

Willingness of the public to share geolocation information in a U.S. Census Bureau survey

Elizabeth Nichols and Erica Olmsted-Hawala
U.S. Census Bureau

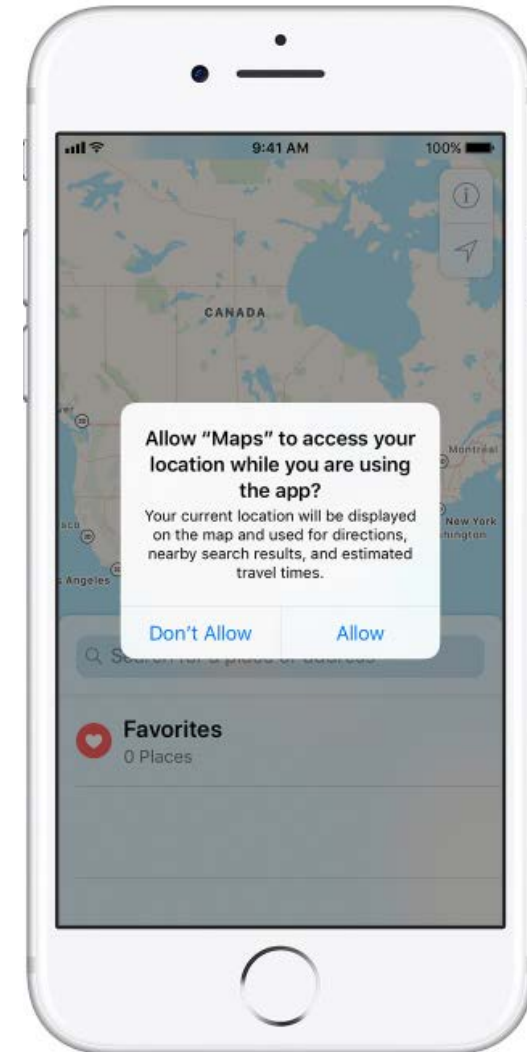
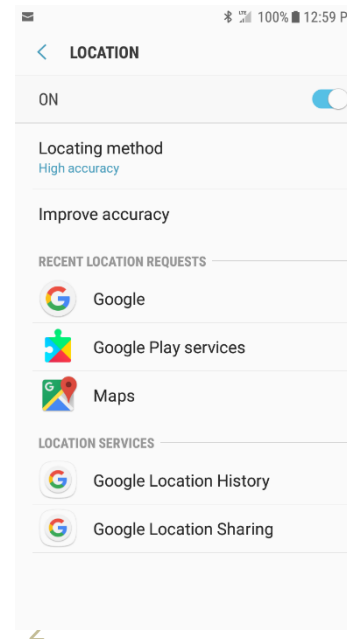
FedCASIC

April 18, 2018

U.S. Census Bureau, Suitland, Maryland

Accessing location information on mobile

- Cell phone location can be determined from:
 - Global positioning system (GPS) data
 - Cell site location information (CSLI)
- Location information includes: Latitude and Longitude and elevation
- A smartphone owner turns on the location in settings.
- Sites ask permission to use this information if it is turned on



Motivation for studying location data sharing

- Household surveys and censuses are location based
- Using technology to provide location information could reduce the burden of answering address questions and possibly increase the accuracy of such location data
- Limited research on either topic
 - Crawford, S. (2017); Gruteser & Liu, 2004; Huang, Matsuura, Yamane & Sezaki, 2005:

Research Questions

- What percent of respondents will share location information in a survey sponsored by the U.S. Census Bureau?
- Does the design or presentation of the location request make a difference in whether respondents will share?
 - How the request is asked – replicating Crawford’s research to an extent
 - Where the request is presented – at the beginning of the survey or at the end of the survey
- How accurate is the location information?

General Method

- Census Bureau nonprobability research panel frame
- 2,000 emails selected
 - 1,285 of them had never been selected for a Census Bureau research study, and 715 had completed an earlier study
- Split panel experiment: 2x2 (2 factors each with 2 conditions)
- Two week data collection – May 16- May 28, 2016
- Email notification (up to 3) with a link to the survey
 - Email stated to use a smartphone or other mobile device with GPS enabled – not a desktop or laptop computer
 - Best if you answer at your home address
 - 5-minute survey

