

---

# Estimating the Effect of Asking About Citizenship on the U.S. Census

Results from a Randomized Controlled Trial

*A Shorenstein Center Discussion Paper*

**March 2019**



**HARVARD** Kennedy School

**SHORENSTEIN CENTER**  
on Media, Politics and Public Policy

*The views expressed in Shorenstein Center Discussion Papers are those of the author(s) and do not necessarily reflect those of Harvard Kennedy School or of Harvard University.*

*Discussion Papers have not undergone formal review and approval. Such papers are included in this series to elicit feedback and to encourage debate on important issues and challenges in media, politics and public policy. Copyright belongs to the author(s). Papers may be downloaded for personal use only.*

Authors

---

**Matthew A. Baum**, Kalb Professor of Global Communication, Harvard Kennedy School

**Bryce J. Dietrich**, Research Fellow, Shorenstein Center on Media, Politics, and Public Policy, Harvard Kennedy School

**Rebecca Goldstein**, PhD Candidate, Harvard University Department of Government

**Maya Sen**, Associate Professor of Public Policy, Harvard Kennedy School

*Authors' names listed in alphabetical order.*

---

# Abstract and Introduction

*The 2020 U.S. Census will, for the first time since 1950, ask about residents' citizenship status. The effect of doing so on census completion across different racial/ethnic groups is, however, unknown. Leveraging a survey experiment (n=9,035 respondents) we are the first to assess the causal effect of this question change. We find that asking about citizenship status significantly increases the percent of questions skipped, with particularly strong effects among Hispanics, and makes respondents less likely to report having members of their household who are of Hispanic ethnicity. Although our study is not designed to address household size effects, when we extrapolate to the general population, our results imply that asking about citizenship will reduce the number of Hispanics reported in the 2010 Census by approximately 4.2 million, or around 8.4 percent of the 2010 Hispanic population.*

In December of 2017, the U.S. Census Bureau announced it would include questions on residents' citizenship status on the 2020 Census. Due to the potential disproportionate effects on Hispanic populations, several jurisdictions filed suit in an effort to block the inclusion of such questions. Two federal district judges have since ordered the Trump Administration to drop the questions, and, at the time of our writing, the Supreme Court has agreed to an expedited hearing.

Research suggests that the litigants' fear may be well founded. A large literature in political science has documented that Hispanics have lower levels of overall political participation (e.g., Verba, Schlozman and Brady 1995) and are less likely to vote (e.g., Shaw, De La Garza and Lee 2000) than other groups of Americans. Although get-out-the-vote campaigns targeting Hispanics can be effective (Barreto 2005), research also shows that foreign-born Hispanics are especially unlikely to vote (e.g., DeSipio 1998; Cho 1999).

Acculturation also seems to decrease Hispanics' trust in government (Michelson 2003; Wenzel 2006), especially when they experience discrimination (Abrajano and Alvarez 2010) or immigration enforcement efforts (Rocha, Knoll and Wrinkle 2015). Since many Hispanics are likely to view the citizenship question as discriminatory and potentially tied to immigration enforcement, it seems likely to repress response rates. Although some scholars have attempted to find ways to ameliorate the anxiety Hispanics face regarding the Census (e.g., Trujillo and Paluck 2012), concern remains that the citizenship question will lead to undercounts. Studies using data from the American Community Survey (ACS) have found some evidence consistent with this expectation (Brown et al. 2018; Van Hook et al. 2014). However, these findings are derived from observational data and cross-survey comparisons, which are ill-suited for estimating causal effects.

By employing a survey experiment that mirrors the form and content of the actual Census, we find that asking about U.S. citizenship significantly reduces the overall share of questions that respondents answer, with suggestive evidence that the effects are more pronounced among Hispanic respondents. We also find that the citizenship question significantly reduces the number of household members reported as being Hispanic. These patterns are particularly stark among Hispanics who report being born in Mexico or Central America. Extrapolating our results to the general population, we estimate that asking about citizenship would reduce the share of Hispanics recorded by the Census by approximately 4.2 million, or around 8.4 percent of the 2010 Hispanic population – a sizable reduction in the share of the U.S. population that would be recorded as Hispanic.

## Research Design

We designed our survey experiment to match the short form of the U.S. Census, with the questions taken verbatim from the Census' proposed questions. A third-party vendor (Qualtrics) recruited the survey panel and implemented the study in two waves. The first wave targeted non-Hispanics (employing an English survey instrument), using self-reported demographic information maintained by Qualtrics. This group (n = 4,104) included 3,413 whites, 246 African Americans, 55 Hispanics, 181 Asians/Asian-Americans, 92 members of other racial groups, and 117 respondents who did not identify their race to Qualtrics. The second wave targeted Hispanics (using English and Spanish survey instruments) in order to facilitate meaningful subgroup inferences. This group (n = 4,931) included 4,562 Hispanics, 13 whites, 3 African Americans, 1 Asian/Asian-American, 3

members of other racial groups, and 349 respondents who did not identify their race.

