City & Community

Relying on the Census in Urban Social Science

John R. Logan
Brown University

Census data have long been a key tool for urban research, and the approaching 2020 Census offers a natural moment to reflect on how we use it. The highly partisan plan to include a citizenship question has recently captured our attention. I suggest that its short-term effects may be modest since immigrant communities already are suspicious of government surveillance and many will prefer to stay hidden regardless of the census questionnaire. I raise several other kinds of questions about the reliance of urban researchers on census data. These include concerns about how we treat census tracts as neighborhoods, how we accept census statistics at face value, and how readily available and increasingly useful quantitative data sources may be crowding out ethnographic research. I also comment on new approaches such as spatial analysis and Geographic Information Systems (GIS), and opportunities for linking individual and place-level data with one another and following both over time.

As I write, the Census Bureau is completing the 2018 End-to-End Census Test in Providence, a test of the readiness of the Bureau to carry out the 2020 census, and I have submitted my own questionnaire online. At the same time, the Secretary of Commerce has decided to add a question about citizenship that has not been pretested, and many doubts have been raised about this decision on technical and political grounds. Aside from concern about its political implications, urban scholars have a stake in this decision because we rely so much on census data to carry out our research. Discussion of this issue reminds me of other concerns about our use of the upcoming census. There are pitfalls here that need to be addressed: reifying the census tract as a proxy for neighborhoods, being misled by sampling variation in small area data from the American Community Survey (ACS), and striking the wrong balance between quantitative (census-based) and ethnographic studies. There are also opportunities worth thinking about, including the growing use of mapping and spatial methods and new possibilities to study individual and place-level data simultaneously.

THE CITIZENSHIP QUESTION

Let us consider the question on citizenship first. There are two sorts of objections: (1) that it is likely to depress response rates by predominantly immigrant groups, and (2) that it is designed to depress turnout and (worse) to open the way to apportioning political representation based on counts of citizens rather than population.

Correspondence should be addressed to John R. Logan, Department of Socioloy, Brown University, Providence, RI 02912; john_logan@brown.edu.

City & Community 17:3 September 2018
doi: 10.1111/cico.12331

The first objection is realistic, but in my view, it is overstated because participation in the census is already in trouble. The Census Bureau has a long-term concern that it is becoming harder to count people despite increasing efforts to be able to find them. Conservatives complain of government overreach; progressives worry increasingly about privacy; and we are all so deluged with junk mail, email scams, and robot phone calls that census takers face an uphill battle to get our attention. But more than this and especially among immigrants, there are well-founded fears of being identified and targeted. In an era with a high rate of deportations and targeting of firms with undocumented workers (beginning well before the Trump Presidency), and now with a President who has such a strong public focus on the exclusion of Muslims and militarizing the border with Mexico, it must be harder to get an accurate count of immigrants. Ethnographic studies show that people are hiding. In this context, answering a question about citizenship on the census form is probably a minor factor. Indeed, the visibility of the issue in this form may, in the end, reinforce advocacy groups' efforts to push up the numbers.

The second objection is not being voiced so publicly, but it concerns me more. Currently, although non-citizens do not vote, the weight of voters in districts where they live is greater because of their presence. A policy not to count non-citizens for electoral purposes would be even more consequential than more familiar methods of voter suppression. There are constitutional obstacles to counting only citizens in the drawing of Congressional Districts (see the Supreme Court ruling in Evenwel v. Abbott in 2016, which applies to federal but not state/local policy). But in practical terms, there is another obstacle. To implement such a policy, there would need to be a census count that can be portrayed as a 100 percent enumeration, and it would have to provide data at the block level. The ACS, which already has a citizenship question, is insufficient for this purpose. This may be the unstated purpose of counting citizens in the census.

In defending the census, we must be aware of its blemishes. The census is imperfect in the way it categorizes people, and it obeys no moral ethic. The selection of information to collect and how to categorize it has always been subject to political influence (Anderson 2015). Before 1870, because of the political compromise that sealed agreement on the U.S. Constitution, the census needed to count slaves (each slave counted as 3/5 of a person for apportionment) but not to learn much about them. Between 1880 and 1920 the census distinguished between Negros and mulattos; subsequently this distinction was lost, and in recent years it has been considered progressive to subsume all black people under a single category. (Note, though, that a new wave of survey research is asking about skin color.) The census was careful to measure native versus foreign parentage through the first half of the 20th century by asking where each person's parent was born. That was in an era when people were supposed to become "American" in the second or third generation. That information is no longer even in the "long form" ACS, and we cannot compare experiences of groups across generations from these sources. In 1980 the census created a new question about Hispanic/Latino origin, the result of a long campaign by advocacy groups (see Mora 2014). Making it separate from the race question—regardless of whether Hispanics view their identities in terms of black and white—has encouraged urban scholars to disregard racial differences among Hispanics. Advocacy to make "Hispanic/Latino" a response category in 2020 on the race question instead (and to create a similar "Middle Eastern" race category) failed. But in 2020 for the first time, whites will be asked about their ethnicity in the race question, which is completed for the full population and not only a sample (like the ancestry question in the recent past). And for the

first time, blacks will be asked to identify as African American, Jamaican, Haitian, Nigerian, etc., so that there will be a 100 percent count of national origin categories among blacks. It will be a challenge for urbanists to be prepared to examine these distinctions.

This history is relevant because in every decade the set of scientific questions that social scientists focused on was affected by the census's list of questions. For example, in 2000 when people were allowed to list two or more race categories, we became interested in where mixed-race persons and married couples lived. I wonder what will be the political and ideological ramifications if census-based studies begin highlighting "ethnic" divisions within categories of black and white. And if a citizenship question is included in the census, how will we take advantage of it?

CENSUS TRACTS AS NEIGHBORHOODS

One of the major uses of census data by urbanists is to treat the census tract (the main geographic unit for which data are published) as a "neighborhood." Often this decision is described as a forced choice—the tract (or for a small number of variables the block) is what we are given. Yet in many metropolitan areas the tract is too small to represent a neighborhood, which for some purposes (e.g., if one is studying an immigrant neighborhood in Los Angeles or gentrifying neighborhood in Brooklyn) very likely extends to a set of adjacent tracts, not a single one. But how can tracts be combined into meaningful larger units that correspond to neighborhoods as people think of them when they decide where to live or describe their local area to others? What is a neighborhood if it is not a census tract?

One well-known study, the Project on Human Development in Chicago Neighborhoods (PHDCN), tackled this problem by starting with the assumption that a neighborhood is an area that is relatively homogeneous with respect to some community characteristics, and distinct from others. This is the same theoretical approach taken by the early Chicago School in mapping "natural areas," but now it can be highly automated. The PHDCN researchers factor analyzed a number of indicators to establish key social dimensions, and they combined nearby tracts with similar characteristics on these factors into neighborhoods (Sampson and Raudenbush 1999). Working with colleagues, I have approached the same issue using GIS methods. In one study (Logan et al. 2002), we used a relatively simple spatial clustering measure (Moran's I) to identify areas that were significantly more similar to their neighbors on a specific measure (in this case, to identify ethnic neighborhoods, such as Chinese or Dominican neighborhoods in New York). The spatial statistic provides reassurance that the pattern of spatial clustering that we observed in the GIS map was not random.

Having to identify neighborhoods rather than take them as given can result in more effective analysis. For example, it can lead to better decisions about how to operationalize the scale of neighborhood effects. Consider a study that examines whether access to ethnic institutions affects co-ethnic solidarity. Imagine measuring access to ethnic churches, markets, restaurants, etc., within the census tract where one lives or within a range of tracts that lie within a 20-minute walk. Which of these is the "neighborhood" for the purpose of this study? It is probably not the tract, but it might be a cluster of tracts, and for some purposes, it might be the whole city.

Consider another study: the effect of the number of other persons that one knows in the neighborhood on sense of security. Imagine measuring number of neighbors known

within the census tract where one lives, or alternatively within a range of tracts that lie within a 20-minute walk. Which is the neighborhood? According to some urban research (Suttles 1970) it is neither—the neighborhood is really a small area along the street where people live and have repeated face-to-face interactions. In that case the tract may not only be a weak proxy for neighborhood, but it may be misleading.

RELIABILITY OF CENSUS DATA

Urban researchers traditionally have understood census data to be infallible because these are the "official" data. They come from a census that counts the whole population, and at the scale of a census tract it may be the only source of information on many population characteristics that are important to our study. In the period 1950 to 2000, most of us were aware that much of the data came from a sample of people, what became known as the "long form" census that included one in six households. These data (income, education, ancestry, country of birth, and much more) are central to much research, but in spite of coming from a sample that was known to be distorted by sampling variation in the estimates, we found it convenient to treat it as truth.

