Research Note: Why Do the Data Synthesis Estimates Differ From Other Published Estimates?

Jewish population estimates from the American Jewish Population Project (AJPP) are based on a data synthesis of a large number of independent samples of U.S. adults (See About Estimates). These estimates are often compared to results from other studies of the Jewish population, such as those conducted by local Jewish Federations and those summarized in the American Jewish Yearbook (AJYB).

Sometimes data synthesis estimates are very similar to other estimates. For example, AJPP estimates the total Jewish population in Las Vegas County to be 72,200, which is nearly identical to the AJYB 2018 estimate of 72,300. Similarly, AJPP estimates the Jewish population in Columbus, OH to be 26,600 compared to the AJYB estimate of 25,500.

In other cases, AJPP estimates look very different from estimates reported in sources such as the AJYB. For example, AJPP estimates the total Jewish population in the Triad region of North Carolina (Greensboro, Winston-Salem, and Highpoint) to be 9,000. The AJYB 2018 estimates a total population of less than half the AJPP estimate, just 4,350 in this area. In Georgia, AJPP estimates a total Jewish population of 6,600 in the northeast corner of the state (26 counties). In the same area, AJYB 2018 estimates a total Jewish population of just 1,000.

There are several reasons why estimates differ. These include:

- Definition of who is Jewish
- Source Bias
- Time Bias
- Error in Estimation
- Undercoverage

Who is Jewish?

The AJPP definition of the Jewish population estimated through data synthesis includes all U.S. adults who identify as Jewish when asked about their present religion. Estimates of Jewish children and Jewish adults who identify as ethnically/culturally Jewish are added to this baseline population. Unlike the baseline population, which is estimated through a data synthesis of hundreds of independent samples of U.S. adults, estimates for children and other Jewish adults rely on single sources of data such as local Jewish community surveys or the national Pew 2013 Survey of American Jews.
For consistency, AJPP defines other Jewish adults for all regions as U.S. adults who consider themselves Jewish aside from religion, have at least one Jewish parent, and do not identify with any other organized religious group (See Tighe et al., 2014). This differs from the definition of Jewish adults reported by DellaPergola (2015) which based on analysis of the Pew Survey of American Jews (2013) excludes nearly half of other Jewish adults because they replied “partly” when asked if they consider themselves Jewish (p. 281). This includes adults who do not just consider themselves Jewish, but were raised Jewish and had two Jewish parents. Without an explicit definition of how respondents such as these might have interpreted the word "partly" we rely on the core definitions of self-identification and parentage. See the figure below for a detailed breakdown of differences between DellaPergola’s and the AJPP estimates of the US Jewish population. The comparison is for the year 2015, the year closest to the release of the Pew Survey of American Jews (2013) that DellaPergola’s estimates are based on. Our definition also differs from some local community studies, which sometimes include other Jewish adults regardless of whether they presently identify with another religion, or if they identify with other non-monotheistic religions. Similarly, AJPP defines Jewish children across all regions as all children being raised Jewish by religion or otherwise.

Source Bias

Estimates from the AJYB represent an amalgamation of multiple methods and sources, including directories, key informants, membership lists, internet searches, news reports, and for some larger geographical areas, local community surveys. It is not the case that these sources are drawn upon to inform each local estimate. Instead, one estimate may be based on a report from a key informant, another estimate on a newspaper article, and another on a targeted survey. This approach is likely to introduce many different sources of error to a larger scope of results.
The AJPP data synthesis provides a snapshot of the Jewish population based on sources of data that are all designed to provide nationally representative samples of the U.S. adult population. The representativeness of these samples allow the data from each source to be evaluated against an external frame of reference such as the U.S. Census. Additionally, the data synthesis includes methodological characteristics of each source—particularly, methodological characteristics that are typically associated with biased estimates—to allow for the direct examination of biases across individual surveys.

**Time Bias**

The variety of sources reported in the AJYB differ not only in terms of methodology but also time frame, spanning years, and sometimes decades. Over such long periods of time, the methods and technologies used to estimate populations change, as do the populations themselves. Changes in birth and death rates as well as in-migration flows are not captured when estimates from different periods are pooled to create a single population profile. Although the AJYB publishes estimates annually, their methods for updating estimates from older sources remain unclear.

The AJPP data synthesis reduces the effect of time bias by standardizing the time frame associated with each release of estimates. Data are aggregated over a period of five (+/-) years to increase the effective sample size required to estimate the Jewish population at the county level, and are then poststratified to data from the most recent Census Population Estimates Program.

