



This document was prepared by and for Census Bureau staff to aid in future research and planning, but the Census Bureau is making the document publicly available in order to share the information with as wide an audience as possible. Questions about the document should be directed to Kevin Deardorff at (301) 763-6033 or kevin.e.deardorff@census.gov

February 8, 2012

2010 CENSUS PLANNING MEMORANDA SERIES

No. 173

MEMORANDUM FOR The Distribution List

From: Arnold Jackson *[signed]*
 Acting Chief, Decennial Management Division

Subject: 2010 Census Item Nonresponse and Imputation Assessment Report

Attached is the 2010 Census Item Nonresponse and Imputation Assessment Report. The Quality Process for the 2010 Census Test Evaluations, Experiments, and Assessments was applied to the methodology development and review process. The report is sound and appropriate for completeness and accuracy.

If you have questions about this report, please contact Cynthia Rothhaas at (301) 763-1896.

Attachment

2010 Decennial Census: Item Nonresponse and Imputation Assessment Report

U.S. Census Bureau standards and quality process procedures were applied throughout the creation of this report.

Final

Cynthia Rothhaas

Frederic Lestina

Joan M. Hill

Decennial Statistical Studies Division



This page intentionally left blank.

Table of Contents

Executive Summary	vii
1. Introduction	1
1.1 Purpose of Study	1
1.2 Background	2
2. Methodology	3
2.1 Universe	3
2.2 Assessment Measures	5
3. Limitations	8
4. Results	9
4.1 Item Nonresponse Rates	9
4.2 Imputation Rates	19
5. Summary and Recommendations for Future Research	31
6. References	32
Appendix A: Classification of Edit/Allocation Flag Variables Into Categories from the Census Edited File	34
Appendix B: Census 2000 INR Rates by Form Type and Response Mode	36
Appendix C: Census 2000 Imputation Rates by Form Type and Response Mode	37
Appendix D: Census 2000 Imputation Rates by Form Source and Response Mode	38

List of Tables

Table 1: 2010 Census INR Rates for Household-Level Items - National.....	10
Table 2: 2010 Census INR Rates for Person-Level Items - National.....	11
Table 3: 2010 Census INR Rates for Self-Response Household-Level Items - National	12
Table 4: 2010 Census INR Rates for Enumerator Return Household-Level Items - National.....	13
Table 5: 2010 Census INR Rates for Self-Response Person-Level Items - National.....	15
Table 6: 2010 Census INR Rates for Enumerator Return Person-Level Items - National	17
Table 7: 2010 Census INR Rates for Rates for GQ Person-Level Items on GQ Forms by form type and by GQ type (when GQ type is known) - National	18
Table 8: Overall “As Reported” and Imputation Rates (in percent) - National.....	20
Table 9: 2010 Census “As Reported” and Imputation Rates (in percent) for Self-Response - National	21
Table 10: “As Reported” and Imputation Rates (in percent) for Enumerator Returns -- National	23
Table 11: “As Reported” and Imputation Rates (in percent) for Group Quarters - National	24
Table 12: Data Completeness Statistic	27
Table 13: Data Completeness Statistic - Number and Percent of Persons in Housing Units by Sum of Non-Imputed Responses to Person-Level Items	28
Table 14: Number and Percent of Totally Allocated Persons by Form Type	30
Table 15: Overall Item Nonresponse and Imputation Rates.....	31
Table A1: 2010 Census INR Rates for Household-Level Items - Puerto Rico	39
Table A2: 2010 Census INR Rates for Person-Level Items - Puerto Rico.....	40
Table A3: 2010 Census INR Rates for Self-Response Household-Level Items - Puerto Rico	41
Table A4: 2010 Census INR Rates for Enumerator Return Household-Level Items - Puerto Rico	42

Table A5: 2010 Census INR Rates for Self-Response Person-Level Items - Puerto Rico.....43

Table A6: 2010 Census INR Rates for Enumerator Return Person-Level Items - Puerto Rico44

Table A7: Overall As Reported and Imputation Rates (as a percentage) – Puerto Rico45

This page intentionally left blank.

Executive Summary

This assessment report provides information on data quality, specifically data completeness, for the person-level and household-level items from the 2010 Census. These items include tenure, relationship, sex, age/date of birth, Hispanic origin, and race. The item nonresponse rates, along with imputation rates, are types of response quality measures. The item nonresponse rate is mainly used as an indicator of respondent cooperation. Imputation rates incorporate respondent cooperation, but also consider inconsistent and unusable responses.

Note that results presented in this report apply to characteristic imputation as opposed to count imputation. The characteristic imputation process assessed in this report begins after the household population is established/resolved through various processes, such as count imputation.

Item nonresponse and imputation rates are calculated for self-response returns and enumerator returns as well as Group Quarters. For item nonresponse rates, the universe excludes vacant and deleted housing units. Note that item nonresponse results for the bilingual questionnaire are included in this assessment and will also be in the *2010 Decennial Census Bilingual Assessment* (see Rothhaas et al., 2011).

Research questions and key results for the 2010 Census related to item nonresponse and imputation are described below.

