Using American Community Survey Estimates and Margins of Error Webinar
Transcript
April 19, 2017
Jennifer Berkley
Decennial Statistical Studies Division

This presentation was presented as a webinar to the general public on April 19, 2017 by Jennifer Berkley of the U.S. Census Bureau's Decennial Statistical Studies Division. A link to the recording is available at https://census.webex.com/census/ldr.php?RCID=714c57425cd24a5b15fb4eea3da438ec

The transcript of the webinar follows. Slide references and links have been added to the spoken text as appropriate.

Coordinator: Thank you for standing by. At this time all participants are in a listen-only mode. During the question-and-answer session you may press Star 1 if you would like to ask a question. Today's conference is being recorded. If you have any objections, you may disconnect at this time. Now I'd like to turn the meeting over to Ms. Jennifer Berkley. You may begin.

Slide 1: Title Slide

Jennifer Berkley: Good afternoon. My name is Jennifer Berkley and I am a Mathematical Statistician in the Decennial Statistical Study Division of the U.S. Census Bureau. Today I will be talking to you about using American Community Survey Estimates and Margins of Error. As a note, the American Community Survey is often called the ACS. I will be using both terms interchangeably in this presentation.

Slide 2: Outline

Today's presentation will start with a discussion of ACS estimates, then we will get into what a Margin of Error is and I will explain why it matters. As a

note, the Margin of Error is often referred to as the MOE. I will be using this abbreviation throughout this presentation.
Next, I will discuss statistical testing or more specifically why statistical testing is important and how it uses Margins of Error. We will then go over a few special cases where the MOE is presented differently. Next, we will walk through an example of how to approximate the MOE for combined estimates. Finally, I will give a brief overview of some available resources before opening the lineup for questions.

Slide 3: Outline

Let us start by talking about ACS estimates.

Slide 4: ACS Estimates

Every year over 3.5 million housing unit addresses are contacted by the U.S. Census Bureau to participate in the American Community Survey or ACS.

The information obtained from this sample is then used to estimate characteristics about the total population in a timely and cost-effective manner. However, these estimates differ from those that would be obtained in a Census, where every household in the nation is contacted.

In other words, we cannot be 100% confident that the sample is truly representative of the entire nation. This results in an element of uncertainty in the data. Users who are interested in more information about ACS sample design should see the ACS Design and Methodology report. A link to this report is provided on this slide.

Slide 5: Availability of ACS Data Products

ACS data products are available in a few different varieties. ACS 1-year estimates are based on data collected over 12 months and are available for geographic areas with a population of 65,000 or more. We plan to release the 2016 ACS 1-year Estimates in September of 2017.

ACS 1-year supplemental estimates are 60 detailed tables that are available for geographic areas with populations of 20,000 or more. They are simplified versions of popular ACS tables mainly univariate or bivariate tables focused on key topics. We plan to release the 2016 ACS 1-year supplemental Estimates in October.

ACS 5-year estimates are based on data collected over 60 months and are available for geographic areas of all sizes down to the Census Tract and Block Group level. We plan to release the 2012-2016 ACS 5-year estimates in December.

As a note in your Data Analysis, you should not compare estimates from single-year data products to estimates from multiyear data products.

Slide 6: Outline

Both 1-year and 5-year ACS estimates are published with an accompanying Margin of Error. We will now get into what exactly that Margin of Error is.

Slide 7-9: What is the Margin of Error?

So as you might know, ACS estimates are available from American FactFinder or AFF. Here we have a typical table, Detailed Table B01001, Sex by Age, as it is displayed on AFF. As you can see, the table has three elements: the characteristics or description, the estimate, and the accompanying Margins of Error. If you have used ACS estimates before you may have seen these numbers and ignored them, not knowing what they are or what their use is.

Today we will learn how these Margins of Error give us more information about the population by telling us how the estimate may vary from the true population value.

Slide 10-11: What is the Margin of Error?

So put simply, the Margin of Error or MOE is a measure of the possible variation of an estimate around the population value.

Margins of error allow data users to be certain that at a given level of confidence, the estimate and the actual population values differ by no more than the value of the MOE. The Census Bureau uses a 90% confidence level as its standard.

All ACS estimates published on AFF have Margins of Error calculated at the 90% confidence level. It is also important to note that the MOEs provided by the Census Bureau are always in the same units of their respective estimate.

For instance, a percent of cement will have a percent MOE and a median income estimate will have an MOE in dollars.

So to understand this better, let us look back at our table D01001 on AFF.

Let us look at the estimate for males under five years. As you can see the ACS estimate is 10,175,713 with a Margin of Error of 3,826. Again since this figure is an estimate we cannot be 100% certain that there are exactly this many males under five years in the U.S. However, using the MOE we can be 90% confident the true number of males under five years is somewhere within 3,826 of this estimate.

