



## THE CHIEF JUSTICE EARL WARREN INSTITUTE ON LAW AND SOCIAL POLICY

UNIVERSITY OF CALIFORNIA, BERKELEY LAW SCHOOL

### BerkeleyLaw

UNIVERSITY OF CALIFORNIA

The Chief Justice  
Earl Warren  
Institute on Law  
and Social Policy

Berkeley Law Center for  
Research and Administration  
2850 Telegraph Avenue  
Suite 500  
Berkeley, CA 94705  
Phone: (510) 642-8568  
Fax: (510) 643-7095

[www.warreninstitute.org](http://www.warreninstitute.org)

#### About the Warren Institute

The Chief Justice Earl Warren Institute on Law and Social Policy is a multidisciplinary, collaborative venture to produce research, research-based policy prescriptions and curricular innovation on the most challenging civil rights, education, criminal justice, family and economic security, immigration and healthcare issues facing California and the Nation.

## REDISTRICTING: ESTIMATING CITIZEN VOTING AGE POPULATION

Jorge Chapa, PhD<sup>i</sup>, Ana Henderson, JD<sup>ii</sup>, Aggie Jooyoon Noah, MA<sup>iii</sup>,  
Werner Schink<sup>iv</sup>, Robert Kengle, JD<sup>v</sup>

*When determining how to draw electoral districts in a way that complies with the Voting Rights Act, many jurisdictions will need to consider proportions of citizen voting age population (CVAP). While the US Decennial Census captures basic demographic information about all individuals in the United States, it doesn't inquire into citizenship status. As such, line drawing officials will need to estimate CVAP through other means. This research brief explains Census Bureau data sources, including the American Community Survey, which estimates CVAP, but has some limitations for redistricting use. The brief then explains a method to use ACS and Census data in conjunction to develop more accurate CVAP estimates that are better for redistricting uses than ACS estimates alone.*

### THE VOTING RIGHTS ACT AND DATA REQUIREMENTS FOR REDISTRICTING

The Voting Rights Act (VRA) prohibits electoral schemes that discriminate against voters on the basis of race or protected language minority status.<sup>1</sup> This includes districts that

dilute voting strength, making it unlikely or impossible for minority voters to have an opportunity to elect a representative of their choice. Vote dilution can occur when a minority population is divided between districts so it cannot form a majority (called “cracking”), or when it is over-concentrated into an unnecessarily small number of districts when it could have been a majority in more (called “packing”). A key question about whether a population’s voting strength is diluted is whether districts pack or crack minority populations, that is, whether districts could be drawn in a way that provides a more fair opportunity for minority voters to elect a representative of their choice. This, in turn, depends on population concentrations within a potential district, among other things.

In 2009, a plurality of the U.S. Supreme Court found that in order to satisfy the first step of a vote dilution claim under the VRA—that the minority population is sufficiently large and compact to constitute a majority in a single member district (also known as the first *Gingles* pre-condition)—minorities must comprise at least 50% of the population of a district.<sup>2</sup>

i. University of Illinois, Institute of Government & Public Affairs, Departments of Sociology and Latina/o Studies

ii. University of California, Berkeley School of Law

iii. Department of Sociology & Crime, Law and Justice, The Pennsylvania State University

iv. President, Community and Local Neighborhood Research, Sacramento and Modesto, California

v. Co-Director, Voting Rights Project, Lawyers Committee for Civil Rights Under Law, Washington, DC

1. Language minority groups protected by the VRA are Asian, Alaska Native, Native American and Spanish heritage.

2. *Bartlett v. Strickland*, 129 S.Ct. 1231 (2009).

Some federal case law requires that a minority population must constitute 50% or more of the citizen voting age population of a potential district in order to qualify for Voting Rights Act protection.

While the Court in that case focused on total population, several cases have held that when determining whether a minority group satisfies the first *Gingles* precondition, the proper reference is potential voters—that is, voting age population<sup>3</sup> (VAP) rather than total population. Furthermore, some federal courts, including the 9<sup>th</sup> Circuit Court of Appeals where California is located as well as three other federal Courts of Appeal, have found that in order to determine whether a population constitutes at least 50% of a district, the proper population to consider is its citizen voting age population (CVAP).<sup>4</sup> This means that in California and other states where courts have so held, in order to show that a population is sufficiently large and compact to constitute a majority in a single member district, its citizen voting age population—citizens aged 18 or over—must constitute at least 50% of CVAP of the target district. A key question for VRA compliance, then, is the determination of citizen voting age population proportions.

## 1. Determining Citizen Voting Age Population from Census Bureau Data:

### A. Census Bureau Data Sources:

Every ten years, the United States must conduct a Census of the population to count all individuals living here. This Decennial Census forms the basis of reapportionment (redistribution of US House of Representative seats among the states based on population) and provides some basic data for redistricting and other governmental uses. The Census, executed by the Census Bureau, collects basic demographic information—place of residence, age, gender, race, Latino ethnicity, and household relationships—and is sometimes referred to as a 100% count because it tries to collect this information about every individual in the United States. Census data can produce counts of total population and voting age population, as well as measures of total and voting age population for various racial groups and Latinos. In the redistricting context, these population data are often referred to as “PL data” or “PL 94-171 data” in reference to the federal law that requires the Census Bureau to provide them to the states for use in redistricting. The Census does not inquire into citizenship status, but some other Census Bureau surveys do.

In Voting Rights Act litigation about electoral districts in which case law regarding citizenship measurement requirements developed,<sup>5</sup> CVAP figures were often estimated through analysis of sample data collected through the Decennial Census Long Form questionnaire. Several of the Censuses conducted in the twentieth century, including the 2000 Census, consisted of a “short form,” with the basic questions listed above, and a “Long Form,” which included the basic short form questions as well as additional questions on a variety of subjects, such as citizenship, ancestry, educational attainment, income, and even the kind of plumbing facilities (e.g., indoor or outdoor) household members used. The Long Form was intended

3. VAP refers to individuals of voting age, that is age 18 or over.

4. *Romero v. City of Pomona*, 883 F.2d 1418 (9<sup>th</sup> Cir. 1989), overruled on other grounds *Townsend v. Holman Consulting Corp.*, 914 F.2d 1136 (9<sup>th</sup> Cir. 1990); *Reyes v. City of Farmers Branch, TX*, 586 F.3d 1019 (5<sup>th</sup> Cir. 2009); *Barnett v. City of Chicago*, 141 F.3d 699 (7<sup>th</sup> Cir. 1998); *Negron v. City of Miami Beach, FL*, 113 F.3d 1563 (11 Cir. 1997). CVAP refers to individuals who are age 18 or over and are US citizens.

