Self-Identification of Race or Ethnicity in the Official California Voter File

Most U.S. states do not require residents to self-identify their race or ethnicity when registering to vote, creating challenges for researchers and policy makers to identify voting behaviors among racial and ethnic groups. In this fact sheet, we outline the limitations of the race or ethnicity data present in the official California Voter File as administered by the California Secretary of State's office.

Self-Identification and Voter Turnout Rates

Self-identification in the registration process is a common method of gathering the race or ethnicity of voters. However, since self-identifying race or ethnicity is optional, an analysis of voter files relying solely on voters who choose to answer the question becomes highly problematic. Self-reported race or ethnicity rates vary across states, with some having response rates over 90% (Georgia) and others having response rates below 30%, including California. In states with low response rates, relying solely on the sample of voters who self-identified their race or ethnicity on registration forms opens the possibility of misrepresenting the voting behaviors of voters.

In an analysis of the California voter file (VoteCal) as of March 2021, CID found that 29.5% of registered voters in California identified their race or ethnicity when registering to vote (Table 1).

Table 1. VoteCal Data: Registered Voters Who Identified Race/Ethnicity in California

	Registered Voters	Count: Self-Reported Race or Ethnicity		Percentage: Self-Reported Race or Ethnicity	
	Total	Yes	No	Yes	No
Total	22,157,164	6,537,198	15,619,966	29.50%	70.50%

Data Source: California Secretary of State Voter File

CID also found that registered voters who self-identified their race or ethnicity had higher turnout rates (84.2%) than those who did not self-identify their race or ethnicity (76.6%) (Table 2).

Table 2. VoteCal Data: Registered Voter Turnout (RVT) – By Self-Reported Race or Ethnicity Status

	Voted	Registered	RVT
Ethnicity Self-Identified	5,504,553	6,537,198	84.2%
Ethnicity Missing	11,958,061	15,619,966	76.6%
All Voters	17,462,614	22,157,164	78.8%

Data Source: California Secretary of State Voter File

Additionally, the relatively small sample of voters who did self-identify their race or ethnicity skewed young and Democratic. Registered voters aged 18 to 24, for example, represented a larger of share of registered voters who answered the race or ethnicity question (18.6% of all registered voters who self-identified their race or ethnicity) than their share of all registered voters (11.9%) (Table 3). Registered voters aged 65 and over, on the other hand, were only 10.7% of registered voters who answered the race or ethnicity question, while representing over 20% of the total number of registered voters.

Table 3. VoteCal Data: Self-Reported Race or Ethnicity in California by Age Group

	Registered Voters (count)	Self-Reporting Race or Ethnicity (count)	Self-Reporting Race or Ethnicity (percent)	Share of Self-Reports (percent)	Share of All Registered Voters (percent)
18-24	2,637,646	1,213,261	46.0%	18.6%	11.90%
25-34	4,134,137	1,691,646	40.9%	25.9%	18.00%
35-44	3,524,798	1,176,377	33.4%	18.0%	15.90%
45-54	3,389,452	898,158	26.5%	13.7%	15.30%
55-64	3,634,942	751,943	20.7%	11.5%	16.40%
65+	4,595,586	698,609	15.2%	10.7%	20.70%
Age Unknown	240,603	107,204	44.6%	1.6%	1.10%

Data Source: California Secretary of State Voter File

Registered Democrats represented 48.0% of voters who marked their race or ethnicity but were only 46.0% of all registered voters (Table 4). Relying on the sample of voters who choose to self-report their race or ethnicity limits analysis to an unrepresentative turnout and potentially skews findings about voting behaviors.

Table 4. VoteCal Data Self-Reported Race or Ethnicity: Percent Reporting in California by Party Affiliation

	Registered Voters (count)	Self-Reporting Race or Ethnicity (count)	Self-Reporting Race or Ethnicity (percent)	Share of Self-Reports (percent)	Share of All Registered Voters (percent)
Democratic	10,184,219	3,140,876	30.8%	48.0%	46.0%
Republican	5,370,890	1,409,131	26.2%	21.6%	24.2%
No Party Preference	5,322,221	1,616,867	30.4%	24.7%	24.0%
Other	1,279,834	370,324	28.9%	5.7%	5.8%

Data Source: California Secretary of State Voter File

Self-Identification by Registration Method

There are many methods to register or re-register to vote in California and the ways in which registered voters can input their race or ethnicity varies by registration method. As shown in Table 5, the rate of self-reporting race or ethnicity varies greatly depending on method of registration. The most used methods to register to vote are online and at the Department of Motor Vehicles (DMV). Online registration (through the California Secretary of State's website) has a drop down menu for the individual to choose from, while paper forms have an open ended question which allows individuals to write in their race or ethnicity. The DMV does not ask registrants to self-report their race or ethnicity when registering to vote. Online voter registration has a notably high rate of self-reporting race or ethnicity by the registrant. Of those who are currently registered through the state's online system, nearly 79% self-reported their race or ethnicity, a much higher percentage than another other registration method (Table 5).