Slide 12: Measures of Sampling Variability

Now you might be wondering where exactly this MOE comes from. Well, it is actually a member of the family of Measures of Sampling Variability. This slide shows you how one can be derived from another. At the top is the variance. The variance is calculated for each estimate. An estimate is tabulated and put simply the higher the variance, the greater the variability in responses. Due to the way it is calculated, variance is not a very useful tool for the everyday user. However, it is an integral part of finding the MOE and it leads us to the Standard Error.

The Standard Error or SE is found by taking the square root of the variance. The standard error gives a clearer picture as to how much variability is in the estimate as it is in the same unit as the estimate itself. The standard error is used to calculate the MOE. This calculation changes depending on the confidence level. To calculate the MOE at the Census Bureau's standard 90% confidence level, the SE is multiplied by 1.645.

Slide 13-15: Alternate Confidence Levels

Sometimes though you might wish to use a confidence level other than the 90% Census standard with your ACS estimate. For instance, you might wish to use the 95% confidence level or a 99% confidence level. Let's go over how to convert from one confidence level to another.

Again, to calculate the MOE of AFF which is at the Census standard 90% confidence level, we multiply the standard error by 1.645. As you can see, the MOE for the 95% confidence level is equal to the standard error multiplied by 1.96.

To convert from the 90% confidence level to the 95% confidence level, simply multiply by the ratio of the new confidence factor to the old confidence factor as shown in the formula.

It might be easier to think about this step-by-step. First, this formula divides the current MOE by the factor used to calculate it, in this case 1.645, to convert it back to the standard error. The standard error is then multiplied by the new factor of 1.96 to find the MOE with a 95% confidence level.

For an illustration, let's look at what our males under five years estimate from Table B01001 looks like at a 95% confidence level. The MOE from AFF is currently at a 90% confidence level. To convert we divide the MOE by 1.645 and multiply by 1.96 to get an MOE with a 95% confidence level. Our equation gives us an MOE of 4,559 at the 95% confidence level.

As you can see, as the confidence level increases, the MOE increases. This is especially clear when we add the 99% confidence level. In other words, to gain more confidence we must allow for more variability in the data.

Slide 16: Outline

So now we know what MOEs are and how they're calculated but you may wonder what this means in a practical context. Why do they matter and what can they be used for? We will get into this now.

Slide 17: Confidence Intervals

One way to use the MOE is to construct a confidence interval. A confidence interval is a range of values where at a given confidence level you can be certain of the population value lies. The confidence level of the confidence interval will always equal the confidence level of the MOEs used to create it.

You might have noticed that our MOEs are always displayed with a plus or minus sign. This is a good clue to remember that the estimate plus or minus the Margin of Error provides the confidence interval. Let's look at an example of a hypothetical Block Group.

Block Groups are the smallest geographies for which ACS estimates are available on AFF. Our hypothetical Block Group has a median income of 37,284 and an MOE of 20,922. The upper bound of the confidence interval is the estimate plus the MOE, that is, 37,284 plus 20,922 or 58,206 and the lower bound is the estimate minus the MOE, or 16,362.

Since ACS MOEs have a 90 confidence level, this confidence interval is at the 90% confidence level, the Census Bureau's standard. Therefore, we can conclude at a 90% confidence level that the true estimate for the population lies somewhere between these two values.

Slide 18-19: Why MOEs Matter

The MOE can also allow you to see if there is enough evidence to conclude the estimates are statistically different from one another. As an example, let's compare five more hypothetical Block Groups. With just the information now on the screen, you would probably conclude that Block Group 1 has the lowest median household income at 37,200, and that Block Group 5 has the highest median household income at 76,850. However, we are missing some of the important information in the MOE columns.

Adding-in the Margins of Error to the table, we see a different picture. As a note since this is a hypothetical example, the MOEs for these estimates are high. Not all of the medians at the Block Group level have such high MOEs, but it is important to be aware of the MOE when looking at an estimate.

Looking back at the table, we can now see that the median we thought was the lowest has an MOE of 20,920. Meanwhile, the seemingly highest median has an MOE of 47,200. Both estimates therefore have substantial variability.

Unless you take into account these MOEs, you cannot conclude that they are statistically different from one another. Instead you have to do statistical testing to check for any differences.

Slide 20: Outline

So now I hope you have an good sense of what an MOE is and why it shoul not be ignored. We will now delve a little deeper into how it can be helpful to you and your data analysis by looking at statistical testing.

Slide 21: What is Statistical Testing

Statistical testing is an important part of data analysis as it can tell us whether or not a difference in estimates is meaningful. In short, a statistical test is a test to determine if a difference is unlikely to have occurred by chance.

