

Studying the relation between data quality and trust in official statistics¹

Joanna Fane Lineback¹, Karthik Vaidyanathan², Benjamin M. Reist³

¹*Census Bureau, USA*

²*State Department, USA*

³*NORC, USA*

Abstract

In this paper, we explore the relation between survey participation and trust in official statistics, sharing findings from a study on factors influencing survey participation. Response rates, a necessary component of data quality, continue to fall for government surveys. Recent studies provide insight into this phenomenon. In a nationally representative survey that tracked public awareness of the 2020 U.S. Census, respondents were asked about their plans to participate in the 2020 Census as well as a variety of other questions we thought might be associated with census participation. Topics included the state of the current political and economic environments, civic engagement, and attitudes toward others. Additionally, there were two questions about attitudes toward government, one about trust in federal statistics, and one about the fear of the government's misuse of survey responses. Common perception is that trust in government is the driving factor of survey nonresponse. Here, however, we have empirical evidence that trust in official statistics and other factors are also important predictors of survey participation. We discuss these findings and the implications on response rates as a necessary component of data quality.

Keywords: response rates, data quality, trust in government, trust in official statistics

1. Introduction

1.1 Increasing survey nonresponse

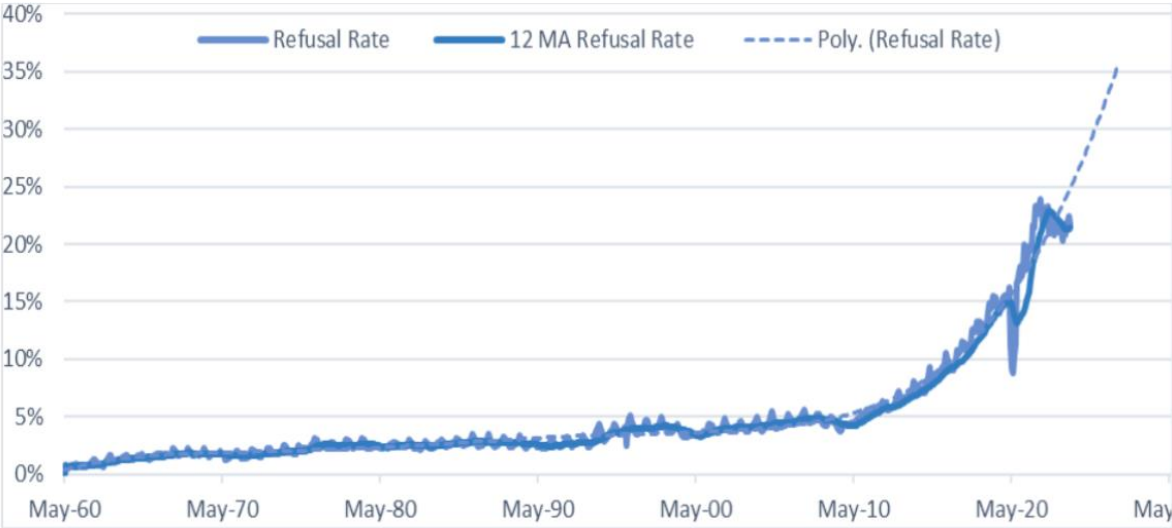
Only until recently in the United States, many major government surveys' nonresponse rates had been dramatically increasing. These increases were driven by survey refusals. For example, Figure 1 below shows Current Population Survey (CPS) refusal rates, or the percent of refusals to eligible cases, from May 1960 through February 2024. Over the entire period, refusal rates have generally been increasing, however an exponential increase started around 2010 and did not end until March 2022, when rates began to fall. As of February 2024, the

¹ The views are those of the authors and necessarily the Census Bureau, State Department, or NORC. The data in the paper are publicly available and therefore not subject to disclosure review: <https://www.openicpsr.org/openicpsr/project/125761/version/V3/view>

CPS refusal rate was at 21.44%, down from the high of 23.91% recorded in March 2022. Over the past few years, the CPS refusal rate has been fluctuating between 21% and 23%.

