Assessing the "Year of Naturalization" Data in the American Community Survey

Characteristics of Naturalized Foreign Born Who Report – and Don't Report – the Year They Obtained Citizenship

Elizabeth M. Grieco and David M. Armstrong
Population Division
U.S. Census Bureau

Presented at the Applied Demography Conference
San Antonio, TX, January 8-10, 2014

Introduction

- American Community Survey includes a question asking foreign-born citizens: "Print year of naturalization."

- Not all foreign born who have naturalized provide the year they became citizens

- Analysis uses logistic regression to determine the characteristics of those who report – and fail to report – a year of naturalization

Survey Nonresponse

Two types of nonresponse behavior:

- Unit nonresponse
  - When the sample unit does not respond to the request to be surveyed

- Item nonresponse
  - Failure to obtain answers to individual survey questions when unit response is obtained

This paper examines item nonresponse by foreign-born citizens to the year of naturalization question.

Results from the Literature

- Literature emphasizes influence of individual characteristics on item nonresponse:
  - Age, sex, education, socioeconomic status
- Household and neighborhood characteristics also considered:
  - Population density, crime rates, household composition, household size, language use, geographic location, culture
- Survey mode also noted:
  - Item nonresponse higher in mail modes when compared with telephone or face-to-face interviewing

United States™
Census
Bureau

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU

4

Analytical Framework

- The ACS is a survey of the resident population
  - The sample is divided into group quarters and household
  - In the household sample
    - Housing units are selected into sample
    - Data collected on each individual who is part of the household
- Information for most household members provided by proxy through primary respondent
  - Most often, primary respondent is the householder
- How/what information moves through the household from each household member to the householder and then to Census Bureau depends on:
  - Survey mode
  - Social distance from primary respondent

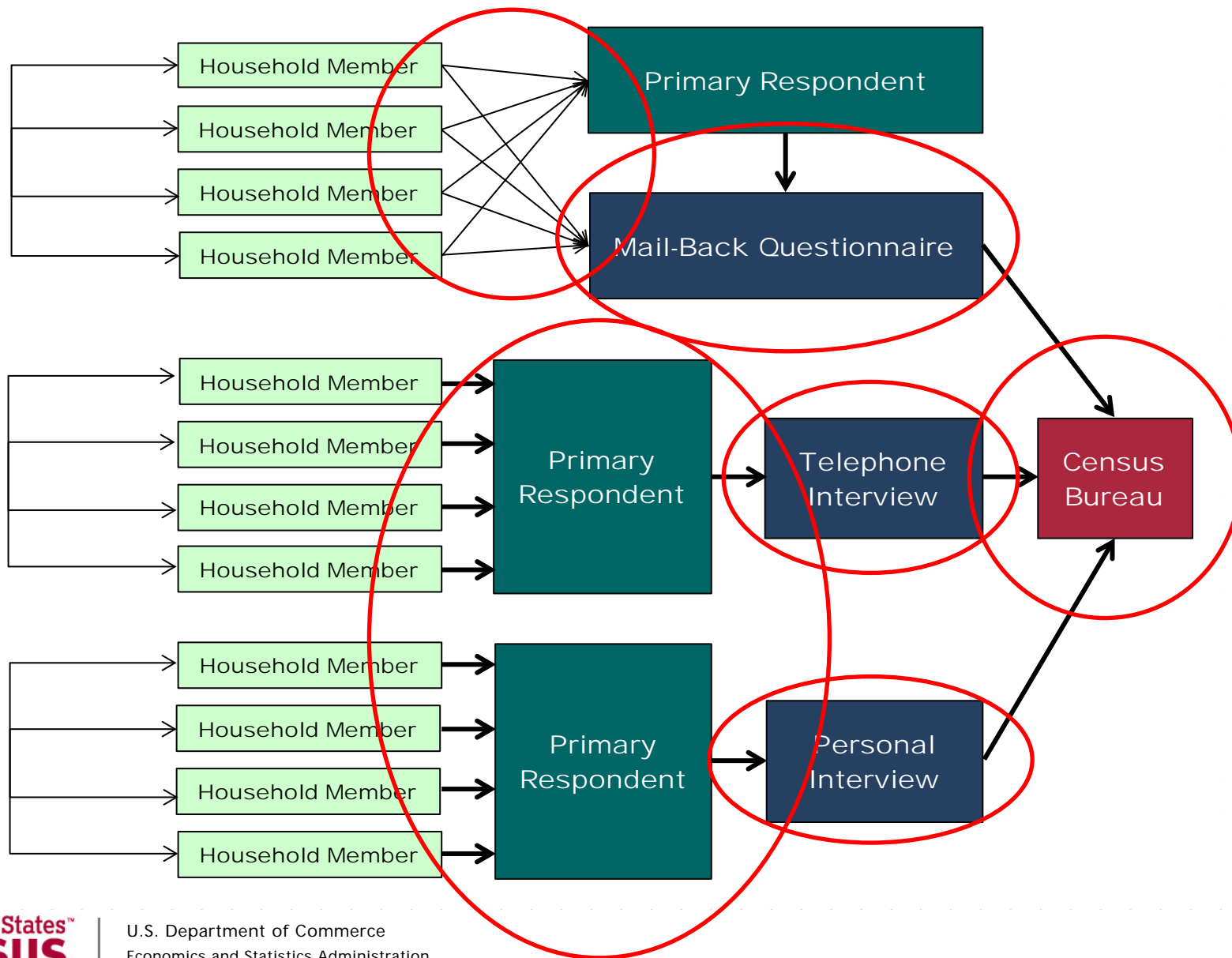# Figure 1. Flow of Information From Household Members to the U.S. Census Bureau
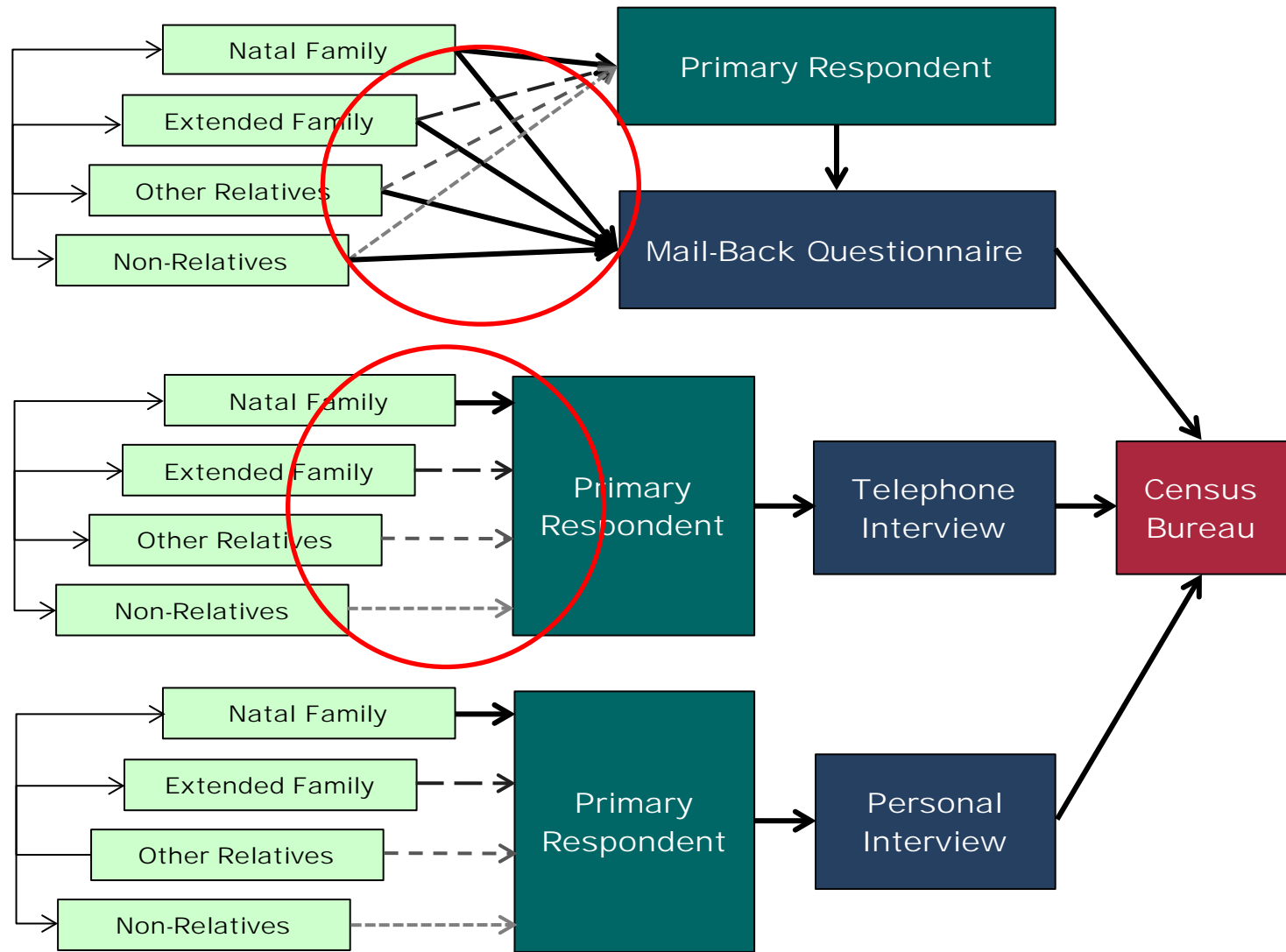
Figure 2. Flow of Information From Household Members to the U.S. Census Bureau, Showing Social Distance From the Primary Respondent

