Issues in Depicting Population Change with Dot Maps

Eugene Turner and James P. Allen

ABSTRACT: Creating dot maps to show changes in racial and Hispanic population distributions between two census periods can be an effective way to examine one of the most important dimensions of change within any metropolitan area. Using dots of one color to show population increase and dots of a second color to show population decrease vividly reveals where changes have occurred within a larger total population. We prepared such maps for the book Changing Faces, Changing Places: Mapping Southern Californians, the text of which analyzes and interprets the population shifts evident on the maps. The maps show the expansion and contraction of racial and Hispanic populations in specific neighborhoods so that community leaders and residents alike can easily relate general trends to their localities. In this article we describe the preparation of these dot maps and explain major problems encountered in linking the 1990 and 2000 census population counts at the tract level. We explain our solutions, which we believe made possible more accurate mapping of neighborhood change.

KEYWORDS: Dot mapping, U.S. census, population change

Introduction

At the end of each decade when the U.S. Census of Population data are released, one of the first questions people ask is, “What has changed?” Some people look at change in terms of counties or metropolitan units, but many elected officials, scholars, and business people are interested in how neighborhoods have changed. By measuring the counts of people gained or lost in census tracts between decennial censuses, we have been able to effectively present this through dot mapping. Just after the release of the 1990 census, we created dot maps that showed, by census tract, gains in population for Los Angeles and Orange Counties with red dots and loss in population with blue dots (Turner and Allen 1991). That work was updated after Census 2000 into a more extensive book, Changing Faces, Changing Places in 2002 (Allen and Turner 2002). In this article, we describe the creation of such maps and the difficulties encountered in trying to resolve the counts of race and Hispanic populations from two successive censuses.

Eugene Turner, Department of Geography, California State University, Northridge, Northridge, California 91330-8249. Tel: (818) 467-0778 Fax: (818) 677-2723. E-mail: <eturner@csun.edu>. James P. Allen, Department of Geography, California State University, Northridge, Northridge, California 91330-8249. Tel: (818) 886-8730. Fax: (818) 677-2723. E-mail: <jallen@csun.edu>.

Representing Change Using Maps

The typical representation of change using maps is to prepare a series of maps at discrete time periods—usually on census dates. Change must be inferred from the apparent visual differences in the individual patterns. For example, a set of three dot maps portrays the distribution of slaves at three census periods in the Historical Atlas of the United States (National Geographic Society 1988, p. 40). Also, in the same atlas, a set of three isoline maps represent a population density surface at three time periods (p. 42). Such maps are helpful in providing a broad overview of population changes, but details are more difficult to see.

An expression of the magnitude of change is often presented through a choropleth map that shows the percent of increase or decrease in population. For example, in the Census Atlas of the United States counties that gained a percentage of population are filled with different shades of green while counties that lost a percentage of population are filled with shades of purple (Suchan et al. 2007, pp. 15, 61, 241). To show the change in the number of persons a graduated symbol map is typically employed. In the Atlas de Franciens, various maps containing graduated squares indicate a numerical population gain in an area with a pink square and a loss with a green square (IAURIF; INSEE 2000, pp. 45, 47, 63).

One obvious solution to portraying change using maps is to create an animation. Thanks to
powerful software such as Adobe Flash, numerous examples of animated maps can be found on the Internet (University of Sydney Archaeological Computing Laboratory 2004; Egwuekwe 2010; Ethington 2001; Lodge 2010). However, to capture detailed spatial changes over time requires frequent data collection that is rarely available. Furthermore, it is usually very difficult for the viewer to gain anything other than the most general impression from map animations. For scholars and other researchers who wish to look closely at patterns of change, animation does not seem to be helpful. Thus, presenting the change in counts of race and Hispanic populations within census tracts between the 1990 and 2000 census by means of a series of static dot-maps seemed especially appropriate.

Dot maps have commonly been used to capture the spatial detail of human population. These have taken a variety of forms that range from a
basic map where each dot represents a number of persons to a mix of graduated symbols and dots (Fawcett 1935; William-Obston 1963, p. 246). Cartographers have employed the latter method in an attempt to deal with the large differences in population density which exist between urban and rural areas where a single dot value would either result in complete coalescence of dots in densely settled areas or virtually empty rural areas.

Dot maps do present a number of advantages for portraying distributions of discrete phenomena such as people, animals, or farms. They provide considerable detail about the nature of the varying density in the underlying distribution and, because they are discrete symbols, they can be overlaid on other data such as shaded terrain, choropleth patterns, or isolines. They also can be combined with other dot distributions to portray the associations of the data (Jenks 1953). For example, a map of soil taxonomic orders across the U. S. illustrates the mixing of ten soil orders, where each dot represents eight square kilometers of a soil type (Barrett 2008).