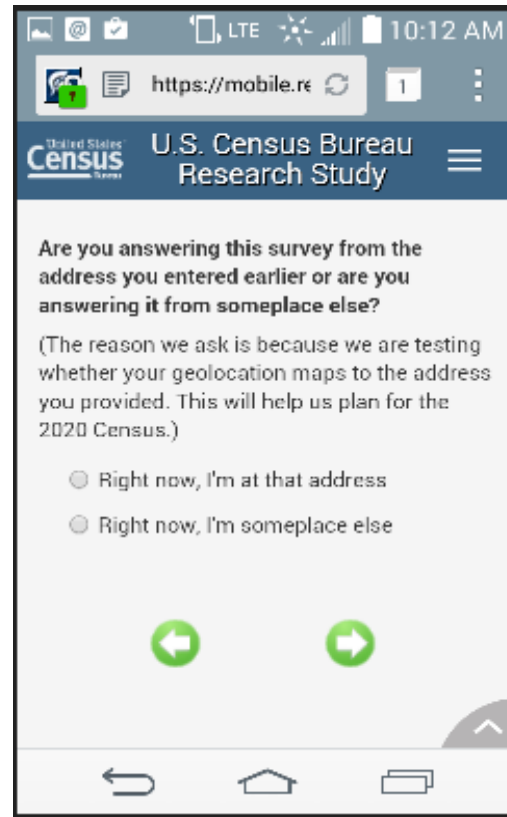
Split-panel Factors

- “How” Factor: How the geolocation request was asked
 - Control: Phone pop-up only
 - Experimental: Survey question asking for permission and then, if granted, the phone pop-up
- “Where” Factor: Where the geolocation request was presented
 - Control: Beginning of the survey
 - Experimental: At the end of the survey

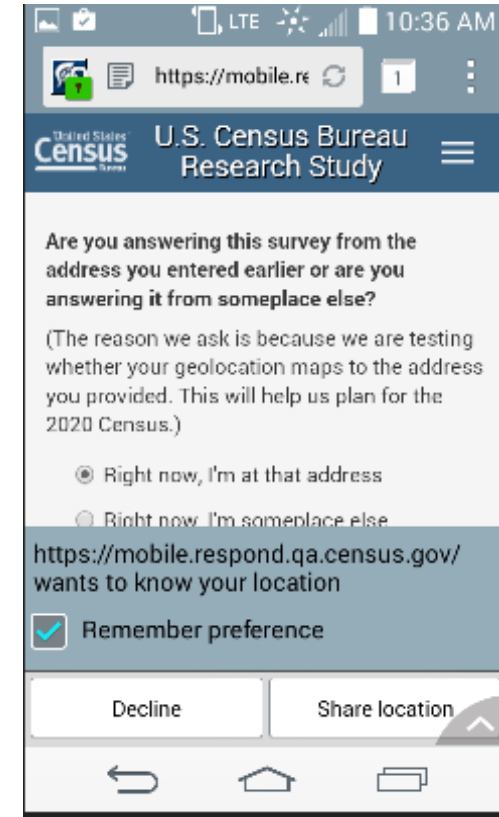
Data collected in the survey

- General demographics of participant
- Some opinion data
- Home address
- Whether respondent was at home when answering survey
- Geolocation data
- Device used (this was from paradata)

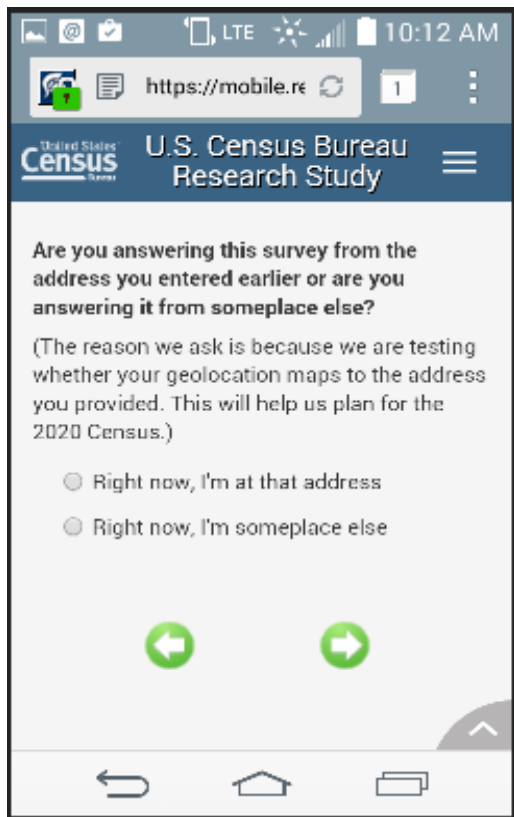
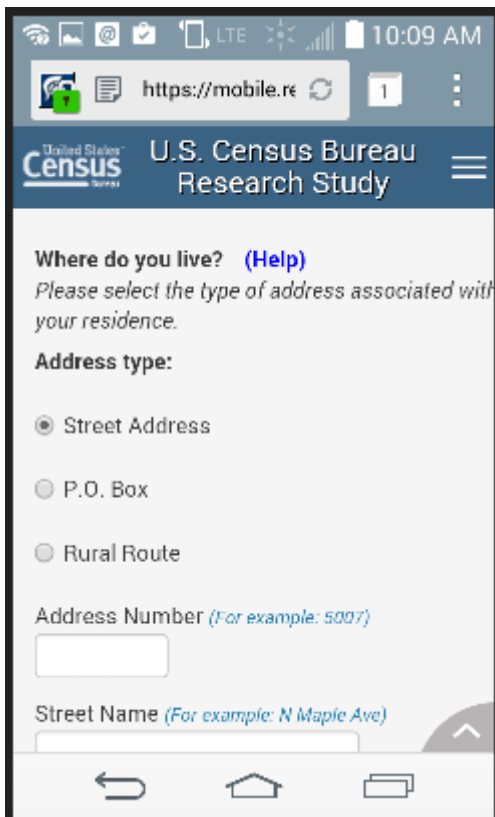
Control: “How” Factor



