John Blodgett, programmer analyst at the University of Missouri's Office of Social and Economic Data Analysis

Interviewed by Batul Hassan

BH: I was hoping we could start by talking through some of the most valuable information journalists can get from the census… Start with your background?

JB: My background is making census type data accessible. How to get it, how to understand it -- basically helping people access the new census data. So I can definitely help you in terms of where to go in terms of the broad spectrum of census data. If you really want to look at income inequality in Columbia versus Kansas City or Minneapolis, you can do that but I think the general wisdom on this is that you're not going to find very much variation. It tends to be more of a … These tend to be kind of global things. And most of the work you see done at this level you see at the national level, frankly. I was just looking at something and they do have statistics available on ol' Wikipedia, if you just do a basic search on the GINI coefficient, which is probably the most widely touted or referenced if you want to pick one number that statisticians are going to say, 'here is the one number we've invented that pretty much is an indicator of the equality of distribution of wealth that we have out there.' Are you familiar with GINI?

It can range from one, which would be the worst -- that would be if the people who had zero was everyone and all of the money went to Bill Gates and we just had to wait for charity. That would be a GINI of one. And then a value of zero would be the opposite -- that would be completely equal distribution. So some people would say, the lower the GINI coefficient, the better off you are because the more equally the wealth has been distributed.

So, they just started putting this in the census -- in the American Community Survey (ACS). Prior to that, there was no such measure in the previous censuses. It sort of became a hot topic, and now they picked up on it and started putting tables up. They did it rather discreetly without any fanfare. It was out there for a couple years before I even realized it was there. And so I saw reference to it but I didn't know it was in the census. They have a thousand tables on the ACS -- you can easily miss something if nobody points it out to you or send a head's up.

So, I could show you how to access that data, but it's just like any other census data. There are four tables in the ACS where they have -- they're so called summary tables, detail tables or base tables -- that go by various names, and I think four of them are really relevant to this topic. You might stretch it and say the table with median household income could also be involved in this but you get to those either through American Fact Finder, which is what most people go through these days, I think most journalists are already familiar with that if they're doing this kind of work. They probably know its quirks even a little better than I do. And we have the data here at the University, so we have our own accesses. But I don't know if I would recommend that -- I don't know if that's what you want.

Now if you were really set on saying, "I want to look for differences in GINI coefficients. Is there a significant difference?" And this what I'm looking at is a set of GINI coefficients by state, which is interesting. So everywhere from the Western states being the most equal, which is sort of surprising -- Utah, Wyoming -- and then the upper mid-West with Iowa, Wisconsin and Nebraska -- and then down at the very bottom, we get the most wealthy places -- Massachusetts, Connecticut and New York -- probably three of the wealthiest states on the other end. And that's because they have the most millionaires. And that's what really makes your GINI coefficient go up.

BH: Recommendations for resources to cover how nationwide income inequality plays out in our own community? Example: Geographical divides within a city's neighborhoods.

JB: Well, that's kind of a different issue. Okay. What we've been talking about here is income inequality without regard to where you live. But just for the whole state of Colorado, here is the … And that's just by looking at aggregate data for a very large area and saying, "Is there an inordinate distribution where a few people have a lot of the money?"

But in terms of -- Okay, what you're talking about -- looking at poor versus well-off neighborhoods …

Or for policy reporters, having the resources to access the data for themselves instead of, for example, relying on a politician's definition of the middle class for the state -- how that might vary from policy to policy… Basically fact checking information…

Okay. That's a lot different. That's more of a -- yeah. Because the other thing is kind of limited. There a few things here, you can look at that, but most of the people accessing this is economists. And they have all these measures … I'm not saying it's junk science, it's just kind of technical and usually most people just want to go to the bottom line and say, "Well, what's our GINI coefficient?" Which is, you know, it is a measure. It's not everything, and they have other things.

But now, we've just switched and you're saying, where can we go to get general data about people's poverty status and income and all that sort of stuff. Well, that is where you really need to be able to use the census. Especially if you're trying to look at neighborhood-type stuff. That's not readily available to most people, they don't … But you can do it by zip code, you can do it by county, you can go down the census tracts and even smaller --

BH: Could you explain what a census tract is?

JB: Those are the -- That's a census geography that is defined by the Census Bureau, it's open to change every 10 years at the decennial census, and they're basically -- ah, they're most useful in urban areas, although they have them in rural areas as well, and they tend to be neighborhoods of four to six thousand people is the usual population. There are committees that draw the boundaries locally, and their instructions are to try to draw them so that it's a somewhat homogeneous area. Better yet, if it has some local designation. So Soulard in St. Louis would be an example of a good census tract. So you have these units

called census tracts and there are maybe four or five thousand that cover the country. Basically, people doing analysis in cities tend to use census tracts in their analysis.

BH: So you can search an area by its census tract?

