Methodology for Preparing the 2019 Demographic Estimates

Xiaojin Wang and Katherine Nesse

Introduction

Demographic data have been an important part of IECAM's early childhood data collection. The main source of demographic data is the U.S. Census's American Community Survey. We have the following demographic data for young children in Illinois: population by age/age groups, by race/ethnicity; poverty, parental employment, and household language use. Available geographies include a wide range of regions in the state of Illinois from counties to school districts, we also have a few special-purpose regions, such as the ISBE and IDHS regions.

Selection of data sources

The Census Bureau's Population Estimates Program (PEP) produces the official estimates of the population for the nation, states, counties, cities and towns, and estimates of housing units for states and counties. For counties, PEP only releases population estimates for five-year age groups (see in PEPAGESEX). For 2010 and other decennial census years, the decennial census provides the official counts of population and housing units. (U.S. Census Bureau, 2018). CDC's National Center for Health Statistics (NCHS) releases single-year of age at the county level data produced by the PEP in collaboration with NCHS. Nevertheless, these estimates have limited reliability. The Census does not release bridged-race or unbridged estimates by single-year of age at the county level due to concerns about the reliability of these estimates. However, these estimates are provided to the National Center for Health Statistics to meet programmatic needs such as the creation of age groupings that differ from the standard groupings used by the Census Bureau. Users of the single-year-of-age county-level bridged-race population estimates should carefully consider the limited reliability of these estimates (NCHS, 2018).

Census Bureau's American Community Survey (ACS) also produces population, demographic, and housing unit estimates. ACS is the recommended source to obtain population characteristics (percents, means, medians, and rates) rather than estimates of population totals. The official population count — including population by age, sex, race and Hispanic origin — comes from the once-a-decade census, supplemented by annual population estimates (the PEP) (Census, 2010). However, ACS data are estimated by using population estimates by sex, age, race, and Hispanic origin, and estimates of total HUs produced by the Population Estimates Program (PEP) to increase precision and ensure consistency. ACS estimates are controlled to official population and housing units at the county level (Census, 2014; 2018). ACS releases 1-year and 5-year estimates[1], which are all "period" estimates that represent data collected over a period of time (as opposed to "point-in-time" estimates, such as the decennial census, that approximate the characteristics of an area on a specific date). 5-year estimates are the most reliable as they have the largest sample size, and have the data for smaller geographic areas down to the census tract and block-group level, compared to 1-year estimates. For this reason, ACS 5-year estimates are used for IECAM's demographic data. However, it's the least current[2] (Census, 2018).

[1] 1 The Census Bureau previously released 3-year estimates based on 36 months of data collection. In 2015, the 3-year products were discontinued. The 2011–2013 ACS 3-year estimates, released in 2014, are the last release of this product.

[2] ACS Data Releases. www.census.gov/programs-surveys/acs/news/data. This Web page includes information about the ACS data release schedule, guidance on using the latest ACS data, and technical information about geography and product changes. Users can also browse the notes from previous years.

Demographic data of previous years use data from PEP and ACS[3]. PEP's annual release of a new vintage of estimates supersedes any previous series and incorporates the most up-to-date input data and methodological improvements. The annual time series of estimates begins with the most recent decennial census data and extends to the vintage year (Census, 2017a). When preparing the new data, IECAM also updates the two year before the vintage year. For instance, the 2015 release IECAM estimates use PEP v2015 and 2009-2013 ACS 5-year for 2015 for each estimate year: 2013, 2014, and 2015. For the ease of data preparation, we use ACS 5-year estimates data only for 2016 demographic data.

There are some caveats about using the ACS data: all ACS data are estimates based on survey data, so they are subject to errors (ACS publishes margins of errors along with estimates); 5-year estimates include data collected over a 60- month period, therefore, ACS multiyear estimates reflect the averages of the full period; there is high undercount of children under the age of 5 (Census, 2017b)[4]. See more how about ACS data in Understanding and Using American Community Survey Data: What All Data Users Need to Know (Census, 2018).