The only other noticeable period of decreasing refusal rates can be explained by the Covid-19 Pandemic and associated changes to data collection procedures. And even though refusal rates were decreasing for about six months during 2020, nonresponse rates were still increasing, but with the increase being driven by noncontacts.

Figure 1: CPS Refusal Rate by Month, January 1960-February 2024



Source: U.S. Bureau of Labor Statistics, Current Population Survey, 1960-2024 (unweighted)

Government surveys are often seen as the gold standard for key demographic and economic data. Therefore, the dramatic increase in survey refusals since the beginning of the CPS is concerning, because response rates are a key component of data quality. Moreover, major government survey programs have many tools at their disposal for mitigating survey nonresponse, but despite access to such tools, we have still seen increases in refusal rates. Additionally, no one event or series of events has pointed to a clear explanation for these increases. In this paper, we explore this topic by modeling reported 2020 Census participation on social, economic, and political explanatory factors.

2. Background

2.1 Previous research

In an earlier study, we used time-series regression to explore the relation between economic and political conditions on CPS refusal rates over the period 1960-2015, an extension of work done by Harris-Kojetin and Tucker (1999). After including a number of relevant predictors in our model, we found that lower refusal rates were associated with increases in presidential approval, and higher refusal rates were associated with an increase in jobs added to the economy, a decrease in the not-in-labor-force population, and whether it was a census year (Larsen et al., 2020).

For that study, we had also hoped to explore the relation between social conditions and CPS refusal rates. (The social aspect is less studied, but we suspect a key component.) However, we were unable to identify time series to represent the social aspect of the social-economic-political construct of survey nonresponse.

Here, we explore the full social-economic-political construct of survey nonresponse. In a survey about public awareness of the 2020 U.S. Census, we were given the opportunity to field questions about one's thoughts on the current economic and political condition in the U.S., one's civic engagement activity, and one's attitude toward others and the government, which could then be tied back to reported 2020 Census participation.

2.2 Theoretical framework

2.2.1 – Response model

If common perception is correct, then we would expect trust in government to be a highly significant indicator of survey participation. Based on a 2018 study by the Census Bureau, distrust in all levels of government was high among all studied demographic groups (Census Bureau, 2019). And in our everyday survey work, we often take for granted that there might be other reasons for survey nonresponse.

In fact, although most people have distrust in government at some level, it has been shown that there is no relation between trust and cooperation (Bauer et al., 2019). Moreover, what does it mean if one does not trust the government? This is a broad statement, without a concrete meaning.

Here we study trust in government and other potential factors of 2020 Census participation to more definitively encapsulate the survey response model. It has been reasoned that one's social-political-economic environment may influence an individual's estimation of the costs and benefits of responding to a survey, which is consistent with many of the theoretical frameworks for understanding the response process: for example, social exchange theory (Dillman et al., 2014), benefit cost theory (Singer, 2011), and leverage-saliency (Groves et al., 2000). We ourselves have shown that the political and economic environment are important predictors of survey participation (Larsen et al, 2020). Other studies have shown that those more socially engaged, for example, those who volunteer, are more likely to respond to a government survey (Amaya, 2015). We expect to find that social factors, as well as attitudes toward others and the government, will be strong predictors of 2020 Census participation.

2.2.2 – Nonresponse paradox

More generally, how does one reconcile official statistics as one of the pillars of democracy with the decline in government survey response rates?

Surveys and censuses are the primary source of official statistics. In Figure 1, we saw that for the first 50 years of the CPS, refusal rates never climbed above 5%. Given that other reasons for nonresponse were negligible over this period, this means that for over 50 years over 95% of the eligible population were represented in the CPS. By contrast, as recently as February 2024, roughly 22% of the eligible, sampled population refused to respond to the CPS. In other words, almost a quarter of the eligible, sampled population were not represented in the CPS...by choice.