Statistical Model

Data

- 2011 American Community Survey

Statistical Method

- Logistic regression analysis

Universe

- Naturalized citizens in the United States aged 18 and over, where citizenship status was not imputed
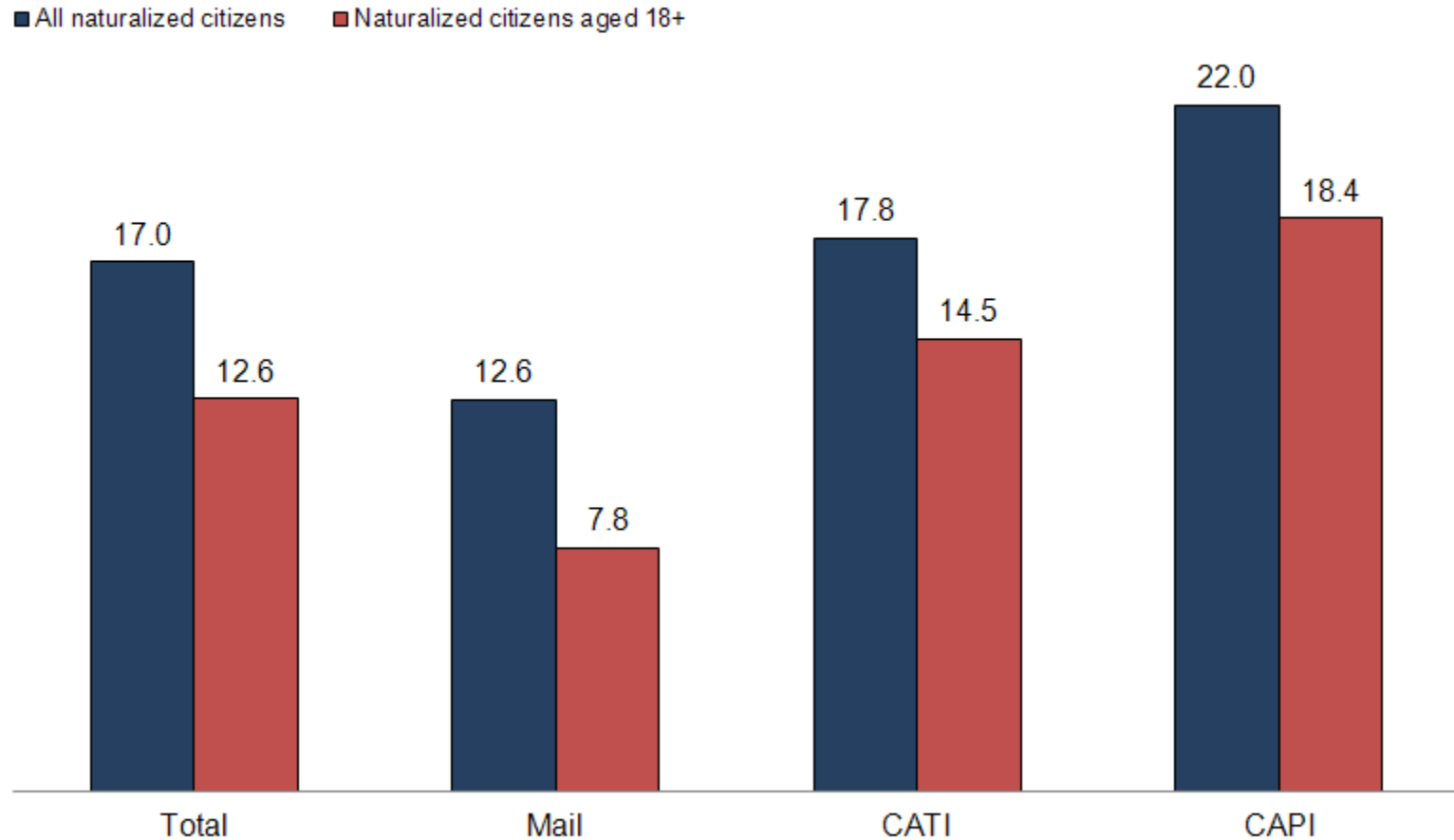
Dependent Variable