To evaluate the impact of asking about household members' citizenship on item non-response and response quality, we randomly assigned half of the respondents ( $n = 4,497$ ) to receive a "Citizenship Treatment" in which we asked, for each member of their household, "Is this person a citizen of the United States?". The other half ( $n = 4,538$ ) did not receive the citizenship question for any household member. Because the Census Bureau has not yet indicated where the citizenship question will be located within the 2020 Census, we also randomly rotated the order in which the citizenship question appeared, conditional on the household member in question.

An obvious difference between our study and the actual U.S. census is our status as academic researchers, which might lead to confidence among respondents that data would not be used for immigration purposes. To assess this, we also randomly assigned half of the respondents ( $n = 4,454$ ) to receive a "Census Prompt" treatment, independently of the first randomization, consisting of a short note at the bottom of their consent form saying "Your responses will be shared with the U.S. Census Bureau," and requiring respondent consent. The other half ( $n = 4,581$ ) received no prompt. (Additional details on survey logistics can be found in Sections S1-S2 of our [Supplemental Information](#) or SI.)

# Results

We now turn to our core results concerning treatment effects on item non-response and the underreporting of Hispanic household members. These results are outlined in Table 1.

Table 1: Treatment Effects on Item Non-Response and Response Quality

	Treatment	Control	Difference	<i>t</i> -statistic	<i>p</i> -value
<i>Percent of All Questions Skipped</i>					
All	22.81	21.32	1.48	2.211	0.027
Hispanic	28.89	26.84	2.05	2.059	0.040
Non-Hispanic	16.52	15.49	1.03	1.203	0.229
<i>Percent of Race/Ethnicity Questions Skipped</i>					
All	23.29	21.61	1.68	2.043	0.041
Hispanic	30.55	28.18	2.38	1.926	0.054
Non-Hispanic	15.79	14.67	1.12	1.079	0.280
<i>Percent of Age/Date-of-Birth Questions Skipped</i>					
All	24.34	22.40	1.95	2.316	0.021
Hispanic	32.26	29.53	2.73	2.149	0.032
Non-Hispanic	16.16	14.85	1.31	1.260	0.208
<i>Percent of Household Members Reported as Hispanic</i>					
All	30.88	32.97	-2.08	2.289	0.022
Hispanic	53.03	55.33	-2.30	1.750	0.080
Non-Hispanic	8.00	9.32	-1.32	1.681	0.093

*Note:* This table shows the marginal treatment effects on item non-response and response quality for the citizenship question. Means under treatment and control are reported in the first two columns. Differences between those columns are shown in the third column. Last two columns report the results from two-sample *t*-tests.

## Treatment Effects on Item Non-Response

We operationalize survey item non-response as the percent of the survey questions for which the respondent submitted a response. Using this measure, we find receiving the Citizenship Treatment increases the overall share of questions skipped by 1.48 percentage points (*t*-statistic = 2.211, *p*-value = 0.027). We also find the Census Prompt increases the share of questions skipped by 1.35 percentage points (*t*-statistic = 2.011, *p*-value = 0.044).

We also find suggestive evidence this effect was more pronounced for Hispanics, who skipped 2.05 points more of the questions when receiving the Citizenship Treatment (*t*-statistic = 2.059, *p*-value = 0.040). We found

even larger effects, albeit only significant at 0.10, among Hispanics taking the survey in Spanish. Here, receiving the Citizenship Treatment increases the percent of questions skipped by 8.91 percentage points (t-statistic = 1.749, p-value = 0.082). For Hispanics taking the survey in English, the corresponding effect is an increase in skipped questions of 1.78 percentage points, also significant only at the 0.10 level (t-statistic = 1.749, p-value = 0.080).

We find strong, significant effects for Hispanics listing Mexico or a country in Central America as their country of birth. For these respondents, receiving the Citizenship Treatment increases the percent of questions skipped by 5.69 percentage points (t-statistic = 2.919, p-value = 0.004). The corresponding effect among Hispanics who listed any other birth country was 1.69 percentage points (t-statistic = 1.627, p-value = 0.104). Among Hispanics who listed the United States as their birth country, we see no significant effect (t-statistic = 1.348, p-value = 0.178). We also see no significant difference for non-Hispanics (t-statistic = 1.032, p-value = 0.229).