To be statistically different, there must be statistical evidence that there is a difference between the two estimates. Statistical testing should be conducted anytime you explicitly or implicitly make a comparison between two estimates.

Slide 22-27: Statistical Testing

So here is our first formula, it's the generic statistical testing formula. As a note, you do not have to remember all these formulas. They are all available through our online resources for statistical testing and I will show you how to access these resources later in the presentation.

This formula uses both the estimate and the MOEs to generate a result. If the result of the formula is greater than one, then the estimates are statistically different at a 90% confidence level. As an aside, this formula is the Z score formula modified to use MOEs instead of standard errors. This is done to make the formula easier to use and because the MOEs are conveniently published on AFF.

Let's walk through a quick example and see how testing is done.

For our testing, we will look at another table from American FactFinder. This time we will look at selected population profile S0201.

Suppose you wish to compare the median age of the total population in the U.S. to the median age of the total population in New York State. This example uses 2015 ACS 1-year Data.

The median age for the U.S. as you can see is 37.8 while the median age for the total population of New York is 38.3. To say that the median age for the U.S. is different than the median age for New York, we must complete Statistical Testing.

So here, we have our two estimates and their Margins of Error. To begin our testing we first take the difference of the estimates so we subtract 38.3 from 37.8. For the second step we take the absolute value of that difference. In Step 3, we square the MOEs for each estimate. For the U.S., .1 squared is .01. For New York, .2 squared is .04. Next, we add these squared MOEs together, that is, we add .01 to .04 to obtain .05.

For Step 5, we take the square root of the sum of the squares. In this case, the square root of .05 is .224.

We then divide the result of Step 2 by the result we found in Step 5, that is, we divide .5 by .224 to obtain a value of roughly 2.24. In Step 7, we compare our result to 1. Since 2.24 is greater than 1, we can say with 90% confidence that the median age of the population of New York is statistically different than the national median age.

To put it all together, here is the generic formula and our worked example. Again, as 2.24 is greater than 1.0, we can conclude with 90% certainty that these two estimates are statistically different.

The nice thing about the Z-score formula, is that you can use to compare a wide range of estimates, including estimates between different years, multiyear estimates between non-overlapping time periods, estimates across geographic areas, and estimates between surveys.

Again, I would like to note that single-year and multi-year estimates should never be compared in your data analysis. Comparisons between 1-year and 5-year data products will not yield very good findings.

Also before comparing ACS estimates to Census estimates, you want to make sure that the estimates are comparable. The question for some estimates have changed so it is not possible to compare those Census data. You can check compatibility via the link at the bottom of this slide.

Slide 28-33: Statistical Testing Tool

Now that I have shown you how to manually conduct statistical testing, I will show you an easier way. The staff here at the Census Bureau released its Statistical Testing tool for public use. This tool allows you to input estimates you would like to test and receive a visual notification of whether the estimates are statistically different or not.

Using the tool, you can compare both pairs of estimates and groups of estimates against each other. The tool is available for download as an Excel spreadsheet via the link at the bottom of the slide. Let's look at a few screenshots of this tool in action.

Here we have an example of the testing for two estimates tab in the statistical testing spreadsheet. The spreadsheet has instructions at the top and column spread testing label, the first estimate, the first MOE, the second estimate and the second MOE below. This tab can compare up to 3,230 pairs of estimates.

As you can see, results are both written and color-coded. Estimates that are not significantly different had a red "No" in the "Statistically Different" column. This is described in text at the top of the spreadsheet.

Multiple pairs of estimates can be compared at once. The same tab can compare estimate pairs of differing types, as shown in the example. Let's look at these comparisons a little more closely.

In the first column of the testing area, I have input our median age estimates and their corresponding MOEs into Row 1. With a "Yes" in the Statistically Different column, the spreadsheet confirms our results since the estimates are in fact statistically different.

You might notice that these estimates are formatted exactly as they are on AFF. There is no need to alter their format. The spreadsheet is configured to handle all table notations used on AFF.

I have also input the hypothetical block level median household income estimates we looked at before. If you remember after looking at the Margins of Error, we could not tell if the Block Group 1 estimate of 37,200 was significantly different from the Block Group 5 estimate of 76,850.

To check this, I have manually input them into the spreadsheet. The red “No” in the “Statistically Different” column tells us that they are not statistically different from each other.

So what if you wanted to compare all five of those hypothetical Block Groups from our example to each other? You could individually put the combinations into the “Two Estimates” tab, or you could use the “Multiple Estimates” tab of the spreadsheet. This tab can compare up to 150 estimates against each other. However, this test can only complete one type of comparison. For instance, medians and counts cannot be compared at the same time.