- Dichotomous variable based on year of naturalization
  - Those who <u>reported</u> a year
  - Those who did not report a year

Statistical Model

Primary Explanatory Variables

- Survey environment
  - Operationalized using mode
  - 3-category dummy variable
  - Mail (reference), CATI, and CAPI
    - CATI = Computer Assisted Telephone Interview
    - CAPI = Computer Assisted Personal Interview
- Social distance from primary respondent
  - Operationalized using relationship
  - 5-category dummy variable
  - Householder (reference), natal family, extended family, other relative, nonrelative
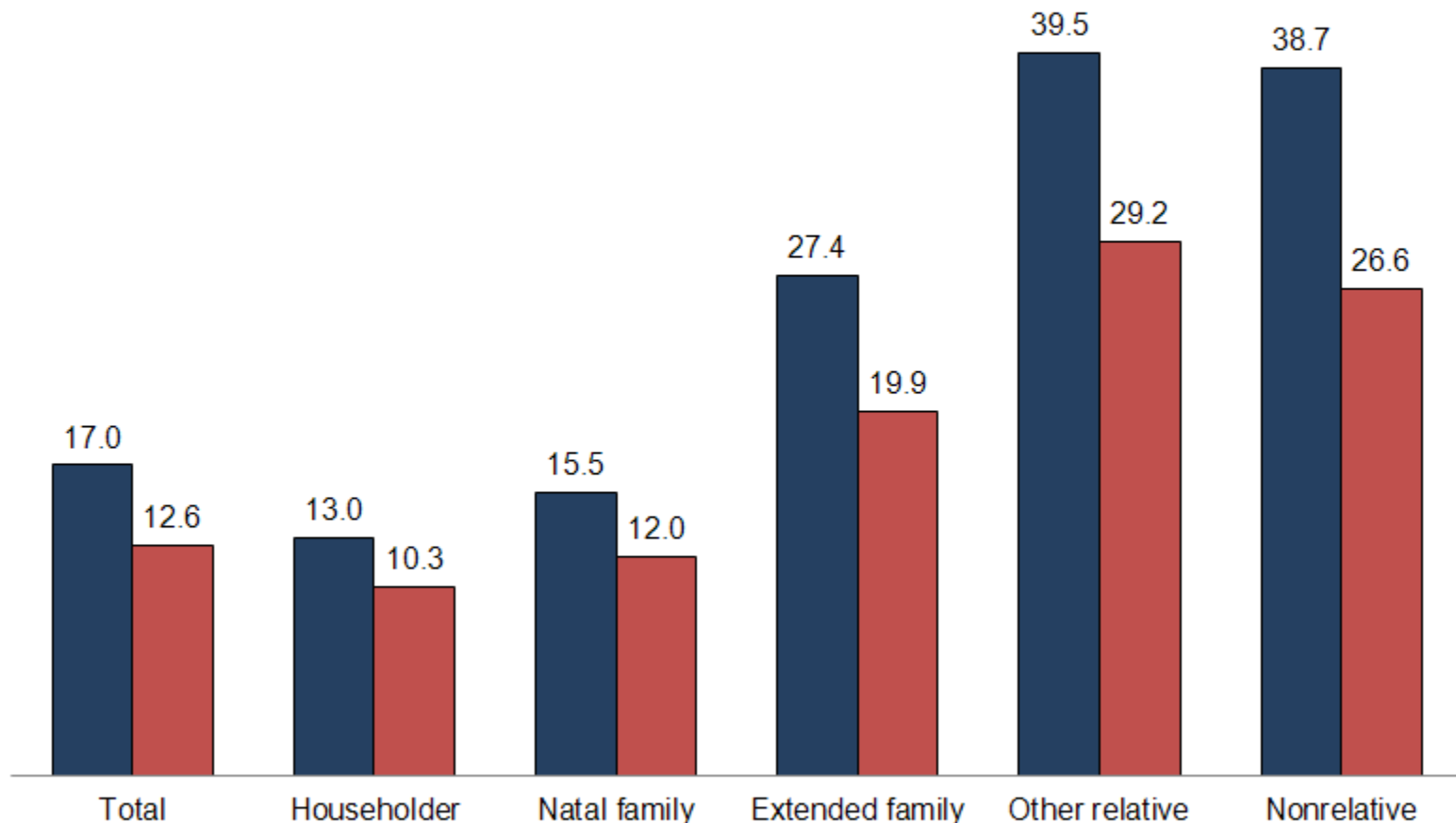
Figure 3.
Item Nonresponse Rate for Naturalized Citizens Reporting Year of Naturalization by Mode: 2011

■ All naturalized citizens   ■ Naturalized citizens aged 18+

Note: CATI=Computer Assisted Telephone Interview. CAPI=Computer Assisted Personal Interiew.
Source: U.S. Census Bureau, 2011 American Community Survey.

