



UNITED STATES DEPARTMENT OF COMMERCE
Economics and Statistics Administration
U.S. Census Bureau
Washington, DC 20233-0001

January 3, 2018

MEMORANDUM FOR: Ron S. Jarmin
Performing the Non-exclusive Functions and Duties of the Director

From: John M. Abowd 
Chief Scientist and Associate Director for Research & Methodology

Prepared by: Stephen Buckner, Patrick Cantwell, Joanne Crane, Melissa Creech, Burton Reist,
and James Whitehorne

Subject: Summary of Quality/Cost of Alternatives for Meeting Department of Justice
Request for Citizenship Data

The Department of Justice has requested census block-level citizen voting-age population estimates by OMB-approved race and ethnicity categories from the 2020 Census of Population and Housing. These estimates are currently provided in two related data products: the PL94-171 redistricting data (PL94), produced by April 1st of the year following a decennial census under the authority of 13 U.S.C. Section 141, and the Citizen Voting Age Population by Race and Ethnicity (CVAP) tables produced every February from the most recent five-year American Community Survey data. The PL94-171 data are released at the census block level. The CVAP data are released at the census block group level.

There are three alternatives for meeting the DoJ request:

Alternative A: Maintain the status quo for data collection, preparation and publication. After the regular PL94 and CVAP data have been published in 2021, prepare a special product for DoJ that combines these tables to produce a Census Bureau best estimate of the block-level citizen voting age population by race and ethnicity. This would be similar to the approach now used to support the Section 203 requirements of the Voting Rights Act.

Alternative B: Add a citizenship question to the 2020 Census questionnaire. Process the citizenship question using the 2020 Census data processing system, including using administrative records where currently authorized and implemented. Produce the block-level tables of citizen voting age population by race and ethnicity during the publication phase of the 2020 Census.

Alternative C: Do not add a citizenship question to the 2020 Census questionnaire. Add the capability to link an accurate, edited citizenship variable from administrative records to the final 2020 Census microdata files. Produce the block-level tables of citizen voting age population by race and ethnicity during the publication phase of the 2020 Census using the enhanced 2020 Census microdata.

An analysis of the cost and quality implications of each alternative follows.

For Alternative A, the cost would be similar to the cost of the every-five-year production of language determinations to support DoJ enforcement of Section 203 of the Voting Rights Act. We estimate that the incremental cost of producing Alternative A is approximately \$200,000. This estimate is based on the total cost of producing the 2016 language determinations. These costs are the salaried time of staff in the Redistricting Office, the Policy Coordination Office, the American Community Survey Office, and the Center for Statistical Research and Methodology. We estimate that delivering block-level citizen voting age population tables by race and ethnicity based on statistical modeling of the 2020 PL94-171 and CVAP tables would be a similar effort in terms of both cost and quality.

For Alternative B, our estimate of the incremental cost is as follows. We estimate that there will be approximately 145 million addresses in the 2020 Census. Recent American Community Survey (ACS) data indicate that 9.8% of households contain at least one non-citizen. This yields an estimate of 14.2 million households with at least one non-citizen. Our analysis of 2010 response data for both the American Community Survey and the Short-Form Census indicates that the presence of a question on citizenship suppressed response by 5.1 percentage points for these households. Based on our working assumption of a response rate of 60.5% in the 2020 census, a reduction of 5.1 percentage points for households containing at least one non-citizen will increase the Nonresponse Follow-up (NRFU) workload by approximately 700,000 households, or approximately 0.5 percentage points. We estimate that for each percentage point increase in NRFU the cost of the 2020 Census increases by approximately \$55 million. Accordingly, the addition of a question on citizenship could increase the cost of the 2020 Census by at least \$27.5 million. It is worth stressing that this cost estimate is a lower bound. Our estimate of \$55 million for each percentage point increase in NRFU is based on an average of three visits per household. We expect that we would make a total of six visits to households containing non-citizens who did not self-respond. We also believe that the decrease in response for these households in 2020 could be greater than the 5.1 percentage points we observed during the 2010 Census.

Alternative B would most likely deliver higher quality block-level citizen voting age population by race and ethnicity data than Alternative A because it is based on obtaining a direct report of citizenship status for each respondent. But it would result in lower quality enumeration data. Because of the estimated 5.1 percentage point increase in NRFU for households with noncitizens and because NRFU is less accurate than self-reports, there is a decrease the coverage quality of the census. We estimate that asking the citizenship question would result in 154,000 *fewer* correct enumerations. This is also a lower bound estimate on the loss of accuracy.

For Alternative C, the cost estimate has not yet been fully vetted. The estimated cost is less than \$1,000,000. In the current system of administrative data acquisition and processing for the 2020 Census, there are ten remaining data acquisitions between January 1, 2018 and April 1, 2020. We include the cost of these data acquisitions in our incremental cost estimate for Alternative C because it would no longer be an option to discontinue the acquisition of these data and use only the already acquired files. In addition, we include the cost of two senior analysts. The analysts would do the required integration of the edited citizenship into the 2020 Census microdata. This estimate does not include the cost of modifying existing Memoranda of Understanding with the agencies that supply the data used to acquire the citizenship variable because those negotiations are already in progress, and would continue. This estimate also does not include the cost of negotiating a new MOU with the United States Citizen and Immigration Services to acquire those data, refreshed appropriately over the 2020 Census life cycle because we have not yet gathered the required information from USCIS to estimate this cost.

Alternative C delivers higher quality data than Alternative B for DoJ's stated uses. Our primary data sources for the administrative record citizen variable require proof of citizenship. For this reason, they are very accurate. There is good evidence that citizenship is accurately reported by citizens, but less accurately self-reported by household responders for noncitizens. This accuracy deficit in the self-responses may be due to the inherent difficulty of the respondent knowing the citizenship status of everyone in the household. Proxy respondents are even less likely to know the citizenship status of all members of the household. It may also be due to the sensitivity of the citizenship question itself. For DoJ's stated purposes, Alternative C produces the highest quality data.

Alternative A is not very costly and does not harm the quality of the census count. Alternative B better addresses DoJ's stated uses; however, it is very costly and does harm the quality of the census count by increasing erroneous enumerations. Alternative C even better meets DoJ's stated uses, is comparatively far less costly than Alternative B, and does not harm the quality of the census count. For these reasons, we recommend Alternative C for meeting the Department of Justice data request.