For Housing Units:

1. What are the item nonresponse rates for the following household-level items: household count, undercount, tenure, and telephone number? What are the item nonresponse rates for the following person-level data items: sex, age/date of birth, Hispanic origin, race, overcount question, and relationship? By self-response mode and enumerator return mode (and the operations within those modes)?

Overall for household-level items, item nonresponse rates are highest for the undercount item and the telephone number item. Enumerator returns have a higher item nonresponse rate for each household item, compared to self-response returns, except for household count. The majority of enumerator returns correspond to households that did not respond to the initial mail census and thus were a harder to enumerate population. The 2010 enumerator return item nonresponse rate for tenure is 9.7, while the Census 2000 enumerator short-form item nonresponse rate for tenure was 6.3 (Norris, 2003 and Appendix B) although the two rates cannot be directly compared due to changes in population and other census implementation methods over time. In addition, the enumerator return item nonresponse rate for tenure is substantially higher than the self-response item nonresponse rate for tenure in 2010. Note that on the enumerator return, the tenure item is placed on the inside page of the questionnaire compared to the self-response return where the tenure item is placed on the first page.

The 2010 Census self-response item nonresponse rates for tenure and household count were 2.4 percent and 1.8 percent, respectively. The Census 2000 Form Replication Experiment, which had Census 2000 content formatted on a 2010 Census mailout/mailback English questionnaire, indicates a tenure item nonresponse rate of 3.1 percent (Standard Error=0.17) and a household count item nonresponse rate of 1.0 percent (Standard Error=0.09) (Reiser et al.). The 2010 Census version of tenure is significantly lower than the Census 2000 version in terms of item nonresponse. The item nonresponse rate is significantly higher for the 2010 Census version of household count, compared to the Census 2000 version, controlling for time and population differences. Note that undercount and telephone number items were not on the Census 2000 form.

The highest person-level item nonresponse rate is for age/date of birth on enumerator returns (10.0 percent). This item also had the highest item nonresponse rate in Census 2000. The difference between self-response returns and enumerator returns is 9.2 percent. This difference is higher than the 6.9 percent difference in Census 2000 (Norris, 2003 and Appendix B).

The item nonresponse rate for self-response returns for the age/date of birth item in the 2010 Census is 0.8 percent. This is consistent with the corresponding Census 2000 Form Replication Experiment age/date of birth item (0.8 percent, Standard Error=0.07). Although the gap between the age/date of birth item nonresponse rate for self-response and enumerator returns widened in 2010, the 2010 Census version of age/date of birth was comparable to the Census 2000 version (with respect to item nonresponse), controlling for time and population differences.

The 2010 Census item nonresponse rate for self-response returns for the Hispanic origin item (4.2 percent) is significantly higher than the Census 2000 Form Replication Experiment item nonresponse rate for Hispanic origin (3.7 percent, Standard Error=0.17). The 2010 Census item nonresponse rate for self-response returns for the race item (3.3 percent) is significantly lower than the Census 2000 Form Replication Experiment item nonresponse rate for race (4.2 percent, Standard Error=0.22) (Reiser et al.). It is interesting to note that, in both Census 2000 and the 2010 Census, enumerator returns had relatively lower item nonresponse rates for Hispanic origin compared to self-response returns.

We also see high person-level item nonresponse rates for replacement forms and Enumeration at Transitory Locations forms. There are relatively higher item nonresponse rates for Hispanic origin on Asian Fulfillment and Be Counted forms, and race on Bilingual forms. Previous research indicates that these results are associated with the specific populations enumerated on these forms. (For more information on the operations that use these forms, please see U.S. Census Bureau, 2011.)

For Nonresponse Followup, the age/date of birth item had relatively high item nonresponse. Research on 2010 Nonresponse Followup using behavior coding methods showed that interviewers often altered the age and date of birth question (both questions were asked together on the form). The age question was asked differently from how it was scripted 78 percent of the time for Person 1 and 33 percent of the time for Persons 2 through 5. In addition, respondents answered the age and date of birth items inadequately approximately 53 percent of the time. The inadequate responses appear to primarily be due to respondents only answering age or date of birth (both needed to be answered to be considered adequate). This was very likely due to the interviewers changing the question to request only a single piece of information during the initial interview. The interviewers omitted either the date of birth item (accounts for 57 percent of the major changes to this item) or the age item (accounts for 25 percent of the major changes to the item). In some (but not all) cases, interviewers proceeded to collect the additional piece of information during the interview (Childs and Jurgenson, 2011).

2. What are the number and percent of totally allocated persons¹?

There were 247,904 (0.1 percent) totally allocated persons in the 2010 Census. Most (90.0 percent) of the totally allocated persons are on forms that were created during Response Processing System activities as a result of various situations/processes in which an original return/person was unavailable, unusable, or insufficient. For example, person records were created to match population counts derived when not all person-level items were received.

3. What are the number and percent of person substitutions²?

There were 5,770,791 persons (1.9 percent) substituted in the 2010 Census. All of the substituted persons are on person form types created during Response Processing System activities.