Figure 4.
Item Nonresponse Rate for Naturalized Citizens Reporting Year of Naturalization by Relationship to Householder: 2011

■ All naturalized citizens   ■ Naturalized citizens aged 18+

	Total	Householder	Natal family	Extended family	Other relative	Nonrelative
All naturalized citizens	17.0	13.0	15.5	27.4	39.5	38.7
Naturalized citizens aged 18+	12.6	10.3	12.0	19.9	29.2	26.6

Source: U.S. Census Bureau, 2011 American Community Survey.

Statistical Model

Household and Individual Control Variables

- Household Level
  - Characteristics of the <u>household</u> in which naturalized citizen resides
  - Number of adults in household, home ownership, poverty status, region of residency
  - Characteristics of the householder: educational attainment, language spoken at home, and nativity

- Individual Level
  - Characteristics of the <u>naturalized citizen</u>
  - Age, sex, region of birth, period of entry, moved in the last year (internationally or domestically)

Statistical Model

Nonresponse Control Variables

- To control for correlation of nonresponse between the dependent and each explanatory variable
  - For example: Are respondents who don't report year of naturalization also likely not to report year of entry?
- Alternative to excluding cases with some missing data
- Variables included:
  - Relationship, householder's level of education, householder's language spoken at home, householder's nativity status, homeownership, poverty status, age, sex, place of birth, year of entry, and moved in last year
  - No nonresponse control for mode, number of adults in household, or region of residency
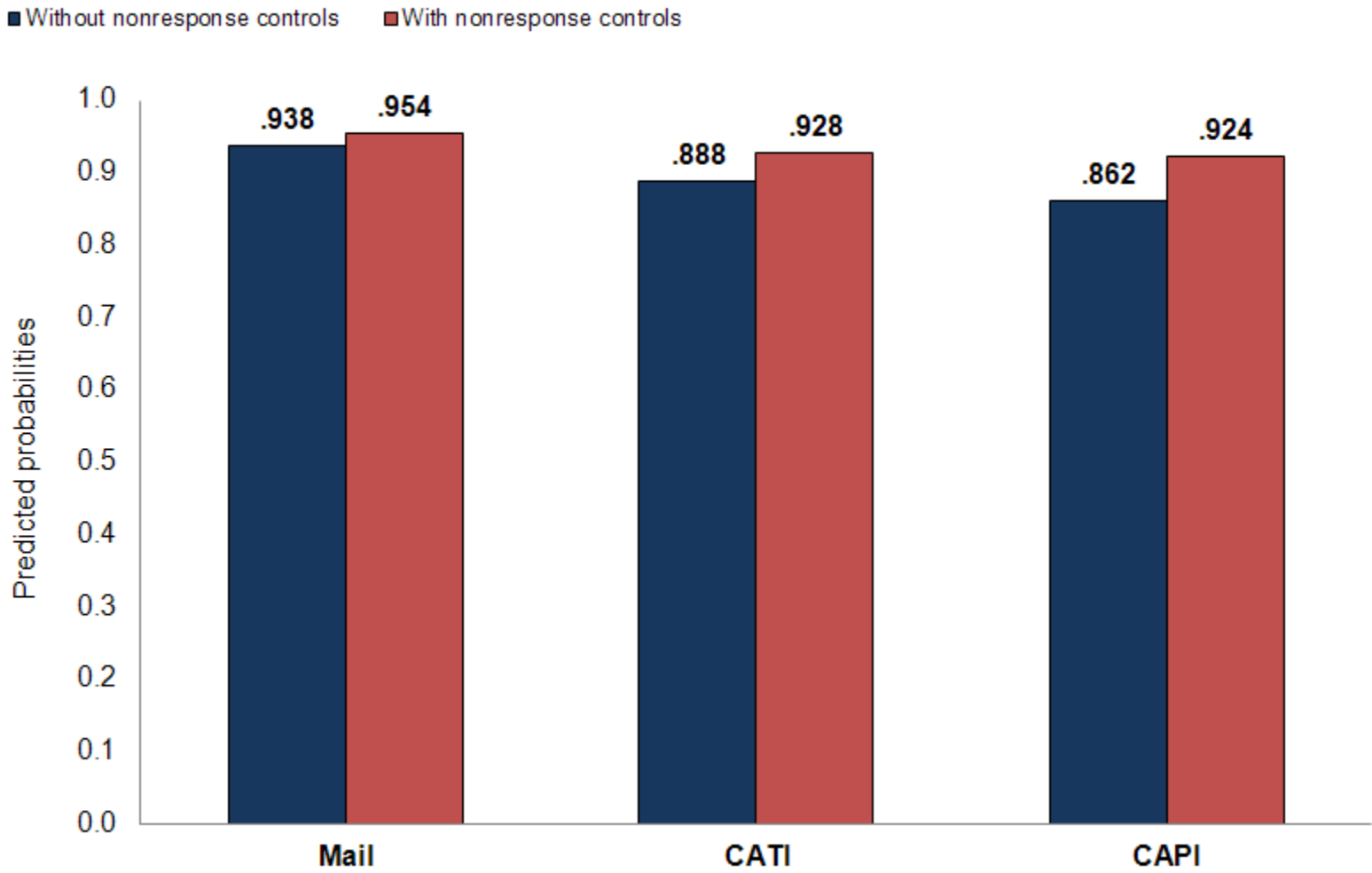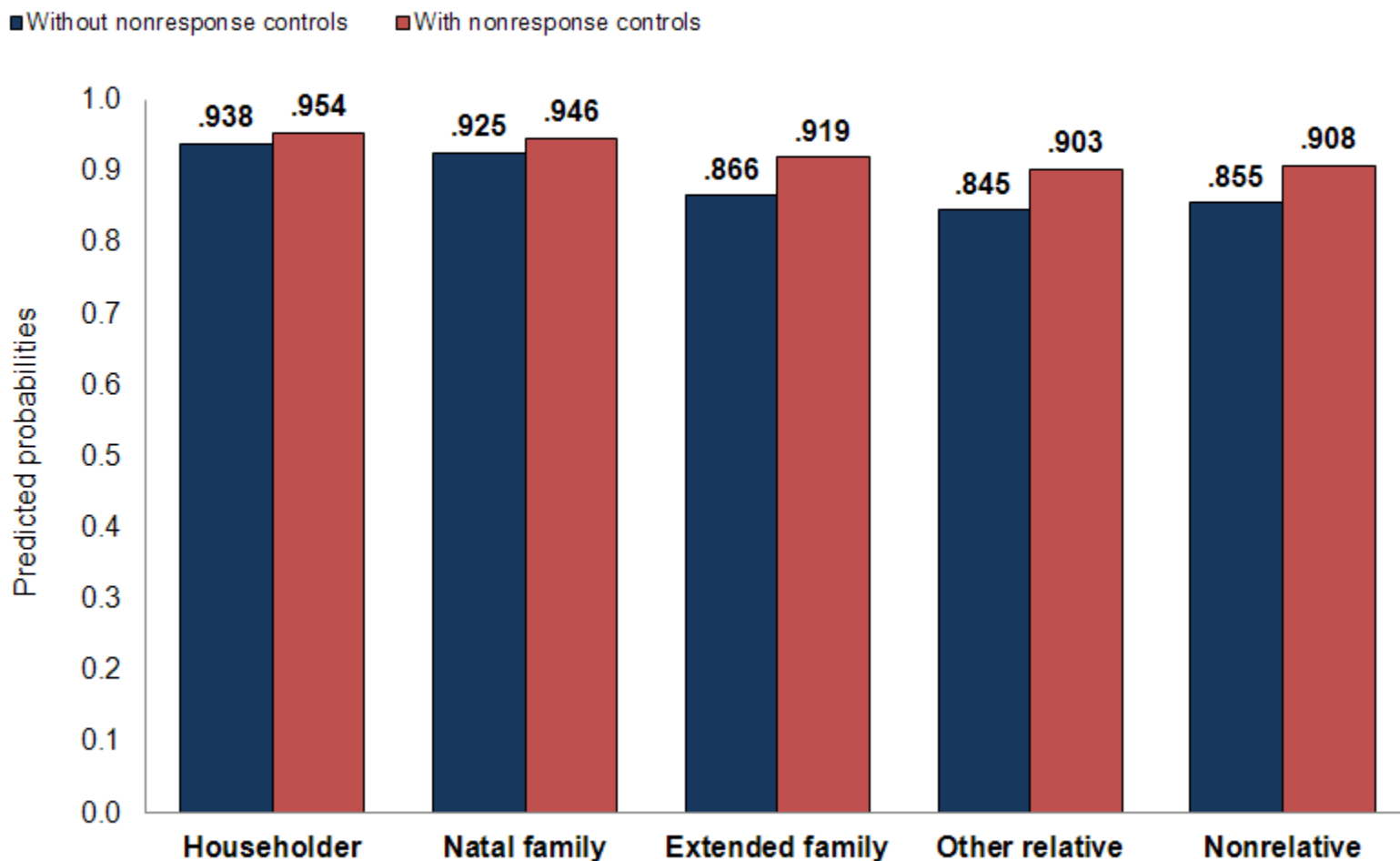
Results

Summary

- Both mode and relationship to householder significantly influence the likelihood of reporting a year of naturalization

- Most household and individual control variables behave as predicted

  - Exceptions: householder's language spoken at home and householder's nativity

- Nonresponse controls alter likelihoods associated with householder's education, region of residence, number of adults, age, sex, and period of entry