Being represented in official statistics *should* contribute to one's trust in official statistics. We suggest that as society has become more individualistic, people see less need to respond to government surveys. Perhaps they do not feel a civic duty to respond, perhaps they see no value in responding, or perhaps they are afraid of responding.

We note that the same factors we are studying in this paper are also important measures of a democracy. We do not attempt to answer this question here, but leave as a topic for future discussion whether the decline in response is tied to a decline in democracy.

3. Methodology

3.1 Data source

The U.S. Census Bureau sponsored a survey that tracked public awareness of the 2020 Census using the 2020 Census Tracking Survey. As part of this survey, we had the opportunity to obtain measures to study the social-economic-political construct of survey participation. Respondents were asked about their plans to participate in the 2020 Census as well as basic demographic information, intended mode of participation, the likelihood of encouraging others to participate in the 2020 Census, political and economic sentiment, civic engagement, and empathetic concern. The latter categories were proposed as measures of one's environment that might be related to one's attitude toward participating in a government survey. A few questions from major government surveys were also added as benchmark measures.

3.2 Building the logistic regression model

To study our hypothesis that social factors and attitudes toward others and the government influence survey participation, we used 2020 Tracking Survey data to build a logistic regression model of known and potentially important predictors of reported intent to participate or having participated in the 2020 Census. The focus of this model was to explore the relation between response and topics such as attitudes toward government and attitudes toward others while controlling for demographics and other known predictors of response. (For more on exact question wording, see Appendix A.)

Model regressors included the demographic variables age, race, education, home ownership, sex, marital status, and language, as well as their interactions. It also included Census Bureau regional divisions as a geographic measure, month in sample as a temporal measure, and government "trust" measures. It also included benchmark variables. Finally, it included the social-economic-political-attitudinal regressors specifically intended to test our hypothesis. A diversity index, population density, and percentage white, were imported from the Census Bureau's Planning Database and matched by census tract to the 2020 Tracking Survey file, but, ultimately, these were dropped from the final models. The full set of regressors is shown in Table 1.

Outside of age, most of the variables in the models were not continuous; most of the variables were already binary dummies, or ordinal scales that were ultimately changed into binary dummies. In addition, the number of "missings" (those who refused to answer questions or

said don't know) were considerable enough that the first regression model had to drop over 30% of observations because of incomplete data. To address this in subsequent models, dummies for the missings were created so that "1" equaled a missing value and "0" was an answer on the given scale.

Table 1: Main effects model regressors

Category	Regressor
<i>Demographic</i>	Race
	Education
	Home ownership
	Sex
	Marital status
	Language
	Age
<i>Spatial</i>	Census division
<i>Temporal</i>	Month in sample (September 2019-June 2020)
<i>Attitude toward government</i>	Trust in federal statistics
	Fear of government using responses against them
<i>Economic view</i>	Economic condition
	Job situation
<i>Political view</i>	Direction of country
	Satisfaction with how nation being governed
<i>Civic engagement</i>	Religious service attendance
	Contacted public official
	Volunteerism
<i>Attitude toward others</i>	Two sides to every question
	Find others' misfortunes disturbing

The age and the demographic variables noted above, as well as geography and time, constituted a "base" model. Previous research has shown that these have an impact on census completion rates. A "reduced" model included the variables from the base model as well as the two government "trust" variables. And, finally, a "full" model included the reduced model, plus nine political-economic-social-attitudinal environment questions (which were turned into binary dummies), including the volunteering question (which was not in our original study list but nonetheless another "civic" variable that was asked on the survey).

Note that unweighted results are shown here. The weights had not been finalized at the time the analysis was being conducted. However, when the survey weights were eventually included, they did not appear to appreciably affect the results.

