American Community Survey (ACS)
Using American Community Survey Estimates and Margins of Error Webinar Transcript

April 6, 2016

Sirius Fuller

Decennial Statistical Studies Division

This presentation was presented as a webinar to the General Public on April 6, 2016 by Sirius Fuller of the U.S. Census Bureau's Decennial Statistical Studies Division. A link to the recording is available at

https://censusevent.webex.com/censusevent/lsr.php?RCID=e993d604bff095314b5c2b03b39a781e

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

Coordinator: Welcome and thank you for standing by. At this time all participants are in a listen-only mode until the question and answer session. The conference is being recorded. If you have any objections, please disconnect at this time. I will now turn the call over to your host, Sirius Fuller. You may begin.

Slide 1 – Title Slide

(Sirius Fuller): Hello, thank you for joining us. My name is Sirius Fuller. I work at the Census Bureau in the ACS Variance Estimation and Statistical Support branch and today we're going to talk about using the ACS estimates and the margin of error and as I said, this is being recorded and it'll be up on the web later on.

So there's going to be some equations. Don't worry, you can reference then later. Oh, and as a note there's going to be a portion of questions and answers at the end and then if there's time there's some bonus slides that do not pertain to the margin of error but they're some useful things to know and if we have time we'll go over them and if not you can reference them later.

Slide 2 - Outline

So to begin here's an outline of the presentation. I'm going to talk about the margin of error and why they matter and as a note the margin of error is frequently referred to as the MOE and so we'll be using the margin of error and MOE interchangeably.

Then I will discuss how to use the margin of error to carry-out statistical testing and I will follow by an example of how to approximate the margin of error should you need to do so and then I'll briefly go over some resources and references and then we'll have questions.

Slide 3 – What is the Margin of Error

So to begin with here's the detail tables for sex by age on the American Community Survey on or AFF. The characteristics, the estimate and the margin of error are shown in the red square, the characteristics on the left followed by the estimate and then the margin of error.

In the upper left you might notice there's a little table ID B01001 and if we have time in the bonus slides I'll explain what that means but if not you can look it up later

Slide 4 – What is the Margin of Error (continued)

So what is the margin of error? Well, the ACS is a survey and it used a sample to estimate the characteristics about the total population in a timely and cost-effective manner.

If multiple samples were taken, each would likely yield a different estimate of population value. The MOE is a measure of the possible variation in the estimates. Data users can be certain at a given level of confidence that the estimate and the actual population value differ no more than the value of the MOE.

Slide 5 – Measures of Sampling Variability

There's a family of measures of sampling variability which are related to one another and so this slide shows you how you can derive one from the othe. So the top is the variance and then following that is the standard error or SE and that's found by taking the square root of the variance.

And then at the bottom is the margin of error which is found by taking the standard error and multiplying it by 1.645 which is a factor associated with a particular level of confidence and this particular level is the 90% confidence level which is the Census Bureau's standard.

Some other common confidence levels are in the 95 and 99% levels and in the next slide I'll show you how to convert from one to the other.

Slide 6 – Alternate Confidence Levels

So it's pretty easy to do. What you do is you take the published margin of error on AFF and you divide it by 1.645 to transfer it back to the standard error and then you multiply it by the applicable number. So if you want to get to the 95% confidence level you take the published 90% confidence level margin of error, divide it by 1.645 and then multiply it by 1.96.

Slide 7 – Why Do MOEs Matter

So now we're going to briefly go over why the margin of error matters and why you should care about it. One reason it matters is you can use it to calculate a confidence interval and the confidence interval is found by taking the estimate value and adding and subtracting the margin of error.

So in this example we have a median household incomes of roughly $37,000 and the margin of error of 20,922. So to find the lower bound of your confidence interval, you take the 37,284 and subtract off the 20,922 to get the roughly 16,000. In the upper bound you take the estimate and add the margin of error to get the 58,000.

Slide 8 – Why Do MOEs Matter (continued)

In addition, the margin of error allows you to conclude if there's enough evidence to conclude that the estimates are statistically different from one another. So as an example, here are five medians from different block groups and block groups are very small geographic areas that the Census Bureau publishes.

And I have these in order so at the top is the first block group which was shown on the last example of roughly $37,000 and it generally increases or appears to increase to the block group Number 5 of $76,833.

Slide 9 – Why Do MOEs Matter (continued)

Now here's the same data although we've included the margin of error. So the so-called lowest median as a margin of error of 20,922 which is what we saw in the example two slides ago and the so-called highest median of 76,833 at the bottom has a margin of error of 47,268.