4. What are the imputation rates for the following data items: tenure, relationship, sex, age/date of birth, Hispanic origin, and race? By self-response mode and enumerator return mode (and the operations within those modes)? What percent of race and Hispanic origin assignments were based on American Community Survey or Census 2000 data?

Short form self-response imputation rates in Census 2000 are similar to self-response rates found in the 2010 Census.

Similar to the item nonresponse rate trend, replacement forms have higher imputation rates for all items, compared to the initial Mailout/Mailback English forms, due to population

¹ See Section 4.2.4 for definition.

² See Section 4.2.5 for definition.

differences. As expected, Be Counted returns have relatively high imputation rates for relationship, Hispanic origin, and race, since the population enumerated on these forms is traditionally hard to enumerate, as evidenced by the use of specialized self-response enumeration procedures (U.S. Census Bureau, 2009d).

Also similar to the item nonresponse trend, imputation rates are substantially higher for the age/date of birth item on Nonresponse Followup and Enumeration at Transitory Locations returns.

Approximately 36.6 percent of all imputed Hispanic origin responses and 28.6 percent of all imputed race responses used information from a Census 2000 or American Community Survey response. This new approach for assigning race and Hispanic origin appears to be beneficial, in terms of the quantity of usable data.

5. What are the numbers and percent of persons with varying non-imputed responses to the person-level imputation data items (data completeness statistics)? This is a distribution of the level of item nonresponse by person (i.e., number of persons with zero to five person-level item responses, counting age and date of birth as one item). This statistic will be broken out by form type and by person number.

On every form type, the majority of persons responded to all five of the person-level items (four for Group Quarters). Over 97 percent of all persons in this universe answered at least four of five data items.

6. How do 2010 item nonresponse and imputation rates compare to 2000 rates.

These results are discussed in questions 1 and 4 above.

For Group Quarters:

7. What are the item nonresponse rates for the following group quarter person-level items: sex, age/date of birth, Hispanic origin, and race? By group quarters type?

Group quarters have relatively high item nonresponse rates for all person-level items. Race and Hispanic origin item nonresponse rates are relatively high across all group quarters form types and all group quarters types. This is possibly caused by:

- difficulty in contacting respondents, especially in noninstitutional group quarters (see table 7 for noninstitutional group quarters categories),
- the limited knowledge of the personnel who might be providing the enumeration data, or
- limited information available from a list provided by the property management (U.S. Census Bureau, 2009c).

8. What are the imputation rates for the following data items: sex, age/date of birth, Hispanic origin, and race? By group quarter type?

Similar to the item nonresponse rate trend, imputation rates are relatively high for Hispanic origin and race across all Group Quarters types.

Summary Table

The table below provides an overall summary of item nonresponse and imputation rates in one table. Note that the item nonresponse rates cannot be directly compared to the imputation rates because the imputation rates are based on valid responses, whereas item nonresponse rates are based on the presence of a response, regardless of validity.

Overall Item Nonresponse and Imputation Rates

	Person-Level Items					Household
	Relationship	Sex	Age/Date of Birth	Hispanic Origin	Race	-Level Item Tenure
Item Nonresponse	1.5	1.5	3.5	3.9	3.3	4.5
Total	100.0	100.0	100.0	100.0	100.0	100.0
“As Reported”	97.9	98.4	95.0	95.5	95.9	96.5
Imputed	2.1	1.6	5.1	4.5	4.1	3.5
Assigned	0.5	1.3	1.5	1.7	1.2	n/a
Allocated	1.7	0.3	3.6	2.8	2.9	3.5
Substituted	1.9 percent of all persons					

See Section 1.2 for definitions of assigned and allocated.

Recommendations for Future Research:

- Age/date of birth and tenure item nonresponse and imputation rates are high for enumerator Nonresponse Followup and Enumeration at Transitory Locations returns; research on source of error should be focused on this area. Since the tenure issue is likely related to the placement of the question on the questionnaire, an automated instrument for the followup process might resolve this issue.
- Research to reduce Hispanic origin and race item nonresponse is recommended.
- Using previous Census and American Community Survey response data for allocation of missing Hispanic origin and race responses proved to be beneficial. Research to expand this allocation by using these sources, as well as other administrative records (including potentially non-primary returns), for all missing items is recommended.

- It is also recommended that a quality check of the previous response data be examined by comparing the previous Census and American Community Survey responses for Hispanic origin and race to the “as reported” Hispanic origin and race responses in the 2010 Census.
- Data for roughly 5.7 million people (1.9 percent) were acquired through substitution in the 2010 Census. Investigation into the operational source of substitutions is recommended. This includes the operations producing cases needing substitution and the operations contributing data for substitution.

1. Introduction

The 2010 Decennial Census is an important opportunity for the Census Bureau to ensure an accurate count of the nation's increasingly diverse and rapidly growing population.

