Finding good stories from the Census Bureau's American Community Survey

November 15, 2006

Presentation by D'Vera Cohn and Cynthia Taeuber
Moderated by Andrew Reamer, Metropolitan Policy Program, The Brookings Institution

D'Vera Cohn: Over the next hour, we'll go over how to find and use numbers from the Census Bureau's American Community Survey, which produces a detailed snapshot of the nation, states and communities that is updated each year. The newest estimates from the survey -- for 2005 -- have just been released. The survey offers details on dozens of topics, including how people live, how they make a living, and what kinds of housing they live in. The idea behind the survey is to replace the once-a-decade snapshot from the Census itself with a yearly video that will help you monitor trends. The Census Bureau has been taking the survey for several years, and past data are available for states, as well as larger communities. What makes the 2005 estimates so useful is that they are available for many thousands more communities than in the past – for places as small as 65,000 people. By learning how to use it now, you will give yourself stories this year, and will be ready to track trends over time when the numbers come out in future years.

We'll look at many ways to use the survey for story ideas, graphics or that last-minute fact you need on deadline. There also are some cautions you need to know. One major one is that these numbers really are "estimates," as with any survey, and you need to keep that in mind as you try to draw conclusions about what they have to say.

This entire presentation will be up on the Brookings website, so you don't have to take detailed notes now. There also will be more in-depth material and examples there. And we'll show you how to reach the Census Bureau or experts in your state who can give you additional help.

Slide 2: Here are some examples of questions you can answer with the American Community Survey. The survey produces the same type of information that we learned from the long form in the 2000 Census. You can find out about people's living arrangements, how much money they make, what their housing is like, and details on special topics such as immigration and poverty. You can compare different racial and ethnic groups, contrast the situations of men and women, rank your state or county, and make maps to display the numbers in easy-to-understand form. And the annual update gives you a fresh picture, unlike the every-10-years census numbers that quickly go out of date.

Slide 3: This gives you a flavor of what the American Community Survey covers. As you can see, it offers something for almost every beat or for the general assignment reporter. There won't be time to go into detail on each of these today, but we'll put more information on some of them up on the Brookings website.

Slide 4: You can look at numbers for many levels of geography, and compare cities, states, counties or other areas. The American Community Survey is a household survey in 2005, but

the intention is that there will be a full population survey from 2006 onwards, including places such as nursing homes and college dormitories. By the 2010 census, assuming that the president and Congress provide adequate funding, it is supposed to have accumulated enough details to offer data at the neighborhood level. The Census Bureau has announced that the survey will replace the long form in the 2010 Census. That means that the 2010 Census will focus on counting people and housing units, and collecting a few basic demographic characteristics.

Slide 5: You can compare racial and ethnic groups in your area or nationally, looking at differences in income, homeownership, education levels and more. You can check out the incomes of single mothers versus single dads, or compare the kinds of jobs men and women hold, or see whether there's any difference in their commuting patterns.

Slide 6: There are some limits to the American Community Survey, and we will go into more detail later about what you can and cannot use it for. The American Community Survey is gradually building up the number of areas available. This year for the first time, it offered data for counties and other geographic areas with 65,000 or more people. The most solid numbers available now are for areas of at least 100,000 people.

Through 2005, the American Community Survey includes people living in households only. The estimates provided so far do not include people living in institutions such as military barracks, prisons, nursing homes and college dormitories. These are called "group quarters." In 2006, the survey began to collect data on people living in these group quarters. We don't know what will happen in 2007 – it depends on whether the federal budget has adequate funding to include people living in group quarters. If people living in group quarters are not included, we will not have estimates about the characteristics of the total population from the American Community Survey.
You should keep in mind who is left out and avoid doing stories about college students, prison inmates, people living in nursing homes or your local military population until the survey includes data for group quarters.

Slide 7: So let's look at some numbers. Here is how to begin. All the survey estimates are available on the Census Bureau website through the American FactFinder page. You can bookmark the American FactFinder page, or you can get there from the Census Bureau home page.

Slide 8: Once you are on the FactFinder page, there are other numbers you can look at, including economic reports and the bureau's annual population estimates. The population estimates, as I'll remind you later, are your best source if all you want is the total number of people living in your county or some other place. But if you want to know how people live and work, turn to the American Community Survey. On this slide, we've added a red arrow to mark how to get data from the survey. Click on "get data" to find your numbers from the 2005 survey.

Slide 9: You can choose among several different ways to display your numbers. You can look at tables, maps or narrative presentations.
The Census Bureau rolled out the 2005 estimates in stages over the summer, and released the last details yesterday (Nov. 14). These are called "selected population profiles" and they

have in-depth information on dozens of racial and ethnic groups. For example, you can ask for estimates about the Asian Indian population in the Washington, D.C., metropolitan area, or the Mexican population in Los Angeles County.
If you really want to dive in, look at "detailed tables," but I wouldn't recommend that for first-time users because you can get overwhelmed. We'll go over the best ones for new users next.