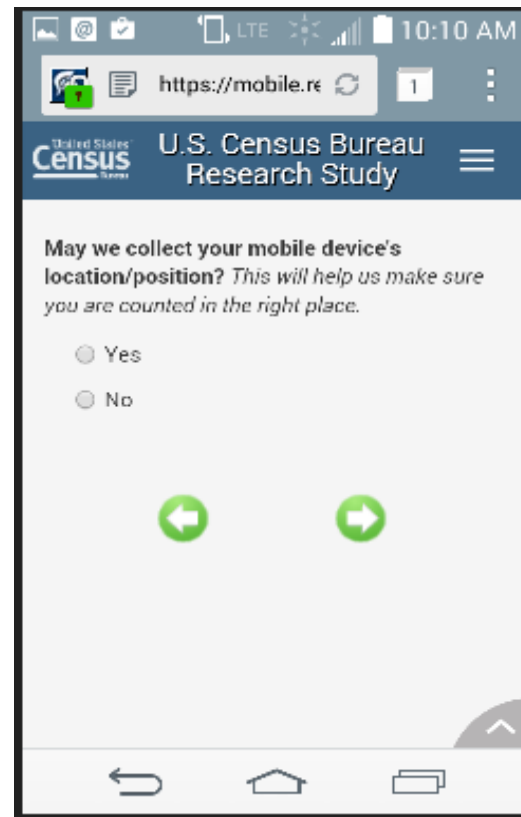
If at home
→



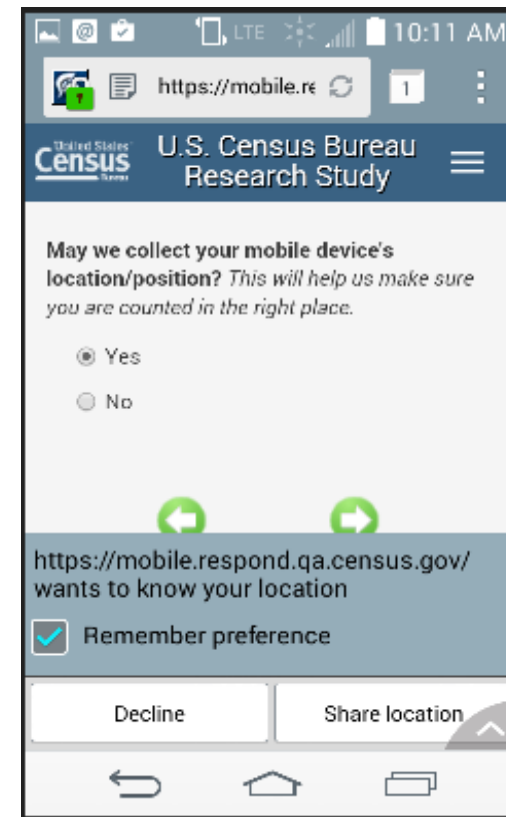
Experimental: “How” Factor



If at home
→



If yes
→



Control: “Where” Factor

Login

Address and geolocation



Demo and opinion

Experimental: “Where” Factor

Login

United States Census Bureau
U.S. Census Bureau Research Study

Please Log In

The email address you enter must match the email address where you received the survey invitation.

E-mail Address:

Login

Note: You may only complete the survey one time per invitation. Once you enter the survey you will not be able to reenter later and the survey will time-out after 15 minutes of inactivity.

Your responses to these questions are protected by law under Title 13 of the United States Code. This is the same law that protects your privacy. No one outside the Census Bureau will see your answers.