When you take into account the margin of errors, you cannot conclude that any of these estimates are statistically different from one another. And as a note these medians are from actual published block group geographies and I chose them for their dramatic effect because of their high margin of errors.

Not all the medians at the block group level have such high margin of errors but it's important for users to be aware of this issue so they do not simply use the estimate to attempt to conclude whether or not two block groups are different from one another.

In addition I took this example from a similar one I saw at the second annual ACS Data User Group conference and if you do not know, there is an ACS Data User Group and you can join it online for free and you can ask questions

and receive answers from other members and at the end of this presentation I'll provide the link on how to get there and how to join.

Slide 10 – Testing for Statistical Significance

Now I will talk about statistical significance. To begin with, you can use the statistical test to determine if a difference is unlikely to occur by chance. So it would be statistically different means that there is statistical evidence that there is a difference between two estimates.

Slide 11 – ACS Comparison Profile: Year to Year Change

There are some published ACS data products on AFF where the statistical comparison testing has been done for you. So the first is the comparison profiles and these are similar to the data profiles except they list the estimates from prior years. And there's also an additional column with an asterisk on them next to the estimate if the two estimates the current year and the prior year are statistically different from one another.

And you don't need to remember what the asterisk means because at the top of the - above the estimates - is a little legend to let you know. It says star indicates that the estimate is significantly different at the 90% confidence level.

Comparisons for one-year data are available for a wide variety of geographic areas so for example you can get them from the nation, the states, each county, school districts and Congressional districts.

Unfortunately, the three-year data products have been discontinued, although the previous three-year data products which were published are still available on American FactFinder and very exciting, there will be a five-year comparison profile available with the next data release at the end of this year.

As an aside as a short note, I will be showing several screenshots of ACS data and the examples may be from prior years and they're just for illustrative purposes, mostly they're from 2013 and 2011.

Slide 12 – ACS Ranking Tables

So the second place where statistical testing is done for you are in the ranking tables and they show you how different geographic areas compare to one another. And these are a little different than the comparison profiles. You have to turn the statistical testing on and you can see that in the red the circle that says without statistical testing. When you turn it on, it flicks back and forth.

You can select the state from the dropdown menu and then the statistical testing will appear. And so you can see that in this one the United States is selected and it has two pound signs or hashtags if you will and then if it's not significantly different, you get a single pound sign. So here Arizona and Colorado are not significantly different from the United States.

Unfortunately, only the nation and states are available for the ranking tables. There are similar tables called the geographic comparison tables. However, statistical testing is not available for those. There is a tool being developed which will allow you to conduct your own statistical testing and we hope to have it available by the end of the year. It's currently under review.

Slide 13 - Statistical Testing

So there are many estimates and in the ranking and comparison profiles, however there are plenty of times when you may wish to make some statements that either directly or indirectly implies a comparison and the testing is not available.

And so I'm going to walk through a brief example in how to conduct your own statistical testing. So the example is suppose you wish to compare the median age of people who report being Asian, alone or in combination versus

the median age for the United States. So the median age for the U.S. is 37.3 and the median age for Asian alone or in combination is 33.5.

And if you're wondering, this example is the one-year ACS from 2011 so it's from a little while ago.

Slide 14 – Statistical Testing (continued)

So first we need to find the appropriate formula and all the formulas are available through online resources and at the end of this I will show you how to get there.

This formula's in a few places, the Compass products and the instructions for statistical testing and I'll mention that again. So we're going to use this formula which uses the MOEs and if the result when we plug-in the numbers is greater than one, then we'll conclude at the 90% confidence level that the two estimates are different from each other.

And as an aside, this is a modified Z score formula. It's modified to use the margin of errors and that's done because it makes the formula easier and the margin of errors are readily available. The regular Z score formula uses the standard errors.

Slide 15 – Statistical Testing (continued)

So bear with me as I walk you through this example. So at the top we have our two estimates, the Asian alone or in combination is 33.5 with a margin of error of .02. The total population is 37.3 with a margin of error of 0.1. So we're going to do it step-by-step with the formula. So you take the difference of the two in Step 1 so 33.5 minus 37.8 gives you negative .38.

And then we take the absolute value of that so you get 3.8 and then we square the margin of errors so .2 squared is .04 and .1 squared is .01. And then in Step 4 is you add them together so you get a .04 plus .01 is 0.05.