CENTER FOR INCLUSIVE DEMOCRACY

Note: The DMV does not ask registrants to self report their race or ethnicity when registering to vote. Self-reported race or ethnicity information is present in the voter file for some DMV registered voters who reported this information when previously registering to vote by a different method. Self-reported race and ethnicity information present in the voter file for some current online and paper form registrants may also have been submitted during a previous registration with a different method.

Table 5. VoteCal Data: Self-Reported Race or Ethnicity by Registration Source

	Registered Voters	Count: Self-Reported Race or Ethnicity		Percentage: Self-Reported Race or Ethnicity	
	Total	Yes	No	Yes	No
Online Registration	5,407,248	4,263,327	1,143,921	78.8%	21.2%
DMV Registration	8,411,184	1,395,787	7,015,397	16.6%	83.4%
Other Registration	8,338,732	878,084	7,460,648	10.5%	89.5%
All Registration Types	22,157,164	6,537,198	15,619,966	29.5%	70.5%

Data Source: California Secretary of State Voter File

Young registrants aged 18 to 24 registered or re-registered to vote online at higher rates than senior registrants aged 65 and over, while seniors registered (or re-registered) through other sources at higher rates than youth. Table 6 shows that as of the 2020 general election, 41.2% of young registrants registered to vote online, compared to 11.5% of senior registrants. Comparatively, 55.8% of senior registrants registered to vote through other sources, while only 23.4% of young registrants did the same. In total, 88.5% of seniors used a method with low race or ethnicity rates (32.7% at the DMV and 55.8% through sources other than the DMV or online), while just 58.8% of young registrants used methods with low race or ethnicity rates (35.4% at the DMV and 23.4% through sources other than the DMV and online).

This difference in registration methods helps explain the gap between older and younger voters self-reporting their race or ethnicity. By only using the subset of voters who self-identified their race or ethnicity when registering to vote, the subsequent voting behavior analysis would overrepresent younger registered voters who registered to vote online.

Table 6. VoteCal Data: Registration Source and Age Group

	Registered Voters	Count: Registration Source			Percentage: Registration Source		
	Total	DMV	Online	Other	DMV	Online	Other
18-24	2,637,646	934,038	1,085,705	617,903	35.4%	41.2%	23.4%
25-34	4,134,137	1,531,737	1,405,835	1,196,565	37.1%	34.0%	28.9%
35-44	3,524,798	1,440,726	979,203	1,104,869	40.9%	27.8%	31.3%
45-54	3,389,452	1,422,702	719,560	1,247,190	42.0%	21.2%	36.8%
55-64	3,634,942	1,504,798	589,469	1,540,675	41.4%	16.2%	42.4%
65+	4,595,586	1,503,830	526,616	2,565,140	32.7%	11.5%	55.8%
Age Unknown	240,603	73,353	100,860	66,390	30.5%	41.9%	27.6%
Total	22,157,164	8,411,184	5,407,248	8,338,732	38.0%	24.4%	37.6%

Data Source: California Secretary of State Voter File

CENTER FOR
INCLUSIVE
DEMOCRACY

As seen in Table 7, registered voters aged 18 to 24 self-identified their race or ethnicity when registering to vote at higher rates than registered voters aged 65 and over. Among voters registering at the DMV, 21.4% of registered voters aged 18 to 24 self-identified their race or ethnicity, while only 9.3% of registered voters aged 65 and over did the same. Among voters registering online, 84.3% of registered voters aged 18 to 24 answered the race or ethnicity questions, compared to 75.9% of registered voters aged 65 and over. Almost 16% of registered voters aged 18 to 24 who registered to vote through sources other than the DMV and online self-identified their race or ethnicity, while only 6.2% of registered voters aged 65 and over answered the question.

