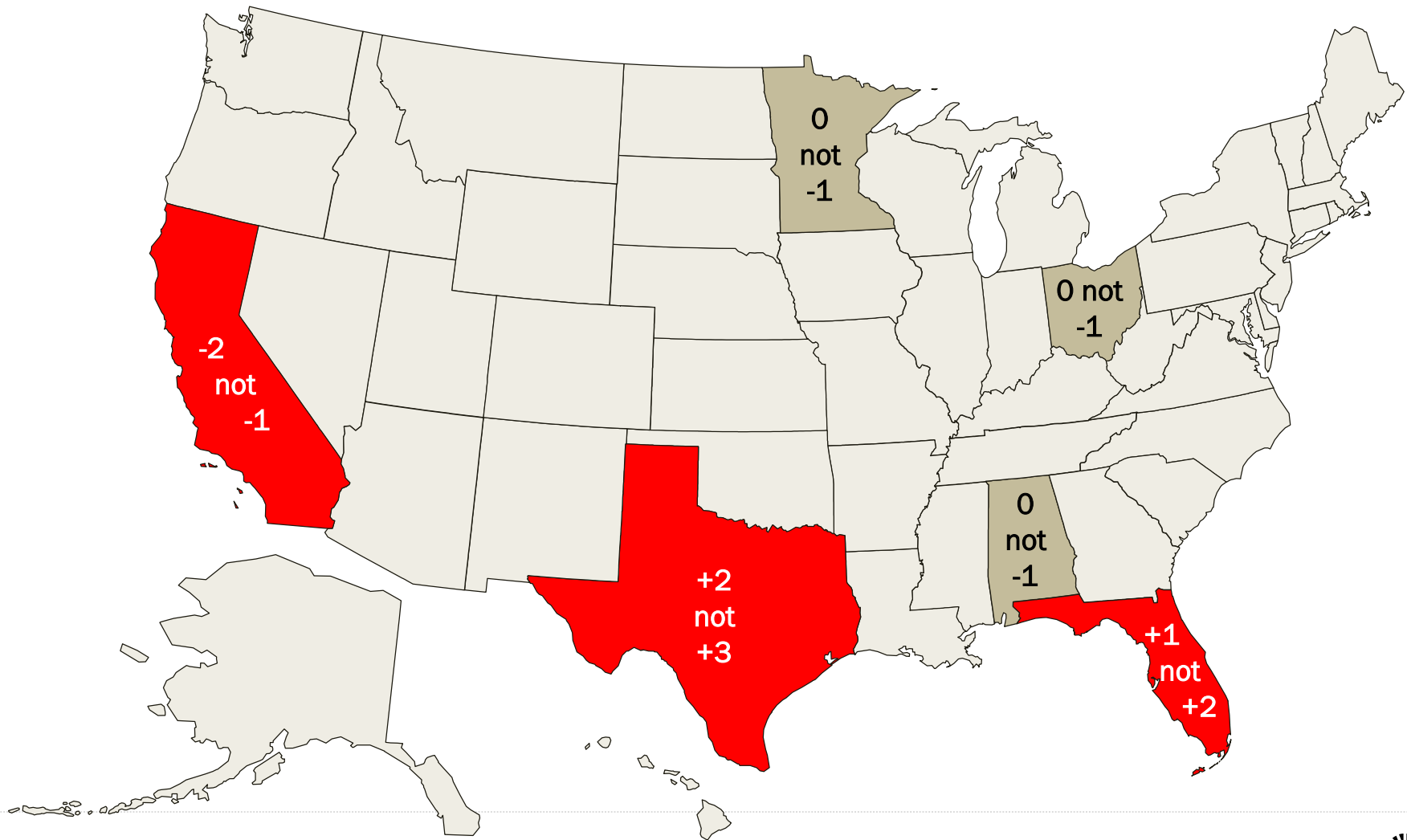 a. Pew estimates for 2017

- *Different from apportionment population*

Apportionment change, 2010-2020
(Projected, total population)

+1 (Oregon)
+1 (Montana)
-1 (Minnesota)
-1 (Michigan)
-1 (Vermont)
-1 (New York)
-1 (Massachusetts)
-1 (Pennsylvania)
-1 (California)
-1 (Illinois)
-1 (Ohio)
-1 (West Virginia)
+1 (Colorado)
+1 (Arizona)
+1 (North Carolina)
-1 (Alabama)
+3 (Texas)
+2 (Florida)

Last 6 seats assigned and next 3 seats

Populations in thousands

Cushion to retain seat:

496K   **126**   **61**   **80**   **2**   **44**

Shortfall for next seat:

10K   **22**   **74**

#430	#431	#432	#433	#434	#435	#436	#437	#438
CA-52	IL-17	NY-26	TX-39	MT-2	FL-29	AL-7	MN-8	OH-16
1.9M	425K	600K	1.5M	5K	775K	15K	80K	95K

Source: Pew Research Center based on projected 2020 populations (Passel and Cohn, July 24, 2020.

Pew Research Center

Additional change for 2020 excluding unauthorized immigrants (estimated)

0 not -1 (Minnesota)

0 not -1 (Ohio)

-2 not -1 (California)

+2 not +3 (Texas)

0 not -1 (Alabama)

+1 not +2 (Florida)

Pew Research Center

Contact Information

Jeffrey S. Passel
Senior Demographer

jpassel@pewresearch.org
~~202-419-3625~~ (direct, not currently in use)
202-527-2146 (mobile)

Global Migration and Demography