A challenge in dot mapping is to select an appropriate dot size and a dot value since these will affect the impact of the patterns on the viewer (Dent et al. 2009). Mackay (1949) and, more recently, Kimerling (2009) have proposed methods to assist in determining these two factors. Bak et al. (2009) sought to improve the visualization of dots by generating a cartogram that could improve the presentation of dots by enlarging the map in very densely populated areas. We found experimentation through trial and error valuable for determining a dot size and value. A dot value of 100 persons provided distribution detail across the counties and enabled quick estimations of populations by simply counting dots. The dot size of about 0.6 mm was clearly visible over the map and presented density variation in most areas of great change. (Note the clustering of blue dots in Figure 1.)

The other requirement for dot mapping is to place the dots in a way that captures the actual pattern of the population being represented. Unfortunately, the underlying population is rarely spread uniformly within the areal unit, and so the cartographer will attempt to determine how population density varies within the reporting units. Traditionally, supporting ancillary information such as aerial photographs or topographic maps were used to make the determination, but this was very time consuming.

Mapping and GIS software have automated the placement of dots, but the software randomly places the dots within the statistical units and creates fairly uniform dot densities within each areal unit. The resulting pattern often looks unrealistic in less dense areas, and especially in hilly terrain when relief is shown in the background. To improve automatic dot placement, researchers have suggested strategies such as using smaller areal units or using GIS to redefine the boundaries of the geographic areas with the aid of satellite imagery, land use data, or street networks (Menmen 2005; Holt et al. 2004; Andresen and Brantingham 2006).

A detailed database covering much of the world, the LandScan Global Population Database, can improve the position of dots in more rural areas (Oak Ridge National Laboratory 2005). LandScan gives an estimate of how many people are expected to be within a roughly 1 square kilometer grid cell at any time of day for any day of the year. The LandScan database is regularly updated, but only the most recent version is made available and it is not recommended for calculating population change over time due to compilation differences. Where we found automated dot placement to be unsatisfactory, we used topographic maps, other large-scale maps, and occasionally aerial photographs to position our dots.

**Representing Population Change with Dots**

We have found dot maps to be especially useful for representing metropolitan population change between census years (Turner and Allen 1991; Allen and Turner 2002). In several maps of race and Hispanic population change in metropolitan Los Angeles census tracts, we used a red dot to indicate an increase of 100 persons and a blue dot to indicate a loss of 100 persons (see Figure 1). The maps vividly capture the areas of change within the metropolitan area and are not influenced by the total count of persons. For maps printed in black and white, black dots represent population gain and population loss is indicated with hollow dots or white dots against a gray background.

An analysis presented in *Changing Faces, Changing Places* of the colored dot map of Black population change between 1990 and 2000 illustrates how important spatial patterns are revealed (see Figure 1). Perhaps the most striking pattern on the map is the decrease in Black populations in the old once-segregated concentrations, shown by the blue dots. The largest and best known of these former ghettos is South Central L.A. or South L.A., but the map also shows Black population
numbers declining in former suburban ghettos of Pasadena, Long Beach, Pacoima in the San Fernando Valley, and Monrovia, as well as in cities beyond Los Angeles County.

In all these areas, the decline of Black population numbers has been due to both a departure of a growing Black middle-class population and their replacement by Hispanic, especially Mexican, immigrants (Allen and Turner 1997). The departure of middle-class and more affluent Blacks has meant that poverty, crime, and poor schools have come to characterize the former ghettos. Forty percent of the remaining Blacks in South Central still own their own homes (Myers 2002), but the lower housing prices in this and other poor areas have attracted Mexicans, with the result that Blacks comprise only about half the population in areas shown in blue. Overall, the out-movement of Blacks from former ghettos into more White neighborhoods has reduced Black-White segregation in Los Angeles County.

The red dots show tracts in which Blacks have increased during the 1990s. Blacks moved to nearby areas west and south of South Central, but high housing prices restricted movement to the north, and dense Hispanic settlement discouraged movement directly east. The arrivals of Black were widespread in more distant areas with moderate housing prices, particularly in older suburbs, such as in Bellflower, Lakewood, and the San Fernando Valley. For the San Fernando Valley, a mostly White area fifty years ago, the new settlement of Blacks represents a major change. Especially significant is the fact that Blacks in the Valley are now widely dispersed in contrast to the former ghettos. Also, Blacks have settled on the urban fringe in a dispersed pattern: Southern Orange County, the 1-15 freeway corridor in Riverside County, and new developments in San Bernardino County.