Slide 16 – Statistical Testing (continued)

And then in Step 5 we'll take the square root of that so the square root of .05 is .224 and then we divide Step 2 by Step 5 so you get 3.8 over .224 or 16.99.

And so that's the formula which I'll show you in a minute and as I said before we compare that to one and 16.99 is obviously larger than one so we can conclude that the median age for Asians alone or in combination is lower than the population median age.

Slide 17 - Statistical Testing (continued)

So just to put it all together, here's the formula again and below it is the what we just went over but altogether. So that we put all the numbers in and we've plugged and (chugged) and we got 16.99 and we compared it to 1.0.

Slide 18 – Statistical Testing (continued)

So the nice thing about the Z score formula is they may use it to compare a wide range of estimates together or estimates between different years, multiyear estimates between different non-overlapping time periods, etc, etc.

And as an aside, if you're comparing ACS estimates to Census estimates, you must make sure that the estimates are comparable and the URL at the bottom of the slide (https://www.census.gov/programs-surveys/acs/guidance/comparing-acs-data.html) allows you to check whether or not they are comparable or if the questions have changed to such a degree that you cannot compare them.

Slide 19 – Statistical Testing Tool

As I mentioned earlier, there is a tool being developed to allow you to conduct your own statistical testing. And this is a screenshot, - a preview, and it's designed to be friendly to use and allow you to handle the special cases that you may see in the from AFF and you just put it in and it does the statistical testing for you.

Slide 20 – Special Case: Controlled Estimates

So as I just mentioned the special cases, let's go over them.

The first one you might see is the controlled estimate. So in the ACS certain estimates are controlled through the official population estimates and if this occurs, you will see an estimate and then the margin of error will have five stars or five asterisks and this is fine. You can use these numbers in statistical testing. Just simply convert the five stars or five asterisks to zero.

Slide 21 – Special Case: Zero Estimate MOEs

Another special case you may notice if you have a zero estimate, they will have a margin of error which is non-zero next to them and the margin of error for zero estimates is created using a different method than for non-zero estimates.

The zero estimate margin of errors are created using a model and zero estimates have a non-zero MOE because you may have a household with a rare characteristic which may not be surveyed but that the people may still exist so we want to account for that.

Slide 22 – Special Case: Medians and Aggregates

There are also special cases where you cannot conduct a statistical test. So if you have only one person, it does not make sense to calculate a median so you'll see an estimate with a dash not a number and then the margin of error will have two stars or two asterisks.

Another case is an aggregate with too few cases. So an example of an aggregate is if you have the total income of all people in a geography. So if you have only one person with a certain characteristic, their aggregate will not be displayed to protect their confidentiality.

And finally if you have a median in the highest or lowest interval of an open-ended distribution then the estimate will have a plus or minus after it and the MOE will have three stars. So for this example you can see the median has an estimate of $2500 with a minus sign after that so that is not actually the median estimate.

All you can say is that the median is below this value. It's not actually a number. And as an aside if you use the API which is also the application programming interface, you may see instead of 2500 with a minus after it you may see 2499 and the MOE will have a value of null and the null value will indicate to you that this estimate is not a number.

So for all these cases, you cannot use them to perform statistical testing.

Slide 23 – Deriving New Estimates: Must Approximate the MOE

There are also instances when you must derive an estimate from the ones that are published on AFF.

So for example you might have to add together some age groups to get a total age. So in this example you want the total population under five years old but

it's broken out by male and female. So you'll have to add those two estimates together and when you do this, you'll have to approximate the margin of error and it's a little more complicated but it's not that much more complicated.

Slide 24 – Approximating the MOE

So what you do is you add the estimates together and then similarly for the margin of error, it is somewhat similar to the Z score in that you square each margin of error, you add them all together and then you take the square root.

Note that this is an approximation to the margin of error and it'll probably not exactly match the margin of error we produce using the raw data, the microdata that we have. The microdata is not available to the public obviously to protect confidentiality, but this is an approximation and it tends to work pretty well and it is useful.

Slide 25 - Approximating the MOE (continued)

So let's go through a worked example.

To obtain the total number of children under the age of five which we just saw, we're going to add-up the total number of males and females. So the estimate for males is roughly 10.2 million and the estimate for females is 9.8 million so you'd add them together and get 20 million, a little over 20 million.

So to approximate the margin of error you would take the margin of error for males 4089 and square it to get roughly 16.7 million and then the same thing for females. You take 3460, square it to get roughly 11.97 million. You'd add 16.7 to 11.97 and then you take the square root and that is at the bottom of the formula in a little bit different form.