Demo and opinion

United States Census Bureau
U.S. Census Bureau Research Study

1. Are you male or female?

Male

Female

2. What is your age?

Less than 18 years old

18-24

25-34

35-44

45-54

55-64

65 or older

High School Diploma or GED

Address and geolocation



United States Census Bureau
U.S. Census Bureau Research Study

Where do you live? (Help)

Please select the type of address associated with your residence.

Address type:

Street Address

P.O. Box

Rural Route

Address Number (For example: 5007)

Street Name (For example: N Maple Ave)

Decline Share location

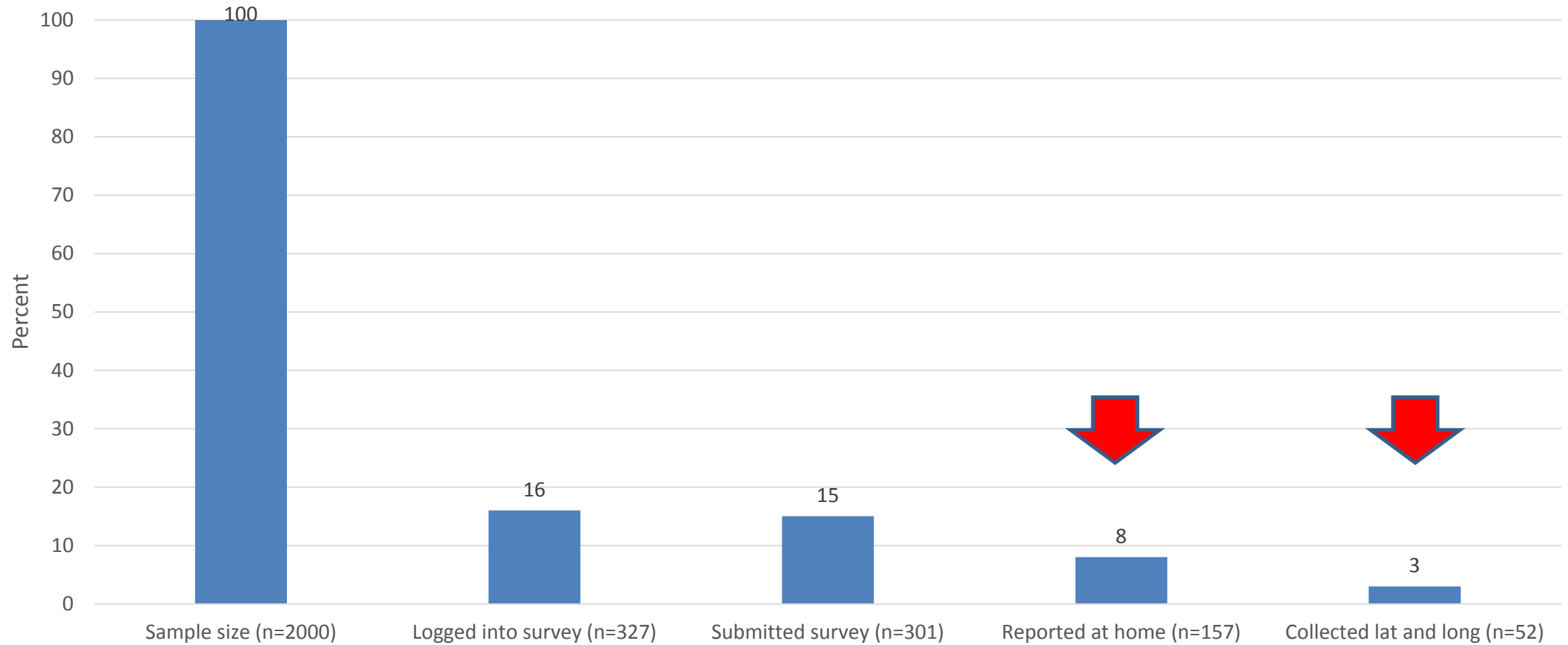
Split Panel

	“Where” Control	“Where” Experimental
“How” Control	500	500
“How” Experimental	500	500