1.1 Purpose of Study

The purpose of this study is to present the item nonresponse (INR) and imputation rates for the questionnaire items from the 2010 Decennial Census. The items include tenure, relationship, sex, age/date of birth, Hispanic origin, and race.³

The INR and imputation rates are types of response quality measures. The INR rate is mainly used as an indicator of respondent cooperation. Imputation rates incorporate respondent cooperation, but also consider inconsistent and unusable responses. Thus, both are useful in determining quality of items and quality of responses to those items.

These rates are documented for informational and historical purposes.

Key research questions related to INR and imputation include:

1. What are the INR rates for the following household-level items: household count, undercount, tenure, and telephone number? What are the INR rates for the following person-level items: sex, age/date of birth, Hispanic origin, race, overcount, and relationship? By self-response mode and enumerator return mode (and the operations within those modes)?
2. What are the number and percent of totally allocated persons?
3. What are the number and percent of person substitutions?
4. What are the imputation rates for the following data items: tenure, relationship, sex, age/date of birth, Hispanic origin, and race? By self-response mode and enumerator return mode (and the operations within those modes)? What percent of race and Hispanic origin assignments were based on American Community Survey or Census 2000 data?
5. What are the numbers and percent of persons with varying non-imputed responses to the person-level imputation data items (data completeness statistics)? This is a distribution of the level of INR by person (i.e., number of persons with zero to five person-level item responses). This statistic will be broken out by form type and by person number.

³ Overall INR rates for additional items such as household count, undercount, phone number, and overcount are presented for completeness.

6. How do 2010 INR and imputation rates compare to 2000 rates?

For Group Quarters:

7. What are the INR rates for the following person-level items: sex, age/date of birth, Hispanic origin, and race? By Group Quarters (GQ) type?
8. What are the imputation rates for the following data items: sex, age/date of birth, Hispanic origin, and race? By GQ type?

1.2 Background

Census 2000 was the first census for which an INR report was produced (Norris, 2003). Prior to that, only imputation rates were produced. This document provides both the INR and imputation rates. Note that results presented in this report apply to characteristic imputation. The characteristic imputation process assessed in this report begins after the household population is established/resolved through various processes (i.e., count imputation and INP processing).

The INR rate is essentially the proportion of missing responses before pre-editing or imputation procedures for a given item (i.e., the respondent did not provide an answer to the item). For INR, missing values are included in the rates, but inconsistent responses (i.e., incompatible with other responses) are considered non-missing responses.

The following are three types of characteristic imputations performed during a census (U.S. Census Bureau, 2010); the results for all are included in this report. Note these imputations differ from count imputation which is an operation performed on the final status of a housing unit and the population count of housing units.

Assignment - a response for a data item is either missing or not consistent with other responses AND an item value can be determined based on other information provided for that same person. For example, if race is missing but a Hispanic origin write-in identifies a race, then race can be assigned from the Hispanic origin response.

For the 2010 Census, a new approach to imputing race and Hispanic origin was implemented. When race or Hispanic origin was left blank or was inconsistent with other responses, and a value could not be determined based on information provided for that same person or from within the housing unit (HU)/GQ, a value was assigned based on matched Census 2000 responses or American Community Survey (ACS) responses. If a value could not be determined based on matched Census 2000 or ACS responses, then a value from a person in a nearby HU/GQ was allocated.

Allocation - a response for a data item is either missing or not consistent with other responses and an item value CANNOT be determined based on information provided for that same

person. An allocation uses a response from another person within the HU/GQ or from a person in a nearby HU/GQ.⁴

Substitution - A “whole-house” substitution occurs when all the characteristics for EVERY person in the HU are missing. The person characteristics include relationship, sex, age/date of birth, Hispanic origin, and race. A nearby HU of up to six people with complete person data is selected to represent the person-level items for up to the first six persons in the HU needing substitution.⁵ When the population of the HU needing substitution is greater than six, all characteristic data for the remaining persons are allocated.

2. Methodology

2.1 Universe

Rates for INR and imputation are based on primary⁶ returns. The rates exclude vacant and substituted HUs/GQs as well as those HUs selected to be in a census experiment (a controlled test of a question item or methodology), although rates from HUs in census experiments are used for comparison purposes. Results for these returns are based on the data in the Census Unedited File (CUF) and the Census Edited File (CEF).

The CUF contains:

- geography for all collection blocks,
- HU and GQ-level information (including address and operational information) for all unique HUs and GQs,
- return data for all housing unit returns selected by the Primary Selection Algorithm (PSA) (an algorithm used to identify a primary return), which do not have a HU status of delete, and
- return data for valid persons in PSA-selected HUs (which do not have a housing unit status of delete).

The CEF contains much of the same data as the CUF but also contains edited and imputed data.

⁴ A person is considered “totally allocated” when every characteristic for a person requires allocation, as long as one person within the unit has some reported data.

⁵ If the HU to be substituted has a return type of a mail return, then the nearby donor HU is a mail return. Otherwise the nearby donor HU is a nonmail return. This stratification of return type only occurs for substitution, not for allocation.

⁶ When more than one return is received from a HU, a primary return is selected. The primary return is the return of record for the census, even though it may be a culmination of data from more than one census return.