Population estimation

Early childhood programs are often designed for different age groups. For instance, Head Start and Preschool for All (PFA) are for children ages 3-5, while Early Head Start and Prevention Initiative are for children ages 0-3. However, the population for children under 6 by single year of age is not available from ACS. Therefore, they need to be estimated. Previous years' data use the prorating or ratio interpolation method, assuming all regions in a county have the same age distribution. ACS does not have detailed single year of ages data, but it has a table B09001 population under 18 years by age has relatively less aggregated age groups for children in households. There are several methods in interpolating single year of ages (see a review of methodology in Judson and Popoff, 2004). One method is widely used is the Sprague's multipliers, due to its ease of use. The method redistributes the total of within a five-year age group into single years of age in a manner that they add up to the original 5-year total (Yusuf, et al., 2014). However, it does not assure monotonicity within five-year age intervals so negative values can occur especially for regions with small population (Baker, et al., 2013). Rizzi, et al. (2016) proposed the Penalized Composite Link Model (PCLM) for single-year-of-age population estimation. They compared 5 methods and found that PCLM outperforms other approaches regardless of sample size. It also solves the non-monotonicity problem that the Sprague's multipliers have. We use PCLM to generate single-year-of-age population for children under 18 in households and then rescale them by using the percentages of single years of age population in the corresponding five-year age groups and data from table B01001 age by sex. Thus, the final estimates include both children in households and group quarters[5].

Race and ethnic data

We estimate and release the following mutually exclusive categories for the racial and ethnic data recommended by the U.S. Department of Education (2007):

- Non-Hispanic White

[3] See methodologies of previous years on IECAM demographic methodology page: http://iecam.illinois.edu/data/methodology/.

[4] Evaluations show that Census Bureau sample surveys like the American Community Survey (ACS), the Current Population Survey, and the Survey of Income and Program Participation also undercount young children, which can result in biased sample survey estimates (O'Hare and Jensen 2014).

[5] Group Quarters/Residence Rules. https://www.census.gov/topics/income-poverty/poverty/guidance/group-quarters.html.

- Non-Hispanic Black
- Non-Hispanic American Indian and Alaskan Native (AIAN)
- Non-Hispanic Asian
- Non-Hispanic Native Hawaiian or Pacific Islander (NHPI)
- Non-Hispanic some other race
- Non-Hispanic two or more races
- Hispanic/Latino of any race

Educational institutions and other recipients of grants and contracts from the Department are required to report the racial and ethnic data in the above categories. The Illinois State Board of Education (ISBE) has been implementing the standards starting with data for the 2010-11 school year (ISBE, 2009).

We use data from ACS Tables B01001 A-I and B03002. For example, the proportion of the Non-Hispanic Black Alone population estimate to the total Black Alone (all racial/ethnic groups combined) population were calculated from the table B03002. Then, the proportions were applied to the total (all racial/ethnic groups combined) population estimates for children under age 5 from the Tables B01001 A-I to produce estimates for the above categories (SEER, 2016). The limitation of this method is that it assumes that the proportions of the Non-Hispanic Race Alone population estimate to the total Race Alone is the same for the total population and children under age 5, which might not be true.

Another potential data source for the race and ethnicity data we have considered is the ACS Selected Population Tables (SPT), which is a special dataset that is released every five years (the latest available year is 2015). The dataset contains detailed race, Hispanic origin, ancestry and tribal groups down to the Census tract geographic level. However, not all population groups for the geographic areas that are available in ACS 5-year estimates are produced by the ACS SPT. That is because SPT has the minimum population threshold: estimates are published for an individual group in a particular geography if it had at least 50 unweighted sample cases, based on 2011-2015 ACS 5-year estimates (Census, 2015). ACS 5-Year Estimates do not have the population threshold (Census, 2016a).

For example, Adams County, Illinois, there are only population data for the following three non-Hispanic race groups in ACS SPT:

- White alone, not Hispanic or Latino
- Black or African American alone, not Hispanic or Latino
- Two or more races, not Hispanic or Latino

However, Table B03002 Hispanic or Latino origin by race from 2011-2015 American Community Survey 5-Year Estimates has population data of other race and ethnic groups too, such as Asian alone, not Hispanic or Latino for Adams County. Using the margin of error, 0 is not in the 90% confidence interval (560-42, 560+42). Therefore, missing population groups for a particular geography in SPT do not imply 0 population, rendering the SPT data unusable for the 2016 estimation.

Children in poverty

Child poverty data are from table B17024 Age by ratio of income to poverty level in the past 12 months, with the universe being the population for whom poverty status is determined. Poverty status was determined for all people except institutionalized people, people in military group quarters, people in college dormitories, and unrelated individuals under 15 years old. These groups were excluded from the numerator and denominator when calculating poverty rates. ACS uses the poverty thresholds published and updated each year by the Census Bureau, which vary depending on three criteria: size of family, number of related children, and, for 1- and 2-person families, age of householder (Census, 2016a). The

Census poverty thresholds differ from the poverty guidelines by the Department of Health and Human Services (HHS), which vary only by family size. The poverty guidelines are a simplified version of the federal poverty thresholds used for administrative purposes — for instance, determining financial eligibility for certain federal programs (HHS-ASPE, 2018a, 2018b).