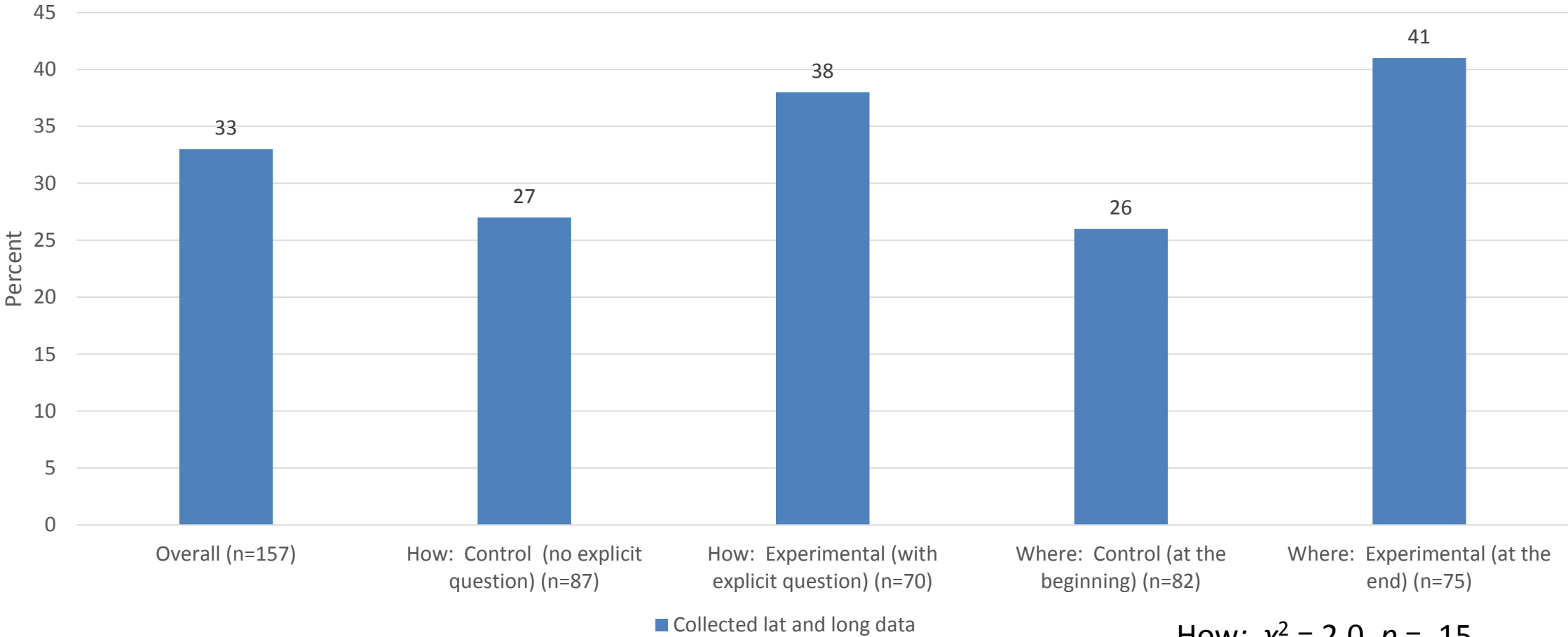
Analysis Methods

- Will respondents share location information in a survey sponsored by the U.S. Census Bureau? ***Percent overall and by factor***
- Does the design or presentation of the location request make a difference in whether respondents will share? ***Chi-square and logistic regression***
- How accurate is the location information? ***Independent geolocation conducted by Census Bureau Geography Division staff with comparison to latitude and longitude collected***

Response rate and sharing rate



Will respondents share geolocation data



How: $\chi^2 = 2.0, p = .15$
Where: $\chi^2 = 4.4, p = .04$

Does how the geolocation request is made matter?

- Answer: Perhaps
- Adding an explicit questions asking for permission (in addition to the default phone pop-up question)
 - Trend is for increased the sharing of geolocation coordinates however, significance levels are borderline
 - Overall chi-square test: $\chi^2 = 2.0, p = .15$
 - Model results:
 - Asking an explicit question to collect coordinates versus not asking an explicit question
 - Odds ratio = 2.0, $p = .08$
 - Survey breakoffs were not impacted by whether there was an explicit geolocation question
 - $\chi^2=0.18, p=0.67$