When you take the square root, you get approximately 5356. I mentioned that this is an approximation. This formula does not take the covariance into account and I'm not going to discuss what the covariance is today.

It's beyond the scope of this presentation so you should know it's not available to the public and if the covariance is very large, then your approximation could differ from the actual margin of error.

Slide 26 – Approximating the MOE (continued)

So one way the approximate margin of error could differ is that if you add together a lot of zero estimates and remember I said if you have a zero in the estimate the margin of error will be non-zero. So this is a bit of a contrived case.

But we're going to add-up the total number of children under 10 in Delaware and Washington, D.C. and the unofficial tip is when you add-up a lot of zero estimates, you take only one of their margins of error and you use the largest one.

So here I'm going to add to get the estimate you add 36 plus zero plus zero plus zero and it gives you obviously 36. To approximate the margin of error you take the 56 which is the non - margin of error for - the non-zero estimate and you square it. And then you select the largest margin of error from the zero estimate so we have 27, 29 and 29 so you use 29 once because it's the largest and you square that and you add them together and take the square root.

So this is the approximation formula that we saw earlier but we're using only one margin of error from the zero estimates. If you do this, you get an approximate margin of error of about 63 and if you'd used all the margin of errors available you'd get 74 which is a little bit larger.

Slide 27 – Estimates with Large MOEs

You may have also noticed that some estimates have a margin of error which is larger than the estimate itself and this can occur when examining smaller geographies, or small groups of people, or households which have a small sample size.

Data users should exercise caution when using these estimates because they have unknown reliability. A margin of error, which is larger than the estimate, can be used as a signal that the sample size is small. The example shown in the slide is for the poverty status.

People above and below poverty at the tract level and tracts are one of the smaller geographies that are published for the ACS and as you can see many of the estimates have margin of error, which are larger than the estimates themselves.

Slide 28 – Estimates with Large MOEs (continued)

When this occurs, there's no ideal solution but there are some ways to deal with it. One way is to use a larger geographic area so if you're using tracts, you may want to use counties instead. Another method is to form larger groups by combining estimates or geographies together and you may have to approximate the margin of error as we saw in the example before.

Slide 29 – Collapsed Tables

Note that for some of the detailed tables, there exists a collapsed version of the table and so here's an example. On the left we have sex by age for black or African-American alone and the alone just means that people reported a single race of black or African-American. So on the left we've got four categories.

We have under five years old, five to nine, 10 to 14 and 15 to 17. And on the right is the collapsed version of the table and there's only one estimate which is for under 18 years.

You may notice again in the upper left of these tables on the one on the left it says B01001B and so that is a table identification number and on the right you have a collapsed version which is C01001B. So the detail table will start with a B and the collapsed version will start with a C.

So the collapsed tables you don't have to approximate the margin of error, it's provided for you and the categories are created by subject experts. They may not be exactly what you want but they'll allow you to combine fewer estimates which will provide you with a better approximation to the margin of error.

Slide 30 – Resources

So now I'm going to go over some useful resources and references and this is a link to the, let's see, the code list and the definitions and the accuracy document, the URL for these is at the bottom of the slide (http://www.census.gov/programs-surveys/acs/technical-documentation/code-lists.html). And so at the just above the URL circled in red are the instructions for applying statistical testing and in those documents you can find how do approximate the margin of error.

The example we went over it but also if you want to calculate percents or ratios or other things, they're there. Sort of in the middle to the right circled in red are the accuracy of the data documents and the accuracy of the data documents provide worked examples as well as the equations and they also have things pertaining to the sampling and the weighting.

So if you want to find the worked examples, they're sort of towards the end of the document but there's an appendix in the - excuse me, not Appendix - a table of contents is the word I'm looking for. On the second page they'll allow you to jump, to tell you exactly where to look.

You may notice that there's several documents. There are some for the multiyear data and some for the single year data and that's because there are issues unique to the multiyear estimates that must be covered but in terms of the worked examples and the approximations, they're the same.

Slide 31 – Compass Handbooks

So here's a link to the handbooks for data users (http://www.census.gov/programs-surveys/acs/guidance/handbooks.html). They're also called the Compass products and there's several, there's a bunch of varieties. There are some for teachers, there's some for Congress, there are researchers, there's a whole bunch.

They also have worked estimates and the approximation formulas that are similar to the instructions for stat testing and the Accuracy document. In the Compass handbooks, appendices are a great resource for referencing when you calculate the margin of error.