5. In previous redistricting cycles, jurisdictions generally did not have access to Census Bureau CVAP data when drawing lines for reapportionment purposes because citizenship data from the Census Long Form were released after district lines were drawn. The CVAP issue emerged in Voting Rights Act cases challenging at-large elections and/or districting plans alleged to dilute minority voting strength, brought after Long Form data, including CVAP, were released. For many jurisdictions, 2011 will be the first time lines are drawn with any Census Bureau CVAP data in hand.

to collect data from a sample of about one-sixth of the nation's households. Unlike the 100% count Census data, the sample-based Long Form data generated estimates of population characteristics and contained margins of error that were sometimes substantial, particularly at small units of geography such as Census tracts or block groups.<sup>6</sup>

As the result of an extensive redesign, the 2010 Census included only a 100% count short form and no Long Form questionnaire. A new annual Census Bureau survey—the American Community Survey (ACS)—has been instituted to replace the Census Long Form.<sup>7</sup> The ACS is a nationwide, continuous survey designed to provide up-to-date and reliable demographic, housing, social, and economic data every year.

The ACS is administered to 250,000 households each month for a total of three million a year. While estimates derived from this sample are very reliable for large geographic areas, such as states and most counties, they are less reliable for small areas where fewer responses are collected. The ACS sampling procedures are designed to produce reliable annual estimates of population characteristics for counties, cities, and other areas with populations of 65,000 or more. For towns and places with populations smaller than this threshold, the ACS sample is designed to permit several years of data to be pooled together (aggregated) to create multi-year estimates for smaller units of geography. The ACS sampling procedure was designed so that survey responses cumulated over three successive years would suffice to estimate the population characteristics of places with populations between 20,000 and 65,000. Similarly, data collected over five years can be aggregated to produce estimates for all census geographic levels down to block groups, the smallest geographic area for which ACS data are available and for which Long Form data were previously tabulated.<sup>8</sup>

...in the context of redistricting, ACS data have some weaknesses that must be addressed to obtain a more reliable and useful measure of CVAP.

The ACS has several advantages over the decennial Census Long Form method of measuring population characteristics. First, the ACS delivers updated data every year rather than every 10 years. Federal, state, and local governments rely on demographic, housing, social, and economic data in their budgeting and planning processes. Also, many governmental funds are annually distributed on the basis of these demographic and economic characteristics. ACS data allow the allocation of these funds based on recent data rather than data that could be up to 10 years old.

Moreover, ACS data is collected by a permanent, highly trained, and experienced staff. In contrast, most of the staff for the decennial Census is typically inexperienced part-time short-term employees with only a few days of training. ACS staff had a better response rate overall and collected more respondent-completed questionnaires than did the 2000 decennial Census staff.<sup>9</sup> However, as we will see, in the context of redistricting, ACS data have some weaknesses that must be addressed to obtain a more reliable and useful measure of CVAP.

6. Summary File 3: 2000 Census of Population and Housing—Technical Documentation—Chapter 8: Accuracy of the Data 8-18, 8-19 (issued: July 2007), available at <http://www.census.gov/prod/cen2000/doc/sf3.pdf#page=933>. Note that the Census Bureau did not publish margins of error along with Long Form estimates, but did provide information to allow independent development of such measures.

7. U.S. Census Bureau, *A Compass for Understanding and Using American Community Survey Data: What General Data Users Need to Know*, U.S. Government Printing Office, Washington, DC, p. 1 (2008).

8. *Id.* at 3. In addition to geography like states, counties, and places, the Census Bureau also releases data at smaller levels of geography that have no independent meaning. These include, in descending order of size, Census Tracts, Census Block Groups, and Census Blocks.

9. *Id.* at 8.

The ACS does not account for aging of the population sampled, but rather reports individuals at the age they were when data were collected. Accordingly, a citizen who was 13 in 2005 when the ACS collected information about her still appears as a 13-year-old today even though in reality she is now over 18 and of voting age.

#### **B. Census Data and ACS Data in the Context of Redistricting:**

The Census is designed to count and provide an enumeration of the entire population of the United States. It collects information about all individuals in the United States on April 1 of each year ending in 0, most recently on April 1, 2010. Although in reality it may not capture information about each individual and some populations, such as homeless individuals, are more likely to be undercounted than others, overall the Census provides a good count of the population. Census data, therefore, are not estimates based on a sample of the population, but actual counts. In addition, due to the large numbers involved in the Census, reliable data are released at all levels of geography down to the Census block level—the smallest geographical unit for which the Census Bureau reports data. Because of their high reliability and their availability at the smallest units of geography that are often used in constructing districts, Census (PL 94-171) data are the best data for redistricting.

In contrast, the ACS does not collect, nor is it intended to provide, a count or enumeration of the population. The ACS is intended to provide information about the distribution of characteristics in a population, such as educational attainment or number of rooms in one's home. Because it

is based on a sample of the population, rather than a total count, the ACS provides population estimates, which are subject to sampling error. Although the ACS reports data as an estimated number of individuals in a given area that display particular characteristics (often called a point estimate), this number is not the same as a number reporting Census-based counts. Rather, it is an estimate of the population, which is provided along with an estimate of margin of error (MOE). The point estimates and MOEs are provided at the 90% confidence level. This means that if the same questions were asked 100 times of a random sample of individuals in the same area, in 90 of those 100 times, the answer would fall within the range of the point estimate plus or minus the margin of error. For example, if the population in a county that does not have indoor plumbing is estimated to be 100 individuals, with a margin of error of  $\pm 23$ , this means that 90 times out of 100 that a random sample of people in the county is asked about plumbing, the population of outhouse users will be between 73 and 123 people. The ACS data reported about citizenship and age, from which CVAP numbers are derived, are an estimate of citizen population aged 18 or over in different geographic areas, each with its own margin of error.