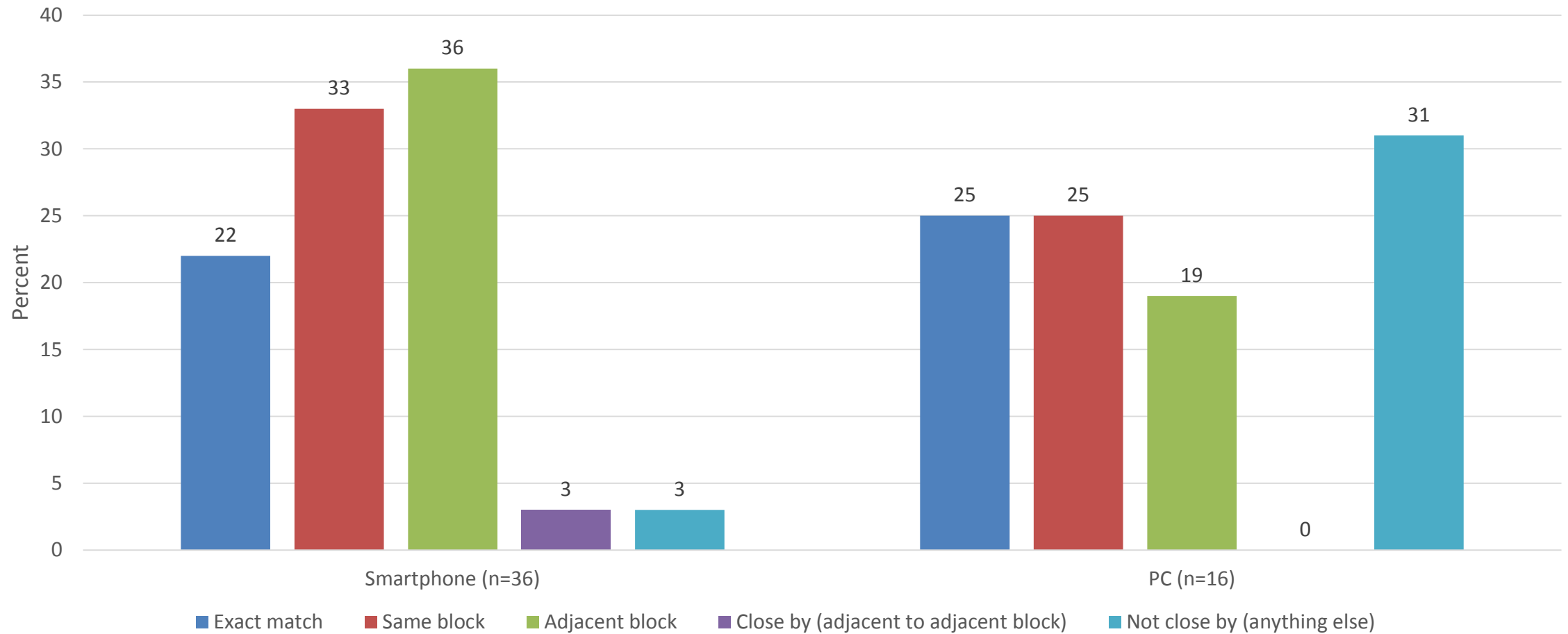
Does where the geolocation request is made matter?

- Answer: Yes
- Requesting address data and geolocation data at the end of the survey
 - Significantly increased the sharing of geolocation coordinates
 - Overall chi-square test: $\chi^2 = 4.4, p = .04$
 - Model results:
 - Placing the address and geolocation question at the end of the survey versus early in the survey
 - Odds ratio = 2.1, $p = .05$
 - Significantly decreased survey break-offs
 - Overall chi-square test: $\chi^2=5.2, p=0.02$

Covariates in model

- Respondents without a college degree were less willing to share their location data than respondents with a college degree ($p=0.05$);
- White non-Hispanic respondents were marginally more willing to share their location data than non-white respondents ($p=0.10$); and
- There was no detectable difference at the 95 percent confidence level in the willingness to share location data by sex, age, type of device used, or by whether they had previously participated in a Census Bureau research study.

How accurate is the information



Summary of Results

- 1/3 of respondents were willing to share geolocation data within the survey.
- Placing the address question and the geolocation request at the end of the survey increases both survey completes and the sharing of geolocation data.
- We found no downside to adding an explicit question about the geolocation request, and some marginal increase in the sharing of geolocation coordinates
- We found some differences in sharing the data by race and education, but we did not find any difference by age, sex, type of device, or whether the respondent had participated in a survey with us previously.
- Half of the geolocation data gathered identified the address in the correct block.
- About 1/3 of geolocation gathered from a PC was very far from the actual location.