We began the process of creating the metropolitan Los Angeles maps by first estimating the 2000 single-race counts of persons for census tracts in the 5-county Los Angeles metropolitan area, a problem that will be addressed later in this paper. This was followed by reallocating the 1990 population counts into 2000 census tracts by aggregating 1990 block data into the 2000 tract boundaries. Change data were generated by subtracting the 1990 population counts for each tract from those of 2000. Positive and negative values were represented as positive numbers in separate columns of a spreadsheet and then a dot map of each category (red dots indicating a gain and blue dots indicating a loss) was generated in ArcMap. The two patterns were then overlaid and the composite map exported to Adobe Illustrator. Because ArcMap randomly places the dots within the statistical areas, the positions of some individual dots within larger tracts were adjusted visually from open to settled areas by consulting background map layers of streets and parks and by referencing the printed Thomas Guide county atlases. While the above steps fairly describe the production of the dot maps considerable data preparation was necessary to make the census tabulations from the two decades comparable. Next we discuss issues relating to the use and comparability of census data.

Problems in Dot Mapping of Population Change

The U.S. Census is constantly being modified by changes in the choices and definitions of variables, collection procedures, and in geographic units (U.S. Census Bureau 1985; Martin et al. 2002; U.S. Census Bureau 2002, Martin et al. 2003; U.S. Census Bureau 2008b). For mapping purposes these issues might be categorized as those related to the collection and reporting of variables such as race and Hispanic identity and those related to geography.

Change in Census Questions Regarding Race

Each decade, changes in the definition of variables present challenges to ensuring that measures of change in populations are comparable (U.S. Census Bureau 2008a). Fortunately, the Census Bureau attempts to document the comparability of different variables to those in the previous census. But even with this documentation not all changes can be confidently accounted for between census decades because the changes may affect where people were counted, how they chose their race or Hispanic identity, or how they interpreted the census questions.

The Office of Management and Budget (1997; 2000) has established standards for presenting race and Hispanic data. Between 1990 and 2000 the OMB modified the terms used to describe some basic race groups. The term "Black" became "Black or African American," "American Indian, Aleut, or Eskimo" became "American Indian and Alaska Native," and "Asian or Pacific Islander" was split into "Asian" and "Native Hawaiian and Other Pacific Islander." The term for Hispanics in 1990, "Hispanic origin," became "Hispanic" in
The Census Bureau uses the separate terms “Spanish, Hispanic, and Latino” interchangeably. In this paper, we use the earlier terms and categories. Generally the Census Bureau defines race as “self-identification by people according to the race or races with which they most closely identify. These categories are socio-political constructs and should not be interpreted as being scientific or anthropological in nature. Furthermore, the race categories include both racial and national-origin groups” (U.S. Census Bureau 2005, p. B-38). For Hispanics the term “origin” is meant to indicate “the heritage, nationality group, lineage, or country of birth of the person or the person’s parents or ancestors before their arrival in the United States” (U.S. Census Bureau 2005, p. B-13).

The questions and categories used in 1990 and the modified questions and categories used in 2000 and 2010 can be seen by examining the race and Hispanic questions placed in the respective censuses. Because race and Hispanic information is gathered from two separate questions, the tabulations of race and Hispanic identity are not mutually exclusive. For example, a person indicating a White race may also be Hispanic and will be counted in the tabulations for each question. However, the Census Bureau does produce a few race tables (e.g., P8 in Summary File 1 from 2000) where race category counts are stratified by those persons who did and did not indicate a Hispanic identity. Thus, it is possible to separate the Hispanic persons from any of the race categories.

The most significant change to the race counts of Census 2000 was that for the first time, people were allowed to mark more than one race on the census questionnaire. People could indicate up to six races on the form and in the PL94-171 data file, 63 race category combinations are reported (U.S. Census Bureau 2001b). In 1990 and prior censuses, people were forced to choose one race, and so the data on race of Census 2000 are not directly comparable to earlier censuses.

In order to determine the change in counts of race groups between 1990 and 2000, one must decide how to handle the count of the mixed-race population. Fortunately, only 2.4 percent of the population nationally and 4.7 percent of the population in the Los Angeles 5-county region reported two or more races and so the majority of the population still falls into a single-race only category. Nevertheless, whether the 2000 single-race only counts are used to calculate the percent change or the 2000 single and mixed-race counts are used can make a large difference in the percent population change reported for the two census periods. For example, in the worst case situation, if the 1990 Los Angeles five-county count of American Indians, Aleuts, or Eskimos is compared to the 2000 single-race only count of American Indians and Alaska Natives, the area experienced a 4.1 percent drop in that population. If the 2000 combined single-race and mixed-race American Indian and Alaska Natives count is used, the group experienced a 118.6 percent increase.