Table 7. VoteCal Data Self-Reported Race or Ethnicity by Registration Source and Age Group

	DMV			Online			Other Source		
	Total Registered	Yes, Self-Reported Race or Ethnicity (Count)	Yes, Self-Reported Race or Ethnicity (Percent)	Total Registered	Yes, Self-Reported Race or Ethnicity (Count)	Yes, Self-Reported Race or Ethnicity (Percent)	Total Registered	Yes, Self-Reported Race or Ethnicity (Count)	Yes, Self-Reported Race or Ethnicity (Percent)
18-24	934,038	199,838	21.4%	1,085,705	915,348	84.3%	617,903	98,075	15.9%
25-34	1,531,737	371,697	24.3%	1,405,835	1,113,020	79.2%	1,196,565	206,929	17.3%
35-44	1,440,726	282,363	19.6%	979,203	740,261	75.6%	1,104,869	153,753	13.9%
45-54	1,422,702	216,570	15.2%	719,560	550,200	76.5%	1,247,190	131,388	10.5%
55-64	1,504,798	180,048	12.0%	589,469	452,595	76.8%	1,540,675	119,300	7.7%
65+	1,503,830	140,220	9.3%	526,616	399,580	75.9%	2,565,140	158,809	6.2%
Age Unknown	73,353	5,051	6.9%	100,860	92,323	91.5%	66,390	9,830	14.8%
Total	8,411,184	1,395,787	16.6%	5,407,248	4,263,327	78.8%	8,338,732	878,084	10.5%

Data Source: California Secretary of State Voter File

As of the 2020 general election, 38.0% of both registered Democratic and Republican voters registered to vote at the DMV, while 36.3% of No Party Preference registered voters registered at the DMV (Table 8). Nearly 21% of registered Republicans registered to vote online, while 24.0% of registered Democrats did the same. Over 41% of registered Republicans registered to vote through a source other than the DMV or online, compared to 38.0% of registered Democrats.

Table 8. VoteCal Data: Registration Source and Party Affiliation

	Registered Voters	Count: Registration Source			Percentage: Registration Source		
	Total	DMV	Online	Other	DMV	Online	Other
Democratic	10,184,219	3,868,288	2,448,691	3,867,240	38.0%	24.0%	38.0%
Republican	5,370,890	2,042,714	1,123,373	2,204,803	38.0%	20.9%	41.1%
No Party Preference	5,322,221	1,933,586	1,543,018	1,845,617	36.3%	29.0%	34.7%
Other Party	1,279,834	566,596	292,166	421,072	44.3%	22.8%	32.9%
Total	22,157,164	8,411,184	5,407,248	8,338,732	38.0%	24.4%	37.6%

Data Source: California Secretary of State Voter File

Republican and Democratic registered voters registering online self-reported their race or ethnicities at similar rates (80.3% and 80.4%, respectively), but both parties self-reported at higher rates compared to No Party Preference (75.4%) and other party (78.2%) registered voters (Table 9). When registering at the DMV, No Party Preference registered voters (13.1%) also self-reported their race or ethnicity at lower rates than Democrats (18.9%) and Republicans (15.6%).

Table 9. VoteCal Data Self-Reported Race or Ethnicity by Registration Source and Party Affiliation

	DMV			Online			Other Source		
	Total Registered	Yes, Self-Reported Race or Ethnicity (Count)	Yes, Self-Reported Race or Ethnicity (Percent)	Total Registered	Yes, Self-Reported Race or Ethnicity (Count)	Yes, Self-Reported Race or Ethnicity (Percent)	Total Registered	Yes, Self-Reported Race or Ethnicity (Count)	Yes, Self-Reported Race or Ethnicity (Percent)
Democratic	3,868,288	732,994	18.9%	2,448,691	1,969,437	80.4%	3,867,240	438,445	11.3%
Republican	2,042,714	317,888	15.6%	1,123,373	902,513	80.3%	2,204,803	188,730	8.6%
No Party Preference	1,933,586	252,419	13.1%	1,543,018	1,163,021	75.4%	1,845,617	201,427	10.9%
Other Party	566,596	92,486	16.3%	292,166	228,356	78.2%	421,072	49,482	11.8%
Total	8,411,184	1,395,787	16.6%	5,407,248	4,263,327	78.8%	8,338,732	878,084	10.5%

Data Source: California Secretary of State Voter File

Research Methods to Identify Race or Ethnicity

There are multiple methods commonly used by researchers to identify registered voters' race or ethnicities, including surname matching, geocoding, and a combination of the two (Bayesian Improved Surname Geocoding (BISG)). While these methods are generally trusted by researchers, each method has limitations as to who can be identified.

CENTER FOR
INCLUSIVE
DEMOCRACY