**Error Estimation**

For many of the sources of data reported in the AJYB, such as those based on reports from key informants or news stories, there is no way to gauge what kind of errors might be associated with the estimate. For the major population areas where local Jewish community surveys have been conducted, there is some degree of error associated with the estimates and this error can be measured (see FAQ). The degree of error in an estimate will vary from community to community depending on how the survey was conducted, the methods of sampling, and other factors. When a variety of sources are combined, as they are in the AJYB, information about the degree of error that might have been associated with the population estimate is lost. It might be that errors from one source counterbalance errors from other sources. It might be that errors are compounded across the sources and result in an increase in bias or inaccuracy in the final estimate. A benefit of the data synthesis approach is that the degree of uncertainty associated with the different sources of data can be examined directly and factored into the final result if needed.

**Undercoverage**

The AJYB only provides estimates where there is a known Jewish community. Although it is true that the majority of the Jewish population resides in major metropolitan areas, the AJYB provides an incomplete picture of the American Jewish population. There are many areas where Jews live that are unknown to informants or outside the target geography of local community studies. Similarly, the DellaPergola estimates (2015-2019) are based to the Pew Survey of American Jews (2013) which
excluded nearly half of counties in the US. Pew provides an adjustment to their estimates of Jewish adults to account for this undercoverage. This adjustment is not reflected in DellaPergola’s estimates of other Jewish adults (those who identify culturally/ethnically but not by religion) and Jewish children.

The AJPP data synthesis looks at distributions across hundreds of independent, representative random samples of the U.S. adult population. One hallmark of this feature is that it covers all areas of the United States, not just large metropolitan areas or areas where there are known Jewish communities.

As an internal check on the validity of our estimates in areas where the data synthesis indicates there are Jews not represented in AJYB estimates, we examined Birthright registration data provided by researchers at the Cohen Center for Modern Jewish Studies. We mapped the location of applicants considered to be eligible for the program, that is, they identify as Jewish and had at least one Jewish birth parent, or had completed Jewish conversion through a recognized Jewish denomination and are recognized as Jewish by their local community or by one of the recognized denominations of Judaism. By all accounts, these individuals are “Core Jews” if one applies the definition proposed by Sergio DellaPergola. The Birthright applicant data are not representative samples for purposes of population estimation, but they do provide convergent validity that there are, indeed, Jews residing throughout the U.S. outside known Jewish population centers and throughout areas where the data synthesis yields estimates that are higher than those reported in the AJYB.

For example, the yearbook reports a total Jewish population of 35,435 in the entire state of North Carolina (AJYB, p. 17). AJPP data synthesis estimates the state population to be between 112,500 with a 95% credible interval ranging from 94,500 to 132,900. Looking at Guilford and Forsyth counties (Greensboro, High Point and Winston-Salem), AJYB 2018 estimates a Jewish total population of just 4,350. The Winston-Salem estimate is based on a Distinctive Jewish Names survey conducted in 2011, the High Point estimate on synagogue membership, and the Greensboro estimate on a 2009 key informant (AJYB, p.62). The data synthesis estimates less than 1% (~0.6% with a 95% CI ranging from 0.3% to 1%) or 5,400 adults in the entire Piedmont Triad area identify their religion as Jewish. This increases to just over 9,200 people if one adds estimates of children and Jewish adults who do not identify their religion as Jewish (Pew, 2013). The distribution of Birthright applicants indicates that the Jewish population is more widely distributed than the "known" population centers reported in the yearbook.
Taking into account the regions not represented in the AJYB estimates, as well as possible underestimates of Distinctive Jewish Name samples for this area or the underestimation of the university town of Greensboro, the data synthesis likely provides a more accurate estimate than is otherwise available.

Many of the limitations of source bias, time bias, error estimation, and undercoverage are resolved using the data synthesis approach. New data are continually added and only a subset of the most recent years are used to provide new estimates of the Jewish by religion population. Thus, estimates for specific time periods are based on standardized sources of data specific to that period. There are, however, limitations to the data synthesis approach. As noted in About Estimates, the synthesis focuses on the most easily measurable or observable portion of the population represented in national surveys: adults who identify their religion as Jewish. Although this group represents the majority of the population, to extrapolate from this base number to the total population requires reliance on data that is less reliable and requires further development. For this reason, we have more confidence in the estimates of the proportion of adults in each area who identify as Jewish in response to questions about religion than we do in the total population estimates. Another limitation is that there are always more data that could be added to the synthesis and, despite our ongoing collection, there are small areas that remain difficult to estimate reliably. It might be that our population models "fit" some areas of the country better than others and more work needs to be done to improve our models in these areas. We continue to improve estimation of these areas.

References