Because the U. S. Census imputes some missing data, in Section S3 of the SI, we show these results are similar after separately imputing respondents in the treatment and control conditions – in fact, the difference increases to 2.65 percentage points and becomes statistically significant for Hispanic respondents receiving the Citizenship Treatment.

## **Treatment Effects on Household Members' Race/Ethnicity**

A key concern is whether asking about citizenship would affect respondents' willingness to report important demographic information, like household members' race or ethnicity. Conditioning on the number of household members initially reported by the respondent, the Citizenship Treatment is associated with a 1.68 percentage point increase in the share of questions skipped about members' race/ethnicity (t-statistic = 2.043, p-value = 0.041). The corresponding effect of the Citizenship Treatment among Hispanics is more pronounced: a 2.38 percentage point increase (t-statistic = 1.926, p-value = 0.054). We found an even larger, 11.57 point, increase, albeit significant at only 0.10, among Hispanics taking the survey in Spanish (t-statistic = 1.867, p-value = 0.064). For Hispanics taking the survey in English, the corresponding effect is an insignificant 2.01 percentage point increase (t-statistic = 1.599, p-value = 0.110).

We again see strong, significant effects for Hispanics listing Mexico or a country in Central America as their country of birth. Here, the Citizenship Treatment is associated with a 6.12 percentage point increase in skipped race/ethnicity questions (t-statistic = 2.395, p-value = 0.017). Finally, among Hispanics who listed any other birth country, those listing the United States as their birth country, or among non-Hispanics, we see no significant effects (t-statistic = 1.534, p-value = 0.125, t-statistic = 1.069, p-value = 0.285, and t-statistic = 1.079, p-value = 0.281, respectively).

## Treatment Effects on Household Members' Ages

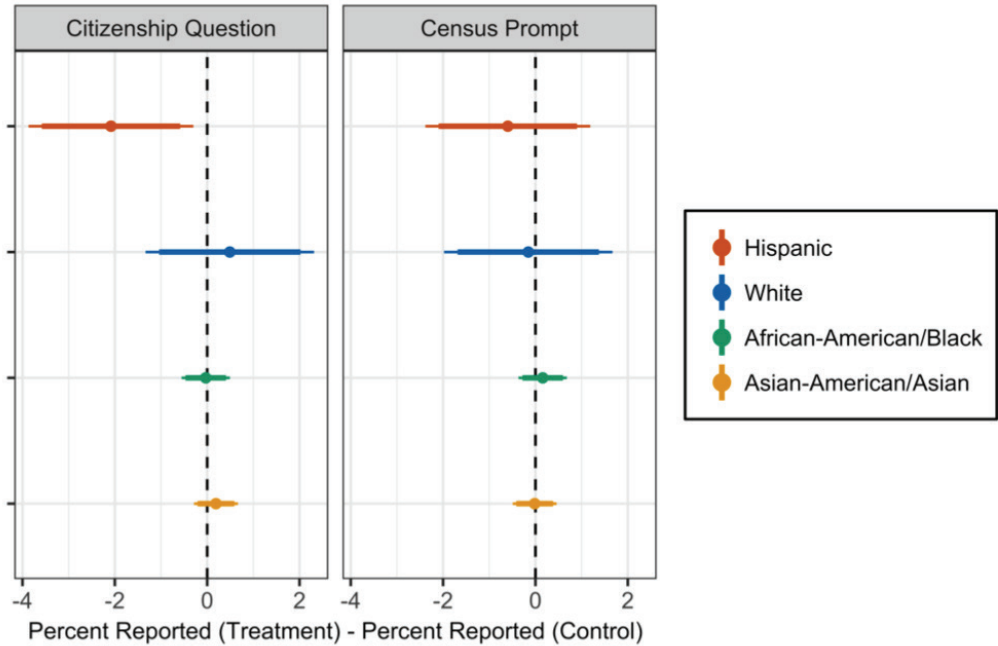
Another concern is the possibility that information about citizenship status could be used against families with young children which is why we also consider whether the Citizenship Treatment leads to more questions skipped regarding household members' ages. Conditional on the number of household members initially reported, respondents who received the Citizenship Treatment are significantly more likely to skip the questions concerning household members' ages, on average by 1.95 percentage points (t-statistic = 2.317, p-value = 0.021). We see suggestive evidence that the effect is stronger among Hispanics, who experience a large, statistically significant increase in questions skipped of 2.73 percentage points (t-statistic = 2.149, p-value = 0.032). For Hispanics taking the survey in Spanish, the Citizenship Treatment increased the percent of questions skipped on household members' ages by 12.11 points, although this effect is significant only at the 0.10 level (t-statistic = 1.885, p-value = 0.061). The corresponding effect among Hispanic respondents taking the survey in English is smaller and also significant only at the 0.10 level – an increase in question-skipping of 2.36 points (t-statistic = 1.819, p-value = 0.069).