In addition, in February 2011, the Census Bureau released a dataset containing CVAP figures for various racial groups and Latinos at the behest of the US Department of Justice; this dataset is commonly referred to as the DOJ Special Tabulation. The Census Bureau used ACS data to perform the DOJ Special Tabulation, which provides CVAP point estimates by race and Latino status at various levels of geography, down to the Census Block Group in many areas. As Table 1 demonstrates, the racial categories in the DOJ Special Tabulation do not match exactly with those in the publicly available ACS data. Notably, the DOJ Special Tabulation provides racial group estimates that do not include Latinos, while the ACS data provide non-Hispanic data only for whites; the DOJ data provide estimates for certain combinations of racial groups, while the ACS provides data only for individual racial groups alone or for “two or more races;” and the ACS provides estimates of individuals reporting “some other race,” while the DOJ Special Tabulation reassigned non-Latino “some other race” respondents to a racial background using a non-public methodology.

**TABLE 1 | Racial Categories Reported in ACS and DOJ Tabulations**

DOJ Racial Categories	ACS Racial Categories
Total	Total
Not Hispanic or Latino	Not Hispanic or Latino
American Indian or Alaska Native Alone (non-Latino)	American Indian or Alaska Native Alone (including Latinos)
Asian Alone (non-Latino)	Asian Alone (including Latinos)
Black or African American Alone (non-Latino)	Black or African American Alone (including Latinos)
Native Hawaiian or Other Pacific Islander Alone (non-Latino)	Native Hawaiian or Other Pacific Islander Alone (including Latinos)
White Alone (non-Latino)	White Alone (including Latinos)
	White Alone (non-Latino)
Not Applicable	Some Other Race (including Latinos)
American Indian or Alaska Native and White (non-Latino)	Two or More Races (including Latinos)
Asian and White (non-Latino)	
Black or African American and White (non-Latino)	
American Indian or Alaska Native and Black or African Am (non-Latino)	
Reminder of Two or More Race Responses (non-Latino)	
Hispanic or Latino	Hispanic or Latino

**C. Challenges in Using ACS Data for Redistricting: the Need for a Methodology to Improve Usability and Accuracy**

As noted above, several courts have required that districts have 50% or greater minority CVAP population in order to make a preliminary showing of VRA protection. Since the Census does not collect citizenship data, the ACS is a useful source of information. Unfortunately, the ACS’s design presents some challenges for redistricting that should be addressed to ensure that CVAP estimates are as accurate and reliable as possible

First, as Table 2 indicates, ACS data are not publicly available at the smallest units of geography that line drawers usually use in redistricting. Redistricting often entails putting small geographical units such as Census Tracts, Block Groups, or Blocks together to form districts. Census data are available and reliable at all these levels, including the Census Block, which is particularly important when attempting to draw districts populations as equal as possible.<sup>10</sup> However, ACS data are never reported at the Census Block level, and are reported at the Block Group and Tract level only by aggregating five years of ACS responses.

Second, ACS data are often less reliable for smaller geographical units than they are for larger units of geography. This is because the samples drawn from smaller units of geography are smaller than those of larger units, and as a general rule of thumb in statistics, small samples tend to generate larger margins of error than large samples. Even with aggregated data, block group estimates sometimes

**TABLE 2 | Smallest Geographical Unit with Data Reported**

	DOJ	ACS (5 year sample)
VAP	County	Tract
CVAP	Block Group	Tract

10. For example, US House of Representatives districts are generally drawn so that district populations within a state vary by only one person. Complying with this level of detail often requires the utilization of Census blocks to equalize populations.

contain large margins of error. Moreover, ACS and DOJ CVAP estimates for block groups come from smaller samples than CVAP block group estimates previously calculated using the Census Long Form data because the Long Form data sample equaled approximately 16% of a block group, compared to 11% for five-year ACS data.<sup>11</sup> Therefore the 2010 CVAP estimates at the block group level likely have larger MOEs than did the 2000 estimates.

Third, in some cases where the population samples are very small, the Census Bureau may suppress the data and not release them at all, in order to protect individual privacy interests. Data suppression can make Block Groups look like they have no population when they may in fact contain population. In some areas, the number of voting age citizens of various racial/ethnic groups may be smaller than the suppression threshold, leading to an underestimate of CVAP due to suppressed data.

Finally, since the ACS estimates available at smaller geographic units are based on data collected over a five-year period, they likely underestimate the CVAP population in some areas today (or on April 1, 2010 when Census data were collected) because the Census Bureau takes no steps to “update” old ACS data. Most notably, the ACS does not account for aging of the population sampled, but rather reports individuals at the age they were when data were collected. Accordingly, a citizen who was 13 in 2005 when

the ACS collected information about her still appears as a 13-year-old today even though in reality she is now over 18 and of voting age. Table 3 shows how five-year ACS data and DOJ Special Tabulation estimates of voting age population differ from the actual counts in the 2010 Census PL data in California.

The failure to address aging has a particularly strong impact on the CVAP rates of racial/ethnic groups that have higher rates of citizenship among children than adults, such as Latinos and Asian Americans. Table 4 shows CVAP estimates from the 2000 Long Form, the DOJ Special Tabulation, and one-year ACS estimates collected each year between 2005 and 2009 for the United States and California. As noted above, the ACS is designed so that one-year estimates are valid and reliable for areas with population greater than 65,000, so the very large samples for the nation and the entire state of California should render highly reliable results. The one-year ACS estimates show growth in CVAP each year. Analysis of the one-year ACS trends as well as comparison with the DOJ five-year estimates reveals some troubling issues for redistricting purposes.

First, the DOJ Special Tabulation CVAP estimates are lower than the most recent (2009) one-year ACS estimates. For example, the 2009 CVAP estimate for California is 3% (660,935 people) greater than the DOJ Special Tabulation CVAP estimate. The disparity is even more pronounced for Latino and Asian American populations. In California, the DOJ Special Tabulation estimate for Latino CVAP is 8.6% less than that of the 2009 ACS, and the Asian CVAP estimate is 7% less.<sup>12</sup> Relying on five-year aggregated data alone for redistricting would underestimate Latino and Asian American citizen voting age populations.