Slide 10: These are my recommendations for people who are new to census numbers. We'll go over the best use for each one, and what to look for. But let me put in a word for maps, which may not be the first thing you think of when you start out on a story that uses numbers. I like maps because they can be a way to test a story idea. Maps can show you patterns more easily than you can see when you are staring at a column of numbers.

Slide 11: Once you decide whether you want a data profile, a map or something else, you need to decide which level of geography you want to look at. Are you doing a national story, an article about your state, an overview of your metropolitan area? A dropdown box prompts you to select something. You have options to look at estimates for the nation, states, counties, places (which include cities), congressional districts, school districts, metropolitan areas and other geographies. [If you choose "county," "place" or another geography that is listed under "state" in the dropdown box, you will be prompted to choose a state and then given a list of what's available to choose from.] It is pretty simple to go back and look at something else.

Slide 12: Here is the first page of a data profile for Tennessee. The data profile is what you should grab on your way out the door to cover an earthquake in a community you've never been to before. It's also a good entry point for writing any kind of story where you want general information about a community.
It has four pages of numbers and one page of narrative. The first page is demographic information, including people's living arrangements. The second page has "social" topics such as education, immigration, marriage, births, disability, grandparents taking care of grandchildren and ancestry. The third page has economic data such as the type of jobs people have, how they commute to work, and how much money they make, including categories such as earnings, retirement benefits and food stamps so you can look at the source of money in more detail. The fourth page offers housing information, including who owns their homes and who rents, housing costs, how many homes have nine or more rooms, which is great for stories about mansions, and how many homes are considered overcrowded.
The last page includes much of this information in plain English, which can help you translate census jargon and put numbers in context.

Slide 13: Rankings are a fun way to look at how your state stacks up compared with others, or to find the lowest or highest rated state for a story about a particular topic – child poverty, for example. [In some earlier years of the survey, rankings are provided for the counties and places for which numbers were produced, but keep in mind that not all counties and places were included.] This chart shows states ranked by the percent of young children whose parents are in the labor force. Maybe you are surprised that South Dakota is first. Utah, which you cannot see on this page, ranks last. Notice that the way I've chosen to display

these numbers shows that each one is an estimate. The real number could be lower or higher. The survey's figures are not as exact as a headcount. We'll come back to that idea later because it's an important detail to know about the American Community Survey.

Slide 14: Subject tables can be helpful for beat reporters: If you write about housing, education, immigration or other topics, the list of tables gives you a quick idea of what's available. Notice that there is a lot of detail on this slide about median income for different groups in Colorado. If you are just looking for a simple income figure for the whole state, I'd look at the data profile. This table has four columns of numbers going up and down. The first shows you the estimated number in a particular group (for example, families). The second is the margin of error for that number. This reflects the fact that these survey numbers are estimates. I'll explain the concept of "margin of error" in a moment. The next column shows you the median income figure for each group, and then there's the margin of error for that.

Slide 15: Here is a map that gives a state-by-state look at veterans. The dark-green states have the highest percentage of veterans among their adult civilian residents. The lighter-colored states have the lowest percentage. This could be helpful if you are doing a story – or graphic – about a military topic. Maps can give you a quick way to see patterns, sometimes in surprising ways. For example, the South is often seen as contributing heavily to military enlistment, but here you can see that several Southern states have low percentages of veterans. The map key, showing what the colors stand for, is off to the left.

Slide 16: The survey does not offer complete coverage yet. Rural areas, places under the minimum of 65,000 people, won't have data available for two to four years. If you try to get numbers, the map or chart will come up mostly blank. That's what will happen if you try to find out which counties in Wyoming have the highest percentage of U.S.-born residents who also were born in the state. You can use this statistic to look at economic vibrancy, or how attractive a place is to newcomers. As this map full of empty space shows, in Wyoming, most counties are not large enough to have survey data yet. There is a way to fill in the blanks, though.

Slide 17: The Census Bureau, along with state planners, has divided each state into areas of 100,000 people or more, called Public Use Microdata Areas, or PUMAs. When you display the numbers that way, as this slide shows, there are dramatic differences between western Wyoming, where resort areas and other attractions bring in many people from out of state, and the rest of the state. If you want to look at a rural or lightly populated area, this gives you a little help.