Here I have input our hypothetical Block Group into the “Multiple Estimates” tab. As you can see, the format is slightly different from the last spreadsheet as each label corresponds to only one estimate and MOE. The spreadsheet then creates a grid of results in which each estimate is compared to the others.

A comparison of an estimate to itself is noted by a gray square with an “X”. When Statistical Testing is inappropriate, a gray square with a dash through it will appear. Otherwise, the notations are the same with significant differences denoted by a white square with “Yes” and non-significant differences denoted by a red square that reads “No”. These notations are given at the top of the spreadsheet.

Here we can see that none of our hypothetical Block Group estimates are significantly different from each other. This is due to their large MOEs as we spoke of earlier.

Slide 34: ACS Comparison Profile

American FactFinder also publishes a few types of ACS data products that include statistical comparison testing.

The first type is the Comparison Profiles. Comparison Profiles are similar to the Data Profiles, except they list the estimates for prior years. The estimates from prior years are inflation-adjusted. There is also an asterisk in the column next to the estimate if the estimate from the current year is statistically different from the prior year. This is explained in the text above the table.

This slide shows a 5-year Comparison Profile. Comparison Profiles for 5-year data are now available for the non-overlapping 2006-2010 and 2011-2015 5-year estimates. Comparison profiles for 1-year data are also available.

Comparison profiles are available for a wide range of geographic areas including for the Nation, States, Counties, School Districts and Congressional Districts.

Slide 35: ACS Ranking Tables

Another product type that includes statistical testing is the Ranking Tables. Ranking Tables show you how different geographic areas compare to one another. I should note that the only geographies provided for ranking tables are the Nation and States.

The Ranking Tables are a little different from the comparison profiles in that you have to turn the statistical testing on. To do so, click “With Statistical Significance” in the top right.

Slide 36-37: ACS Ranking Tables

After clicking with statistical testing, you can select a state or the nation using the dropdown menu and compare it to the other states and the nation. Two-pound signs appear next to the geography, which you select. If an estimate from another state is not significantly different from the selected state, then a single pound sign is shown next to it. For example, here the national estimate is selected, and we see that it is not statistically different from Arizona or Colorado.

The notation is a little different from the comparison profiles we found in the last slide, but you do not need to remember this as it is explained in the circled text above the table.

Slide 38: Outline

Now I hope you have some familiarity with Statistical Testing and some of the various products and tools that are available to you. However, there may be times when you come across something that might look a little odd on AFF. You might not know how to proceed. We will explore how to deal with some of these special cases now.

Slide 39: Special Case, Controlled Estimates

The first special case we will look at is a controlled estimate. Certain estimates in the ACS are controlled to match the official population estimate. If this has occurred, then the MOE will have five stars instead of a number. You may use this estimate in a statistical test. If using the formula simply set the MOE to be zero. If using the statistical testing tool, you can simply paste the five-star notation into the tool. The spreadsheet will convert it automatically.

Slide 40: Special Case, Zero Estimate MOEs

Another special case is for estimates which are zero. Here we have an example of Table B01001E, which includes many zero estimates. As this table demonstrates, zero estimates have non-zero Margins of Error. This is because the ACS is a survey. Households with rare characteristics may not be surveyed but may still exist.

Slide 41-44: Special Case, Medians and Aggregates

There are also special cases for which Statistical Testing is not possible. Table B6001, median income by place of birth in the U.S. for the Census-designated Haena, Hawaii, gives us an illustration of these cases.

Medians and aggregates from too few observations are noted on AFF as an estimate with a dash and an MOE of two stars. As you can see for the characteristic “Native born outside the United States”, values are not displayed to protect confidentiality.

Medians in the highest or lowest interval of an open-ended distribution also have a different notation on AFF. These estimates will be followed by a plus or minus sign with an MOE of three stars. Looking back to the table, we see that the foreign-born characteristic is at the lower median interval. We can tell this because the estimate is not a number but rather a combination of a number and the subtraction sign, in this case 2,500 minus. The minus sign tells us that the median is less than the number that precedes it.

For foreign-born person, we know that the median income is less than 2,500. The actual median is not provided. A median at the upper interval would be displayed similarly but with an addition sign following the number. The MOE would still be three stars.

All you would be able to say about a median with an estimate on AFF of 250,000 plus, sorry about that, is that it is above 250,000. Due to the methodology we used, the actual median for these cases cannot be calculated.

As a note if you are using data from the Census API, you may see this estimate displayed as 2,499 instead of 2,500 minus and the MOE will be null. A null value in the Margin of Error will alert you to the fact that the estimate is not a number.

As a reminder, explanations of table notations can be found at the bottom of the table on AFF. Also, statistical testing cannot be performed for either of these cases; however, all AFF notation is recognized by the statistical testing tool that I demonstrated earlier. For instance, these cases will tell you that testing is not appropriate.