Second, examination of the annual ACS CVAP estimate percentages for different groups in Table 3 shows that the racial composition of total citizen voting age population is changing. Most striking, annual ACS estimates indicate that the Latino proportion of citizen voting age population is increasing, while the non-Hispanic white

**TABLE 3 | Select California Voting Age Population Reports from Three Census Datasets**

	ACS (2005-2009) Estimate	DOJ Special Tabulation Estimate	2010 Census Count
Latino	8,490,040 (31.6%)	8,490,040 (31.6%)	9,257,499 (33.1%)
Asian Alone	3,516,607 (13.1%)	3,503,460 (13%)	3,809,082 (13.6%)
Non- Hispanic White	12,577,310 (46.8%)	12,646,350 (47.1%)	12,409,858 (44.4%)

**Note:** The 2010 and DOJ Asian figures do not include Hispanic Asian Americans, while the ACS figures do. Also, including the Asian + white category would add 117,210 to the DOJ Asian category

11. National Research Council, *Using the American Community Survey: Benefits and Challenges*. The National Academies Press, Washington, DC, at Table 2-3a (2007).

12. Note that the DOJ Special Tabulation estimate for Native Americans in California is 47.4% less than the 2009 ACS estimate. Interestingly, the DOJ Special Tabulation estimate is also significantly less than the 2000 Census count for this population.

**TABLE 4 | CVAP Estimates by Race and Latino Status for the United States and California**

CITIZEN VOTING AGE POPULATION (CVAP): UNITED STATES						
	Total	Latino	Asian American	Non-Hispanic White	Black	American Indian/ Alaska Native
2000 Census Long Form	193,376,975	14,300,581	4,696,448	147,768,945	22,614,559	1,557,130
	100%	7.4%	2.4%	76.4%	11.7%	0.8%
DOJ Special Tabulation	208,196,325	18,512,565	6,730,455	153,938,560	24,962,205	1,451,430
	100%	8.9%	3.2%	73.9%	12.0%	0.7%
2005 ACS	197,004,322	16,896,498	6,309,701	147,254,400	22,892,954	1,603,981
	100%	8.6%	3.2%	74.7%	11.6%	0.8%
2006 ACS	206,287,902	17,891,634	6,701,987	153,015,764	24,757,732	1,607,600
	100%	8.7%	3.2%	74.2%	12.0%	0.8%
2007 ACS	208,186,178	18,427,267	6,828,445	153,734,679	25,132,367	1,622,077
	100%	8.9%	3.3%	73.8%	12.1%	0.8%
2008 ACS	210,710,906	19,401,013	6,970,157	154,598,294	25,633,395	1,644,200
	100%	9.2%	3.3%	73.4%	12.2%	0.8%
2009 ACS	213,020,665	20,101,592	7,202,612	155,322,801	26,139,525	1,643,644
	100%	9.4%	3.4%	72.9%	12.3%	0.8%

CITIZEN VOTING AGE POPULATION (CVAP): CALIFORNIA						
	Total	Latino	Asian American	Non-Hispanic White	Black	American Indian/ Alaska Native
2000 Census Long Form	20,011,574	3,888,220	1,850,180	12,085,427	1,495,075	187,604
	100%	19.4%	9.2%	60.4%	7.5%	0.9%
DOJ Special Tabulation	21,942,930	5,117,250	2,515,395	12,164,955	1,563,045	130,400
	100%	23.3%	11.5%	55.4%	7.1%	0.6%
2005 ACS	20,736,632	4,675,574	2,391,672	11,698,453	1,437,534	174,480
	100%	22.5%	11.5%	56.4%	6.9%	0.8%
2006 ACS	21,846,683	4,989,806	2,509,868	12,183,038	1,561,271	185,918
	100%	22.8%	11.5%	55.8%	7.1%	0.9%
2007 ACS	22,042,254	5,089,309	2,554,657	12,202,749	1,589,309	189,253
	100%	23.1%	11.6%	55.4%	7.2%	0.9%
2008 ACS	22,417,194	5,396,279	2,617,919	12,185,021	1,608,369	207,541
	100%	24.1%	11.7%	54.4%	7.2%	0.9%
2009 ACS	22,603,863	5,558,160	2,692,708	12,107,216	1,607,039	192,170
	100%	24.6%	11.9%	53.6%	7.1%	0.9%

ACS estimates, particularly five-year aggregated data at the block group or tract levels, are best considered a conservative estimate of citizenship rates and CVAP. The actual rates are likely higher...

share is decreasing, nationally and to a greater degree in California. Accordingly, the five-year aggregated data used for the DOJ Special Tabulation overestimate the percentage of non-Hispanic whites and under-estimate the percentage of other groups, especially Latinos, citizen voting age population.

Differences between five-year aggregated data and the most recent 1-year ACS data are explained in large part by the failure to account for age changes, noted above. The failure to “age up” teenage citizens results in five-year aggregated ACS data underestimating Latino and Asian CVAP, in particular.<sup>13</sup> Accordingly, ACS estimates, particularly five-year aggregated data at the block group or tract levels, are best considered a conservative estimate of citizenship rates and CVAP. The actual rates are likely higher due to the natural aging of the population, the higher rate of citizenship among Latino and Asian children as opposed to adults, and the low mortality rate among teenagers versus older individuals.

The DOJ Special Tabulation has two additional challenges to note. First, as noted above, the Census Bureau reattributed the racial identity of non-Latino individuals who reported their race as “Some Other Race.” While

most “Some Other Race” respondents are Latino<sup>14</sup> and therefore remained allocated in the Latino category in the DOJ Special Tabulation, a small number of “Some Other Race” respondents who are not Latino have been attributed to other racial groups, e.g., non-Latino white, non-Latino black, etc. Unfortunately, this process cannot be reproduced because the Census Bureau used non-public information to perform the attribution. Second, due to privacy concerns, estimates of voting age population (VAP) at geographic units smaller than the county level, i.e., Census tracts and block groups, were suppressed. Therefore, there is no way to compare VAP and CVAP at the Census Block Group using DOJ Special Tabulation data; the smallest geographical unit for which a CVAP to VAP comparison is possible is the County level, as seen in Table 2.