Slide 18: Here are a few hints if you need technical help as you go along. There is some census jargon that you will want to learn – for example, "tenure" at a college means you are a professor with job security, but "tenure" in census-land is a way of saying whether you own or rent your house. If you look up "room" to find out how the Census Bureau decides what is included in adding up the number of rooms in a house, you find that an enclosed porch that is used year-round is included. But bathrooms and unfinished basements are not. On the Brookings website, we will show you more about downloading numbers into a spreadsheet so you can do your own calculations or use the numbers for graphics without having to retype them.

Let's pause here for questions. The first part of the presentation told you the great things you can do with these numbers. The next part of the presentation will talk about what you cannot do.

Slide 19: Numbers are like any source of information. You need to know how far you can push the data in the direction you want it to go. You may have to "interview" them a little to find out what numbers are trying to tell you. Here is a list of some red flags. We'll go over each of them individually.

Slide 20: Lesson one is that you should not use the American Community Survey to find out whether a city's population is up or down, or how many folks in your community are Hispanics or old people or whatever other group you are interested in. The bureau's annual population estimates are the best source for this, and the official source. They are on the American FactFinder page. The survey was not designed to count people, but to describe them. The way demographers put it, the survey is designed to give us a sense of distributions by population characteristics.

Its estimates are fuzzy around the edges, as with any survey.
When you use an estimate from the survey, try to hedge it by saying "about" or "roughly" and round the number up or down. Otherwise, you fall into the trap of what's sometimes called "false precision" by trying to say that something is exactly a certain number, when it may not be.
So instead of "12.3 percent," for instance, say about "12 percent" or even "roughly one in eight people." Keep in mind that it is best to use percentages from the American Community Survey – for example, 20 percent of children in Community X live with single mothers. Sometimes the census website gives you the percentage and sometimes you have to calculate it yourself. You can use numbers to give a sense of magnitude, but if you do, round those up or down, too.

Slide 21: In census terms, the "universe" is the particular group that is included in a table or map. On this slide, we've put a red arrow that points to the universe of this map, which is "own children under 6 in families and subfamilies." So this is not all children under 6, but only those living with their own parents. Children being cared for by others, including grandparents or other relatives, are not included. Each table has a universe – for example, all households, adults age 25 and over, or homeowners. BE SURE you always check the universe so the table is for the population group you want. Some tables are for specific race groups or for people of Hispanic origin. Some are for specific age groups. Some are for families but others are for households that may include people unrelated to each other. The glossary will tell you about the difference between a household and a family. It makes a big difference in the data if the universe of the table is "poor children" or "poor families with children." And keep in mind that the American Community Survey universe is people in households through 2005. The 2005 tables say "total population," but it really is only the total "household population." That covers 97 percent of the national population, but that will vary from place to place locally.

Slide 22: In some areas, it's relatively easy to make comparisons between 2005 survey numbers and earlier years to see whether anything has changed. There are American Community Survey estimates for 2000 through 2004 for all states, for about 100

congressional districts and for many areas of 250,000 or more. You can make some comparisons between 2005 survey numbers and the 2000 Census, but you need to remember that the survey includes only people in households and the census covered everybody, including people in group quarters. There also are some differences in the way questions were asked, and the time period covered. There's some detailed information about this on the Brookings website.

Slide 23: Now here is where we try to give you a beginner's overview of margin of error. Some of you are familiar with this from political polls. If one candidate is leading another by three percentage points, but the margin of error is five points, then the margin of error is bigger than the estimated difference between the two candidates and you cannot say that one of them is ahead. It's the same with other kinds of survey numbers.
These numbers are from a sample and you need to pay attention to those margins of error. We don't want to see you make a mistake in saying, for example, that income has increased in your area, and then have someone else point out there has been no change because they looked at the sampling error. Compute the ranges of the estimates before you say one area is different from another, or that your area has changed from one year to the next. The apparent difference in the estimates may be due to sampling error, not real change.
I want to emphasize to you that the single-year estimates are based on a sample of 2.5 percent of household addresses. The survey will have three-year averages based on 7.5 percent of addresses, and the five-year averages are based on a 12.5 percent sample – that's 1 in 8 households. That means the margins of error are smaller for the multi-year averages than for the single-year estimates. We expect single-year estimates to bounce around more than estimates based on the larger, multi-year sample. Starting in 2008, you will have the more reliable 3-year estimates. It is the 5-year averages that are most like the larger sample of the Census 2000 long form (a 16 percent sample – about 1 in 6 households).