Slide 45-46: Estimates with Large MOEs

Some estimates have an MOE that is larger than the estimate itself. This occurs mainly in estimates for small geographies or small groups of people ,or households which have a small sample size. As an example, we have Table B17001, which is showing the number of people above and below the poverty level by Age and Gender, at the tract level.

Tracts are among the smallest geographies that are available on AFF for ACS data. As you can see, many of the MOEs here are larger than their respective estimates.

Data users should exercise caution when using these estimates because they have questionable reliability. Large MOEs can also signal that the sample size for the estimate is small.

There is not an ideal solution to dealing with large MOEs; however, a few workarounds are available. One method is to use a larger geographic area. For example, you could use County estimates in lieu of Tract estimates.

Another method is to form larger groups by combining estimates or geographies. When doing this, you will have to approximate the MOE for your derived estimate.

Slide 47: Outline

So now we will go into how to approximate Margins of Error for these derived estimates.

Slide 48: Deriving New Estimates

Let's start with a simple example of how to derive a new estimate. Here are some age groups by gender. Supposed you want the total population who are less than five years old. You have to add-up the estimates for males and females who were under five to obtain the desired estimate. When you do so, you must also approximate the MOEs to find this estimate from the estimate.

Slide 49-51: Approximating the MOE

Approximating the MOE takes a little more work. Similarly to what we did with the example with the Z score, you must square the individual MOEs, sum them together and then take the square root.

Note that this is just an approximation of the MOE. It probably will not exactly match the MOE were we to produce this estimate using the ACS microdata. While the microdata is not available to the public, this approximation is available and provides a method of obtaining a usable MOE.

Here we have a worked example. To obtain the total number of children under the age of five, we sum the estimates for males and females, that is, we add

10.2 million to 9.8 million to obtain a total of about 20 million. To approximate the MOEs, we first square the MOEs, then sum the squares.

In this case, we are adding the square of 3,826 to the square of 3,377. Finally take the square root of the sum to obtain the MOE. Share MOEs roughly 5103.

One way to approximate - one way the approximate - MOE can be different from the actual MOE is if you are aggregating a zero estimate.

In this case, the best method to approximate the MOE is to include only the largest zero estimate MOE in the calculations. Summing multiple zero estimate MOEs will overstate the MOE of your derived estimate.

For example, let's say we want to know the total number of children under 10, who identify as Native American or other Pacific Islander alone in Maine and Rhode Island. To find the estimate we sum the four estimates, that is, we would add 41 plus the three zeroes.

To approximate the MOE, square 37, which is the MOE for the non-zero estimate, and square 29, which is the largest zero estimate MOE. Take the square root of their sum to obtain an approximate MOE of about 47. If you were to include all of the non-zero MOEs in this example, you would obtain an MOE of about 56 which is somewhat larger.

Slide 52: Variance Replicate Tables

Advanced users may also be interested in using the variance replicate tables. The variance replicate tables include estimates, Margins of Error and the corresponding 80 variance replicates for selected ACS 5-year detailed tables. These tables can be used to calculate MOEs for derived estimates in the place of the MOE approximation method that I just demonstrated. Using these tables, the calculated MOE will include the covariance, which is the element that is missing from the approximation method.

I will not get into what covariance is now, which is more advanced than the scope of this presentation. However, you should know that the covariance must be calculated individually for each derived estimate, which makes it impossible to be provided to the public.

If the covariance is large, the approximation could be very different from the actual MOE. The variance replicate tables give users a means to calculate this actual MOE. Using these tables requires a fair understanding of statistical programming. For more information please follow the link at the bottom of this slide.

Slide 53: Collapsed tables

I would also like to note that for some detailed tables, a collapsed version is available. Here we have an example. On the left is Table B01001B while on the right is the collapsed version, C01001B. The detailed table will have a table ID which begins with a B, while the collapsed version table ID begins with a C.

This example is for sex by age for persons reporting a race of black or African-American alone. The alone means that black or African-American is the only race reported. As you can see, you do not have to approximate the MOE for collapsed tables, as it is provided for you. The categories were created by subject experts.

The categories may not be exactly what you want, but they will allow you to combine fewer estimates together. Here we see that the age categories have been collapsed into those under 18, those 18 to 64, and those 65 years and older.

As a note, the table ID is a shortcut to the table you are interested in. For example, you could find this table on AFF by searching by Age, Sex and for the Race Group for Black or African-American alone. However, if you search by the table ID B01001B, you would obtain the above table. The collapsed version of the table is found by using the table ID C01001B.

Slide 54: Outline

Now we will go over a few resources available to you before opening-up the line for questions.

Slide 55: Resources

The ACS Website includes many great resources for users looking for more information. I will highlight a few specific pages that you might find helpful.