That said, the DOJ Special Tabulation may be a better source of redistricting data in the VRA context than ACS data due to the racial information it provides. Aggregated ACS data are not readily available in some of the racial groupings of interest in some VRA-related redistricting issues. Federal directives call for the grouping of racial categories in particular ways in civil rights enforcement contexts, so in redistricting dealing with Asian Americans,<sup>15</sup> for example, the proper population would include both individuals reporting Asian race alone as well as those reporting Asian and white background. The DOJ Special Tabulation provides CVAP estimates for both “Asian alone” (non-Hispanics reporting Asian racial background alone) as well as “Asian and white” (non-Hispanics reporting Asian and white racial background) categories that can be combined to derive an Asian American estimate, but the ACS aggregated estimates do not.<sup>16</sup>

Finally, ACS data and the DOJ Special Tabulation of those data pose a challenge for redistricting because they are reported using different geographical units than 2010 Census (PL 94-171) data. The Census geography used to report data (Tracts, Block Groups, Blocks) can change between Censuses. As a result, data from the same physical

13. Nevertheless, these data will likely have to be used because although the 2009 one-year ACS data provide a more up-to-date and accurate estimate of CVAP, they are not reported at units of geography that are useful in most redistricting.

14. For example, according to 2010 Census PL data for California, only 85,587 of California's 6,317,372 “some other race” respondents were not Latino. See PL 94-171 Summary File.

15. OMB BULLETIN NO. 00-02 - Guidance on Aggregation and Allocation of Data on Race for Use in Civil Rights Monitoring and Enforcement, available at [http://www.whitehouse.gov/omb/bulletins\\_b00-02](http://www.whitehouse.gov/omb/bulletins_b00-02)

16. See Table 1



address could be reported in one Census Tract in 2000 and a different Tract in 2010. Unfortunately, ACS data are reported using 2000 Census Geography, while Census (PL 94-171) data are reported using 2010 Census geography. For redistricting in 2011, line drawers will use 2010 Census geography, rendering ACS and DOJ data reported on 2000 Census geography inaccurate when geographies conflict. Where the 2000 and 2010 Census geographies are different, additional steps are needed to bring the ACS data into 2010 Census geography.

## 2. Method to Address ACS-based CVAP Estimate Challenges:

Because of the issues identified above, ACS data reports, as well as tabulations based on them such as the DOJ Special Tabulation, should undergo additional analysis in order to be more useful for redistricting. What follows is a method to use publicly available Census data and ACS data in conjunction to develop a reliable, albeit conservative, estimate of CVAP for various population groups at the smallest level of geography for which the data are available.

Since several aspects of a population can change in the period over which ACS data were collected (2005-2009) and the time the Census enumeration occurs in 2010, we argue that the best use of ACS data is to establish a ratio or rate of citizenship. This citizenship ratio is applied to 2010 Census data to produce estimated CVAP populations using the most up-to-date data that is reliable at small geographical units. This method produces a conservative estimate of CVAP populations since the ratio does not directly correct for the age issues mentioned above that affect the accuracy of CVAP data. Accordingly, a determination of 50% CVAP estimated under this method should be sufficient to satisfy the first *Gingles* precondition that a minority group is sufficiently large and compact to constitute a majority in a single member district.<sup>17</sup>

ACS data reports, as well as tabulations based on them such as the DOJ Special Tabulation, should undergo additional analysis in order to be more useful for redistricting.

This method has several benefits as well. First, using ACS data to develop a fraction or ratio of citizenship is a more appropriate use of ACS data than using its point estimates as if they were absolute counts. ACS data are best used as descriptors of the population rather than a tally or count. Using these data to determine rates of voting age citizenship for various racial/ethnic groups is a proper use of ACS data. Other options, such as using 1-year ACS data to make group-specific corrections to five-year ACS data, while demographically sound, are not a use for ACS data of which the Census Bureau approves.

Second, by applying the CVAP ratio or fraction to 2010 Census data, we address some of the shortcomings of ACS aggregated data and produce a more reliable point estimate. The citizenship rate established with ACS data is applied to 2010 Census voting age population figures to derive an estimate of 2010 citizen voting age population. This provides a partial update of aggregated ACS data by using the most recent and complete VAP data from the 2010 Census.

17. In fact, due to the underestimate of Latino and Asian American communities in particular, a five-year aggregated ACS-based CVAP estimate of less than 50% could very well still represent an actual CVAP population over 50%. When considering districts dealing with these populations, line drawers and courts should carefully consider where to place the threshold to avoid disenfranchising Latino and Asian American voters due to data problems.

Although the Census is a 100% count of the population, providing an enumeration of certain population characteristics down to the smallest units of Census geography, the American Community Survey, which reports citizenship data, provides only estimates of population characteristics based on population samples and is not available and/or accurate for all groups, particularly at small units of geography.

Finally, this method does not produce any systematic bias that might advantage or disadvantage a particular group.<sup>18</sup> The method uses Census Bureau data for all groups in the same manner. Unlike a group-specific “correction” to ACS data, our method should be used to produce 2010 CVAP estimates for all racial groups. Since any bias in the method would therefore affect all groups, this method does not systematically advantage or disadvantage any one group.

In what follows, we outline two different methods for estimating CVAP through a combination of ACS and Census PL data. The first applies in areas with reasonably homogenous CVAP rates within racial groups and/or when an entire county is contained within a district. This method uses county-level data. The second is a method to develop more fine-grained analysis of CVAP, where CVAP rates are not homogenous within racial groups<sup>19</sup> and/or when a jurisdiction wants a smaller unit of analysis, such as a city drawing districts. For both methods, we recommend analyzing the CVAP proportion of each racial/ethnic group present in a potential district. That is, CVAP

estimates should be generated for all racial/ethnic groups at issue and not just Latinos or Asian Americans.<sup>20</sup>

#### A. County-based Method

Although preliminary analyses conducted for this paper showed variance in Census tract CVAP rates for all racial groups, in many counties these rates may be relatively homogeneous. In addition, if an entire county is included within the same district, sub-county variations in citizenship rates will not matter. In these cases, jurisdictions can use DOJ Special Tabulation data to develop a county-wide CVAP rate for each racial group that can then be applied to Census PL 94-171 population counts to calculate 2010 CVAP estimates for each block. Since there were very few changes in county boundaries between 2000 and 2010, county-level CVAP ratios can be applied to every block reported for that county in 2010 PL tabulations.