4. Findings

4.1 Demographic, geographic, spatial, and temporal variables

Among demographic, geographic and temporal items, date was statistically significant. There was a clear uptick in the percent of respondents who said they had completed or intended to complete (hereafter intent to complete) the 2020 Census as 2020 progressed, especially after mid-March compared to prior dates. Age was also statistically significant. Young people were less likely to report intent to complete the 2020 Census. Education, marital status, and home ownership were all statistically significant, and in the direction one would expect. Those who were more educated, married, or owned their homes were more likely to express intent to complete the 2020 Census. Depending on the model, non-white race and speaking a language other than English at home were statistically significant. A couple of interactions between the demographic variables were statistically significant, the New England states showed a higher than average intent to complete the 2020 Census, and the West South Central states showed a less than average intent to complete the census.

4.2 Social-economic-political module and related variables

For the social-economic-political question module, attitudinal questions, and volunteer question, we found that the following political and economic factors were consistently associated with intent to complete the 2020 Census: trust in federal statistics, pessimistic feelings about the current job situation, fear responses could be used against them, and dissatisfaction with the current government (although only marginally statistically significant).

Those who said the job situation was good were less likely to say they intended to complete the 2020 Census, and those who said the jobs situation was bad were more likely to say they intended to complete the 2020 Census. This is consistent with our prior work and suggests that people see the value of government/official statistics more during a difficult job market. Those who said they did not trust federal statistics and who expressed concern about the 2020 Census being used against them were less likely to say they intended to complete the 2020 Census. Those who said they were satisfied with the way the country is governed were marginally less likely to say they intended to complete the 2020 Census, and those dissatisfied were more likely to say they intended to complete the 2020 Census.

And we found that the following social and emotional factors were consistently associated with intent to complete the 2020 Census: attended at least some religious services, contacted a government official during the past year, volunteered, try to see both sides of an issue, and being disturbed by others' misfortunes (although only marginally statistically significant).

Table 2: Key outcomes

Category	Factor	Intent ... Likely
Attitude government	Don't trust federal statistics	Less
	Fear census responses will be used against them	Less
Economic view	Good job situation	Less
	Economic condition	Not significant
Political view	Direction of country	Not significant
	Satisfied with way country is governed	Less
Civic engagement	Attended religious services	More
	Contacted a government official	More
	Volunteer	More
Attitude others	Don't strongly say try to see both sides of issue	Less
	Others' misfortunes don't disturb them	Less

Those who said they attended at least some religious services a year and those who said they contacted a government official in the past year were more likely to say they intended to complete the 2020 Census. Those who said they don't volunteer, were less likely to say they intended to complete the 2020 Census. Those who did not strongly say they try to see both sides of issues were less likely to say they intended to complete the census. And those who said others' misfortunes don't disturb them were marginally less likely to say they intended to complete the census. (For full model output statistics, see Appendix B.)

5. Discussion

Response rates, a necessary component of data quality, have continued to fall for many major government surveys. Said another way, although informed decision-making is a pillar of democracy, more and more of us have been making the conscious choice not to be represented in official statistics. Here we explored factors related to reported 2020 Census participation.

Our findings support a complex social-economic-political model of survey participation. In addition to the demographic, temporal and geographic variables we expected would be associated with intent to complete the 2020 Census, there seemed to be four overarching categories of intent to complete: lack of trust in federal statistics/government, satisfaction with the economic/political condition, and not seeing things from others' perspective were associated with being less likely to report intent to participate in the 2020 Census, while civic/community engagement was associated with being more likely to report intent to participate in the 2020 Census. This suggests a lack of participation from those who report some type of complacency or mistrust, but higher participation among those who are civically engaged.

When interpreting the results, one should keep in mind that the survey respondents were anonymized, so we do not know if they actually participated in the 2020 Census. To validate our findings, ideally, we would have followed this up with a study of whether the same people who expressed intent to participate are the same as those who participated. And while the results here generally support previous findings, it might be that the results do not extend to other countries or other types of government surveys. Along these lines, it would be good to replicate these findings in other settings.