Slide 24: Here is a simple way to compute the range of an estimate. You start with the estimate, and then do one subtraction and one addition to figure out the range of the estimate. The number that you add to and subtract from the estimate to get its range is the margin of error. The Census Bureau gives you the margin of error for each estimate, and the math is pretty simple from there.
If you are comparing two different numbers, compute the range of the estimates to see whether they overlap. If they do overlap, you cannot say the numbers are different.
Sometimes you won't even have to compute the margin for the second number because you'll be able to see an overlap after you compute the margin for only one number.
Using the example on the slide, if the estimate is 12 and the margin of error is 2. So you add 12 plus 2 to get 14, the upper range of the estimate. Then subtract 2 from 12 to get 10, which is the lower range of the estimate. So your range is 10 to 14. [These ranges reflect a 90 percent confidence interval, which is a measure that most statisticians would be comfortable with. It means that nine chances out of 10, the true value falls within the range of the confidence interval. That reflects the fact that these are estimates, as is the confidence interval, so that is one more reason not to make a big deal out of small differences.]
You do not need to put these calculations into your story or graphic, but you do need to know what they are.

Slide 25: Here is an example of how numbers may look different, but when you take into account the margin of error, the apparent difference is not meaningful and you cannot say

one is higher or lower than the other. Here we have a table of median household income by state. If you add and subtract to get the ranges of the first three states on the list, you can see that their median household income may not be any different from each other. New Jersey cannot brag that it's richer than Maryland, and Maryland cannot brag that it's richer than Connecticut. But sometimes officials do. And if you know your margin of error, you can truth-squad that and say: They are all about the same.
It's pretty easy to calculate error margins on your own, but why not let the Census Bureau do it for you? The bureau does so on the state ranking tables and also on thematic maps. The trick is to click on the phrase "with statistical significance" to display whether two numbers are really different.

Slide 26: "Statistical significance" does not mean something is "important," in the way we think of something being significant. It means that two numbers probably are different. This slide shows what happens when you look at the median incomes of the first three states with statistical significance calculated for you. In this case, we've used Connecticut as the reference state – that is, we've asked whether the median incomes of other states are different from Connecticut's. There are single red dots next to New Jersey and Maryland on the median income chart, indicating that their numbers may not be any higher than Connecticut's.

Slide 27: But even if two numbers are different, please do not make a big deal out of small differences…It's just not worth it. And though journalists often like to make stories out of unusual or uncommon things, the survey's estimates for small groups may be too fuzzy to let you do that. For example, say that you want to look at the population in Little Rock, Arkansas, of people claiming Greek ancestry. The estimate is 266, but the margin of error is 271 either way. That means the range of the estimate is from zero to about 540. It's not that you cannot use this number, but you have to be careful. In fact, the Census Bureau won't make some estimates available because they are so small.
But there are alternative ways to get a good story.

Slide 28: If your first approach does not work, try another. Maybe you cannot say a place is No. 1, but it's among the top-ranked areas. Or you could try comparing different groups in one place. And, if good trend numbers are available, you could look at one place over time. If you do that, check whether the geographic boundaries of a place have changed. Especially in fast-growing areas, it's common for cities to annex nearby suburbs. The American Community Survey estimates are shown for the most recent geography.

Slide 29: Put this on a post-it note next to your desk to remind yourself to check all the bases when you do a story with the American Community Survey. This is the way to make sure you are on firm ground when you quote from the survey. Make sure of your definitions, and know which groups are included in the table you are looking at – the "universe," Be careful if you write about college students, prisoners or other groups where the lack of group quarters information makes a difference. Check whether the boundaries of an area have changed if you are comparing estimates over time. Compute the range of an estimate using the margin of error, so you can be sure how much you can say about it. And think big: Don't dwell on small differences.

Slide 30: Here is the data release schedule for next year and beyond, assuming that Congress and the administration continue to provide sufficient funding for the American Community Survey. You have single-year averages now for larger areas. Eventually you will have 3-year averages and 5-year averages for smaller areas. That's still a few years away, but if you are interested in smaller areas, be sure to learn about how to use these multi-year averages. They are more reliable and give you more data points to look at so you can see whether trends are going up or down. Some multi-year averages for previous years of the survey are up on the Census Bureau website.

Slide 31*:* Here are some links to more help, from the census press office and from your state data center. Each state has a data center, part of a partnership with the Census Bureau. Some data centers are located in planning departments, others in economic development offices or other government agencies. They can help walk you through numbers for your state or locality. Some of them publish reports or analysis based on the survey. They are used to answering questions from people who do not have a lot of experience with numbers. But Census Bureau officials and many state demographers will not give you sweeping policy conclusions. You will have to look elsewhere for commentary.

Slide 32: Here is information if you want to dig deeper by enrolling in an online course or reading a book on this subject by my collaborator, Cynthia Taeuber.

Slide 33: I hope we've left you with the idea that though the American Community Survey has some limits, it can be a rich and interesting source of information to help you do your job. The slides and transcript from this presentation, as well as a reference guide with more detail, will be available on the Brookings website at the address on this slide. [We've also provided contact information for Michelle Daniels of the Metropolitan Policy Center at Brookings.]