Slide 56-58: ACS Documentation

The code lists, definitions and accuracy page of the ACS Website is a great resource for more information.

Here you can find the instructions for applying statistical testing documents, which provides information on how to derive your own estimates, and how to approximate their MOEs.

The accuracy of the data document is also housed here. The accuracy document provides the same equations as the instruction for statistical testing document as well as some worked examples.

Note that there is an accuracy document for the single-year and for the multiyear estimates. Two are needed because the multiyear document covers issues that are different from the single year; however, their examples are basically the same.

This page also contains a link to the statistical testing tool that I demonstrated earlier.

Slide 59: Compass Handbooks

Compass Handbooks for data users are also available and can be reached via the link at the top of the slide. These handbooks provide an overview of the ACS, and are tailored to fit the needs of different groups. There is one for the general public, one for teachers, one for members of Congress and several others. Choose the one that is most relevant to your needs.

The Compass Handbook appendices are great resources to reference when calculating derived Margins of Error. The Compass Handbooks are currently in the process of being updated and new versions will be posted once available.

Slide 60: Training Presentations

Past training presentations on various aspects of ACS data are also available online. Some topics covered include how to access ACS data and how to use the Public Use Microdata Sample files.

The slides, the recording, and transcript for this webinar will be added to the webpage soon.

Slide 61: Crosswalk

A crosswalk to enable you to see which tables in ACS match-up to which Census 2000 long-form tables is also available. Note that the 2000 Census data is also available on AFF. Again, you can reach the crosswalk by going to the top of the slide.

Slide 62: Design and Methodology

As I mentioned earlier, for more technical information on the ACS and how the MOEs are created, you can look at the design and methodology report. Due to its large size, chapters of the report are individually available for download as shown here. Information concerning variance estimation can be found in Chapter 12.

Slide 63: Source Us!

Finally, if you are using ACS estimates, make sure to source us. It helps people know that the information they are using is powered by the American Community Survey. Shown here are three real-world examples.

Slide 64: Continue the Conversation #ACSData

I encourage you to connect with us. You can sign-up for and manage alerts on the ACS via GovDelivery. You can visit our website, or you can connect with us on the various social media platforms using the hashtag #ACSData.

Slide 65: Data Users Group

You can also connect with other ACS data users via the ACS data users' group.

The Data Users Group was formed in partnership with the Population reference bureau. It is a great way to learn from your peers about how to use ACS data for all kinds of applications. Visit acsdatacommunity.prb.org to learn more including how to sign-up to be one of the over 1,800 users in the ACS online community.

I would also like to mention that the 2017 ACS Data Users conference is scheduled to take place on May 11th and 12th. More information on the conference can be found on this ACS data user group Website.

Slide 66: Need Local Stats?

As a reminder, we have data dissemination specialists throughout the country who can provide data workshops locally. If you are interested in a workshop, please reach out via the phone number or e-mail on this slide.

Slide 67: Outline

So we have now reached the end of this presentation.

Slide 68: Questions?

We will now open-up the line for questions. If you have specific data questions after the Webinar, feel free to e-mail them to the link provided on this slide.

Questions & Answers Section

Coordinator: At this time if you have questions from the phone line, you may press star 1. Please remember to unmute your phone and record your first and last name clearly when prompted. Once again, if you have a question, please press star 1 and record your name. One moment as we wait for questions.

Deborah Rivera: Hey, thank you everybody, and while we wait for the questions to come onto the queue, I would like to bring your attention to the training evaluation form that is on your screen. We would certainly appreciate it if you would take five minutes of your time to fill it out. It is not very long, but your feedback helps us assess the information that we are providing you.

If there is anything that you would like to hear in the future, or any future presentations that we can put together for you, so your feedback is very important to us. It only takes a few minutes.

We appreciate it if you would participate, and also, as a side note, the presentation that is being recorded today will be made available through the American Community Survey Webpage through the Census.gov Website.

Coordinator: And we do have a few questions in queue and the first one comes from (name removed). You have an open line, sir.

(Question #1): Yes, how do you calculate the MOE when averaging individual estimates across geographies?

Jennifer Berkley: If you are using estimates published on AFF, you can use the approximation formulas that we went over earlier, where you square the individual MOEs, add them together, and take the square root of that sum. For more information on the approximation formulas, you can go to the Compass Handbooks, or the Accuracy of the Data document, on that code list accuracy and definitions page of the ACS Website. Is that…does that answer your question?

(Question #1): Yes, so the answer is it is the same whether you sum individual estimates or you average individual estimates, correct?