Finally, what if anything is actionable based on these findings? We suggest that trust in official statistics is not manageable at the level of the statistical agency (aside, perhaps, from offering monetary incentives for survey response, which could introduce a different kind of bias). Instead, it would require a larger paradigm shift only obtainable through education on the importance of representation and the right to information.

Acknowledgments

The authors would like to thank our reviewers at the Census Bureau, as well as the 2020 Census Tracking Survey design team for supporting this research.

References

- Amaya, A. (2015). *Enhancing the Understanding of the Relationship between Social Integration and Nonresponse in Household Surveys*. [Doctoral dissertation, University of Maryland]. <https://doi.org/10.13016/M29D90>
- Bauer, P. C., Keusch, F., & Kreuter, F. (2019). Trust and Cooperative Behavior: Evidence from the realm of data-sharing. *PLoS ONE*, 14(8): e0220115. <https://doi.org/10.1371/journal.pone.0220115>
- Census Bureau (2019). *2020 Census Barriers, Attitudes, and Motivators Study Survey Report: A New Design for the 21st Century*. <https://www2.census.gov/programs-surveys/decennial/2020/program-management/final-analysis-reports/2020-report-cbams-study-survey.pdf>

- Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, Phone, Mail, and Mixed Mode Surveys: The tailored design method* (4th ed). Hoboken, NJ, US: John Wiley & Sons Inc.
- Groves, R., Singer, E., & Corning, A. (2000). Leverage-Saliency Theory of Survey Participation: Description and an illustration. *Public Opinion Quarterly*, 64(3): 299–308.
<https://doi.org/10.1086/317990>
- Harris-Kojetin, B. & Tucker, C. (1999). Exploring the Relation of Economic and Political conditions with Refusal Rates to a Government Survey. *Journal of Official Statistics*, 15(2), 167-184.
<https://www.scb.se/contentassets/ca21efb41fee47d293b-bee5bfbe7bf3/exploring-the-relation-of-economic-and-political-conditions-with-refusal-rates-to-a-government-survey.pdf>
- Larsen, L. J., Lineback, J. F., & Reist, B. M. (2020). Continuing to Explore the Relation between Economic and Political Factors and Government Survey Refusal Rates: 1960–2015. *Journal of Official Statistics*, 36(3), 489-505. <https://doi.org/10.2478/jos-2020-0026>
- Singer, E. (2011). Toward a Benefit-Cost Theory of Survey Participation: Evidence, Further Tests, and Implications. *Journal of Official Statistics*, 27(2): 379–392.
<https://www.scb.se/contentassets/ca21efb41fee47d293bbe5bf7be7fb3/-toward-a-benefit%20cost-theory-of-survey-participation-evidence-further-tests-and-implications.pdf>

Appendix A – Key study questions

For purposes of this work, the outcome we were interested in was the intent to participate/having participated in the 2020 Census. Before the 2020 Census was fielded, the Tracking Survey asked about intent to participate this way:

- *How likely are you to participate in the 2020 Census? By participate we mean answer the questions on the census form. Would you say you definitely will, probably will, might or might not, probably will not, or definitely will not participate?*

After the Census was fielded, the 2020 Census Tracking Survey asked about Census participation this way:

- *Have you or someone in your household answered the 2020 Census questions, or has your household not answered them yet?*

If the answer was no, then the respondent was asked about intent to participate. Intent to participate and having participated were treated as positive outcomes.

For purposes of our study, eight questions were included on the 2020 Census Tracking Survey to examine the effect of various aspects of one's perception of their environment on intent to participate in the 2020 Census. Topics included the state of the current political and economic environments, civic engagement, and attitudes toward others. We also used existing questions about volunteerism and attitudes toward the government.