The situation is worse now because although data on a handful of variables are still gathered through the full-count decennial census (including race and Hispanic origin), most are gathered through a separate process in the annual ACS. And the ACS samples for census tract estimates, despite being based on samples that are pooled across five successive years, are only about half as large as the previous census long-form. Now we face a new problem: not only is the census tract possibly not a neighborhood, but we realize that we cannot trust the published estimate. A tract that is estimated to have a median income of $80,000 may now be reported to have a median of $80,000 plus or minus $25,000. That is a wide range, evidently it can result in erroneous conclusions about any single tract, and despite the fact that there may be no bias in these estimates (so in the long run, on average, they are on the mark) they may mislead.

We have found this to be an issue in the analysis of changes in income segregation across census tracts between 2000 and 2010 (Logan et al. 2018). Although the observed values of some measures of income segregation rose strongly in this period, about half of the increase appears to be attributable to the fact that income data for individual tracts were based on smaller samples in the ACS than in the 2000 decennial census. The point is that census data, like all data, are fallible and we all now have to deal with sampling as a practical problem rather than leaving it to the sampling experts.

Fortunately, for the purpose of estimating values for individual tracts, advances in statistics offer useful alternatives. This research area is called small area estimation (SAE), and Bayesian models of SAE have been developed to pull information from other tracts to reduce the volatility of estimates. Bayesian SAE has been well-known to specialists for decades, but it is a new challenge for urbanists to become familiar with them.

THE STATUS OF URBAN ETHNOGRAPHY

The more I understand the limitations of census data, the more I yearn for other sources of information. One direction that is natural for quantitative researchers is to find non-census data on characteristics like social class (e.g., from home prices and rents), race and

ethnicity of children (e.g., from birth records and school enrollments), as well as other kinds of information that we never could find in the census (like public health, crime, and school achievement). My inclination is in another direction, urban ethnography.

On most questions that we study quantitatively, there is a parallel literature based on extended fieldwork, observation, and interviews. For example, students of Burgess, McKenzie, and Park wrote dissertations about specific kinds of neighborhoods (such as *The Gold Coast and the Slum* by Zorbaugh, 1929), and later generations in this Chicago tradition continued to probe the nature, boundaries, and impacts of neighborhoods (Small 2004; Suttles 1970). The best theoretical guidance for census-based studies is found in this work. At its best, urban social science deploys a mix of methods, each learning from the other. In practice, this is more aspiration than reality, though there are exceptional cases (Desmond 2016; Desmond and Gershenson 2016). I was struck by reading the literature review of research on gentrification by Brown-Saracino (2017), which emphasized the discontinuities between interpretations of neighborhood change by quantitative studies and ethnographers. The latter, for example, emphasizes processes of privilege and displacement, while the former mostly concludes that gentrification is mainly not based on forcing poor people out. To form an opinion requires being familiar with each.

I am aware that scholars who share a large subject area may actually be interested in different dimensions or facets of the phenomenon, so this is not a case of one approach proving the other wrong. We do a disservice to our readers and students, and to ourselves, when we are unaware of or fail to think about or push aside work using different methods. Long ago I stopped being surprised that urban economists rarely cite other urban scholars (and vice versa). "Economics is so self-absorbed!" My concern as an urbanist who mostly relies on analyses of census or survey data is that I see so few references to qualitative works in the core studies that are most similar to mine—and perhaps this behavior is reciprocated by many urban ethnographers.

What makes this a larger problem is that there have been such tremendous advances in the last two decades on access to census data, inclusion of geographic identifiers in large-scale longitudinal surveys, and financial support for these efforts. I suspect that urban ethnography is being crowded out rather than emerging—as it should—as a more crucial support for our ability to frame questions and interpret findings.

SPATIAL ANALYSIS

Mapping has played a large role in urban research from the beginning, as illustrated by DuBois's (1899) careful map of Philadelphia's Seventh Ward in the 1890s. As GIS technology and understanding of how to use it have become readily accessible, we are relying more on maps both to display information in a visual format and to analyze spatial data. I once did a study of how the distance of Long Island suburbs from New York City affected their composition and growth, for which a graduate student and I used a road map and a ruler to calculate distances. Now distance is measured instantly, and one can choose whether to use a straight-line distance or distance through the road network, or even expected duration of travel between two points at a given time of day and using a given mode of transportation. We can rely on spatial clustering algorithms to measure the overall extent of clustering, and also to identify the boundaries of specific areas where adjacent areas are similar enough to deem the cluster to be statistically significant.

GIS is closely tied to census data for two reasons. First, accurate GIS maps of census tracts are readily available as far back as 1940 for many cities. Second, the census provides data for every tract, a comprehensive coverage that most other sources offer only at the county level, or perhaps only at the level of states or regional groupings of states. Even when other data are also available for tracts, as in the case of some vital statistics or crime data, their analysis typically depends on incorporating other information from the census.

The relative ease of using GIS techniques (along with the fact that maps can make a powerful impression on readers) is leading to a rapid expansion of the number of urban scholars who take advantage of them. I have also noticed a tendency (Logan 2012) to simplify and standardize the approaches to spatial analysis. (1) Often a key outcome or predictor is proximity or exposure to something, measured by a combination of its level and intensity and its distance away. (2) Spatial autocorrelation (which is the degree to which areas near one another are similar on some measure) is widely used to characterize a spatial pattern (describing clustering as intense or light) or to simplify data (e.g., creating a dichotomy between tracts that are inside or outside of clusters). (3) True to the roots of spatial analysis in multivariate models, spatial autocorrelation is also often invoked as a concern (a problem of the correlation among error terms that can violate assumptions of the statistical model), and spatial regression routines are deployed to account for (or control for) it.

These are reasonable steps. My only reservation is that the easier it is to carry out the computation and the clearer is the template for the usual steps of analysis, the more scholars need to think about what they are doing. How exactly does "distance" (so readily measured) translate into exposure, familiarity, or interaction? How do we know when strong spatial boundaries are widely understood and enforced even among people who live near to one another? When do political boundaries (the boundaries of aldermanic wards or school districts or police jurisdictions) count more than distance? Once asked, these are fundamental questions that can be examined with spatial methods, although often they require much more real-world knowledge than is available in a spatial data set.

INDIVIDUALS AND PLACES

We have come a long way from the "ecological fallacy." It would be antiquated today to suggest that ecological correlations could be directly interpreted as reflections of correlations at the level of individual residents. Familiarity with multilevel models has taught us to distinguish more clearly between associations that are best measured at the level of individuals and involve processes at that scale, versus phenomena that occur at a larger ecological scale (e.g., the work place, the neighborhood, the ethnic community). Many urban questions are framed now in terms of how characteristics of larger units influence individual outcomes, controlling for processes at the individual level. Neighborhood effects, once considered cautiously (Mayer and Jencks 1989), are now a central urban question (Sampson 2012). The neighborhood is newly relevant (or relevant in new ways) to research on stratification, child development, public health, criminology, structural racism, immigrant assimilation, and education.

Again, census data are deeply implicated because in most of these studies—even when outcomes and important inputs are measured at the individual level in surveys or other

sources—we rely on key place characteristics drawn from the census and we typically use census administrative areas (like tracts) to establish the boundaries of places.

The demands for information are growing. We are becoming more aware that the effects of the neighborhood one lives in have to be estimated independent of the characteristics that resulted in living there in the first place (residential selection). We know that places change over time, partly as a result of who is moving into and out of them. Residential mobility is common enough that we need to follow people over time, rather than assume that the "neighborhood effect" is based on the neighborhood where they live today or when they were a toddler. To some extent, this can be accomplished with long-term longitudinal surveys like the Panel Study of Income Dynamics. The major limitation of these surveys is that they are best used to study individual-level outcomes (they are, after all, based on samples of persons, not of neighborhoods, and they provide information about where people live but not how their neighborhoods are evolving). The great advantage of the census is that it provides information about all localities, and it links large samples of persons to places every decade.