Karen King: Hi, this is Karen King. I work with Jennifer. So your question is whether you are adding averages that already exist together and then …

((Crosstalk))

(Question #1): If I am taking the average of let's say two geographies, and I present the average of those two geographies, let's say two Census Tracts, the Margin of Error would be the sum or the square root of the sum of the squares, is what I was just told.

Karen King: I think …

(Question #1): But that is the same as if I added the two estimates for the two geographies together.

Karen King: Well, I am suddenly drawing a blank, but if you send the question to that Website, that address, then we will get right back to you on that.

(Question #1): Okay, thank you.

Karen King: I apologize; I cannot give you a good answer at the moment. Like I said, I've kind of gone blank for a minute here.

(Question #1): Okay, thank you.

Coordinator: Our next question comes from (name removed).

(Question #2): Jennifer, two questions if I may. So the estimate and the corresponding Margin of Error is always in the same unit, whether it's a hard number or in some cases a percentage, so you had one sample where you are running the statistical tool, median age in years, the first estimate was 37.8, so almost 38 years of age. The first Margin of Error was plus or minus 0.1 so is that one-tenth of one year?

Jennifer Berkley: Yes.

(Question #2): It is not a ratio, it s …

Jennifer Berkley: No, it would be in years.

(Question #2): … so it is in years, okay, okay. And then most of the time we are doing in numbers and not percentages, is that correct, most estimates just tend to be that way?

Karen King: There are definitely more estimate counts, I mean, available on the detailed tables than there are but the percentages are available on some of the data profiles.

(Question #2): So but whenever you see a Margin of Error and a percentage, then that means the estimate was in the same fashion percentage in the account?

Karen King: The Margin of Error and the estimate are always given in the same denomination or whatever, yes.

(Question #2): Okay, great, thank you. And then, one other question if I may. So I happen to deal with local areas. I try to do analysis at the Block Group level, you know, income, 10-year, whatever so Margin of Errors tend to be kind of high.

That said, I am dealing with ACS data pretty much for 5-year estimates, so I don't think this counts statistically but the fact that the estimates are over a period of five years, does that give a little more solidity to the estimates even though it may not be reflected in the Margin of Error?

Karen King: 5-year estimates are definitely stronger than you would with a 1-year, and you should consider it as kind of an average over that period.

(Question #2): But that does not, the 5-year doesn't, the fact that the estimate was taken over a period of five years does not lower the Margin of Error statistically.

Jennifer Berkley: Well, it certainly is lower than it would have been if you had used one year.

(Question #2): So the duration of the survey is calculated within the Margin of Error statistic?

Jennifer Berkley: Well, there is more data that is being pulled in the five year so …

(Question #2): Right, yes, so again I do not think it is reflected in the actual Margin of Error but it gives you more confidence?

Karen King: Well, the Margin of Error is calculated using the five years of data just like the estimate itself so it is impacted by the, you know, by the peers.

(Question #2): By the duration, okay.

Jennifer Berkley: Yes.

(Question #2): All right, thank you.

Coordinator: And our next question comes from (name removed).

(Question #3): Yes, I had a question about, or a clarification about the Margin of Error approximation. If the examples you gave were approximating the Margin of Error for various estimates with a variable, I was wondering, and maybe I misheard this, if you could approximate that same Margin of Error across different variables?

Jennifer Berkley: By different variables you mean different types?

(Question #3): Right, so yes, different types of estimates?

Karen King: I am just trying to think of…could you be a little more specific about what you are …

(Question #3): Sure, sure, what our group is actually doing is creating an index of multiple variables, and so using percentages, but if you were trying to add those percentages, we are looking for some measure to get a Margin of Error, or some sort of error across those summative variables.

Karen King: I think might be under the category of e-mailing us directly and then, because sometimes trying to explain how you are creating things can …

(Question #3): Right, right. It might be to in-depth here. Okay, thank you very much.

Coordinator: At this time there are no further questions in queue. Once again, if you have a question, please press star 1 and record your name.

Another question coming-in, one moment, please. Question comes from (name removed).

(Question #4): Hi, yes, if I want to use metro areas and compare them to non-metro areas, do I need to compute the standard error after aggregating Census Tracts, or counties, or are those available to me on the Website? You've got a very extensive Website here. Do you understand the question, was that clear?

Karen King: Say it again, what is it you are trying to do?

(Question #4): You have been talking about Blocks and Census Tracts, Block Groups and Census Tracts. I am doing an analysis at the metro/non-metro, so I am aggregating Counties to look at metropolitan areas. Are the standard errors available, or the Margins of Errors available for metro counties on the

Website somewhere, or do I have to compute them in the way that you showed us?

Jennifer Berkley: There are estimates available for metropolitan statistical areas if those might better suit you. Otherwise, if you want to specifically choose the geographies that you want to put together, then you would have to use the approximation method.

(Question #4): Right, that is exactly what I want, for the metropolitan statistical areas and where can I find those?