Slide 32 – Training Presentations

There are also training presentations which walk you through various aspects of the ACS data and the slides and recording and speaker notes for this webinar will be added to this webpage soon.

There are also recordings for recent public webinars which you may have attended called Accessing the ACS Block Group Data and the Introduction to the PUMS or Public Use Microdata Sample so the URL for that page is at the

bottom of the slide (http://www.census.gov/programs-surveys/acs/guidance/training-presentations.html).

Slide 33 – Crosswalk

Here's the at the bottom of the slide is the URL for the crosswalk to compare ACS detail tables to Census data and as a note the Census 2010 data is also on American FactFinder (http://www.census.gov/acs/www/guidance/comparing-acs-data/acscensus-table-lookup).

Slide 34 – Design and Methodology

Finally if you want a very technical description of the ACS and how the margin of errors are created, you can look at the design and methodology documentation and the relevant chapter is Chapter 12 for the variance estimation. This is probably a lot more technical than you're looking for but should you be interested, here it is and once again the URL's at the bottom (http://www.census.gov/programs-surveys/acs/methodology/design-and-methodology.html).

Slide 35 – Source Us!

And now I've got a few slides just for information only purposes and then after this we're going to open-up to questions and if there's time at the end, there'll be some bonus slides.

So it's important if you're using ACS data to please source us and in order to source us you would say the U.S. Census Bureau, the year and then say American Community Survey and then the time and then you can use estimates, statistics, data releases. It's sort of interchangeable.

So for example if you're using the five-year 2010 to 2014 it would say U.S. Census Bureau's 2010-2014 American Community Survey 5-year estimates.

Slide 36 – Continue the Conversation #ACSdata

In addition, I also encourage you to connect with us either conventionally or using social media. We love to hear from you. If you want, you can use the hashtag #ACSdata

Slide 37 – American Community Survey Data Users Group

And as I mentioned earlier, there's a data user group which you can join for free.

The URL (http://www.acsdatausers.org) is at the bottom and you can post questions that you have about the ACS data. There are several different forums for, you know, there's one for the margin of error, one for the PUMS data, if you're interested and there are several others and it's free as I said.

Slide 38 – Need Local Stats?

And finally as a reminder we have data dissemination specialists throughout the country who can provide data workshops for you locally and if you're interested in a workshop, you can contact them. The e-mail address is there. it's clmso.ddb.questions@census.gov.

Slide 39 – Questions?

So thank you for joining us. That concludes my presentation and we're going to open it up to questions now and I'm going to show you where the evaluation of the presentation is.

Coordinator: Great, thank you.

Sirius Fuller: So you should see the, yes, the evaluation.

Coordinator: At this time if you'd like to ask a question, please press Star 1 on your touch-tone phone and record your name clearly when prompted. Again if you'd like to ask a question, please press Star 1. Thank you. Please standby for the first question. (Dennis Holton), your line is open.

(Dennis Holton): Thank you. I use ACS five-year data and on occasion I have measures that have zero margin of error such as the number of Hispanics. I believe these are coming over from Census Population Estimates. Can you tell us more about that connection and how we handle that from an estimate of uncertainty?

Sirius Fuller: Yes, so those Hispanic is one of the estimates that could be controlled. So there's the official population program which releases estimates of the population between the decennial Censuses and so those estimates in the ACS are controlled so they are made to equal the exact official population estimates.

If you look at the Population Estimate Program on the Census website and look at the ACS estimates, they should be exactly equal so you can consider them to have a margin of error of zero. There's no variance associated with them and you'll see if it's on the API you'll have a margin of error of zero or you may see as I said the five asterisks, the five stars and I hope that answers your question.

(Dennis Holton): Thank you.

Coordinator: Thank you. Our next question comes from (Emily Goldman). Your line is open.

(Emily Goldman): Yes, hi. Thank you very much for the presentation. I still have some sort of weird block when it comes to margin of error. It's just my own thing but I have a question which is that I need to get over obviously but my question is are margins of error always the same unit as the estimate? It's probably a stupid question but there you go.

Sirius Fuller: No, no stupid questions, that's a good question. Yes, they should be the same unit so if you have an income estimate, you know, it's in dollars then the margin of error will also be in dollars and, you know, if you have a count of the total number of households say the estimate will be plus or minus the number of households.

(Emily Goldman): Thank you.

Sirius Fuller: You're welcome.

Coordinator: Thank you. Our next question comes from (Claudia Meltanga). Your line is open.