We now have two routes through which to exploit these advantages. Each is promising, but there are limitations to both. One is to work directly with the original but still confidential data gathered by the Census Bureau through a Federal Statistical Research Data Center. The other is to analyze the 100 percent data from earlier censuses (1880 or even earlier, through 1940) that have been transcribed by volunteers and by commercial firms and made available as harmonized data files by the Minnesota Population Center (MPC). In both cases researchers have direct access to millions of individual records, with geographic identifiers that potentially allow them to know exactly where people live and therefore to create local contexts based on any definition of neighborhood. Further, it is becoming more feasible to link data on people over time, so that as neighborhoods are changing one can also follow people's movements among them. Having struggled with the published tract-level tabulations over several decades, and having looked for creative ways to take advantage of 5 percent samples of microdata where the smallest level of location is a district (sometimes a county or more) with over 50,000 residents, I am pursuing both directions with relish.

Consider first the Research Data Center (RDC) route. With sufficient permissions, a researcher can use records from any recent decennial census and long-form or ACS sample, learn the longitude/latitude of each residence, link records among these sources over time, and link to other administrative sources that are cooperating with the Census Bureau. There are two necessary caveats: (1) It can require a high level of familiarity with census data collection, sampling, and processing approaches to use these data appropriately, and (2) Studies conducted in the RDC must have benefit, broadly defined, to the Census Bureau, proposals have to be vetted, and users must obtain sworn status as a volunteer census staff person.

Other conditions are reasonable but impose limits that deserve more consideration because they impose barriers selectively for different kinds of scholars. Many universities now have an RDC, but most do not. Although the Census Bureau is thinking about how to arrange more distributed access, currently the most favored researchers are those who have a local branch available at no cost, and least favored are those who have to travel to an RDC and pay a user fee. Another barrier is technical knowledge. Many scholars depend on a consultant or graduate student to assist in the management of large data files and analyses, but only persons with sworn status can fulfill this role in an RDC, and a

typical RDC does not have staff available for this purpose. Another impediment to collective use of this resource is that although many files are maintained in a central location, most working files and files created for specific projects are stored only in that specific project folder, unavailable to anyone else. They cannot be exported. This means that reports of findings from an RDC cannot readily be replicated by others even if the author wishes to share the working files and computer code. The high utility of working in an RDC is reduced by the ways in which effective access is channeled.

The RDC can be avoided entirely by using historical census data that are no longer confidential. Through 1940 there is no problem with sampling (although in 1940 for the first time a handful of questions, like parents' birthplace, were asked only for a 5 percent sample of persons). The currently available data are for 100 percent of the population, the files include names and addresses, and it is possible both to map and geocode address locations and to link people's records over time. None of these steps is without difficulties. I began working with 1880 census data several years ago and only recently was I able to make public a geocoded data file for most of the residents of 39 major cities (Logan et al. 2011; see also the Urban Transition Historical Geographic Information Systems (HGIS) website, https://s4.ad.brown.edu/Projects/UTP2/39cities.htm). All the information is subject to transcription errors, omissions, and other mistakes. Yet it is surprisingly complete, and MPC has already disseminated linked data between 1880 and samples from several other years (Ruggles 2002). I find that many of the substantive questions that I previously approached with contemporary data are also relevant in an earlier time, and we gain a better understanding of them when we can see how patterns and processes have changed or persisted over the decades. I will cite as examples studies of racial residential segregation in Southern cities and in Philadelphia in 1880 (Logan and Bellman 2016; Logan and Martinez 2018). This is admittedly a very specialized topic, perhaps one that urban historians or historical geographers would more likely find interesting. It has contemporary relevance because urban scholars continue to be interested in black–white segregation, and our interpretation of its sources hinges partly on what we believe about when it emerged (e.g., is it a product of the era of racial zoning and redlining, or the Great Migration, in the 1920s and 1930s). Further, working with the 1880 geocoded data allowed us to experiment with different spatial scales (leading to the insight that African Americans were highly segregated in the South and in Philadelphia's back alleys and side streets, but still in close proximity to whites in the same part of the city) and with alternative spatial indices of segregation (including distance-based exposure and "next door neighbor" measures).

LOOKING FORWARD

The inclusion of a citizenship question on the 2020 census has drawn our attention to the census. It reminds us of the inherent politicization of the census, not for the first time. It raises practical concerns to the extent that it will likely affect reporting rates and requires imputation of missing data at the local scale. At the same time it creates another data point for research into differences between immigrants and the second and later generations that was lost when "ancestry" replaced "parents' country of birth" in 1980. How can we best use this information, if we have it, and are there concerns about its use? Similarly how will we use information about categories among whites and blacks? More

generally, how do choices made by the Census Bureau affect the questions that scholars pay attention to?

We tend to reify the census tract as a proxy for neighborhoods because that is the main scale at which the census provides areal data. That is not the only choice though it is the most convenient one. As we become more aware of sampling variation in the ACS, we are pushed to reconsider. Is the tract too small (so it should be replaced by clusters of tracts) or is it too large (so we should turn to geocoded data to identify spatial patterns more finely, requiring innovative research in RDCs or a turn toward historical census data)? The main problem is that we accept census data too readily at face value. The answer is not to avoid these data but to be more aware of the substantive and measurement issues that they pose.

I also raise the concern that public census data have become so readily accessible that they may be crowding out alternative evidence. To use them wisely requires a level of understanding of the urban process (e.g., what is a neighborhood and how would we know if it has boundaries?) that can be informed by fieldwork. Intensive observation and interviewing can help us understand what phenomena we need to measure and provide a yardstick with which to assess how well we are measuring them from the rough indicators at our disposal. When our attention is called to a new topic because of a change in the census questionnaire (e.g., the multiple race item) or due to major social trends (e.g., the foreclosure crisis), historical and ethnographic studies offer the first and best thoughts on what to look for. I also feel sure that ethnography will lead to better contextualized interpretations if it is informed by analysis of the census, survey data, and other quantitative sources. As much as I support mixed methods research, however, I would settle for more awareness of what people are learning on both sides.

There are new approaches and possibilities for using census data. GIS mapping and spatial analysis is one arena. Bringing together information over time about individuals (emplaced in locations) and places is another. The possibilities are multiplying through the availability of confidential census data and 100 percent transcriptions of historical census records. Many scholars are pursuing these avenues, not only with census data but also with innovative applications of longitudinal surveys. As urbanists move to exploit the opportunities, we face some new questions, including whether and how much we care about historical patterns and how access to confidential data can be expanded.

The payoff from using census data has been great. From the 1920s when the founders of the Chicago School brought attention to how urbanism could be studied through settlement patterns, we have been mapping those patterns mainly with census data. We have inquired how social boundaries across communities reflect the boundaries among people by race, class, and other dimensions and how inequalities among groups and between neighborhoods reinforce and reproduce one another. In recent years we have learned how to connect census data about places to census microdata, survey data, and other kinds of information about people. More than ever before we can now study urban inequality from the perspective of how people move among places, how places evolve over time, and how places affect the well-being and futures of residents. Hence my intention is to encourage reliance on census data to study core urban questions, while pointing out both the limits of its usefulness and the prospects for its creative use.

REFERENCES

Anderson, Margo. 2015. *The American Census: A Social History*, 2nd edition. New Haven, CT: Yale University Press.

Brown-Saracino, Japonica. 2017. "Explicating Divided Approaches to Gentrification and Growing Income Inequality." *Annual Review of Sociology* 43:515–39.

Desmond, Matthew. 2016. *Evicted: Poverty and Profit in the American City*. New York, NY: Crown.

———, and Carl Gershenson. 2016. "Who Gets Evicted? Assessing Individual, Neighborhood, and Network Factors." *Social Science Research* 62:362–77.

DuBois, W. E. B. 1899. *The Philadelphia Negro: A Social Study* (1995, Paperback). Philadelphia, PA: University of Pennsylvania Press.

Logan, John R. 2012. "Making a Place for Space: Spatial Thinking in Social Science." *Annual Reviews of Sociology* 38:507–24.

———, Richard D. Alba, and Wenquan Zhang. 2002. "Immigrant Enclaves and Ethnic Communities in New York and Los Angeles." *American Sociological Review* 67:299–322

———, and Bellman, Benjamin. 2016. "Before *The Philadelphia Negro*: Residential Segregation in a Nineteenth-Century Northern City." *Social Science History* 40:683–706.

———, Andrew Foster, Jun Ke, and Fan Li. 2018. "The Uptick in Income Segregation: Real Trend or Random Sampling Variation?" *American Journal of Sociology* 124:185–222.

———, and Matthew Martinez. 2018. "The Spatial Scale and Spatial Configuration of Residential Settlement: Measuring Segregation in the Postbellum South." *American Journal of Sociology* 123:1161–203.

Mayer, Susan E., and Christopher Jencks. 1989. "Growing Up in Poor Neighborhoods: How Much Does it Matter?" *Science* 243:1441–45.