We apply the poverty rates for children age 5 and under to all the single year of ages generated in the population estimation stage, assuming that poverty rates for each age cohort are the same and equal to that for all the children under age 6. Also, as the population estimates include children in group quarters, another assumption is that the poverty rate for children living in families is the same as all the children. 130% Federal Poverty Level (FPL)[6] is not included in the ACS table as one of the cutoff points. To calculate the numbers of children living in families with income below 130% FPL, we assume that population distribution by income are rectangularly (evenly) distributed within the interval to be subdivided. Thus, children living in families with income below 130% FPL is the sum of children living in families below 125% FPL and a quarter of the 1.25-1.49 FPLs population[7].

Another potential method to be considered for future use is to use ACS and other data, such as the administrative records and third-party data, to produce more useful and informative data products[8]. For instance, Chicago's Early Care and Education Programs: Supply and Demand program (ECSD) uses administrative data from the Illinois Department of Human Services (DHS) on young children receiving the Supplemental Nutrition Assistance Program (SNAP) benefits (commonly referred to as "food stamps"), and the Integrated Public Use Microdata Series (IPUMS) release of ACS data (ECSD,2016). They derive their basic estimates for all Chicago community areas by combining the complementary features of these data sets.

Language use

Household language data are from ACS table C16002 Household language by household limited English speaking status[9]. The number of households speaking other non-English languages at home that are linguistically isolated is the sum of households "Other Indo-European languages: - Limited English-speaking household", "Asian and Pacific Island languages: - Limited English-speaking household", and "Other languages: - Limited English-speaking household". A "limited English-speaking household" is defined as one in which no member 14 years old and over (1) speaks only English or (2) speaks a non-English language and speaks English "very well." In other words, all members 14 years old and over have at least some difficulty with English. By definition, English-only households cannot belong to this group. Previous Census Bureau data products referred to these households as "linguistically isolated" and "Household where no one age 14 and over speaks English only or speaks English 'very well.' (Census, 2016b).

[6] According to the Department of Health and Human Services, "The poverty guidelines are sometimes loosely referred to as the 'federal poverty level' (FPL), but that phrase is ambiguous and should be avoided, especially in situations (e.g., legislative or administrative) where precision is important." (HHS-ASPE, 2018a). However, FPL is widely used by federal and state agencies and programs for eligibility determination.

[7] Another method to use is the midpoint method by von Hippel, et al. (2017).

[8] Census's Center for Administrative Records Research and Applications (CARRA) Working Papers: https://www.census.gov/library/working-papers/series/carra-wp.html.

[9] Detailed Table C16002 replaces B16002 in 2016, with the same format. ACS change the name of the table in 2016 https://www.census.gov/programs-surveys/acs/technical-documentation/table-and-geography-changes/2016/5-year.html.

Parental employment

Parents' working conditions data are from the ACS table B23008 Age of own children under 18 years in families and subfamilies by living arrangements by the employment status of parents. Children 5 and under living in one-parent families with a working parent is the sum of children living with father - father in labor force and living with mother - mother in labor force.

Geographies

In addition to the geographies that are readily available in ACS:

Summary level code	Geographies
040	State
050	County
060	County subdivision, township
160	Place within the state, municipalities
500	Congressional district
610	State Legislative District (Upper Chamber)
620	State Legislative District (Lower Chamber)
860	5-Digit ZIP Code Tabulation Area (ZCTA) 2010
950	School District (Elementary)
970	School District (Unified)

we also produce data for the following regions:

- Chicago community areas (CCAs) (estimates are generated by aggregating tract-level estimates[10]).
- IDHS Regions[11]
- ISBE Regions[12]

[10] Chicago Community Area and Census Tract Census tract boundaries in Chicago can be found on Chicago Data Portal: https://data.cityofchicago.org/Facilities-Geographic-Boundaries/Boundaries-Census-Tracts-2010/5jrd-6zik.

[11] IDHS Regions- Region Offices Coverage (Adult Services): http://www.dhs.state.il.us/page.aspx?item=85694.

[12] ISBE Regions – Directory of Regional Offices of Education and Intermediate Service Centers, ISBE. https://www.isbe.net/Documents/roedirectory.pdf.

References:

HHS-ASPE. 2018a. U.S. federal poverty guidelines used to determine financial eligibility for certain federal programs. Retrieved from https://aspe.hhs.gov/poverty-guidelines.