(Claudia Meltanga): Hello, thank you for the - can you hear me?

Sirius Fuller: Yes.

(Claudia Meltanga): Okay, thank you for the webinar. My quick question is when you're using ACS data, is it important how should we include disclaimers kind of when presenting this information to non-data users, should we or does it confuse people? How do you suggest we cite it, the margin of error?

Sirius Fuller: That's a very broad question and it kind of depends. So if you are from our point of view, we love the margin of error so we always try and, you know, you have the estimate and then the margin of error but so it really depends on your audience.

We have web apps and some of the web apps you may know. There's one called EasyStats and then EasyStats, you'll see the estimates and then there's a little box so you can hide the margin of error or you can show it.

There are QuickFacts where we just show the estimates and then you can click, there's a little circle with an eye in it usually and you can click on that and that'll take you to a brief description about the ACS or there are several different datasets. There's a population estimates in there as well and there's a link to it so it depends on your audience.

They may care to see it, you have the estimate, you could show the estimate as, you know, smaller font in gray with a plus or minus or it really depends. It's a very broad question but there's different ways, it depends on your audience is …

((Crosstalk))

(Claudia Meltanga): Well, that's what I was kind of saying for non-people who just are looking at data overall, they're really not statistical people. They're really non-data users so just kind of laymen. Is it important to show the margin of error when, you know, just presenting information to non-data users?

Sirius Fuller: I'm going to bunt on this question and say that maybe you should contact us offline for your specific estimate.

(Claudia Meltanga): Okay.

Sirius Fuller: But it really depends what you have.

(Claudia Meltanga): Okay. Thank you.

Sirius Fuller: Yes, thank you.

Coordinator: Thank you. Our next question comes from (Gene Shackman). Your line is now open.

(Gene Shackman): Hi, yes, I have a question about the margin of error but I want to respond to the previous person. I would say, you know, I always include the margin of error because then you can tell the audience whether the differences, whether the two estimates are statistically different.

But my question is I learned using the margin of error to construct confidence intervals for the two estimates and if confidence intervals of the two different estimates overlap, then they are not statistically significant, they're not statistically different and if they don't overlap they are different. Is that another way to use them?

Sirius Fuller: Yes, that's a common rule of thumb and actually in the accuracy document we have a short example. That is a good rule of thumb but it's not always true so we recommend calculating the Z score which is the example we did in the presentation because sometimes that rule of thumb although it's a good one can be misleading.

(Gene Shackman): Oh, yes, all right. You have a pointer to that, to explaining why that doesn't always work?

Sirius Fuller: Yes, so in the accuracy of the data document, I think we provided the there is one slide with the instructions and the other with the accuracy of the data. If you go down to the worked examples I think just before the worked examples, somewhere right around there, there's a short, you know, example with some made-up data but here's the estimate in the margin - two estimates - and their margin of error.

And it has the confidence intervals and then it has the Z scoring. You get different results.

(Gene Shackman): All right, thank you.

Sirius Fuller: You're welcome.

Coordinator: Thank you. The next question if from (David Chu). Your line is open.

(David Chu): Oh, hi, yes, my question is about the new statistical testing tool to come out, whether or not it will allow people to calculate the margin of error for, you know, the sum of a number of published estimates and if true whether or not it's still going to be an approximation.

Sirius Fuller: Yes, so the statistical testing spreadsheet, it will not do the approximations for you. There are other outside of Census there are some statistical testing tools, some spreadsheets where you can put the estimates in and it will sum them up and do the approximation for you but this specific statistical testing tool does not do that. It just does the statistical testing, it does the Z score test for you.

(David Chu): Okay, thank you.

Coordinator: Thank you. The next question comes from (J.J. Chung). Your line is open.

(J.J. Chung): Hi, thank you very much for the webinar. My quick question is on the MOE aggregation so you mentioned if we aggregate MOE for a lot of zero estimates, we would probably get a biased result.

So my question is from previous practice we had trouble with aggregating MOEs for non-zero estimates. We still get very different results. Just wonder if it matters how many estimates I aggregate together, how many MOEs I aggregate together or it's mostly about the zero estimates?

Sirius Fuller: Yes, so that's an excellent question. Anecdotally and this is just anecdotally I haven't done - there hasn't been research - as far as I know but it seems if you had four or fewer estimates together, you'll be all right so there's if you add a lot of zeroes together, your approximate margin of error can be off from the actual margin of error.