Figure 5.
Likelihood of Reporting Year of Naturalization by Mode

Note: CATI=Computer Assisted Telephone Interview. CAPI=Computer Assisted Personal Interview.
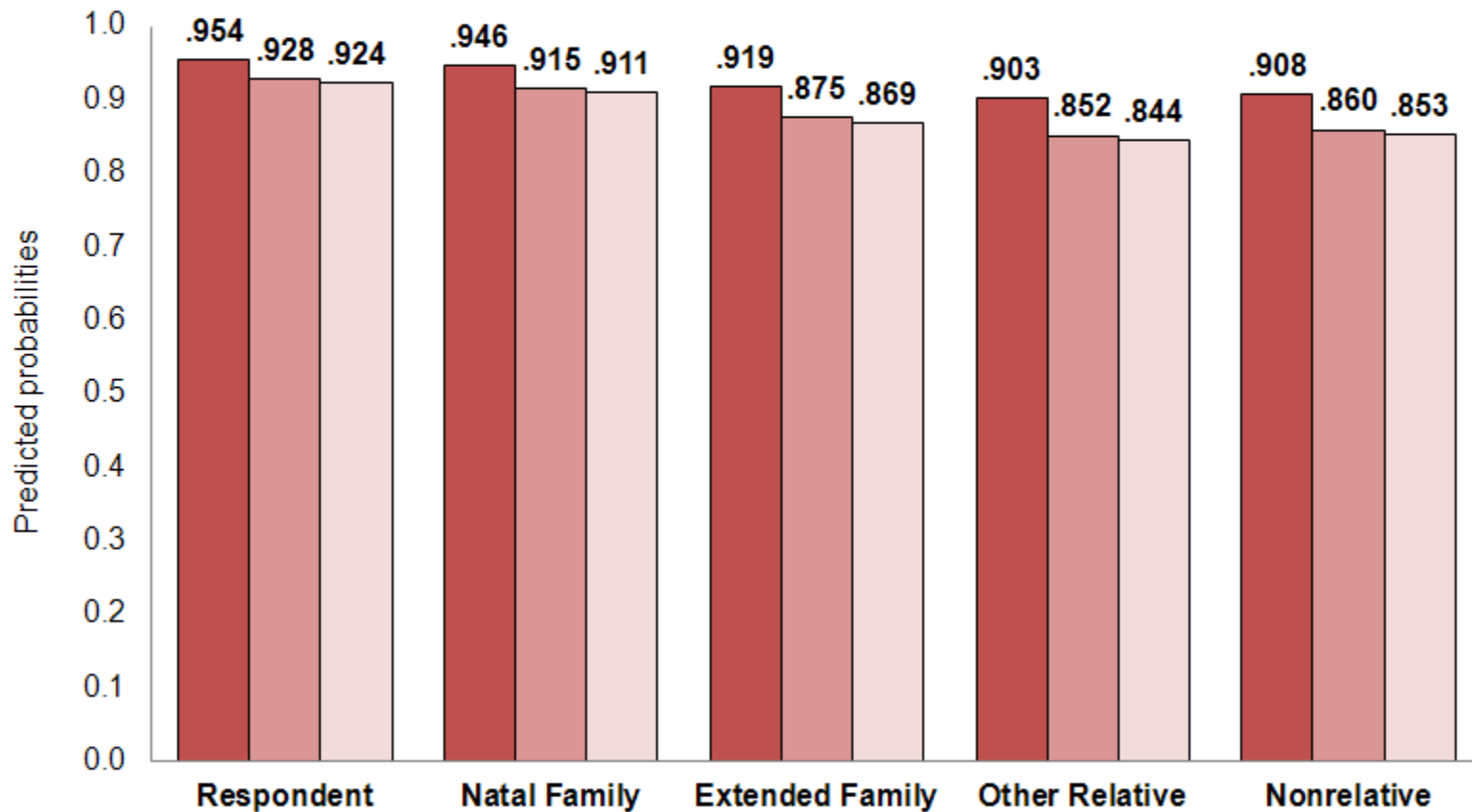Source: U.S. Census Bureau, 2011 American Community Survey.

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU

Figure 6.
Likelihood of Reporting Year of Naturalization by Relationship to Householder

■ Without nonresponse controls   ■ With nonresponse controls

Source: U.S. Census Bureau, 2011 American Community Survey.

Figure 7.
Likelihood of Reporting Year of Naturalization by Mode and Relationship to Householder

HOUSEHOLD CHARACTERISTICS:
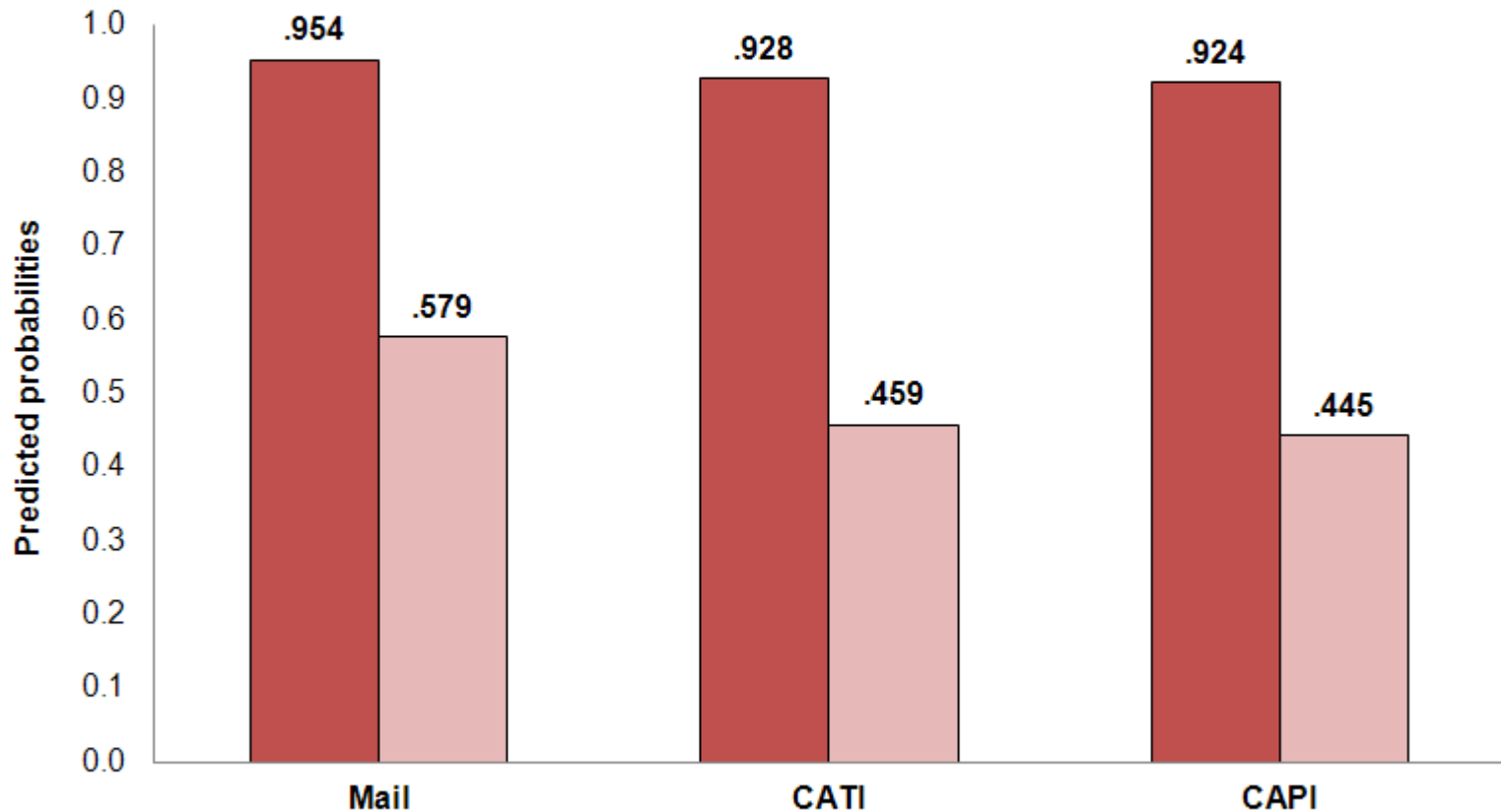## Association with Reporting Year of Naturalization