JB: Yes. And the Bureau provides maps and mapping files. A lot of people working with this have desktop GIS systems so they can display the data. *City planners are really big on this stuff*. Now, one of the things that has happened however, is most of the census tract data comes from the decennial census, traditionally. For the last 50 years. 2010 census, you can still get census tract data but the sort of stuff you used to get --- the census for many decades had two sub-genres: short form and long form. So one out of six households got a long-form, and that's where they ask people the questions they don't like to answer sometimes, about income, education, occupation, even things like how you get to work and all those things. So that's referred to as long-form data, and it includes income. And of course from that comes poverty status. In 2010, they no longer had long-form. Everybody got short-form. So the 2010 census is still good for basic population count like age, race, sex and Hispanic-origin. Total count of households and some household-type things. But other than that, no income, no poverty designations, no education. No socioeconomic indicators. That has all been moved to be reported and collected as part of the American Community Survey. Okay. The nice thing about the ACS is that it gets refreshed every year -- we get a new set of data every year. It comes in three pieces: right now, we're in the 2013 vintage ACS data. They've released the data based on a single year, the 2013 data, back in September, then in October they released the three-year data which is ACS data based on the most recent three years (2011-13). So they publish estimates. It's a one percent sample per year, so it takes three years worth of data before they get enough statistical significance for them to put out data for geographic areas as small as 20,000 people. So the data we have out there -- vintage 2013 -- now is only for areas of 20,000 and above. So you can get Columbia, Boone County, but you can't get Fulton. You have to wait, and then in December they're releasing the five year estimates, which will be based on 2009 through 2013. That's the set that ends up getting the most use because it's got all the geographies. There's no population threshold. So even a city of 200 people -- they put out data. It's statistically garbage, but it's out there. Here, on the MO Census Data Center site, which is what I'm involved with, mcdc.missouri.edu. It's the only state data site pointed to by the Census Bureau.

BH: Nice.

JB: I was going to show you the ACS profiles -- there are actually two websites now. They just haven't really advertised it yet. So if I go to the profile, we get some background information… But basically you select a period, you see we have one year, three and five year periods. So if you need the latest data you can select the 2013. But you notice on five year there are no 2013 to 2009 yet. So the latest five years is 2008-2012, which is what most people use, even though it's fuzzy. I mean, if you were trying to look at what's going on with income inequality -- have we come out of the recession. If we want to compare now to 2008 -- you could do that with the one-year data, but then you have to keep in mind that

it's only published for areas of 65,000 and more. So, ah… Only the big picture stuff. You can't use that to analyze trends in neighborhoods.

For example, the tract which I mentioned were so important to planners and people who actually work in the nitty-gritty of a city, you don't get any tract data except at the five year level. Even if a tract should slip over the 20,000 threshold -- and there are handful, they're kind of mistakes, but they do. They're not supposed to be that large, but even if they were, they still wouldn't publish any data about them. So we have this data. If you wanted to see what's happening in MO, you select 2013, then select state. And within this thing there are four categories, which are all checked initially. But if you didn't want to waste space on demographic or housing and you just wanted to see economic and social, you can deselect whichever you want.

And there's your report. Household income starts right off. So then the question becomes, out of all these different measures of income and such, which is the best? Okay. I have something I can point you to there.

BH: A lot of people have been saying the Federal poverty line isn't the best, but the supplemental poverty measure is better.

JB: Yeah, that's another thing. Of course, there's no data with the supplemental other than at the very high, national level. The Census Bureau, which is rather adventurous on their part, actually goes out of their way to publish a report where they come out and say, 'By the way, here's the poverty we think is more relevant,' and then they go ahead and publish their alternate figures. But they aren't official, and no federal funds are distributed based on that data. And we can't get that data for anything but basically the nation and regions and possibly states.

BH: Could you figure it out from this information?

JB: No, it's not here. That data isn't as widely distributed the way this is. I mean, the data that goes into these reports is significant, it's huge, it's gigabytes of data you have to wade through it already. So, I don't know, I've never really looked into the alternate things other than to read the reports. I understand what they're doing, and it's too bad we can't actually make those official. Basically the people who understand all of it sort of agree that our poverty definition was set in 1964, and there are so many new things out there that obviously aren't included in the definition because they didn't exist. Food stamps didn't exist in 1964. So various programs and things that affect peoples' actual well-being. Also, another one of the things that is obvious and should be looked at: poverty is pegged strictly to household income, and we all know that a person living in rural Missouri with an income of $30,000 is a lot different than a person living on $30,000 in Manhattan because the rent and cost of living is so much higher. And this is apparently what they do in the alternates -- they do make adjustments for those kinds of discrepancies and costs of livings. That's a group within the Census Bureau.

On the other hand, even though that poverty definition should be improved, it could be, and there are some problems with it, it's still… If you go out and say, here's a map of the Detroit metropolitan area and put in the census tract and say here is a map of the poverty levels, it will show you what the poor people live. It's just degree of poverty.

BH: So of the data we do have, what is the most useful?

JB: That's what I'll show you now. [goes to homepage… goes to 'all about'... there are 'measures of income in the census] That's what I'd suggest people read. So it goes through measures of income of poverty. It's a few years old but hasn't changed.

BH: Are there state poverty lines?