Similar to Census 2000, the universe for person-level imputation rates in the 2010 Census refers to reported responses, assignments, and allocations (i.e., substituted persons are excluded from the universe). This was done in Census 2000 strictly for ease of reading (Zajac, 2003). For the 2010 Census, substituted persons will be excluded to avoid artificially affecting the person-level item rates since whole persons are substituted from 2010 Census person-level item data, which is already factored into the rates prior to any substitution. In addition, the computation is kept consistent for historical comparability.

Items for totally allocated persons are included in the imputation rates for person-level items.

The INR and imputation rates are presented separately for various self-response and enumerator return modes (and the operations within those modes). Note that the mode reported is the mode for the primary return. An Update/Enumerate form, for example, may have Be Counted data appended to the return; however, the Update/Enumerate form will be reported as the primary return. Also note that Telephone Questionnaire Assistance forms are classified as enumerator returns due to the agent interaction, despite the interview being initiated by the respondent.

Self-response mode includes:

- Mailout/Mailback (MOMB) (including bilingual),
- Update/Leave (UL),
- Be Counted, and
- Fulfillment.

Enumerator return mode includes:

- Coverage Followup (CFU)⁷,
- Telephone Questionnaire Assistance (TQA),
- Nonresponse Followup (NRFU) , NRFU Reinterview, NRFU Residual, Update/Enumerate (UE), UE Reinterview, Remote Alaska (RA), Rural Update/Enumerate (RUE), and Vacant Delete Check⁸, and
- Enumeration at Transitory Locations (ETL).

GQ mode includes:

- GQ Enumeration.

Note that INR and imputation rates for the bilingual questionnaire are included in this assessment (in the self-response mode) and also will be included in the *2010 Census Decennial Bilingual Questionnaire Assessment* (Rothhaas et al., 2011).

⁷ Although CFU primary returns are enumerator returns, it should be noted that the majority of the person-level data items on CFU returns are actually carried over from other returns (often self-response returns).

⁸ In the remainder of this document, this mode will be referred to as NRFU and Others for simplicity.

2.2 Assessment Measures

The formulas used to calculate the INR and imputation rates are presented in this section.

Note that the denominator of the person-level imputation rates is the number of non-substituted persons. Similarly, the denominator of the household-level rate is the number of all HUs in the universe (See Section 2.1). The formulas for imputation and substitution rates are located in Section 2.2.2.

Also note that for occupied HUs, tenure could not be assigned because there are no data on the 2010 Census form to provide any insight to the tenure status. During Census 2000, tenure could be assigned because there were data on the long form that could help determine the tenure status. For the 2010 Census, tenure in occupied HUs was allocated, but it was not assigned or substituted.

Note that INR rates may have been affected by the Decennial Response Integration System (DRIS) Reprocessing Operation, which was conducted after data capture. This operation attempted to improve the accuracy and quality of responses based on the respondent's intent. Sometimes respondents wrote notes in the margin of the questionnaire to clarify their responses. If the intent of the respondent was clear, a reviewer keyed in the intended response. For example, if the respondent filled out a person panel and then wrote "deceased" in the margin, the person was voided out under the reprocessing rules. If the respondent filled out every person number on the form with the same information, only the first instance might be keyed under a reprocessing scenario. Sometimes a respondent used an 'X' to indicate what they did not want as a response and then used blanks, different color ink, or circles to indicate what they wanted as a response (Coon and Osborne, 2011).

Nationally, 842,170 forms in the INR universe for this report went through DRIS Reprocessing. In Puerto Rico, 455,939 forms in the INR universe went through DRIS Reprocessing. DRIS Reprocessing captured the intent of the respondent, as it pertains to form completion. We included reprocessed cases in this report since the INR analysis also attempts to capture respondent intent.

2.2.1 Item Nonresponse Rates

The INR rates include all occupied, non-substituted units and persons within these units and are computed using the following formula:

$$\frac{\text{number of missing item responses in universe}}{\text{total records in universe}} * 100$$

“Missing” refers to responses that were not reported by the respondent. “Total records” refers to total data item records.

The universe for the INR analysis excludes substituted persons (See Section 2.2.2.4 for more details) and persons created by headquarters (HQ) during processing. Households and persons in 2010 Census Experiments are not included in this report but can be found in the individual experiment reports. The exception is the Census 2000 Form Replication Experiment (see Section 2.2.4 for more detail). Note that supplemental NRFU forms (continuation forms) on the HU CUF also are excluded because these forms do not ask the household-level questions.

2.2.2 Imputation Rates

When computing imputation rates, the assignment and allocation rates are considered. The “as reported” rates reflect the items NOT imputed and are included only for completeness. These rates, as well as substitution rates, are defined below.