Mora, G. Cristina. 2014. *Making Hispanics: How Activists, Bureaucrats, and Media Constructed a New American*. Chicago, IL: University of Chicago Press.

Ruggles, Steven. 2002. "Linking Historical Censuses: A New Approach." *History and Computing* 14:213–24.

Sampson, Robert J. 2012. *Great American City: Chicago and the Enduring Neighborhood Effect*. Chicago, IL: University of Chicago Press.

———, and Stephen W. Raudenbush. 1999. "Systematic Social Observation of Public Spaces: A New Look at Disorder in Urban Neighborhoods." *American Journal of Sociology* 105:603–51.

Small, Mario Luis. 2004. *Villa Victoria*. Chicago, IL: The University of Chicago Press.

Suttles, Gerald D. 1970. *The Social Order of the Slum; Ethnicity and Territory in the Inner City*. Chicago, IL: University of Chicago Press.

Zorbaugh, Harvey. 1929. *The Gold Coast and the Slum: A Sociological Study of Chicago's Near North Side*, Chicago, IL: The University of Chicago Press.

City & Community

Going Small: Urban Social Science in the Era of Big Data *City & Community* Forum on Census Data

Robert M. Adelman*
University at Buffalo, State University of New York

I appreciate John Logan's thoughts about census data; they mirror my own in many respects. Logan remarks on six issues regarding census data that incorporate two dimensions: measurement and improvement. Rather than repeat his commentary, I offer my own suggestions for the advancement of urban social science. From the outset, though, acknowledging the strengths and uses of census data should encourage the Census Bureau to refrain from making the census political—either as a decennial collection or in smaller forms such as the American Community Survey (ACS). At their best, census data are the optimum estimates we have of the U.S. population. My goal is to continue Logan's discussion about the ways urban scholars can improve our conceptualizations and study of urban life.

THE 2020 CENSUS

Logan first takes up the issue of the emerging citizenship question on the 2020 census. Clearly, this is the current administration's attempt to make political hay out of the impending data collection. In practice, as Logan points out, adding a citizenship question will dampen response rates from immigrant groups; in terms of apportionment and politics, the potential (ab)uses are endless. Among others, a lawsuit has already been filed by the California Attorney General implicating the question as a violation of the U.S. Constitution (see, e.g., Da Silva 2018). Any social scientist who wants to collect quality data knows that to add this question—regardless of how the data will be used—is to contribute instability into the Census Bureau's measurement of the country's population in 2020. Even before the announcement of the citizenship question, Census Bureau tests showed respondents "giving false names or incorrect birthdates, leaving family members out of questionnaires, or abandoning interviews before they were finished" (Capps 2018: Para 1). In fact, a National Advisory Committee on Racial, Ethnic, and Other Populations 2017 report quotes one Census Bureau interviewer as saying, "Three years ago was much easier to get respondents compared to now because of the government changes … and trust factors. … Three years ago I didn't have problems with the immigration questions" (Myers 2017: Slide 13). Although concern about undercounts did not start with the Trump administration, the overall milieu toward migrants in the administration (and beyond), and the citizenship question will exacerbate the situation.

*Correspondence should be addressed to Robert M. Adelman, Department of Sociology, University at Buffalo, SUNY, 430 Park Hall, Buffalo, NY 14260; adelman4@buffalo.edu.

City & Community 17:3 September 2018
doi: 10.1111/cico.12328

My recent work with colleagues on the relationship between immigration and crime across metropolitan areas underscores the political nature of such issues (see, e.g., Adelman et al. 2017; Adelman and Reid 2017; Reid et al. 2005). Emails and commentary from readers have branded a relatively simple research question into a political hot potato. From accusations such as "Obama funded research" or the supposed defense of "illegal alien murderers," politics is almost inescapable in today's partisan climate when studying immigration. Although the citizenship question data may be used by urban scholars for a range of important and reasonable research, it is the larger (often unknown) Machiavellian forces that concern me the most. Undercounting migrants (legal and otherwise), for instance, may lead to an uncalled-for shift in political and economic power from urban centers to rural areas (Scherer 2018).

NEIGHBORHOODS

Logan also discusses neighborhood definitions. One of my favorite activities in my undergraduate and graduate urban courses occurs on the first day of class when I ask students to describe their neighborhoods. Following much of what urban sociologists have learned across the years from studying neighborhoods, some students define them as relatively large areas, probably including multiple census tracts. Other students, though, define their neighborhoods as tiny, only encompassing two to three streets. How, then, should urban social scientists measure neighborhoods?

In this era of "big data," I suggest going small in at least two ways. First, asking people about their neighborhoods and their experiences in them can reveal important insights. Of course, this is not an original idea on my part as it goes back to Du Bois (1899), Liebow (1967), and others all the way through to Anderson (1990) and Pattillo (2007). However, it is worth reminding ourselves of the importance of asking residents straightforward questions about their residential and spatial lives. Whether we call them ethnographies or in-depth interviews, collecting these first-hand accounts balances and enhances lessons learned from census data. No longer should urban scholarship (nor social science generally) be force-fed an outdated and unrealistic dichotomy between qualitative and quantitative data.

Along with two student coauthors, we asked residents of a predominantly poor, African American neighborhood in Buffalo, New York, how they felt about their neighborhood and what they believed could improve their quality of life (Richardson et al. 2014). Although we received some positive comments, the majority said something along the lines of, "Ain't nothing here in Buffalo. … Ain't no real jobs, so the only thing here in Buffalo, is nothin' but to be out there in the streets and doin' something you ain't supposed to be doin' " (2014:11). However, when we asked respondents if they would be willing to move to a new neighborhood, only about half of them said they would, indicating the strength of their social ties, social networks, and social capital in their current neighborhood. Thus, such qualitative data help balance the largely quantitative literature about neighborhood poverty in important ways, including the potential improvement in survey and variable construction.

Second, thinking about the long term and the uses of mixed methods data can provide better and more holistic accounts of neighborhoods and communities. As Logan points out, individuals are embedded in local places, and I want to know about the practical details of their everyday lives. There is much to learn by diving into the local, whether it is the well-known South Side of Chicago or the rapidly changing neighborhoods of Los

Angeles or the smaller world of second- and third-tier cities. Moreover, while census data are surely part of studying these places, there is more to learn about the individuals and families who dwell in them beyond the standard census questionnaire.

Buffalo, again, provides a good example in this case of studying small-scale neighborhood change. A team of graduate students and I studied neighborhood change in Buffalo's West Side neighborhood by using an assortment of data including census tract, crime, in-depth interview, and content analysis data. Although much is known about Buffalo's West Side locally, by combining different types of data about the neighborhood—and by using longitudinal census data—we documented much more change than was originally thought to have occurred. Through the 1970s, for instance, the West Side was known as an Irish and Italian mainstay. The census tracts throughout the area were, on average, 94 percent white in 1970. Thus, it is not an exaggeration to say that this was an overwhelmingly white-ethnic area. Over the last 50 or so years, the community, made up of about 18 census tracts, has changed considerably; the area has become two separate (and separated) neighborhoods. One of the neighborhoods is now in late-stage gentrification composed mostly of highly educated, professional, white residents while the other houses many refugees, immigrants, and other minority groups (Adelman et al. 2018). Make no mistake: The census data were crucial elements in our study because they allowed us to examine change over time but so too were the local, qualitative data because they contextualized the broader findings.

HISTORICAL AND SPATIAL WORK

A number of federally funded projects have made census data more usable in order to study historical urban issues in the United States. Perhaps the most widely known is the Integrated Public Use Microdata Series (IPUMS; www.ipums.org). IPUMS seeks to "harmonize variable codes and documentation to be fully consistent across datasets," especially across individual-level U.S. census data. Today, IPUMS also includes data from other nations and aggregate levels of data for analysis. For example, the National Historical Geographic Information System is a dynamic research tool that includes geographic boundary information. And IPUMS Terra now offers integrated population and environmental data. Similarly, Social Explorer (www.socialexplorer.com) exploits census tract data to make the visualization of the data, for instance, more reproducible. These and other research engines (see Logan's own federally-funded "Urban Transition Historical GIS Project"; https://s4.ad.brown.edu/Projects/UTP/index.htm) have made the manipulation of census data much easier. In addition, there are any number of geo-coded datasets, such as the Panel Study of Income Dynamics, that allow for spatial analyses (see https://psidonline.isr.umich.edu/). All of these projects—and many others—have enriched census data beyond the imaginations of urban scholars just 50 years ago.