JB: No, just Federal. And they put out new tables every year. It's based on household size, a number of children under a certain age, 18, and some things about 65 and over. There's fine print, but basically it's family of four with five kids and then here's the number for this year: 32,000 or something. But it's all taken care of. All of that goes out there, and when the Census Bureau is processing their surveys, they look at the household income and the structure of the household and they assign a number to that household and all the people within it. And that's your poverty ratio. It says, okay, we've looked at your household and seen four people: two adults and two kids, we know what year it is, we know the threshold, and we take your household income over that threshold and then that's your poverty number. So, that's very good. If you're looking at the PUMS data -- the Public Use Microsample Data -- that's the data that the bureau publishes where they strip off geographies, more or less, so they don't let you know -- for privacy issues -- they're not going to give you any identifiers. There is a geography code called a PUMA, a Public Use Microsample Area, but it's at least 100,000 people so it's pretty large. But there are two levels: household level and person level, and its micro data, meaning it's not a description of a country or a census tract or a zip code but it's this person, or this household. And it actually has data about how they responded to that. Age, gender, various measures of income, race, country of origin, are you Hispanic, basically you've got the responses to the questionnaire. A few things have been derived. So they don't ask you, "What is your poverty level? Or "are you poor?" but they're able to calculate it because they have your information. So there is a single variable called poverty. And it's one of the more useful numbers in there. Because you can then go out and use that number to classify people. And the beauty of that number -- we're talking about all these numbers and why are they good and what are their limitations -- and when looking at aggregate data, the number that most people tend to settle on, and I tend to agree with them, if you had to pick one, is median household income. So, you look at a household, get their income, and from that you can tell if they're well off or not. But the trouble with that is you use the same measure for a student living in an apartment by themselves versus a family of five. So if they both have an income of $20,000, they're still not the same thing -- at all. You can live comfortably as one person on $20,000, but a family of five is in big trouble. At any rate, that's the trouble with household income: not all households are the same. But you've also got family income, which is a subset of household income, but it only includes households that are classified as families. And some people think that's better. And it's okay, except you sometimes have

neighborhoods that have very few families. So you go into a neighborhood right here off of College Ave, you might find that 80 percent of the households are not families. So the data isn't very good. And it may even be hiding things. I mean, the families may be doing well and the non-families not so good. So that's one of the problems with using the families in the household income. This is where the poverty ratio is so good, because the poverty ratio takes that all into effect. When you assign your poverty ratio, you know that the person living alone has a poverty threshold of 18,000. So if their income is 20,000, their poverty ratio is 20/18. It's not poor, I mean it's not well, but it's at least above the poverty line. But if you have a family of five with 20,000, they're definitely going to be below the poverty line. And not only does it tell you that they're below, you actually have a number with which you can see how far below they fall or how well off they are. If your number is at 100, you're at the poverty level. If you're 200, you're twice the poverty level. And that's generally regarded as a key thing, if you're above twice the poverty level, then economically you're not struggling. Obviously that's an arbitrary number, but people between 100 and 200 are sometimes called the working poor. They typically find a lot of two people with two low-income or three low-income jobs. They're getting by, but -- no savings or retirement. That's another thing. If you're an elderly couple and you own your house -- you don't have to buy a lot of things -- you don't travel much so you don't have gas -- it's possible that the income can be misleading. It's also doesn't measure how much wealth you have. You might have assets that if you had to use them, you could. But except for income, that's not reflected. The Census does not ask about wealth.

BH: So aren't they missing a lot, especially in upper classes?

JB: Well, they're missing a lot of wealth. They're missing -- but most people with a lot of wealth also have a lot of income. Probably evens out pretty well. But a lot of people have low wealth -- if you don't own a home, a car, don't have investments, or much savings. There's plenty of people with practically zero wealth. But in terms of common surveys, it's measured in terms of income.

BH: So what are things you see miscommunicated about the census?

JB: I'm sure I see things, but not any one specific thing.

BH: Maybe concepts?

JB: Well, one of the things, it's not a big deal, except for when you live in a college town it kind of comes up. One of the ways they define poverty -- well, they include students if they're not living in dorms. They're considered part of the poverty universe. If you live in a group quarters: so dormitories, state mental health facilities, prison, military barracks, you're excluded from poverty statistics. But everyone else, 90-odd percent of the population is part of the poverty universe -- they're assigned a poverty level. But a student in the dorm is excluded. Once they move out to their own apartment, they're included. But the income used to measure whether or not you're poor is at the household level. So it doesn't take into account that you have a family back in St. Louis or Kansas City sending you, you know, $5,000 a semester to live off of. So, students are poor, and Columbia has a

high poverty rate. And if you list the medium-sized cities with the highest poverty rates, they are Columbia, Lawrence, Kansas, Bloomington, In., Ann Arbor, Mi.,… state college towns. It's hard to adjust because there is no table that says "poverty by student status." If I were running the show, I would just say, if you are enrolled in college, exclude you from the poverty. Because we don't know. A lot of people really are poor and basically survive on ramen noodles and have a job flipping burgers somewhere and just get by. It depends on your family support. Usually working, though, isn't your primary source of income. Especially in a college town like this one, where most people are from out of town. It's not the same if you're in a community college and living at home, because then you're included in that same family household.