Another way is if you add a bunch of estimates together and they sum up to an estimate that is controlled or nearly controlled and so the estimate with the official estimate will be controlled and so you'll have a zero margin of error or you'll be very close to controlled estimate and so it could be you have a margin of error which is smaller than what you'd normally get and so your approximation could look off as well.

Yes, and I mentioned earlier that there's the covariance, yes, so there's no covariance published with the margin of error. That's not - to do all covariances for all, you know, the two estimate and three estimates and four estimates together - that would make the tables on AFF basically unusable.

If you so I don't know the sophistication of different data users but one method you could use is the Public Use Microdata Sample and that allows you to calculate the actual margin of error and that's sort of a separate subject and there are plenty of caveats and issues that go with that but yes, if you approximate a, you know, add-up a large number of estimates, you can your approximate margin of error could diverge from the actual margin of error.

(J.J. Chung): Thank you.

Sirius Fuller: Yes, and just one more, just as I said before, you can use a larger geographic area. We often say instead of tracts use counties but there are smaller geographic areas that you could use instead of counties. The idea is just to use

a larger area so you have a larger sample together. That might help you as well so thank you for your question.

(J.J. Chung): Thank you.

Coordinator: Thank you. Our next question is from (Alicia). Your line is open.

(Alicia): Yes, I was wondering if there's a sort of like a little rule of thumb you would use to determine whether or not a particular estimate was statistically unstable based on the margin of error?

Sirius Fuller: Yes, so there's a thing called the coefficient of variation and so you can take the margin of error and divide it by the - what am I thinking of - yes, the margin of error divided by the estimate*.

There's no official cutoff of what makes a of what is considered unreliable. It depends on the context but if you have as I said on the slide before, if you have margin of errors that are equal to or larger than the estimate, then those are considered in general to be unreliable. It makes sense, you know, if an estimate of 100, plus or minus 120, you know, you kind of look at that sideways.

(Alicia): Uh huh.

Sirius Fuller: So I hope that answers your question.

(Alicia): Yes, thank you.

Coordinator: Thank you. The next question comes from (Max Crogalecki). Your line is now open.

*Correction: The coefficient of variation is actually the standard error divided by the estimate.

(Max Crogalecki): Hi, thank you and this is basically along the lines of the last question I think is if say if I'm looking at a series of five-year estimates and in one of those years the MOE takes me to a zero value that the negative MOE takes me to a zero value, my question is basically what am I to make of that and I think that goes along with the previous question? Am I correct about that and then in that case, what do I do, ignore that, that estimate then?

Sirius Fuller: Yes, so the first part of your question, you could, you know, if you're reporting a confidence interval or something, it doesn't make sense to have negative people, you know, if you have a 100 people plus or minus 120, you could, you know, go down to the lower bound would be negative 20. That makes no sense.

So you'd have to, you know, you'd change that to zero, you know, it doesn't make sense to have negative people so that if you're you have to be cautious in using that so if you wanted to draw a conclusion, you know, or there are 100 people there, you might, you know, put a little it depends on what you're doing but I would be very cautious in using that. You might want to, yes.

(Max Crogalecki): I'm sorry. Let's say the years around this zero, I get that you would never report a negative for a number of people. You know, it's zero, but let's say that zero value occurs in a series of estimates that again I'm saying looking at five-year estimates but, you know, there's positive counts all around it, you know, for years beyond it and years before it. What am I to make of something like that, just add data?

Sirius Fuller: No so …

(Max Crogalecki): Go ahead.

Sirius Fuller: … so if you have five-year data, you shouldn't really be comparing overlapping five-year periods so if you have 2005 to 2009 and I don't know,

2008 to 2012, they overlap and so, you know, they're not we recommend not comparing them.

But it is true, you may have because of this five-year data you could have a very small sample in one of the five-year periods and there's no sample in the next, you know, that people sampled were in 2005 and then when you move to a different say the 2010 to 2014 there was nobody in the sample.

In general I would suspect that the non-zero estimate would be very small and have a large margin of error so …

(Max Crogalecki): Yes, and if you were looking at a series and if I'm taking-up too much time, just let me know, I'll shut up but if you're looking at these series of five years and I understand what you mean by not because they overlap so in that sense you would never average a series of five-year estimates. Is that correct?

Sirius Fuller: You can't consider the five-year time period to be the average. You can't say it's the middle year. It's an over the average, that five-year time period is what the estimate is for. If that doesn't answer your question, perhaps we could talk offline because I feel like you've got more questions.