Future Work

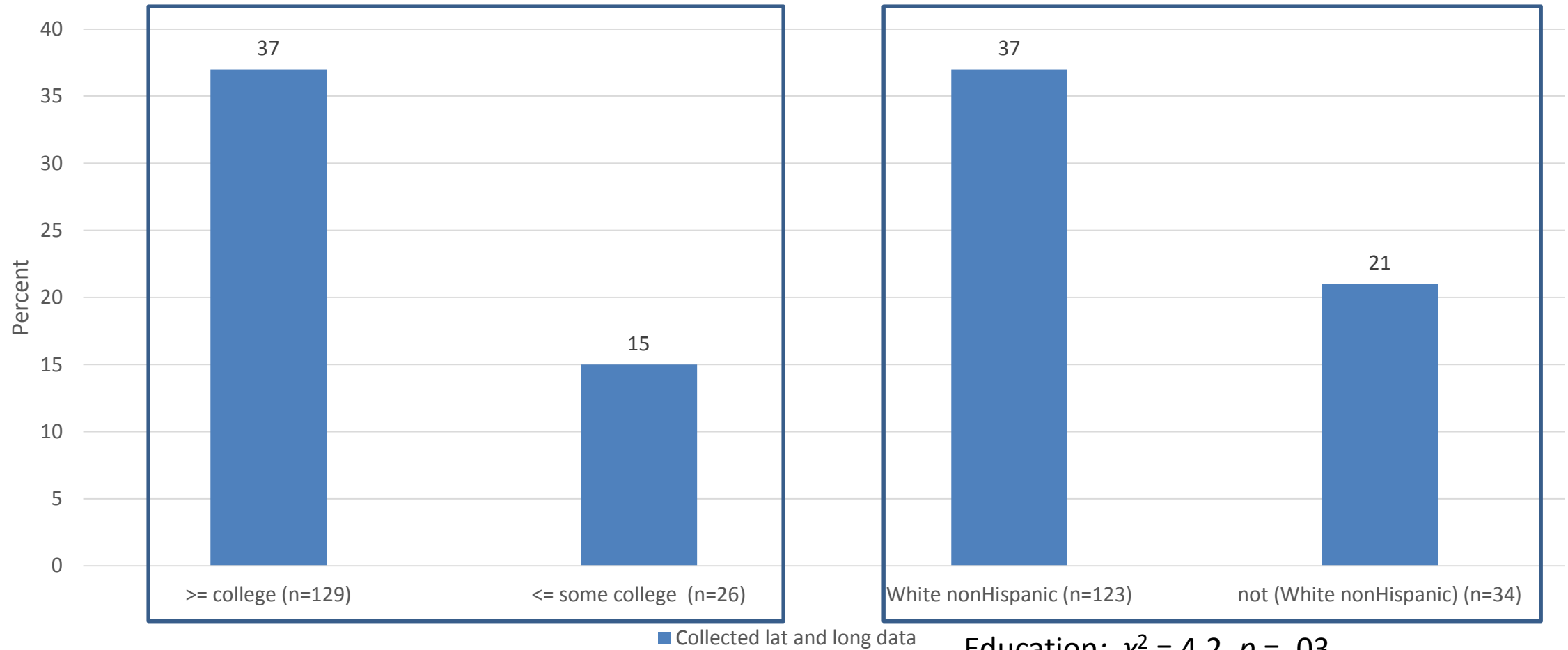
- Repeat experiment with different population
- Determine sample size needed to confirm differences in main effects and covariates
- Determine if there are any covariates that can help determine the accuracy of the geolocation coordinates to the block level.