We again find large, significant effects among Hispanic respondents who report being born in either Mexico or a Central American country, where receiving the Citizenship Treatment increases the percent of questions skipped on household members' ages by 7.72 points (t-statistic = 2.542, p-value = 0.012). The corresponding increase among Hispanic respondents listing any other birth country is 2.25 (t-statistic = 1.708, p-value = 0.088). We again find no significant effect among Hispanics listing the United States as their birth country (t-statistic = 1.274, p-value = 0.203) or among non-Hispanics (t-statistic = 1.260, p-value = 0.208).

## Treatment Effects on Percent of Household Reported as Being Hispanic

Because accurately counting racial/ethnic minorities has substantial implications for federal resource allocations, we also consider any treatment effects on the share of household members identified by the respondent as being of “Hispanic, Latino, or Spanish Origin.” That is, we consider the percent of household members identified as Hispanic (as opposed to other ethnicities or non-responses) by each respondent.

Figure 1: Citizenship Question Leads To Underreporting Hispanics



Note: Difference in the percent of the household reported being various racial/ethnic groups by treatment and control for the citizenship question (left panel) and Census prompt (right panel). Thicker (—) and thinner (—) lines represent 90 and 95-percent confidence intervals, respectively.

Figure 1 shows those receiving the Citizenship Treatment reported fewer Hispanic household members (30.88 percent of households) compared to those in the control condition (32.97, t-statistic = 2.289, p-value = 0.022). Hispanic respondents receiving the Citizenship Treatment reported 53.03 percent of their household members were of Hispanic origin compared to 55.33 percent in the control condition, but the difference of 2.30 percentage point difference is significant only at the 0.10-level (t-statistic = 1.750,

p-value = 0.080). Non-Hispanic respondents receiving the Citizenship Treatment reported that 8.00 percent of their household members were Hispanic as compared to 9.32 percent in the control condition. This comparatively smaller difference of 1.32 percentage point was also significant only at the 0.10-level (t-statistic = 1.681, p-value = 0.093). We see no significant differences associated with the Census Prompt on these dimensions.

## Extrapolating to the 2020 U.S. Census Context

We now consider what, more precisely, our study might predict for the 2020 U.S. Census context in terms of the share of Hispanics that might be undercounted. Our survey purposefully oversampled Hispanics (51.10 percent of our sample) relative to the U.S. population (16.35 percent, as reported by the 2010 U.S. Census). To produce more nationally representative estimates, we created post-stratification weights using a raking algorithm, based on the available information about our respondents (their race/ethnicity, whether they were Democratic identifiers, and their state of residence, based on zip codes). (We provide more details in [Section S2-S3 of the SI.](#))

We then re-estimated the Citizenship Treatment effect using these nationally representative survey weights. The results indicate that, assuming our weights are accurate, we would expect to see a statistically significant national drop in the share of Hispanic household members reported, from 13.02 percent to 11.65 percent (t-statistic = 2.391, p-value = 0.017). This suggests that our previous findings were not simple artifacts of our oversampling of Hispanics.

Applying this estimated national-level treatment effect to the U.S. population, as reported by the 2010 U.S. Census (308,745,538), we estimate that asking about citizenship will reduce the number of Hispanics reported in the 2010 Census by 4,229,814 or 8.38 percent of the 2010 Hispanic population (50,477,594). The 95-percent confidence interval of our estimate is 3,890,194 to 4,600,308, which represents a decrease of 7.71 to 9.11 percent relative to the 2010 Hispanic population. In 2016, six percent of all respondents and 7.40 percent of Hispanic respondents did not respond to the ACS citizenship question (Brown et al., 2016), which suggests that our estimates are reasonable, though admittedly suggestive.

Although we cannot say with certainty how the predicted undercounted Hispanics will be distributed, in some configurations congressional ap-

portionment could be affected. The New York Times estimated that a 15% undercount of non-citizens would result in California and New York both losing congressional seats (Badger 2016). In Table S6 of the SI we estimate the effect of the citizenship question across all 435 districts. There we find in some districts where Hispanics represent a larger share of the population – especially in Southern California – our Citizenship Treatment would have a larger impact, which suggests apportionment could also be affected by introducing the citizenship question.