2.2.2.1 “As Reported” Rates

The “as reported” rates include occupied units and persons within those units. Substituted persons are excluded. An item is determined “as reported” by the edit/allocation flag variables. Appendix A shows the values used to determine “as reported” for each item. The “non-imputed values” category in the appendix is considered “as reported”. The “as reported” rates will be computed using the following formula:

$$\frac{(\text{“as reported” records})}{(\text{“as reported” records})+(\text{assignments})+(\text{allocations})} *100$$

2.2.2.2 Assignment Rates

The assignment rates include occupied units and persons within those units. Substituted persons are excluded. An item assignment is identified by the edit/allocation flag variables. Appendix A shows the values used to determine an assignment for each item. Note that the proportion of race and Hispanic origin assignments from previous data (i.e., Census 2000 and ACS) are presented separately. The assignment rates are computed using the following formula:

$$\frac{(\text{assignments})}{(\text{“as reported” records})+(\text{assignments})+(\text{allocations})} *100$$

2.2.2.3 Allocation Rates

The allocation rates include all occupied units and persons within those units. Substituted persons are excluded. An item allocation is identified by the edit/allocation flag variables. Appendix A shows the values used to determine allocation for each item. The allocation rates for

totally allocated persons are reported separately. The allocation rates are computed using the following formula:

$$\frac{(\text{allocations})}{(\text{“as reported” records})+(\text{assignments})+(\text{allocations})} *100$$

2.2.2.4 Substitution Rates

The substitution rates include occupied HUs and persons within those units. A housing unit of up to six persons is considered substituted when all persons within that unit have been substituted. When the population of the HUs needing substitution is greater than six, all characteristic data for the remaining persons are allocated. The “substituted values” category in Appendix A show the values used to determine substitution. The substitution rates are computed using the following formula:

$$\frac{(\text{substitutions})}{(\text{“as reported” records})+(\text{assignments})+(\text{allocations})+(\text{substitutions})} *100$$

2.2.3 Data Completeness Statistic

The data completeness statistic is a person-level summary of the total number of “as reported” responses to the data items. The distribution of this statistic indicates the level of INR or modified response by person. It ranges from zero to five since five person-level items are considered (relationship, sex, age/date of birth, Hispanic origin, and race). This statistic is disaggregated by form type and by person number.

2.2.4 Comparison of Census 2000 and 2010 Census INR Rates and Imputation Rates

The INR rates and imputation rates for the 2010 Census and for Census 2000 are documented in the results section. The limitations to these comparisons include:

- Census 2000 only reported INR rates by self-response mode and enumerator return mode. Thus, we can only compare the 2000 rates to the 2010 rates for those same modes (however, the specific operations that make up those modes have changed).
- The 2010 Census questionnaire content is not the same as the Census 2000 content. The 2010 Census questionnaire has undergone changes to question wording, format, response categories, instructions, and other design features.
- A confounding factor is that there is a ten-year difference between the 2010 Census and Census 2000. This likely resulted in real change to the survey environment and thus respondent compliance. In fact, Census 2000 and 2010 Census response rates are different and, with that, likely incur a different self-selection population.

Because form types changed between Census 2000 and the 2010 Census, the categories are not equivalent. For example, the Remote Alaska form in 2000 used the NRFU/UE form in 2010. Also note that there is no List Enumerate operation for the 2010 Census.

Note that an experiment was conducted during the 2010 Census which was designed to evaluate the cumulative effects of all questionnaire content changes from the previous decade. This study is called the Census 2000 Form Replication Experiment (Reiser et al.). This experiment was designed to compare the 2010 Census mailout/mailback questionnaire content to a mailout/mailback questionnaire that replicated the Census 2000 wording, categories, order, and other essential design features. By comparing the Census 2000 and 2010 Census mail questionnaire INR results in the same time frame, the impact of real changes to the population over time should be eliminated.

The INR results for the experiment are referenced in this assessment in order to evaluate the INR differences from Census 2000 and the 2010 Census questionnaires while time and responding populations are held constant.

Although these comparisons yield valid, meaningful results, given that this study is based on a randomized experimental design, an important caution should be noted when considering the source of item-level differences between 2010 and the experimental 2000 estimates. Prior to implementing item-level changes on the 2010 Census questionnaire, the impact of those changes was tested during mid-decade tests. However, the 2010 Census form was the first time that we evaluated all of the changes together on the same form. The numerous item-level changes made over the decade (in particular, the addition of the undercount and overcount questions) resulted in the 2010 Census form containing considerably less white space, and thereby had a more crowded appearance. Although we attempt to experimentally control for various factors, such as time and population differences, we cannot differentiate how the overall appearance of the final form may have affected respondent behavior with regard to individual items. As such, the combined effects of all changes to the form must be considered when assessing the causal nature of item-level differences.

3. Limitations

This assessment report includes documentation of the 2010 Census INR and imputation rates; the Census 2000 INR and imputation rates; and the INR rates from the 2010 Census experimental panel that replicates the 2000 short-form design. Details of this documentation, including limitations to any comparisons, are in Section 2.2.4 of this report.

4. Results

4.1 Item Nonresponse Rates

Nationally, there were 115,655,995 occupied HUs in the universe for the INR analysis. These HUs contained 293,537,999 data-defined⁹, non-substituted persons. On GQ forms, there were 7,894,078 data-defined persons. In the Puerto Rico INR analysis universe, there were 1,357,708 occupied HUs, 3,608,373 data-defined persons in those HUs, and 37,666 data-defined persons in GQs.