Jennifer Berkley: It would be available on AFF with all the other tables.

(Question #4): Okay.

Karen King: I believe there are instructions on AFF on how to locate things, but there is also information on any of the compass products about how to access that data, and I think that we also mentioned how to find tables and stuff like that in the presentation itself so…

Jennifer Berkley: Yes, I think there is training presentations can help you access ACS data and learn how to get to the different geographies and products.

(Question #4): Thank you.

Jennifer Berkley: Okay, yes.

Coordinator: Our next question comes from (name removed).

(Question #5): Hi, yes, this is (name removed). My question is about 5-year estimates. I know that there is some pending changes to the SOC, and possible changes to ethnic and racial demographic categories. How would bridging be handled for those 5-year estimates going forward and would there be any guidance provided?

Karen King: I believe that there will be some guidance provided but we are not as familiar with the changes in the questions and the resulting estimates that come-out of that. I think that there will be a lot more guidance on that in the future but it is not something we can probably help you with today.

(Question #5): So historically if you could talk about how bridging has been done in the past when there has been a change with respect to 5-year estimates?

Karen King: I am not sure if I can help you right now but again, if you want to e-mail us, then we can try to direct you to the appropriate people to help you.

(Question #5): Okay, thank you.

Coordinator: Once again, there are no further questions in the queue. If you would like to ask a question, please press star 1.

Another question coming-in, one moment, please. Question comes from (name removed).

(Question #6): Yes, I was wondering if you could speak briefly about coefficients of variation and other ways of interpreting Margin of Error, specifically when it comes to ACS data at the Block Group level?

We like to use the smallest unit of analysis possible but we also want to be using good data so is there any other resources you can direct us to that have some guidance about coefficients of variation?

Karen King: The coefficient of variation is not something we covered in this presentation so I do not want to get too technical but there is some guidance. I think it is in the accuracy statement, which is part, which was on the code list webpage that Jennifer went through, that will help you.

Also, if there is something specific that you would want to talk about with us, if you use the e-mail address that's on the screen, we can try to help you. It is not something that I could probably explain to you or give you guidance here and now so …

(Question #6): I guess can I ask just a quick question then? I will probably follow-up on e-mail. When we are calculating the coefficients of variation if we find a low reliability estimate, how do we use those estimates?

Jennifer Berkley: Well, one thing you could do is if you are trying to compare it to something or use the benchmark against it. Even if an MOE is high, it might still be statistically different than a certain benchmark you're looking at. You could look into completing some of the statistical testing to kind of see what that falls. You could also look at the confidence interval and see if that is acceptable for your needs.

(Question #6): Okay, thanks.

Coordinator: And our next question comes from (name removed).

(Question #7): Hi. In the example, you used subgroups so the two estimates were the subgroups of male and female within an age group. So my question is, if the estimate is if you are looking at an estimate and it is a result of the total estimate, and then you're subtracting the estimate of a subgroup of the total, would your calculation for the MOE change very much?

Karen King: Okay, if I understand your question, is that you have a total and you are trying to subtract a piece of that total away so you would get the balance? Is that the case?

(Question #7): Yes, so let's say if you were interested in looking at persons who did not identify as African-American, so you get the total, you subtract-out the number of who identified as African-American and that, you know, that estimate that you get, how would you calculate the MOE for that?

Karen King: Well, I know it is going to be a surprise, but you use the same formula. You would square the margins there, add them together, even though you are taking a difference, you would add the Margins of Error together just like we showed you in the example for a sum, and you take the square root ,and that will give you the Margin of Error for the balance after you subtract-off a piece of a total.

(Question #7): Okay, so that's great, so that makes me think that it does not matter if you are doing a sum or a difference, that formula for that MOE is pretty robust in that way. You can use the same formula.

Karen King: Yes, yes, exactly and if you go to the Accuracy document, that is one of the items on the resource slide that we showed you, if you go there it is talked about. It even shows you an example of a difference in how to do the Margin of Error.

Jennifer Berkley: And there is also equations for if you are looking for ratios or percent, how to do those as well and those documents.

Karen King: Yes, there is a lot of information on those in those documents that will be very helpful.

(Question #7): Thank you for that, and it would not matter if you were working with counts or percentages for that scenario that I asked about, the MOE calculation would still be the same?

Jennifer Berkley: It should be.

(Question #7): Okay, great, thank you.

Jennifer Berkley: All right, so thank you for the great questions, and thank you for joining us today.

Karen King: I think that is, we are at the end, we are going to wrap-up and say goodbye.

Coordinator: Are you ready to conclude then?

Jennifer Berkley: Yes, we are.

Coordinator: Okay, thanks everyone for joining us. Participants may now disconnect.

END