## Discussion and Conclusion

This study presents the first explicit causal exploration of the impact of asking about citizenship on the 2020 U.S. Census. We find that asking about citizenship status significantly increases the percent of questions skipped, with particularly strong effects among Hispanics, and makes respondents less likely to report having Hispanic household members. Aggregating this to the national level suggests that asking a citizenship question may lead to an undercounting of Hispanics of around 3,890,194 to 4,600,308 in the 2020 Census (based on 2010 figures).<sup>1</sup>

We note two caveats regarding our study. The first is that we likely underestimate the effect of asking about citizenship status on the 2020 Census. Not only are we university affiliated academic researchers – and not the U.S. Government – and so respondent concerns over providing the government with personal information may not have inhibited participation in this survey, but our respondents were paid panelists and thus financially incentivized to complete the survey. The second is that our study was ill suited for estimating the causal effects of citizenship questions on household member undercounts. In order to mirror the actual census form, all citizenship questions appeared in our survey after respondents listed the number of members of their household. Additionally, unlike the Census Bureau, we had no “baseline truth” against which to compare reported household size. Future research might be better equipped to assess the potential impact of asking about citizenship status on household size reporting.

---

<sup>1</sup> In [Section 3 of the SI](#), additional analyses are reported including (1) further subgroup analyses, (2) the direct and indirect budgetary costs, and (3) an exploration of race-based opinions towards the U.S. Census. We also assess the interaction between our Citizenship and Census Prompt Treatments. Although we find no evidence of a significant interaction, this is partially due to our inability to recontact individuals who dropped from the survey after receiving the Census Prompt.

These caveats aside, the key takeaway is that including a citizenship question will likely result in undercounts of Hispanics and less complete information collection by the 2020 Census – above and beyond the household non-response effects already well-documented by the U. S. Census

## Acknowledgements

This research was supported by the Shorenstein Center on Media, Politics, and Public Policy at the John F. Kennedy School of Government, Harvard University. We thank Rafael Carbonell, Kyla Fullenwider, Candelaria Garay, Horacio Larreguy, Michelle Lyons, O’Reilly Miami, Nicco Mele, Noelia Negri, Nathan Richard, David Lazer, John Thompson, and Sunshine Hillygus for helpful comments and suggestions. We also thank Amaris Taylor for her help at Qualtrics.

## References

Supplemental Information available here: [https://shorensteincenter.org/wp-content/uploads/2019/03/census\\_supplemental\\_information.pdf](https://shorensteincenter.org/wp-content/uploads/2019/03/census_supplemental_information.pdf)

Abrajano, Marisa A and R Michael Alvarez. 2010. “Assessing the causes and effects of political trust among US Latinos.” *American Politics Research* 38(1):110–141.

Barreto, Matt A. 2005. “Latino immigrants at the polls: Foreign-born voter turnout in the 2002 election.” *Political Research Quarterly* 58(1):79–86.

Brown, J David, Misty L Heggeness, Suzanne M Dorinski, Lawrence Warren, Moises Yi et al. 2018. Understanding the Quality of Alternative Citizenship Data Sources for the 2020 Census. Technical report U.S. Census Bureau.

Cho, Wendy K Tam. 1999. “Naturalization, socialization, participation: Immigrants and (non-voting.” *The Journal of Politics* 61(4):1140–1155.

DeSipio, Louis. 1998. *Counting on the Latino vote: Latinos as a new electorate*. University of Virginia Press.

Michelson, Melissa R. 2003. “The corrosive effect of acculturation: How Mexican Americans lose political trust.” *Social Science Quarterly* 84(4):918–933

Rocha, Rene R, Benjamin R Knoll and Robert D Wrinkle. 2015. "Immigration enforcement and the redistribution of political trust." *The Journal of Politics* 77(4):901–913.

Shaw, Daron, Rodolfo O De La Garza and Jongho Lee. 2000. "Examining Latino turnout in 1996: A three-state, validated survey approach." *American Journal of Political Science* 44(2):338–346.

Trujillo, Matthew D and Elizabeth Levy Paluck. 2012. "The devil knows best: Experimental effects of a televised soap opera on Latino attitudes toward government and support for the 2010 US Census." *Analyses of Social Issues and Public Policy* 12(1):113–132.

Van Hook, Jennifer, Frank D Bean, James D Bachmeier and Catherine Tucker. 2014. "Recent trends in coverage of the Mexican-born population of the United States: Results from applying multiple methods across time." *Demography* 51(2):699–726.

Verba, Sidney, Kay Lehman Schlozman and Henry E Brady. 1995. *Voice and equality: Civic voluntarism in American politics*. Harvard University Press.

Wenzel, James P. 2006. "Acculturation effects on trust in national and local government among Mexican Americans." *Social Science Quarterly* 87(5):1073–1087.