Questions

Elizabeth Nichols

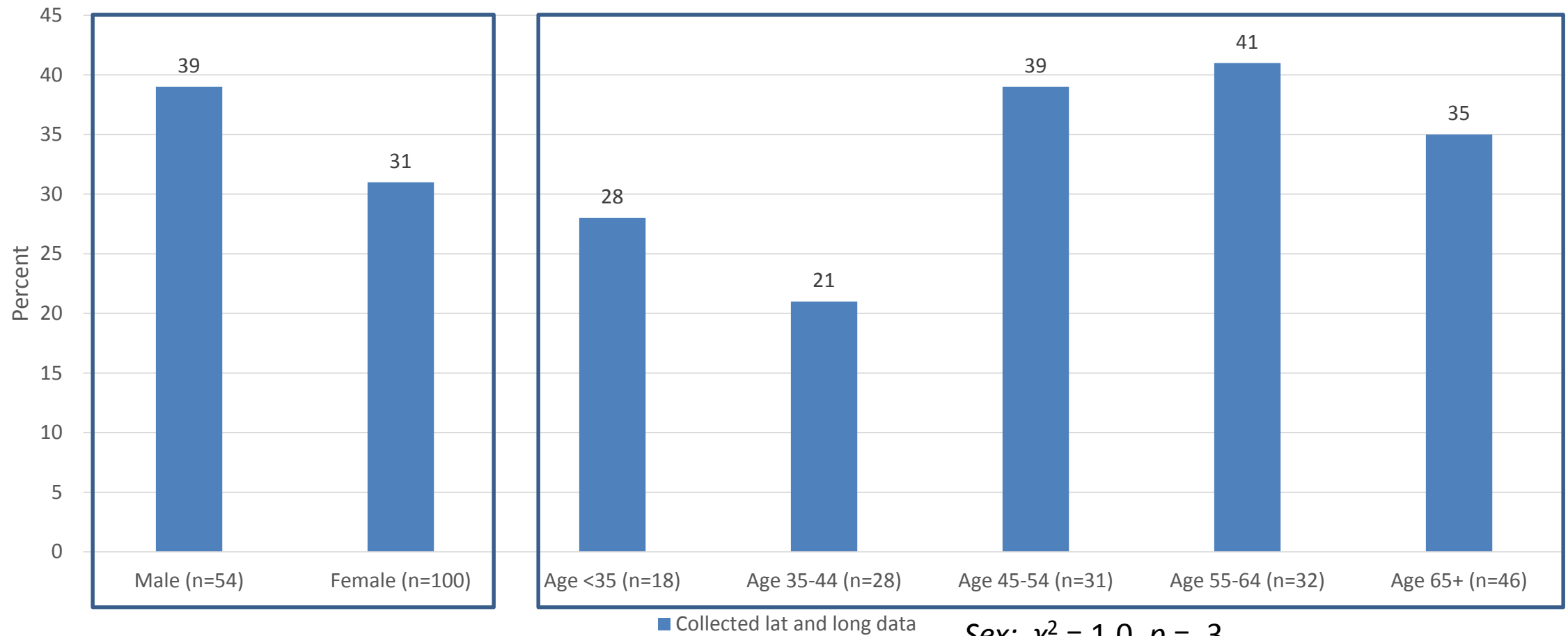
Elizabeth.May.Nichols@census.gov

Addendum: Education and race were significant in model



Education: $\chi^2 = 4.2, p = .03$
Race: $\chi^2 = 3.1, p = .08$

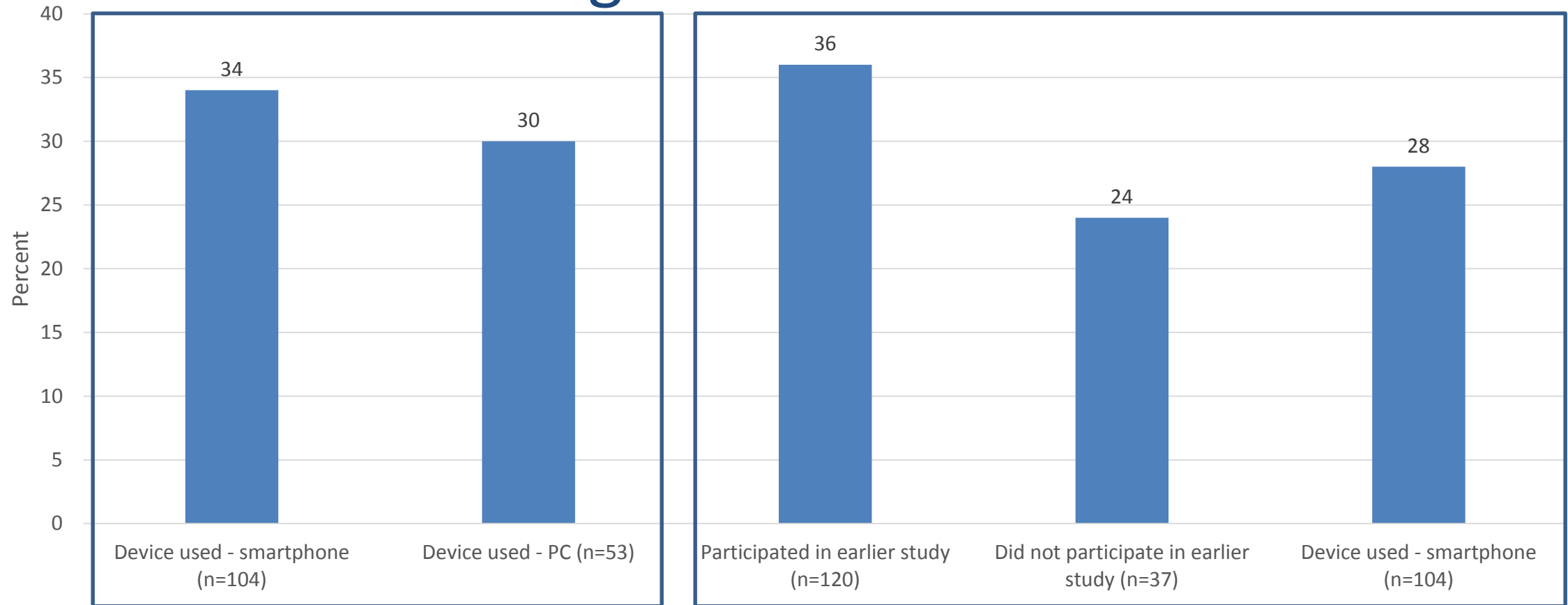
Addendum: Sex & age were not significant at 95% CI



Sex: $\chi^2 = 1.0, p = .3$
Age: $\chi^2 = 3.2, p = .5$

Addendum:

Whether they participated earlier & device were not significant at 95% CI



■ Collected lat and long data

Device: $\chi^2 = .3, p = .6$

Participated before: $\chi^2 = 1.7, p = .2$