Variable	Without Nonresponse Controls	With Nonresponse Controls
Householder's level of education	High school degree or higher education, more likely to report	Not significant
Householder's language spoken at home	Speaks only English at home, less likely to report	Speaks only English at home, less likely to report
Householder's nativity status	If native, less likely to report	If native, less likely to report
Region of residence	Compared with West, Midwest then Northeast less likely to report; South not significant	Compared with West, Midwest then Northeast less likely to report; South not significant
Number of adults aged 18 and older	More adults, <u>less</u> likely to report	More adults, <u>more</u> likely to report
Homeownership	Own home, more likely to report	Own home, more likely to report
Poverty status	Not in poverty status, more likely to report	Not in poverty status, more likely to report

INDIVIDUAL CHARACTERISTICS:
Association with Reporting Year of Naturalization

Variable	Without Nonresponse Controls	With Nonresponse Controls
Age	Compared with 60+, 18-29 most likely to report, followed by 30-44 then 45-59	Compared with 60+, 45-59 more likely to report; 18-29 and 30-44 not significant
Sex	Males <u>less</u> likely than females to report	Males <u>more</u> likely than females to report
Region of birth	Compared with respondents from Europe, those from other regions, Asia, and Latin America less likely to report, in that order	Compared with respondents from Europe, those from other regions, Latin America, and Asia less likely to report, in that order
Period of entry	Compared with respondents who entered in 1996 or earlier, those who arrived 1997-2001 and 2002-2006 more likely to report while 2007-2011 less likely	Compared with respondents who entered in 1996 or earlier, those who arrived 1997-2001 and 2002-2006 were more likely to report; 2007-2011 not significant.
Moved in the last year	Did not move last year, more likely to report	Did not move last year, more likely to report

Figure 8.
Likelihood of Reporting Year of Naturalization by Mode and Year of Entry Imputation Status

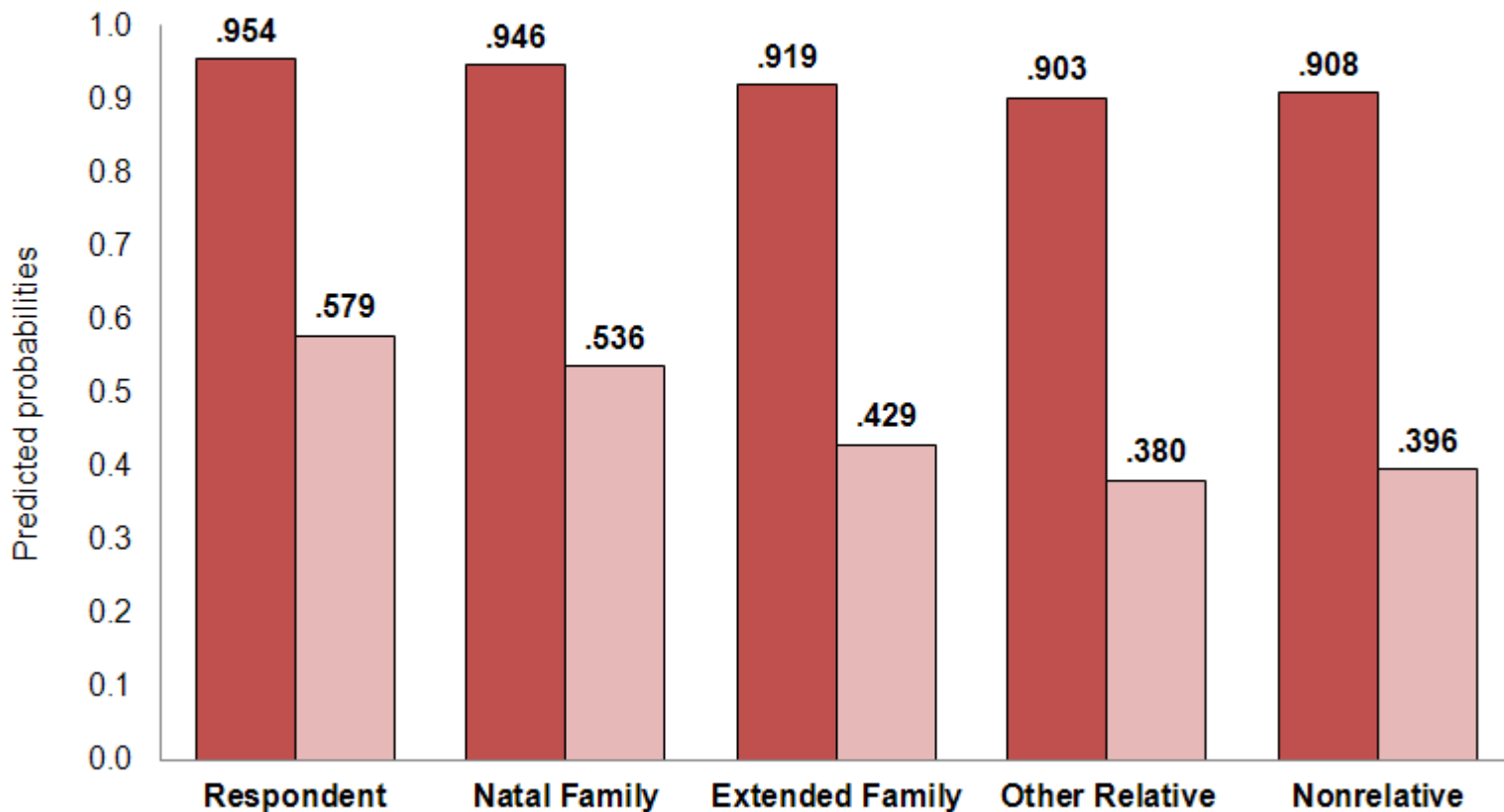■ Year of entry not imputed   □ Year of entry imputed