This approach is much easier than the tract-level analysis described below. Also, by using county data to calculate the CVAP fraction, we minimize the margin of error issues that arise for smaller units of geography, since ACS

18. In fact, due to the underlying underestimate of Latino and Asian American CVAP in ACS data, this method, although an improvement, likely is still biased against these groups.

19. We have analyzed CVAP rates for all racial groups in many different parts of the U.S. and found that these rates can vary greatly, even within racial groups. For example, analyses of a national sample of ACS data not presented in this paper show that almost all Latinos who identify as Puerto Rican are U.S. citizens compared to 60% of those reporting Mexican origin. There are also significant variations between Asian national origin groups. For example, about 55% of

the Asian Indians are citizens compared to 79% of Vietnamese. We can expect similar differences between native-born African Americans and African immigrants as well as non-Hispanic whites and white immigrants

20. The state of Texas recently took the opposite approach, calculating a CVAP estimate only for Latinos while assuming that citizenship rates were uniform for all other groups, e.g., blacks, Asians, non-Hispanic whites, etc. See email from David R. Hanna, Senior Legislative Counsel, Texas Legislative Council sent on April 13, 2011 (on file with authors).

In addition to the Census-based method discussed above, CVAP is sometimes estimated from voter registration rolls...we strongly caution that voter registration and turn out methods include drawbacks that may be very difficult or impossible to address.

and DOJ CVAP estimates are more reliable at the county level due to the larger samples sizes. Finally, the county-level approach allows the utilization of the DOJ Special Tabulation data set. If a jurisdiction must consider the representation of racial groups that are available from the DOJ Special Tabulation but not the ACS, such as “African American and white” or “Asian and white,” this method produces block-level CVAP estimates for these groups. When these considerations apply, we propose the following steps:

**1. Produce a County CVAP Ratio for Each Racial Group at Issue:**

First, use data about citizenship and age to determine a ratio or fraction of CVAP for each racial/ethnic group. This means dividing the CVAP for a particular group by the VAP for that group. Here we use DOJ Special Tabulation data at the county level<sup>21</sup>:

$$\text{DOJ CVAP} / \text{DOJ VAP} = \text{CVAP Ratio}$$

**2. Produce County 2010 CVAP Numeric Estimate:**

After confirming that 2000 and 2010 county boundaries are closely comparable, apply the County CVAP Ratio for each group to its voting age population in the 2010 Census PL 94-174 data set (hereinafter “2010 PL”) to produce a county level CVAP numeric estimate.

$$\text{DOJ County CVAP Ratio} * \text{2010 PL VAP County} = \text{2010 CVAP County Numeric Estimate}$$

**3. Calculate Census Block 2010 CVAP Numeric Estimate:**

In addition, the County CVAP Ratio can be applied directly to 2010 Census geography within the County, most notably Census Blocks, that are used in redistricting:

$$\text{CVAP Ratio} * \text{2010 PL Block VAP} = \text{2010 Block CVAP estimates}$$

To verify results, these 2010 Block CVAP numeric estimates can be added up and compared to the 2010 CVAP County Numeric Estimate.

We conducted this method to produce CVAP estimates for Alameda County, CA; results are set forth in Table 5.

**B. Areas Requiring More Fine-Grained Analysis**

In some areas, jurisdictions may want or need to develop CVAP estimates at units smaller than the county level. This includes areas where variations in citizenship rates within racial groups exist. For example, the citizenship rate for non-Hispanic whites may be very high in one neighborhood, but lower in a neighborhood with a large proportion of Russian immigrants. Where such variations exist, applying a citizenship rate for an entire county may produce misleading results, artificially high in some areas while artificially low in others. Accordingly, a jurisdiction may want to use ACS data<sup>22</sup> to produce a more fine-grained analysis than is possible using DOJ Special Tabulation data.

21. We suggest using the DOJ Special Tab because the Census Bureau developed it specifically to comply with OMB directives regarding calculation of racial groups laid out in the OMB guidance and in response to a request from the

US Department of Justice. It is likely that the DOJ Special Tab data will be what the DOJ will use. However, the method could also be run using ACS data, if a jurisdiction prefers.

**TABLE 5 | CVAP Estimates using County-based Method for Alameda County, CA**

Racial Groups	DOJ VAP	DOJ CVAP	2010 PL VAP	CVAP/VAPRatio	2010 CVAP
<b>Total</b>	<b>1,115,765</b>	<b>919,100</b>	<b>1,169,650</b>	<b>0.8237</b>	<b>963,487</b>
Not Hispanic or Latino	905,195	796,415	938,477	0.8798	825,697
American Indian or Alaska Native Alone (non-Latino)	3,665	3,500	3,341	0.955	3,191
Asian Alone (non-Latino)	277,700	196,455	307,657	0.7074	217,648
Black or African American Alone (non-Latino)	141,105	136,225	143,194	0.9654	138,242
Native Hawaiian or Other Pacific Islander Alone (non-Latino)	7,790	5,995	8,887	0.7696	6,839
White Alone (non-Latino)	452,385	432,465	436,886	0.956	417,648
American Indian or Alaska Native and White (non-Latino)	3,920	3,920	3,482	1	3,482
Asian and White (non-Latino)	7,145	6,890	13,850	0.9643	13,356
Black or African American and White (non-Latino)	3,135	3,090	4,195	0.9856	4,135
American Indian or Alaska Native and Black or African American (non-Latino)	2,420	2,420	1,767	1	1,767
Remainder of Two or More Race Responses (non-Latino)	5,935	5,450	12,342	0.9183	11,333
Hispanic or Latino	210,570	122,685	231,173	0.5826	134,689

Although the DOJ Special Tabulation contains CVAP data for the racial groups often at play in VRA-related districting,<sup>23</sup> it does not provide VAP data at geography smaller the county level. In contrast, five-year aggregated ACS estimates are available at the tract level for both CVAP and VAP, permitting the production of a tract-level CVAP ratio. Since ACS tract geography may differ from 2010 Census tract geography, this method includes a step to analyze geography and if necessary a technique to address changes.