The tables in this section present the overall national INR rates. See Attachment 1 through 7 for similar tables for Puerto Rico.

Note that different types of questionnaires can have different content. For example, the Be Counted form does not ask the household count, undercount, or overcount questions. In this situation, the form is excluded from analysis. Also, the same questions are not asked of all persons. Race, for example, is not asked of Persons 7 and higher on MOMB forms. These persons are excluded from the race analysis.

Table 1 shows the INR rates for the household-level items. Overall, INR rates are highest for the undercount item and the telephone number item. Enumerator returns have a higher INR rate for each household-level item, compared to self-response returns, except for household count. The majority of enumerator returns correspond to households that did not respond to the initial mail census and thus were a harder to enumerate population. In addition, the traditionally higher use of proxy responses for enumerator returns may have contributed to the higher item nonresponse.

Considering the limitations described in Section 2.2.4, the 2010 Census enumerator return INR rate for tenure is 9.7 percent, while the Census 2000 enumerator short-form INR rate for tenure was 6.3 percent (Norris, 2003 and Appendix B). In addition, the enumerator return INR rate for tenure is substantially higher than the self-response INR rate for tenure in 2010. Note that on the enumerator return, the tenure item is placed on the inside page of the questionnaire compared to the self-response return where the tenure item is placed on the first page. Similar to the tenure item, the high INR rate for the undercount item on enumerator returns could be the result of the question being placed on the inside flap of the questionnaire compared to the self-response return where it is placed on the first page.

The 2010 Census self-response INR rates for tenure and household count were 2.4 percent and 1.8 percent, respectively. The Census 2000 Form Replication Experiment, which had Census 2000 content formatted on a 2010 Census MOMB English questionnaire, indicates a tenure INR rate of

⁹ A person is considered data defined, or valid, if they have at least two of the person-level data items filled. The person-level data items considered for the data-defined determination include name, relationship, sex, age/date of birth, Hispanic origin, and race (Alberti, 2008).

3.1 percent (Standard Error (SE)=0.17) and a household count INR rate of 1.0 percent (SE=0.10) (Reiser et al.). The 2010 Census version of tenure is significantly lower than the Census 2000 version in terms of INR. The INR rate is significantly higher for the 2010 Census version of household count, compared to the Census 2000 version, controlling for time and population differences. Note that undercount and telephone number items were not on the Census 2000 form.

Table 1: 2010 Census Item Nonresponse Rates for Household-Level Items - National

Form Type/Operation	Percent Item Nonresponse*			
	Household Count	Undercount ****	Tenure	Telephone Number
Overall	1.3	8.5	4.5	8.3
Self-Response	1.8	7.3	2.4	7.8
Enumerator Return	<0.1**	12.2	9.7	9.7***

Source: Census Unedited File

*Note that the INR rates in Tables 1 through 7 cannot be directly compared to the imputation rates in Tables 8 through 11 because the imputation rates are based on valid responses, whereas INR rates are based on the presence of a response, regardless of validity.

**Enumerators were required to provide a response for household count regardless of respondent participation. Household count was automatically set for CFU and TQA returns so this cell includes only NRFU and ETL returns.

*** Telephone number is required for the CFU operation so these cases are excluded from this cell.

**** Undercount is not asked in CFU but is passed through from the original form .

Table 2 provides the INR rates for person-level items. The majority of enumerator returns correspond to households that did not respond to the initial mail census and thus were a harder to enumerate population. The highest INR rate for HUs is age/date of birth on enumerator returns. This item also had the highest INR rate in Census 2000. The difference between self-response returns and enumerator returns is 9.2 percentage points. This difference is higher than the 6.9 percentage point difference in Census 2000 (Norris, 2003 and Appendix B).

The INR rate for self-response returns for the age/date of birth item in the 2010 Census is shown in Table 2 (0.8 percent). This is consistent with the corresponding Census 2000 Form Replication Experiment age/date of birth item (0.8 percent, SE=0.07). Although the gap between the age/date of birth INR rate for self-response and enumerator returns widened in 2010, the 2010 Census version of age/date of birth was comparable to the Census 2000 version (with respect to INR), controlling for time and population differences.

The 2010 Census INR rate for self-response returns for the Hispanic origin item (4.2 percent) is significantly higher than the Census 2000 Form Replication Experiment INR rate for Hispanic origin (3.7 percent, SE=0.17). The 2010 Census INR rate for self-response returns for the race item (3.3 percent) is significantly lower than the Census 2000 Form Replication Experiment INR

rate for race (4.2 percent, SE=0.22) (Reiser et al.). It is interesting to note that, in both Census 2000 and the 2010 Census, enumerator returns had relatively lower INR rates for Hispanic origin compared to self-response returns.

The GQ INR rates are relatively high on all applicable items compared to other enumerator returns (see Table 7 discussion).