MULTITUDES OF DATA

Trained as an urban sociologist and demographer, I cannot overstate the uses of census data in my research (and more recently, in my teaching, too). I worry about the 2020 census collection, the ACS data, and what will happen if and when the Census Bureau becomes a political weapon. I am concerned about the impact of the citizenship question not only on apportionment but also on the reliability and validity of the decennial

data. However, this is not the first time the census has come under political and legal fire. As recently as the 2010 census, the Obama administration was accused of a "power grab" (see, e.g., Sullivan 2009); for the 2000 census, the U.S. Supreme Court ruled that statistical sampling could not be employed for the population count (Prewitt 1999).

In fact, Kenneth Prewitt, Director of the U.S. Bureau of the Census at the time, wrote regarding the 2000 collection: "When partisanship intrudes, a shadow falls across U.S. science" (1999:935). More than the public may know, there is much at stake in counting the population. Nonetheless, the plethora of data sets and geo-coded data mean that there is more than one way to skin a cat. I agree with Logan that rather than crowd-out different types of data, our research questions should nourish a proliferation of studies using multitudes of population data. But there is no replacing the valid and reliable data collected by the decennial census as a cross-sectional and cumulative accounting of the American population.

REFERENCES

Adelman, Robert M., Aysegul B. Ozgen, and Watoii Rabii. 2018. "Buffalo's West Side Story: Migration, Gentrification, and Neighborhood Change." University at Buffalo, SUNY, unpublished manuscript.

———, and Lesley W. Reid. 2017. "Immigration and Crime: What Does the Research Say? Across 200 Metropolitan Areas." *The Conversation*, February 1. Available at: https://theconversation.com/immigration-and-crime-what-does-the-research-say-72176.

———, Gail Markle, Saskia Weiss, and Charles Jaret. 2017. "Urban Crime Rates and the Changing Face of Immigration: Evidence Across Four Decades." *Journal of Ethnicity in Criminal Justice* 15:52–77.

Anderson, Elijah. 1990. *Streetwise: Race, Class, and Chicago in an Urban Community*. Chicago, IL: University of Chicago Press.

Capps, Kriston. 2018. "Census Report Found 'Unprecedented' Fears About Privacy Last Year." Available at: https://www.citylab.com/equity/2018/03/census-report-found-unprecedented-fears-about-privacy-last-year/556700/ [Accessed 4 June 2018].

Da Silva, Chantal. 2018. "Trump Sued by California's Attorney General Over Census Citizenship Question." Available at: http://www.newsweek.com/california-attorney-general-take-trump-court-over-2020-census-citizenship-861378 [Accessed 31 May 2018].

Du Bois, W.E.B. 1899. *The Philadelphia Negro: A Social Study*. Philadelphia, PA: University of Pennsylvania Press.

Liebow, Elliot. 1967. *Tally's Corner: A Study of Negro Streetcorner Men*. Lanham, MD: Rowman & Littlefield Publishers, Inc.

Myers, Mikelyn. 2017. "Respondent Confidentiality Concerns and Possible Effects on Response Rates and Data Quality for the 2020 Census." Presented at the 2017 National Advisory Committee on Racial, Ethnic, and Other Populations Fall Meeting, November 2, Suitland, Maryland. Available at: https://www2.census.gov/cac/nac/meetings/2017-11/Meyers-NAC-Confidentiality-Presentation.pdf [Accessed 4 June 2018].

Pattillo, Mary. 2007. *Black on the Block: The Politics of Race and Class in the City*. Chicago, IL: University of Chicago Press.

Prewitt, Kenneth. 1999. "Census 2000—Science Meets Politics." *Science* 283:935.

Reid, Lesley W., Harald E. Weiss, Robert M. Adelman, and Charles Jaret. 2005. "The Immigration-Crime Relationship: Evidence Across U.S. Metropolitan Areas." *Social Science Research* 34:757–80.

Richardson, William, Lori Glantz, and Robert M. Adelman. 2014. "Ain't Nothin' Here in Buffalo': Residents' Perceptions About Living in a Racially Isolated, High-Poverty Area." *Journal of Children and Poverty* 20:73–90.

Scherer, Michael. 2018. "Potential Citizenship Question in 2020 Census Could Shift Power to Rural America." Available at: https://www.washingtonpost.com/politics/potential-citizenship-question-in-2020-census-could-shift-power-to-rural-america/2018/01/23/c4e6d2c6-f57c-11e7-beb6-c8d48830c54d_story.html?noredirect=on&utm_term=.14dfbf3eb356 [Accessed 4 June 2018].

Sullivan, Amy. 2009. "Why the 2010 Census Stirs Up Partisan Politics." Available at: http://content.time.com/time/nation/article/0,8599,1879667,00.html [Accessed 4 June 2018].

City & Community

Census Data and its Use in the Study of Residential Inequality

Samantha Friedman*
University at Albany, SUNY

At the time of this writing, it was still unclear whether the citizenship question will be included in the 2020 census. On June 6, civil rights lawyers from the American Civil Liberties Union sued the U.S. Department of Commerce to try to stop the U.S. Census Bureau from adding the citizenship question to the 2020 decennial census. This situation is an important reminder of the impact that the U.S. government and larger political system have on our research as urban scholars. Given the lack of clarity on whether the citizenship question will be included on the 2020 census, I would like to focus my response to John Logan's essay on other important issues that he discusses, particularly as to how they relate to residential inequality, a core area of urban sociology.

Logan's essay focuses on the use of census tracts as proxies for neighborhoods, the reliability of census data, and the value of census data at the aggregate level when merged with individual-level data. The data and the issues that he covers are extremely critical to the study of residential inequality in metropolitan America. My response to his essay focuses on the importance of the use of census tracts and decennial census data, as compared to data from the American Community Survey (ACS), in the study of residential inequality. Then I will focus on two aspects of the 2020 census that are not given as much attention in Logan's essay—questions on race and Hispanic origin and new response categories in the relationship question. The two questions on race and Hispanic origin will remain similar to previous decennial censuses, despite the fact that tests have revealed that making Hispanic origin a separate category within the race question and eliminating the Hispanic origin question was shown to yield higher quality data (Porter and Snipp 2018; U.S. Census Bureau 2017a). For the first time in decennial census history, individuals in same-sex relationships will be able to explicitly identify themselves as such through the relationship question on the 2020 census (U.S. Census Bureau 2018). I discuss the implications that the data yielded by these questions will have for the study of residential inequality.

CENSUS TRACTS, DECENNIAL CENSUS DATA, AND THE STUDY OF RESIDENTIAL INEQUALITY

As Logan discusses in his essay, urban scholars have used census tracts as proxies for neighborhoods for decades in urban sociological research and particularly in studies

*Correspondence should be addressed to Samantha Friedman, Department of Sociology, University at Albany, SUNY, Albany, NY 12222; samfriedman@albany.edu.

City & Community 17:3 September 2018
doi: 10.1111/cico.12329

focused on residential inequality. While he correctly points out that census tracts may be smaller or larger than residents' geographic conceptions of their neighborhoods, there are important reasons for urban scholars to continue using this geographic definition of neighborhoods. Census tracts allow researchers to measure residential segregation—one of the most important indicators of residential inequality—uniformly across metropolitan America. In addition, with the increasingly widespread availability of restricted data, microlevel data sets that are available by special permission often provide researchers with census-tract definitions of the locations where respondents live rather than their actual addresses. This is also discussed at great length in Logan's essay as an important tool for urban researchers to advance their knowledge on neighborhood effects. Data providers often use census tracts to protect the confidentiality of their respondents. Thus, in studies of residential inequality at the micro level, including those on locational attainment (e.g., percent white, black, and owners in the neighborhood) or the impact of neighborhoods on migration and access to home ownership, there is no choice but to use census-tract level data.

Most studies that examine alternative definitions of neighborhoods focus on a single or handful of cities or metropolitan areas, making it difficult and time-intensive to employ such definitions for a larger number of areas over time (e.g., Coulton et al. 2001; Grannis 1998; Sampson 2012). Even in using alternative definitions, Coulton et al. (2001) find that the correlations of values between these newly defined neighborhoods and census-tract definitions for percent poverty and percent female-headed families exceed 0.80, suggesting that variation in substantive characteristics based upon alternative neighborhood definitions is quite similar to that found among census tracts. More recently, researchers have found that individuals spend a substantial amount of time away from their homes in "activity spaces"; however, it has been shown that they experience substantial levels of segregation in those places, sometimes as much as they experience in their home environment (e.g., Jones and Pebley 2014; Li and Wang 2017). In sum, census-tracts should remain an important proxy for neighborhoods moving forward in urban quantitative research, and particularly research that focuses on residential inequality.