(Max Crogalecki): I'm sorry, I didn't mean to take-up so much time, thank you.

Sirius Fuller: No, no, that's fine, so thank you, thank you for your question.

Coordinator: At this time I show no further questions.

Sirius Fuller: All right, we'll pause for a minute and then I'm going to show you a few - okay - since we have a couple of extra minutes, I'm going to show you the bonus slides and these do not pertain to the margin of error but I found them to be useful and they may be useful to you as well.

Slide 40 – Bonus Slide: Finding Your Geographies

So the first bonus slide is to find your geographies so you can go to American FactFinder and click on geographies and then click on address and put in your specific address and once you hit go, it'll show you the different geographic summary levels where you're in so you can only see at the bottom that it's in the region is the south region and the division is South Atlantic and then to Maryland for this example.

But if you go further down, it'll tell you such things as what Congressional District you're in, what tract you're in so that can be very useful because you don't need to guess what geography you're looking for, if you're looking for your own personal geography.

Slide 41 – Bonus Slide: How to Determine if You Are in a Survey

Another bonus slide is how to determine if you're in the survey.

And so I've had relatives who have gotten the ACS and they call me up and say "Is this for real?" and it's important to note that if it is a legitimate survey, you can go to you can call the Census Bureau or contact us and we can say yes, this is a legitimate survey. Please fill it out so the URL for that is at the bottom of the slide (http://www.census.gov/programs-surveys/are-you-in-a-survey.html).

Slide 42 – Bonus Slide: Why We Ask the Questions on the ACS

In addition there are questions about why are we asking the questions in the ACS and this links to the link at the bottom (http://www.census.gov/acs/www/about/why-we-ask-each-question/) provides a page that breaks the questionnaire down question-by-question and gives a

very brief synopsis of the history and the laws that are applicable and why it's relevant that we're asking the questions in case you are interested.

Slide 43 – Bonus Slide: Explanation of Detailed Table Code

And then the final bonus slide is as I mentioned before, there was a detailed table IDs and you might see something. This is a little oddly color-coded at the top but you'd see something like B01001APR and so if you go to AFF you could search, you know, one of our examples was sex by age for black alone.

You could search for that and get the table. You can also enter the table ID directly into AFF and it'll just take you directly to the table so the first letter will be a B or a C and that's for the B is for detailed table and C is for collapsed.

The second two digits is a two-digit code here, it's 01 and that pertains to the subject and below there's a URL that's an ask (https://ask.census.gov/faq.php?id=5000&faqId=1687). It's an old version of the website but it provides you the different subject groups for each of the tables and there's a new relatively new subject group, the B28 series for computers and Internet which isn't on there.

Then you have a three-digit 001 which is the table number and then if you have a race or Hispanic iterated table, you'll have a letter A, B, C, et cetera, through I and then finally if the table is produced for Puerto Rico specifically for the Puerto Rico community survey, you'll have a PR at the end of the at the detailed table code so that's just for your information.

And that concludes everything I have and if there are any more questions, I will take them at this time.

Coordinator: Thank you. Again if you would like to ask a question, you may press Star 1 on your touch-tone phone. Please record your name clearly when prompted. Again press Star 1 if you'd like to ask a question. Thank you, our question comes from (Steve DeCosta). Your line is open.

(Steve DeCosta): Hello, I use ACS to compare different time periods sometimes and I just want to know like if when I find like a difference between them, you know, I can say there's an increase in '10, is there a way to determine the margin of error of that number, the difference number?

Sirius Fuller: Yes, so you would approximate the margin of error so it's the you would take the - we did it in the earlier examples for the Z score - it's just the each individual margin of error squared, add it all up and then square root so that would be the margin of error used.

(Steve DeCosta): All right, thank you.

Sirius Fuller: You're welcome.

Coordinator: Thank you. (Emily Goldman), your line is open.

(Emily Goldman): Hi, yes, with regard to this first bonus slide, if you put in your address, can you get all the way down to your block?

Sirius Fuller: I believe so. I've it's a whole lot of some I think it's all the published summary levels. I'll leave it up to you an example to try. I'm pretty sure there's tract and I think there's block group but I'm not certain.

(Emily Goldman): Thank you.

Coordinator: Thank you I show no further questions at this time.

Sirius Fuller: All right, well thank you all for attending and once again this'll be up shortly on the website with the transcript and the recording.

Coordinator: Thank you and this concludes today's conference call. You may disconnect at this time. Speakers, please standby.

END