**1. Produce Census Tract-level CVAP Ratios for Each Racial Group at Issue:**

As above, divide each population’s CVAP by its VAP. Here we use five-year aggregated ACS data at the Census tract level.<sup>24</sup>

$$\text{ACS Tract CVAP} / \text{ACS Tract VAP} = \text{ACS Tract-level CVAP Ratio}$$

22. Note that the ACS provides non-Hispanic estimates only for whites, so estimates of other groups (African Americans, Asians, Native Americans, etc.) include Latinos who are also African American, Asian, Native American. In most jurisdictions these numbers will be very small, but should be kept in mind for redistricting.

23. The racial grouping issue is most pronounced when dealing with VRA issues concerning Asian American and African American populations. In both of these cases, a complete tabulation, in accordance with federal guidelines, should include both citizens of that minority racial background alone as well as those with both minority and white backgrounds. However, when dealing with Latinos and non-Latino whites, the ACS and DOJ Special Tabulation produce the same estimates for the same groups (Latinos and non-Latino white alone). See Table 1

24. Tract VAP and/or CVAP data were not published (suppressed) by the Census Bureau to prevent the possible disclosure of personal information. The populations of these tracts do have a CVAP rate but we do not know what it is. We suggest using the average CVAP rate for the appropriate group over the entire area being analyzed. For example, for the analysis in Table 6, we used the race-specific county-wide average CVAP ratio for the tracts where data were suppressed. In other analyses using this method for entire states, we have used the state-wide average CVAP rate.

**TABLE 6 | CVAP Estimates using Tract-based method for Alameda County, CA**

Racial Group	ACS VAP	ACS CVAP	2010 PL VAP	2010 CVAP
<b>Total</b>	<b>1,115,765</b>	<b>919,100</b>	<b>1,169,650</b>	<b>963,055</b>
Not Hispanic or Latino	905,195	796,415	938,477	825,627
American Indian or Alaska Native Alone	5,819	4,972	3,341*	3,104
Asian Alone	277,987	197,048	307,657*	218,028
Black or African American Alone	141,634	136,773	143,194*	138,412
Native Hawaiian or Other Pacific Islander Alone	7,985	6,220	8,887*	6,970
White Alone (non-Latino)	449,213	430,058	436,886	418,194
Hispanic or Latino	210,571	122,686	231,173	137,428

\*Does not include Latinos

**2. Determine Correspondence Between ACS Tracts and 2010 Census Blocks.**

Since ACS data are reported using 2000 Census geography while Census PL data, as well as this decade’s redistricting efforts, use 2010 Census geography, the next step is to determine whether there have been any changes in the area included in each Census tract. The correspondence between ACS Tracts and Census PL blocks can be performed using GIS, analysis of Census geographic correspondence files, visual map inspection, etc. (We will detail these procedures in a forthcoming paper.)

**3. Associate each 2010 block or block portion with a 2000 Census Tract.**

Many 2010 Census blocks can be identified with one 2000 Census tract, so this step entails correctly associating each ACS 2000 Census Tract with the geographic area that is now a 2010 block.

However, some 2010 blocks will be split between 2000 Census tracts and therefore not match perfectly. In those cases, use the proportion of the 2010 block area within the 2000 tract to assign the same proportion of the block’s 2010 PL VAP to the tract for each racial group.

**4. Apply the ACS Tract CVAP Ratio for each Racial Group to its 2010 PL Block VAP**

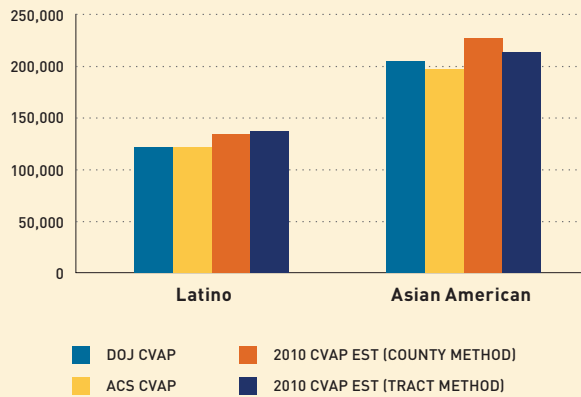
$$\text{Tract CVAP Ratio} * \text{2010 PL Block VAP} = \text{2010 Block CVAP estimates.}$$

This should be done for each racial/ethnic group at issue in the district. For 2010 blocks that are split between 2000 tracts, multiply the portion of the block located within a given Tract by the Tract CVAP Ratio to calculate the 2010 CVAP estimate for each block portion.

We have conducted full-scale trials of this method on Alameda County, California and found that the sum of the block-level estimates differed from the published county totals by only a very small amount, attributable to very small rounding errors.

Tables 6 sets forth our analysis using this method for Alameda County, California. Figure A provides a comparison of the Tract-based method for Latino and Asian American populations with the ACS and DOJ estimates for these populations in Alameda County, California.

**FIGURE A | Latino & Asian CVAP Estimates for Alameda County, CA**



	Latino	Asian American
DOJ CVAP	122,685	203,345
ACS CVAP	122,686	197,048
2010 CVAP EST (County Method)	134,689	231,004
2010 CVAP EST (Tract method)	137,428	218,028

**Note:** County-based 2010 CVAP estimates for Asian Americans include individuals who identify as “Asian” and “Asian and White.”