The collection of the 2020 census data, as well as future decennial data, is extremely important in advancing our understanding of urban residential inequality. Logan's essay as well as a recent publication of his (Logan et al. 2018) discuss the fact that data from the ACS are problematic in producing reliable estimates at the census-tract level because of large levels of sampling variability. Simply put, the decennial census collects data on a significantly larger number of American households than the ACS. Napierala and Denton (2017) show that measures of racial and ethnic residential segregation produced from the 2005–2009 tract-level ACS data differ substantially from those calculated from the 2010 decennial census tract-level data, particularly in metropolitan and micropolitan areas that have smaller population sizes.

There is now a trade-off that urban scholars must contend with—having a wider range of data at the census-tract level but with lower levels of accuracy (from the ACS) or having a much more limited range of data at the census-tract level but with higher levels of accuracy (from the decennial censuses). Moving forward, I echo Logan's call for urban scholars to utilize techniques in small area estimation so that urban researchers can adjust their data to be more sensitive to the sampling variability present in the ACS data. Until recently, the majority of urban researchers have used these data as if they are error-free. Urban scholars should plan to use the 2020 decennial census data that will include

information on the age, sex, race, and Hispanic origin of persons in American households. In addition, there will be a relationship question that will allow the U.S. Census Bureau to release more accurate data on same-sex households (discussed more below) as well as a question on housing tenure.

Decennial census data are clearly essential for accurately measuring racial and ethnic residential segregation in metropolitan America because they utilize racial and ethnic data at the census-tract unit of analysis collected from a large number of households and persons in each census tract. The precise measurement of residential segregation is also critical in understanding many other urban- and nonurban-related outcomes, including residential mobility, access to home ownership, health, and education. In addition, studies of variation in groups' locational attainment will rely on accurate, census-tract level data from the 2020 decennial census.

RACE AND ETHNICITY IN THE 2020 CENSUS AND RESIDENTIAL INEQUALITY

For the first time in census history, the 2020 census will make modifications to the racial question by asking whites and blacks to record their origins in a box below the checkbox. For the race question, the questionnaire asks whites to print their origins and gives the following groups as examples—German, Irish, English, Italian, Lebanese, Egyptian, etc.; for blacks the examples are—African American, Jamaican, Haitian, Nigerian, Ethiopian, Somali, etc. (U.S. Census Bureau 2018). As Logan discussed, previously on the decennial census (before 2010) origins were only asked on the long-form questionnaire for a sample of the U.S. population and not on the short-form questionnaire.

These new data will provide urban scholars an opportunity to conduct a more nuanced analysis of residential inequality. Previous research has shown that foreign-born blacks, who are comprised of immigrants from Africa and the Caribbean, fare better in residential outcomes than native-born blacks (e.g., Rosenbaum and Friedman 2007) although they do experience more residential segregation from whites than their native-born black counterparts (Iceland and Scopilliti 2008). Although these new origin data do not include nativity status, they will allow researchers to examine residential segregation between specific groups of blacks and whites as opposed to the majority of previous research that has examined blacks together as one category, relative to whites.

Urban scholars can also examine segregation within the white category to assess the extent to which those from Northern African, Asian, and Middle Eastern origins that are predominantly Muslim (e.g., Egypt, Morocco, Turkey, Iraq, Iran, and Palestine) are segregated from whites of non-Muslim origins (e.g., English, German, Irish). These segregation scores may be contrasted to the segregation of whites of non-Muslim origins from blacks, Hispanics, and Asians. Given that Muslim whites and blacks experience worse neighborhood outcomes than non-Muslim whites and blacks, respectively, in Philadelphia, this addition to the 2020 census can add to the lack of research that exists on this topic (Friedman et al. 2017). In the future, however, it would be even better if the decennial census questionnaire included a distinct "Middle Eastern or North African" response category in the race question, as had been recommended in the report following the 2015 National Content Test (NCT) (U.S. Census Bureau 2017a).

The modified race question on the 2020 census, however, will not include Hispanic as a new racial category, contrary to the recommendations made after the 2015 NCT (U.S. Census Bureau 2017a). Instead the 2020 census will continue to collect data on Hispanic origin via a separate question, as has been practiced in previous decennial censuses. Preserving the continuity in the way Hispanic origin is asked will make the measurement of residential segregation consistent over time, but higher levels of reliability are not always preferred over higher levels of validity. Research has shown that the two-question version—asking respondents about their race and Hispanic origin—results in an unusually high prevalence of persons reporting that they are of "some other race" (Porter and Snipp 2018; U.S. Census Bureau 2017a). Moreover, this practice results in higher levels of missing values on the racial question (Porter and Snipp 2018; U.S. Census Bureau 2017a). By keeping the questions on race and Hispanic origins separate, urban scholars also run the risk of underestimating the U.S. population that self identifies as Hispanic and white, which is a sizeable and growing share of the young population (Alba 2018).

SAME-SEX COUPLES AND RESIDENTIAL INEQUALITY

A major advancement in the 2020 census that departs from the ACS and previous decennial censuses is the addition of response categories to the question on relationship to the householder that will explicitly identify different- and same-sex relationships. More specifically, the "husband/wife/spouse" and "unmarried partner" categories that were present in the 2010 census and on the ACS will be expanded to distinguish between different- and same-sex relationships (U.S. Census Bureau 2017b). The direct availability of such data will make it much easier and more accurate to gauge the residential inequality that exists between same-sex and different-sex households. Previously, same-sex households were identified by whether householders lived with spouses or unmarried partners of the same sex (Spring 2013). However, studies have revealed that in the 2000 and 2010 censuses persons in same-sex couples failed to identify themselves as being in a same-sex relationship (Badgett and Rogers 2003; Gates 2010), and there was coding error that resulted in the misclassification of a sizeable share of same-sex households (DiBennardo and Gates 2014; O'Connell and Feliz 2011).

Few studies have examined the residential segregation of same-sex versus different-sex households due to the complicated nature of identifying and measuring same-sex households, and these studies come to different conclusions in their assessment of the levels of segregation between these household types (Poston et al. 2017; Spring 2013). The 2020 census data will make it much easier for urban scholars to study residential inequalities between same-sex and different-sex households. Given that housing discrimination is higher for same-sex couples than different-sex couples (Friedman et al. 2013; Levy et al. 2017), such research will be important. Moreover, research that can link more accurate census-tract level counts of same- and different-sex households to individual-level data that also identifies same-sex and different-sex couples may be used to study the locational attainment of these households, providing urban scholars with an important direction for future research. Individual-level data sets that could be potentially used, with restricted data access, would include the National Study of Adolescent Health, the National Health Interview Survey, and How Couples Meet and Stay Together (Joyner et al. 2017; Umberson et al. 2015).

CONCLUSION

John Logan's essay provides us with important insights on the usefulness and limitations of census data as applied to the study of urban social science. He calls for new directions and possibilities for using these data. I echo Logan's enthusiasm as I believe that the 2020 census will offer researchers exciting opportunities to enhance the study of residential inequality in new directions. However, like Logan, I offer cautions about using census-based data, particularly data from the ACS. While we await the outcomes of pending court decisions concerning the addition of the citizenship question to the 2020 census, we as urban scholars must continue to educate and inform our elected representatives about the importance of the data collected by the U.S. Census Bureau. Our vitality as scientists and as participants in American democracy relies on these data being collected carefully and accurately.

REFERENCES

Alba, Richard D. 2018. "There's a Big Problem with how the Census Measures Race." Monkey Cage—Analysis. *The Washington Post*, 6 February. Available at: https://www.washingtonpost.com/news/monkey-cage/wp/2018/02/06/theres-a-big-problem-with-how-the-census-measures-race/?noredirect=on&utm˙term=.721aea4029fb [Accessed 10 June 2018].

Badgett, M. V. Lee, and Marc A. Rogers. 2003. *Left Out of the Count: Missing Same-Sex Couples in Census 2000.* Amherst, MA: MA Institute for Gay and Lesbian Strategic Studies. Available at: http://ourfamiliescount.org/wp-content/uploads/2010/01/LeftOut.pdf [Accessed 14 June 2018].

Coulton, Claudia J., Jill Korbin, Tsui Chan, and Marilyn Su. 2001. "Mapping Residents' Perceptions of Neighborhood Boundaries: A Methodological Note." *American Journal of Community Psychology* 29(2):371–83.