A better estimate for race counts for 2000 that more closely approximate the single-race counts of 1990 can be achieved by reallocating the mixed-race responses to single-race categories through a process called “bridging” (Office of Management and Budget 2000; National Center for Health Statistics 2000). The OMB reviewed several methods of assigning the mixed-race counts to single-race counts. Generally, methods that fractionally assigned the mixed-race counts to the groups which were part of the mixed-race category were preferred, though no single bridging method was selected.

For Changing Races, Changing Places we developed a fractional assignment method by comparing ancestry responses to the race responses in 1990, using the U.S. Census Public-Use Microdata Sample (PUMS) file for California (Allen and Turner 2001). By aggregating ancestries that clearly indicated White, Black, Asian and Pacific Islander, and American Indian, Aleut, and Eskimo races, we determined for 1990 those persons who might have indicated a mixed-race response had the questionnaire been like that used in 2000. That is, someone who indicated a White race and a Japanese or Chinese ancestry could be considered a mix of Asian and White races. This research suggested that 67.1 percent of the California Black-White population identified their race as Black and that 32.8 percent of the California Asian-White population identified their race as Asian and Pacific Islander. We used these figures for apportioning those two mixed-race counts for the Census 2000 data, but for other race combinations we assigned mixed-race counts in equal fractions to each of the groups that comprised their category.

After assigning the mixed-race populations to the race categories, we reduced double-counting of persons in race and Hispanic total counts. To do this we chose to count non-Hispanic Whites and Non-Hispanic Blacks as the significant race categories, thus leaving Hispanic Whites and Hispanic Blacks to be counted as Hispanics. This was because we considered the Hispanic identity to be the more important one for people who also identified racially as either White or Black. In contrast, we considered the Hispanic identity to
be secondary among people whose race identity was American Indian, Aleut, or Eskimo or Asian and Pacific Islanders. This meant that Hispanic members of those groups (primarily Filipinos and Guamanians) were included in the counts of these race groups but were subtracted from the Hispanic total count.

The Census Bureau publishes in its Summary File 1 even more detailed race counts, but mixed-race counts of the various categories are not given for these more detailed groups. For this reason, population change maps of specific races like Chinese, Japanese, Filipinos, and Koreans in Changing Faces, Changing Places used the single-race counts from Census 2000, resulting in a slight underrepresentation on our maps.

Change in Census Questions Regarding Hispanic Identity

Sometimes subtle changes in the wording of questions can cause a change in the response rate and either an increase or decrease in the numbers of counted persons. For example, in Census 2000 the question on Hispanic origin was changed so that a lengthy list of countries of origin did not appear after the Other Hispanic write-in question. This apparently was the cause of a large drop in the reported counts of some Hispanic groups such as Dominicans, Salvadorans, and Ecuadoreans, though, the total count of Hispanics was somewhat higher than expected possibly due to a better count in 2000 (Suro 2002; Ramirez 2008). For this reason, maps showing population change for Hispanic subgroups were not included in Changing Faces, Changing Places.

Changes in Tract Boundaries and Group Quarters Locations

At the metropolitan level, the census tract serves as a useful geographic unit for analyzing and displaying census data. However, between census dates, the boundaries of many tracts have been modified and, as a result, the population counts for the non-matching tracts from one decade must be reallocated to match boundaries of the second decade (Holt et al. 2004). Particularly challenging for establishing common census tract areas for Changing Faces, Changing Places were situations where the 1990 census block boundaries crossed Census 2000 tract boundaries. This typically occurred in new suburbs constructed in the 1990s. Prior to 2000 this was open country without the physical boundaries normally required by the Census Bureau to define blocks, but with large-scale development new blocks and tracts had to be created. In those cases, the street and park overlays used in ArcMap and the printed Thomas Guide maps provided information as to whether or not there was settlement in those areas.

One unexpected outcome of mapping population change data with dots in Changing Faces, Changing Places was the revelation of shifts of some group quarters populations (e.g., non-housing units such as prisons, hospitals, dormitories, shelters, etc.) between adjacent tracts between 1990 and 2000. This was suggested by the presence of a census tract filled with red dots adjacent to a census tract filled with blue dots. In most cases these occurrences happened where an institution such as a college campus or prison was present and seems related to the way census mailing addresses were allocated to blocks. The most notable example occurred in Los Angeles County in tract 9292 near the community of Castaic where the count of 7,115 prisoners in the Peter Pitchess Detention Center (the entire tract population) was shifted into a residential tract to the west. Other shifts occurred at tract 5500 in the community of Norwalk where the count of 918 patients of the Metropolitan State Hospital was moved; at tract 2227 where the count of 2500 college students in dormitories at the University of Southern California was moved; and at tract 4024.04 where the count of 1200 students attending California State Polytechnic University, Pomona, was moved. Another error in tract 1152.02 in 1990 misplaced the count of 1400 students attending California State University, Northridge. It is likely the students used a campus mailing address, which placed them in a tract containing no dormitories.