Questions about the economic environment were asked this way:

- *How would you rate economic conditions in this country today – as excellent, good, only fair or poor?*
- *Thinking about the job situation in America today, would you say that it is a good time or a bad time to find a quality job?*

Questions about the political environment were asked this way:

- *Thinking about the way things are going in the country today, do you think things are generally going in the right direction or the wrong direction?*
- *On the whole, would you say you are satisfied or dissatisfied with the way the nation is being governed?*

Questions about civic engagement were asked this way:

- *Aside from weddings and funerals, how often do you attend religious services... more than once a week, once a week, once or twice a month, a few times a year, seldom, or never?*
- *Please tell me whether or not in the last 12 months you have contacted or visited a public official-at any level of government-to express your opinion?*

Questions about one's attitude toward others were asked this way:

- *Do you believe that there are two sides to every question and try to look at them both?*
- *Do other people's misfortunes usually disturb you a great deal?*

Finally, there were two questions about one's attitude toward the government and one on volunteerism that were already being included for other purposes that we also used for our study:

- *Personally, how much trust do you have in the federal statistics in the United States? Would you say that you tend to trust federal statistics or tend not to trust them?*
- *How concerned are you, if at all, that the answers you provide to the 2020 Census will be used against you?*
- *We are interested in volunteer activities, that is, activities for which people are not paid, except perhaps expenses. We only want you to include volunteer activities that you did through or for an organization, even if you only did them once in a while. Since **[current month] 1st** of last year, have you done any volunteer activities through or for an organization?*

Appendix B – Full regression model estimates and fit statistics

<u>Effect</u>	<u>Point Estimate</u>	<u>95% Wald</u>	
		<u>Confidence Limits</u>	
AGE	1.023	1.02	1.026
RACE_NONHISPWHITE_m	0.752	0.577	0.978
LANG_ENG_m	1.721	0.747	3.964
MARITAL_STATUS_GROUP	0.637	0.393	1.031
OWN_RENT_GROUP_m	0.55	0.432	0.702
EDU1_m	1.199	0.645	2.227
TRACKINGMO	1.222	1.196	1.25
DIVISION_1	0.801	0.627	1.023
DIVISION_2	0.963	0.807	1.148
DIVISION_3	1.125	0.943	1.342
DIVISION_4	0.911	0.733	1.131
DIVISION_6	0.979	0.781	1.227
DIVISION_7	0.825	0.694	0.981
DIVISION_8	0.849	0.696	1.034
DIVISION_9	0.838	0.713	0.984
DIVISION_m	0.659	0.529	0.821
TRUST_FED_STATS_notr	0.506	0.457	0.559
TRUST_FED_STATS_m	0.675	0.565	0.808
USED_AGAINST_SomeCon	0.736	0.668	0.81
USED_AGAINST_m	0.453	0.313	0.656
CPS_VOL_no	0.699	0.627	0.78
CPS_VOL_m	0.318	0.201	0.502
CIVIC1_EconFair	0.935	0.836	1.046
CIVIC1_m	0.973	0.699	1.354
CIVIC2_GoodTimeJobs	0.74	0.661	0.827
CIVIC2_m	0.719	0.59	0.876
CIVIC3_RightDirectio	1.026	0.906	1.163
CIVIC3_m	0.919	0.774	1.09
CIVIC4_Satisfied	0.823	0.727	0.931
CIVIC4_m	0.692	0.571	0.838
CIVIC5_LittleEmpathy	0.839	0.762	0.925
CIVIC5_m	0.75	0.525	1.071
CIVIC6_MisfortuneDis	0.817	0.74	0.903
CIVIC6_m	0.766	0.574	1.021
CIVIC7_Religious	1.332	1.196	1.484
CIVIC7_m	1.064	0.719	1.575
CIVIC8_ContactGovtYe	1.776	1.549	2.037
CIVIC8_m	1.64	0.785	3.427

Association of Predicted Probabilities and Observed Responses

Percent Concordant	77.3	Somers' D	0.545
Percent Discordant	22.7	Gamma	0.545
Percent Tied	0	Tau-a	0.087
Pairs	49411880	c	0.773