DiBennardo, Rebecca, and Gary J. Gates. 2014. "Research Note: US Census Same-Sex Couple Data: Adjustments to Reduce Measurement Error and Empirical Implications." *Population Research and Policy Review* 33:603–14.

Friedman, Samantha, Joseph R. Gibbons, Colleen E. Wynn, and Recai Yucel. 2017. "Muslim/non-Muslim Locational Attainment in Philadelphia: A New Fault Line in Residential Inequality?" *SocArXiv* [Preprint]. Available at: http://doi.org/10.17605/OSF.IO/N62RT [Accessed 24 August 2018].

———, Angela Reynolds, Susan Scovill, Florence Brassier, Ron Campbell, and McKenzie Ballou. 2013. *An Estimate of Housing Discrimination against Same-Sex Couples.* Washington, DC: U.S. Department of Housing and Urban Development, Office of Policy Development and Research.

Gates, Gary J. 2010. *Same-Sex Couples in US Census Bureau Data: Who Gets Counted and Why.* Los Angeles, CA: UCLA School of Law, The Williams Institute. Available at: https://williamsinstitute.law.ucla.edu/wp-content/uploads/Gates-Who-Gets-Counted-Aug-2010.pdf [Accessed 14 June 2018].

Grannis, Rick. 1998. "The Importance of Trivial Streets: Residential Streets and Residential Segregation." *American Journal of Sociology* 103(6):1530–64.

Iceland, John, and Melissa Scopilliti. 2008. "Immigrant Residential Segregation in U.S. Metropolitan Areas, 1990–2000." *Demography* 45(1):79–94.

Jones, Malia, and Anne R. Pebley. 2014. "Redefining Neighborhoods using Common Destinations: Social Characteristics of Activity Spaces and Home Census Tracts Compared." *Demography* 51(3):727–52.

Joyner, Kara, Wendy Manning, and Ryan Bogle. 2017. "Gender and the Stability of Same-Sex and Different-Sex Relationships among Young Adults." *Demography* 54:2351–74.

Levy, Diane K., Doug Wissoker, Claudia L. Aranda, Brent Howell, Rob Pitingolo, Sarale Sewell, and Robert Santos. 2017. *A Paired-Testing Pilot Study of Housing Discrimination against Same-Sex Couples and Transgender Individuals.* Washington, DC: The Urban Institute.

Li, Fei, and Donggen Wang. 2017. "Measuring Urban Segregation based on Individuals' Daily Activity Patterns: A Multidimensional Approach." *Environment and Planning A* 49(2):467–86.

Logan, John R., Andrew Foster, Jun Ke, and Fan Li. 2018. "The Uptick in Income Segregation: Real Trend or Random Sampling Variation?" *American Journal of Sociology* 124:185–222.

Napierala, Jeffrey, and Nancy Denton. 2017. "Measuring Residential Segregation with the ACS: How the Margin of Error Affects the Dissimilarity Index." *Demography* 54:285–309.

O'Connell, Martin, and Sarah Feliz. 2011. *Same-sex Couple Household Statistics from the 2010 Census.* Working Paper Number 2011–26, Washington, DC, U.S. Census Bureau, Social, Economic, and Housing Statistics Division (SEHSD), released 27 September 2011. Available at: https://www.census.gov/content/census/en/library/working-papers/2011/demo/SEHSD-WP2011-26.html [Accessed 14 June 2018].

Porter, Sonya R., and C. Matthew Snipp. 2018. "Measuring Hispanic Origin: Reflections on Hispanic Race Reporting." *The Annals of the American Academy of Political and Social Science* 677(1):140–52.

Poston, Dudley L., Jr., D'Lane R. Compton, Qian Xiong, and Emily A. Knox. 2017. "The Residential Segregation of Same-Sex Households from Different-Sex Households in Metropolitan USA, Circa-2010." *Population Review* 56(2):1–29.

Rosenbaum, Emily, and Samantha Friedman. 2007. *The Housing Divide: How Generations of Immigrants Fare in New York's Housing Market.* New York, NY: New York University Press.

Sampson, Robert J. 2012. *Great American City: Chicago and the Enduring Neighborhood Effect.* Chicago, IL: University of Chicago Press.

Spring, Amy L. 2013. "Declining Segregation of Same-Sex Partners: Evidence from Census 2000 and 2010." *Population Research and Policy Review* 32:687–716.

Umberson, Debra, Mieke Beth Thomeer, Rhiannon A. Kroeger, Amy Caroline Lodge, and Minle Xu. 2015. "Challenges and Opportunities for Research on Same-Sex Relationships." *Journal of Marriage and the Family* 77(1):96–111.

U.S. Census Bureau. 2017a. *2015 National Content Test: Race and Ethnicity Analysis Report.* United States Census 2020. Washington, DC: U.S. Department of Commerce.

———. 2017b. *2020 Research and Testing: 2015 National Content Test Relationship Question Experiment Analysis Report.* United States Census 2020. Washington, DC: U.S. Department of Commerce.

———. 2018. *Questions Planned for the 2020 Census and American Community Survey.* Washington, DC: U.S. Department of Commerce.

City & Community

Problems, Puzzles, and the Production of Knowledge: Harnessing Census Data in the Age of Trump

Karyn Lacy*

We tend to think of the Census Bureau as merely a bean counter, but the institution performs another, less apparent, role: signaling which demographic shifts carry the most weight in society. Trump's insistence that the Census Bureau include a controversial citizenship question on the 2020 census would mark a decisive shift in the Bureau's ability to count unauthorized immigrants accurately and in the distribution of federal resources to communities where immigrants settle in large numbers. This essay considers what these consequences, should Trump prevail, would mean for social scientists who study immigration. This distressing prospect presents an opportunity for demographers to consider how the work of ethnographers could be utilized to circumvent the data limitations a citizenship question would likely impose.

The stated purpose of the U.S. Census is to count every person residing in the country. Determining which residents are citizens of the United States has not been a part of the calculus since the 1950 Census. To be sure, from 1890 through 1950, each decennial census included a question about citizenship status, a direct response to the marked increase in immigration principally from southeastern Europe (Cohn 2018). By continually including this question decade after decade, the census signaled to researchers and the general public that disparities between citizens and immigrants were an important indicator of social inequality. When mass migration from Europe tapered off, and when the disparities between immigrants and the native-born disappeared as white ethnics were incorporated into the American mainstream, the census dropped the citizenship question from the decennial census. Thus, while some may think of the Census Bureau as simply a bean counter, in practice, the institution does not merely seek to provide an accurate count of the population. It is also a tool that the government employs to demonstrate which demographic shifts carry the most weight in society and which no longer warrant consideration. For example, "when the census asks [respondents] about race, it means race matters" (Stone 2018:5).

From 1960 on, the decennial census counted unauthorized immigrants and those awaiting citizenship along with U.S. citizens. The census drew no distinction between citizens and non-citizens. That could all change if a controversial citizenship question, proposed by the Trump administration, is included on the 2020 census. It would not be the first time that the census revised the way that the federal government measures a concept. For too many years, the census included the pejorative "Negro" among the

*Correspondence should be addressed to Karyn Lacy, Department of Sociology, University of Michigan, Ann Arbor, MI 48109; krlacy@umich.edu.

City & Community 17:3 September 2018
doi: 10.1111/cico.12330

list of options available to respondents who identify as Black, arguing that elderly black respondents indicated a preference for the term. The 2020 census marks the first time that "Negro" will not be included among the labels designating black ancestry. There may also be changes to the Hispanic category in 2020. From 1980 to 2000, the census required Latino respondents to select a race after identifying their ethnicity because the census did not consider "Hispanic" a racial category. Latino respondents' reaction to this designation ranged from confusion about the question's intent to outright anger that the census was forcing them to choose an identity they felt was not representative of them. Mexican American and Puerto Rican Americans know that they are not white, but they are unwilling to classify themselves as black when completing the census because they are aware that blacks occupy the bottom rung of the racial hierarchy in the U.S. For these respondents, Hispanic *is* a race (Lacy 2011). The shift in 2000 from a binary conception of race to a multiracial category represents another sea change in how the census measures race and signals the ways in which the bureau believes established racial classifications are changing. Rather than requiring respondents to settle on a single race, respondents may now "mark one or more" when they complete their census form, choosing to belong to as many racial categories as they like (Williams 2006). As Logan observes, this reorganization of the way that the federal government measures race points to new directions in research on racial inequality and raises penetrating questions about how scholars might leverage these new racial classifications in their scholarship.