HHS-ASPE. 2018b. Frequently asked questions related to the poverty guidelines and poverty. Retrieved from https://aspe.hhs.gov/frequently-asked-questions-related-poverty-guidelines-and-poverty.

Baker, J.D., Alcantara, A., Ruan, X., Vasan, S. and Nathan, C., 2013. An evaluation of the accuracy of small-area demographic estimates of population at risk and its effect on prevalence statistics. *Population health metrics*, *11*(1), p.24.

ECSD. 2016. Method for Estimating Chicago Community Area Population. Retrieved from http://dfss-ecsd.chapinhall.org/understanding_data.php.

ISBE. 2009. New U.S. Department of Education Race and Ethnicity Data Standards. Retrieved from https://www.isbe.net/Pages/Data-Analysis.aspx.

Judson, D.H. and Popoff, C.L., 2004. Selected general methods (Appendix C). Edited by Siegel Jacob, S. and Swanson David, A., 2004. The Methods and Materials of Demography. *Elsevier Academic Press, California, USA*. P677-732.

NCHS. 2018. Bridged-Race Population Estimates - Data Files and Documentation. Retrieved from https://www.cdc.gov/nchs/nvss/bridged_race/data_documentation.htm.

O'Hare, W. P., and Jensen, E. B. 2014. The Representation of Young Children in the American Community Survey, presentation at the ACS Users Group Conference, Washington, DC. May 29-30. Retrieved from https://acsdatacommunity.prb.org/p/2014_acs_conference.

SEER. 2016. Single Year of Age County Population Estimates, 1969-2016. Retrieved from https://seer.cancer.gov/popdata/singleages.html.

Rizzi, S., Thinggaard, M., Engholm, G., Christensen, N., Johannesen, T.B., Vaupel, J.W. and Lindahl-Jacobsen, R., 2016. Comparison of non-parametric methods for ungrouping coarsely aggregated data. *BMC medical research methodology*, 16(1), p.59.

U.S. Census Bureau. 2010. ACS Key Facts. Retrieved from https://www.census.gov/content/dam/Census/programs-surveys/acs/news/10ACS_keyfacts.pdf.

U.S. Census Bureau. 2014. ACS Design and Methodology Report, 2014. Retrieved from https://www.census.gov/programs-surveys/acs/methodology/design-and-methodology.html.

U.S. Census Bureau. 2015. 2011-2015 ACS Selected Population Tables and American Indian & Alaska Native Tables Accuracy. Retrieved from https://www.census.gov/programs-surveys/acs/technical-documentation/race-ethnicity-aian.html.

U.S. Census Bureau. 2016a. American Community Survey and Puerto Rico Community Survey 2016 Subject Definitions. Retrieved from https://www2.census.gov/programs-surveys/acs/tech_docs/subject_definitions/2016_ACSSubjectDefinitions.pdf.

U.S. Census Bureau. 2016b. Note for Language Spoken at Home from the 2016 American Community Survey. Retrieved from https://www.census.gov/content/dam/Census/programs-surveys/acs/tech-doc/user-notes/2016_Language_User_Note.pdf.

U.S. Census Bureau. 2017a. Population and Housing Unit Estimates Methodology Vintage 2017. Retrieved from https://www.census.gov/programs-surveys/popest/technical-documentation/methodology.html.

U.S. Census Bureau. 2017b. Investigating the 2010 Undercount of Young Children – Analysis of Census Coverage Measurement Results. Retrieved from https://www2.census.gov/programs-surveys/decennial/2020/program-management/final-analysis-reports/2020-2017_04-undercount-children-analysis-coverage.pdf.

U.S. Census Bureau. 2018. Understanding and Using American Community Survey Data: What All Data Users Need to Know, U.S. Government Printing Office, Washington, DC, 2018. Retrieved from https://www.census.gov/programs-surveys/acs/guidance/handbooks/general.html?eml=gd&utm_medium=email&utm_source=govdelivery.

U.S. Census Bureau. 2018. Population and Housing Unit Estimates. Retrieved from https://www.census.gov/programs-surveys/popest/about.html.

U.S. Department of Education. 2007. Final Guidance on Maintaining, Collecting, and Reporting Racial and Ethnic Data to the U.S. Department of Education, *Federal Register*, 72 FR 59266

von Hippel, Paul T., David J. Hunter, and McKalie Drown. 2017. Better Estimates from Binned Income Data: Interpolated CDFs and Mean-Matching. *Sociological Science* 4: 641-655.

Yusuf, F., Martins, J.M., Swanson, D.A., Martins, J.M. and Swanson, D.A., 2014. *Methods of demographic analysis*. Springer. P269-278.