Note: CATI=Computer Assisted Telephone Interview. CAPI=Computer Assisted Personal Interview.
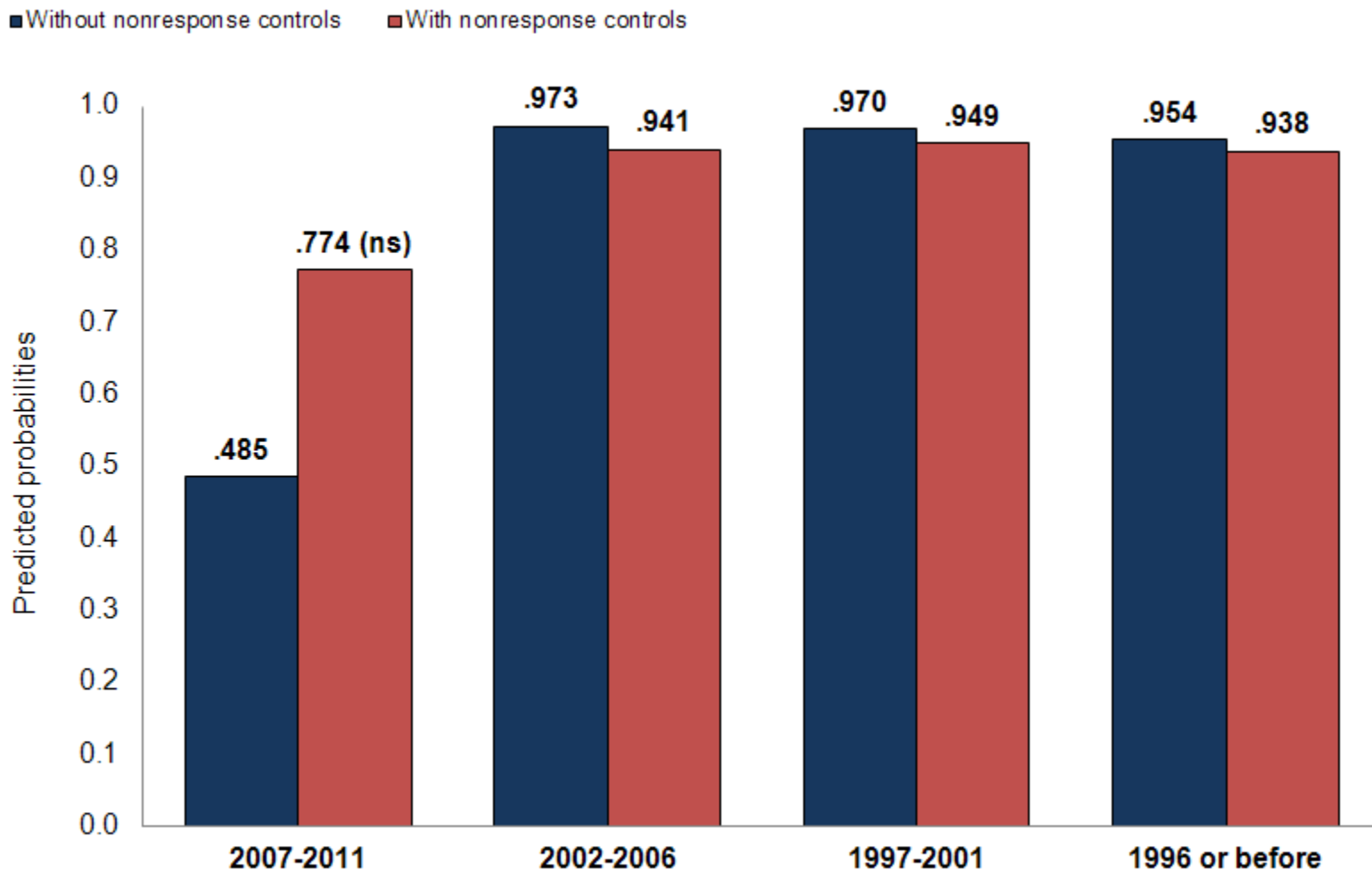Source: U.S. Census Bureau, 2011 American Community Survey.

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU

Figure 9.
Likelihood of Reporting Year of Naturalization by Relationship to Householder and Year of Entry Imputation Status

Source: U.S. Census Bureau, 2011 American Community Survey.

Figure 10.
Likelihood of Reporting Year of Naturalization by Period of Entry

Note: NS = not statistically significant.
Source: U.S. Census Bureau, 2011 American Community Survey.

Conclusion

- Mode and relationship to householder significantly influence year of naturalization reporting
  - Association remains even after inclusion of nonresponse controls, although strength is tempered
- Results of analysis support model of survey environment and social distance from respondent
- Using nonresponse control variables is an alternative to excluding cases with some imputed data

Contact Information

Elizabeth M. Grieco

301-763-5275

Elizabeth.M.Grieco@census.gov

David M. Armstrong

301-763-5671

David.M.Armstrong@census.gov