These revisions in the way the census measures race are designed to ensure an accurate count of the population. But the citizenship status question proposed by the Trump administration would be different. To begin with, while the census solicits input from the public, scholars in particular, as they make determinations about which questions to revise, omit, or add, it is unusual for the president to weigh in, demanding that the census add a category of his choosing. Population data collected by the census are used in the apportionment of the House of Representatives and to distribute federal funds to communities based on population growth or decline. The Trump administration claims the proposed citizenship question is needed to protect voting rights for minorities. Given Trump's political attacks on racial minorities and general disregard for civil rights legislation, it is unlikely that his administration's push for inclusion of the citizenship question on the 2020 census is, in fact, motivated by a desire to enforce the Voting Rights Act. It is more likely that the citizenship question is designed to discourage unauthorized immigrants from completing the census, thereby reducing the number of politicians in the House of Representatives drawn from states with large immigrant populations, such as California, as well as the percentage of federal funds allocated to these populous states. Critics believe Trump's proposed citizenship question is designed specifically to harm blue states, traditional destinations for immigrants. However, by weaponizing citizenship status, Trump could hurt red states too, as many of these new destinations experienced the highest rates of immigration in recent years (Lacy 2016; Waters and Jimenez 2005).

Trump's citizenship question coincides with a rise in immigration too, with one key distinction: This time, the majority of immigrants are people of color, not white ethnics. The Hart–Celler Act of 1965 eliminated the restrictions established under the National Origins Act of 1924, a policy that limited immigration from the eastern hemisphere, regions populated by people of color. With the restrictions removed, the immigrant population in the United States grew from a practically inconspicuous 5 percent in 1970 to 14 percent by 2017. About 20 percent of these Hart–Celler immigrants arrived in the 2000s. At

the end of that decade, people of color comprised the majority of the immigrant population in the United States, migrating from Mexico, Asia, Latin America, and Africa. Only 12 percent of new arrivals came from Europe. Due to immigration and reproduction, demographers predict that by 2050, the United States will be a "majority–minority" nation, meaning racial minorities will comprise more than 50 percent of the population in the United States. Whites, the group to whom Trump has consistently pledged his loyalty, will no longer constitute the majority group. One way to hold on to power as your share of the polity declines is to deprive the growing majority of political, economic, and civil rights, just as the apartheid regime did in South Africa. Trump's insistence that the census include a citizenship status question is the first step in this process.

Finally, the addition of a citizenship question could also increase undercounting and other enumeration errors. The census is already deeply concerned about undercounting. Black men, highly-mobile children and grandchildren, members of large non-nuclear families, and Latino Americans annoyed by the sequence of the race question are the groups at the highest risk of being undercounted (Lacy 2011). These characteristics are manifest in unauthorized immigrant populations as well; therefore this group is already at high risk of being omitted from the census count. These risks multiply in the current political climate, in which few unauthorized immigrants would trust that the census is serious about protecting confidentiality. Many immigrants will assume the rogue Trump administration would demand that confidential census data be turned over to the Department of Justice. A citizenship status question might also lead unauthorized immigrants to misrepresent their status for fear of retribution. The census found that when immigrants fill out the American Community Survey, which does include a citizenship question, 30 percent of immigrants who are not citizens yet falsely report that they are (Cohn 2018). In the end, unauthorized immigrants may decide that they have little to gain and far too much to lose if they allow the census to intrude in their lives.

If the Trump administration prevails and the citizenship question is included in the 2020 census and if the question has the intended effect of intimidating unauthorized immigrants such that they avoid participating in the census, what would this undercounting mean for our understanding of American life in a diverse society? Logan presents the dilemma as an opportunity for scholars to think differently about how we conduct research. He argues we have privileged demography over and above ethnographic evidence when we should be thinking about how ethnography could enrich demographic studies. So, let's give some thought to this. There are at least three possibilities.

First, demographers could turn to ethnographic studies to gain a comprehensive understanding of the processes underlying outcomes of interest. For example, Lacy (2016) assesses three emerging trends revealed by the work of demographers, then turns to an analysis of ethnographic studies to explore how these trends impact the everyday lives of people. Concerning suburban poverty, she argues demographers identify new sites for the concentration of poverty, while ethnographers reveal how the poor make sense of their lives in a space built for families in possession of middle-class resources (on suburban poverty, see Murphy forthcoming). Regarding immigration, demographers show *how* the spatial patterns of post-1965 immigrants differ from those of their 19th- and 20th-century white ethnic counterparts. Ethnographers chronicle the everyday experiences of immigrants, including undocumented immigrants, to illustrate precisely how their assimilation trajectory is fraught with obstacles (on suburban immigrants, see Mahler 1995). In studies of the return migration, demographers pinpoint which groups left and where

they resettled, while ethnographers live among these internal migrants to understand why they left and what is gained from the relocation (on black suburbanization, see Lacy 2007). Ethnographers could shed light on the study of unauthorized immigrants too. Demographers tell us that the citizenship status question will likely lead to undercounting, but we must turn to ethnographers for insight into how unauthorized immigrants negotiate this exclusion day in and day out.

Second, as Logan argues convincingly, demographers could turn to ethnographic studies as the theoretical foundation for the studies they conduct using census data. For example, much of what we know about neighborhoods, he writes, comes from the Chicago School's ethnographic tradition. In that sense, demographers who study neighborhoods today are building on the work of Park, his students, and urban scholars like DuBois. Could this model help us to wrestle with the limitations of the citizenship question? One potential drawback is that so few ethnographic studies of unauthorized groups have been conducted due to the ethical dilemmas of studying this group. But we should leverage what we do know from ethnographic studies about the challenges unauthorized groups confront and use those findings to generate new research questions that could be studied among a broader swath of the population by demographers.

Third, Logan suggests an analysis of context (whether a neighborhood or some other site) is important for demographers as a way of determining what outcomes they need to study and with what variables. He contends this process is especially important when a change to the census generates a new research topic. We may think we know why unauthorized immigrants will either decline to participate in the census, lie about their status, or simply skip the citizenship status question while completing the remainder of their census form, but as Duneier et al. (2014) make clear, we cannot answer these questions with certainty in the absence of analysis of the daily life of immigrants as they experience their environment. Ethnography allows researchers to uncover mechanisms which may not be apparent to the respondents under study. The method of data collection provides a framework for scholars to "understand things that are taken for granted but never defined." Going into the communities where unauthorized immigrants live for an extended period of time will not only aid the census in their efforts to reduce miscounts in future censuses, it will help us to think about strategies to create a more inclusive society, one in which immigrants are not pushed to the margins.

The work of scholars concerned with inequality will surely be affected if the citizenship question is included once again on the decennial census. It is useful to begin to think now about how we will adapt.

REFERENCES

Cohn, D'Vera. 2018. *What to Know About the Citizenship Question the Census Bureau is Planning to Ask in 2020.* Washington, DC: Pew Research Center.

Duneier, Mitchell, Philip Kasinitz, and Alexandra Murphy. 2014. *The Urban Ethnography Reader.* New York, NY: Oxford University Press.

Lacy, Karyn. 2007. *Blue-Chip Black: Race, Class, and Status in the New Black Middle Class.* Berkeley, CA: University of California Press.

———. 2011. *An Ethnographic Study of Enumeration Methods and Coverage in a 2010 Census Coverage Measurement Survey Operation Primarily Among Blacks in a Chicago Field Site.* Washington, DC: U.S. Census Bureau.

———. 2016. "The New Sociology of Suburbs: A Research Agenda for Analysis of Emerging Trends." *Annual Review of Sociology*, 42:369–384.

Mahler, Sarah. 1995. *American Dreaming: Immigrant Life on the Margins.* Princeton, NJ: Princeton University Press.

Murphy, Alexandra. (forthcoming). *Where the Sidewalk Ends*. New York, NY: Oxford University Press.
Stone, Lyman. 2018. "Why It's Perfectly Fine to Ask about Citizenship Status on the Census." *Vox*, 5 April. Available at: https://www.vox.com/the-big-idea/2018/4/5/17202092/census-question-citizenship-immigration-trump-deportation-undercount-latino-hispanic. [Accessed May 22, 2018].
Waters, Mary, and Tomas Jimenez. 2005. "Assessing Immigrant Assimilation: New Empirical and Theoretical Challenges." *Annual Review of Sociology* 31:105–25.
Williams, Kim M. 2006. *Mark One or More: Civil Rights in Multiracial America*. Ann Arbor, MI: University of Michigan Press.