Once we discovered a few errors of this type, we used table PCT16, Group Quarters Population by Group Quarters Type, from California Summary File 1 to locate all tracts with large group quarters populations and examined them for gains and losses of population between adjacent tracts (U.S. Census Bureau 2001a). We shifted the group quarters populations of these tracts back to the tract that contained the institution, which in nearly all cases was the 1990 tract. Visually, the corrections removed isolated population losses or gains within more general patterns of gain or loss.

Such errors may be common nationally and any map showing population change at the tract level should be checked in areas where group quarters are present (Scott 2001). These shifts are a real
problem for those using the census for a statistical analysis intended to measure change over time as the abrupt shift in tract population can lead a researcher to note significant changes when, in fact, little has occurred. The problem is part of the larger issue of defining residence especially for those in group quarters that are counted as part of a local population (National Research Council 2006). Such tracts tend to present atypical socioeconomic data, and it may be useful to subtract them from the total tract populations when they represent a large proportion of the total.

Census 2010

The race and Hispanic counts in Census 2010 essentially will be like those in 2000 so that measures of change can be made confidently for the questions that are mailed to all households on the “short form” questionnaire (U.S. Census Bureau 2008b). One small change that will occur is that the Other Asian, Other Pacific Islander, and Other Hispanic write-in categories included a list of example answers similar to those presented in 1990. The lack of those examples in 2000 was thought to be partly responsible for a reduction in write-in responses previously recorded in 1990, and their inclusion in 2010 likely could be a factor in increasing the reported counts of the write-in tabulations for the listed groups. Thus, any calculation of population count differences between 2000 and 2010 for the groups listed under the “Other” categories may be inflated in 2010.

There is a major change for some variables such as ancestry, country of birth, or language spoken at home which were previously collected on a more detailed “long-form” questionnaire sent to about one in six households. The sample questionnaire will not be distributed at the 2010 decennial census; it has been replaced by an annual survey of 3 million addresses called the American Community Survey. Data will be released annually for geographic areas containing over 65,000 persons. They will be released as a three-year average for geographic areas containing over 20,000 persons and as a five-year average for areas containing fewer than 20,000 persons. So far, one and three-year counts have become available, with the five-year counts scheduled to begin release in 2010.

Although the American Community Survey will provide population and socioeconomic variables similar to those of Census 2000, the estimates are from a smaller sample and will carry a greater sampling error, especially for small-area tabulations such as census tracts (National Research Council 2007; U.S. Census Bureau 2009). Other differences include the fact that the five-year averages reported for census tracts represent data collected over a five-year period rather than during a brief census data collection period. Therefore, a longer sampling time frame is involved. The Census Bureau suggests using the five-year data to represent change, and that mixing of data from one, three, and five-year samples be avoided. Even though five-year average data will be released annually, measures of change should be reported at no less than five-year intervals to avoid using overlapping years in adjacent time periods, though, statistically, some overlap can be accommodated. Most reported values include a 90 percent sampling error or margin of error which will be larger for small-area geographic units such as census tracts. These appear in tables along with an estimate of the data value. Even less precision exists for smaller subgroups of populations within census tracts (Wombold 2008a; Wombold 2008b). Thus, for smaller geographic units such as census tracts it may be reasonable to produce inset maps displaying the error associated with census tracts along with the map of the census variable.

Conclusion

Dot maps have proved a most effective approach for capturing detailed variations in population density. This same advantage applies when they are used to capture change data by means of dots of two different colors indicating a gain or a loss in the population. Though, the approach is relatively rare, recent examples of its use have been presented in East-West Gateway Coordinating Council (2002), Center for Urban and Regional Affairs (2003), and U.S. Department of Agriculture (2007). Dot maps showing change allow one to see in greater detail the size, location, and trend of changes occurring within a large population. They may also be an effective visualization tool for researchers analyzing population change within metropolitan areas since dot maps can suggest where errors in the data have occurred.

One of the challenges of mapping change is accounting for differences in the methods of data collection and reporting which have occurred between census dates. The cartographer needs to check census reports for modifications that might affect the data of interest and to be suspicious.
REFERENCES


196