### C. Alternate Methods of Estimating CVAP: Voter Rolls Approach

In addition to the Census-based method discussed above, CVAP is sometimes estimated from voter registration rolls. While we present information on this method, we strongly caution that voter registration and turn out methods include drawbacks that may be very difficult or impossible to address. In addition, in many courts, Census Bureau data enjoy a presumption of validity.<sup>25</sup>

Voter registration rolls provide information that may help estimate CVAP for certain populations. Voter rolls indicate the name, address, date of birth, and in a few states the race, of each individual registered to vote in a jurisdiction. Since only citizens may vote in federal and most state/local elections and individuals must be 18 or older to register, individuals on the rolls are assumed to be citizens of voting age. By using racial data on the rolls in states that collect this information, one can generate a racial count of registered voters, which can be used as a proxy to estimate CVAP for each racial group.

In states that do not collect racial information on voter registration, like California, voters’ surnames can be analyzed for membership in certain minority groups

in order to estimate a racial distribution of CVAP. The process of surname analysis compares voters’ surnames with a database of surnames associated with certain ethnic groups, most prominently Latino and Asian American groups. This surname-matched registration data may be used as a proxy to estimate Latino or Asian American CVAP. While this method will produce some false positives, e.g., a non-Asian citizen with a traditionally Asian last name, and some false negatives, e.g., a Latino citizen lacking a traditionally Latino surname, surname matched registration data can provide an approximation of Latino and Asian citizens who are registered to vote.

However, voter registration data has limitations in estimating actual citizen voting age population in an area. First, surname analysis cannot produce a reliable estimate of white or black populations due to the lack of surname databases for these groups. Second, registration-based methods will underestimate CVAP due to high rates of under-registration among citizens of color, particularly Latino and Asian American citizens. According to another Census Bureau data source—the Current Population Survey (CPS)—Asian American and Latino citizens have

25. See, e.g., *Valdespino v. Alamo Heights Independent School Dist.*, 168 F.3d 848, 853-4 (5th Cir. 1999).

...due to quirks in the ACS data, such as reporting individuals at the age they were when data were collected rather than what their current age actually is, our method produces a conservative estimate of actual CVAP, particularly for Asian American and Latino populations because it does not directly or completely correct for underlying problems with CVAP estimates in the ACS data.

the lowest rates of voter registration of all groups. The November 2008 CPS indicated that while 74% of non-Hispanic white citizens and 70% of non-Hispanic black citizens reported being registered to vote, only 55% of Asian and 56% of Latino citizens reported being registered.<sup>26</sup> In fact, courts have rejected arguments that the VRA requires majority-minority voter registration in order to warrant a district, recognizing that for a variety of reasons, including discriminatory voting practices, minority voter registration is often suppressed and therefore not an accurate reflection of minority voter potential.

Voter-based methods of estimating CVAP are also logistically difficult in several states. Although in California, surname matched registration and turn out data are publicly available free of charge on the internet,<sup>27</sup> in many states, interested parties must request and pay for voter rolls and then run the surname analysis themselves.

Another method some propose is to analyze voter turnout rolls in a similar way. However, as with registration analysis, a population of color does not have to demonstrate that they constituted a majority of actual voters in a particular area in order to state a VRA claim. Moreover, reliance on voter turnout would underestimate minority CVAP even more than reliance on voter registration, since

citizens of color participate in the electoral process less often than their white counterparts.<sup>28</sup>

Voter-roll based methods of estimating CVAP must take into consideration the fact that voters of color often register and participate at rates lower than their actual presence in the population. This registration and participation gap is most pronounced among Latinos and Asian Americans, precisely the groups that are generally the focus of citizenship inquiries. Thus, any registration, or certainly turn out, analysis should be carefully balanced by an inquiry into local registration and participation rates, so that this method does not significantly underestimate the Latino or Asian CVAP population. That said, data about under-registration and participation are difficult to obtain, particularly at small geographical areas. CPS data are not provided at levels of geography smaller than the state, and racial data is only reported at the national level. Accordingly, a finding of 50% Latino or Asian registration or turn out, given the lower rates of turnout, surely indicates that a given district satisfies the first *Gingles* precondition and any requirement that the population be “effective,” because the actual CVAP in the district is most likely significantly higher than that indicated by voter registration or participation rates.

26. Thom File and Sarah Crissey, “Voting and Registration in the Election of November 2008,” US Census Bureau, Population Characteristics. (May 2010)

27. California’s Statewide Database contains all of California’s redistricting data, including precinct-level statistics on Asian Americans, Latinos, and others derived from surname-matched voter registration rolls. See <http://swdb.berkeley.edu/>

28. U.S. Census Bureau, Statistical Abstract of the United States: 2011, p. 259, Table 415, available at <http://www.census.gov/compendia/statab/2011/tables/11s0415.pdf>.

## CONCLUSION:

Some federal case law requires that a minority population constitute 50% or more of the citizen voting age population of a potential district in order to qualify for Voting Rights Act protection. Although the Census is a 100% count of the population, providing an enumeration of certain population characteristics down to the smallest units of Census geography, the American Community Survey, which reports citizenship data, provides only estimates of population characteristics based on population samples and is not available and/or accurate for all groups, particularly at small units of geography. This brief presents a method to use ACS data about population characteristics to inform 2010 Census count data. ACS data, including the DOJ Special Tabulation, can be used to produce a fraction or ratio of CVAP for various racial/ethnic groups. This CVAP fraction can be applied to 2010 Census voting age population data, and with the help of mapping software, be disaggregated to small units of geography while maintaining the reliability of larger level ACS estimates. This is a proper use of ACS data and provides the most accurate and reliable estimates of CVAP possible without complicated statistics to try to “correct” ACS data directly.

That said, due to quirks in the ACS data, such as reporting individuals at the age they were when data were collected rather than what their current age actually is, our method produces a conservative estimate of actual CVAP, particularly for Asian American and Latino populations because it does not directly or completely correct for underlying problems with CVAP estimates in the ACS data. Accordingly, a determination that a target district contains a minority CVAP population of at least 50% should satisfy not only the first *Gingles* precondition’s requirement that a minority group be large and compact enough to constitute a majority in a single member district, but also any concerns about whether the population is large enough to be “effective” in the district.

---

*This brief was made possible by a grant from the James Irvine Foundation. The conclusions in this brief are those of the authors.*