Table 2: 2010 Census Item Nonresponse Rates for Person-Level Items - National

Form Type/Operation	Percent Item Nonresponse					
	Relationship*	Sex	Age/Date of Birth	Hispanic Origin	Race	Overcount**
Overall	1.5	1.5	3.5	3.9	3.3	2.3
Self-Response	1.1	1.7	0.8	4.2	3.3	1.4
Enumerator Return	2.4	1.0	10.0	3.3	3.4	4.9
GQ Enumeration	n/a	3.0	6.5	25.0	18.1	n/a

Source: Census Unedited File

n/a - item is not asked for this questionnaire type.

* Excludes Person 1 (PNUM=1) since relationship uses first person as reference; for TQA, REF_PERSON=1 is excluded.

** For CFU cases, data for all items are passed through from the original form or asked if missing, with the exception of overcount, which is not asked if missing.

Table 3 shows the household-level INR rates by type of self-response return. The MOMB returns make up the vast majority of forms in the self-response category. In the MOMB category, the INR rates for replacement forms (English only) were higher for each household-level item, compared to the INR rates for the initial English forms. Update/Leave was higher in each category than MOMB initial questionnaires.

Although the undercount item has high INR rates for Fulfillment forms, the number of returns represented by those rates was small.

Table 3: 2010 Census Item Nonresponse Rates for Self-Response Household-Level Items - National

Form Type/Operation	Percent Item Nonresponse			
	Household Count	Undercount	Tenure	Telephone Number
Self-Response	1.8	7.3	2.4	7.8
MOMB	1.8	7.3	2.4	7.8
- Initial English	1.5	6.8	2.1	7.6
- Initial Bilingual*	4.3	11.5	4.7	8.3
- Replacement**	2.8	9.5	3.8	11.7
Update/Leave	2.4	7.7	3.2	9.0
Be Counted	n/a	n/a	2.3	6.3
- English	n/a	n/a	1.9	6.5
- Spanish	n/a	n/a	5.8	4.9
- Chinese	n/a	n/a	2.7	3.0
- Korean	n/a	n/a	3.5	1.5
- Russian	n/a	n/a	6.2	+
- Vietnamese	n/a	n/a	5.4	+
Fulfillment	2.5	13.1	4.4	4.4
- English	1.5	6.4	2.9	5.8
- Spanish	3.8	21.8	6.4	2.7
- Chinese	4.9	21.1	4.0	1.7
- Korean	+	13.4	2.8	1.7
- Russian	4.0	10.4	5.9	+
- Vietnamese	2.9	21.0	9.3	2.5

Source: Census Unedited File

n/a - item is not asked for this questionnaire type.

*Note these rates differ slightly from the INR rates in the 2010 Bilingual Assessment Report because those rates represent all bilingual returns where the rates in this report only represent returns that made it to the CUF.

** Replacement forms were English only.

+ Cell statistic involves less than 10 cases.

Table 4 presents the household-level INR rates by type of enumerator return. The NRFU forms made up the vast majority of forms in the enumerator return category.

Table 4: 2010 Census Item Nonresponse Rates for Enumerator Return Household-Level Items -- National

Form Type/Operation	Percent Item Nonresponse			
	Household Count	Undercount	Tenure	Telephone Number
Enumerator Return	<0.1	12.2	9.7	9.7
CFU	**	****	0.2	*
- English	**	****	0.2	*
- Spanish	**	****	0.1	*
TQA	**	0.8	0.9	0.9
- English	**	0.8	0.9	1.0
- Spanish	**	0.7	0.5	+
NRFU and Others***	<0.1	12.3	11.2	9.7
ETL	0.3	n/a	6.0	13.2

Source: Census Unedited File

n/a - item is not asked for this questionnaire type.

* Telephone number is required for the CFU operation.

** Household Count is automatically set for CFU and TQA returns based on number of person records.

*** These forms were used for NRFU, UE, RA, and RUE and did not have a Spanish version.

**** Undercount is not asked in CFU but is passed through from the original form .

+ Cell statistic involves less than 10 cases.

Table 5 provides person-level INR rates by type of self-response return. Again, MOMB returns represent the vast majority of forms in the self-response category. Replacement forms have higher INR rates for all items compared to MOMB initial English forms. This could be an indicator that these respondents are reluctant to provide data.

Chinese, Korean, and Vietnamese Be Counted and Fulfillment forms had relatively high INR rates for Hispanic Origin, compared to the other form types, though it should be noted that these form types accounted for less than 0.1 percent of returns. Cognitive research has shown that this question is difficult for Asian respondents, which may have caused them to skip this item.

- Many respondents do not understand what the word “Hispanic” means. The Chinese translation for the word “Hispanic” is translated as “western language descendant.” Some respondents interpret this to mean English speakers (Pan et al., 2005). In Korean, the word “Hispanic” is translated as “Latin America,” which may have led some respondents to assume this meant people of America. In Vietnamese, “Hispanic” is translated as “Spanish-speaking Latin American person” (Pan et al., 2009).
- The instruction for the Hispanic origin and race question can be confusing for Asian respondents. Many Asian respondents are used to seeing instructions after a question, not before. The instruction has the word “BOTH” in uppercase letters which cannot be shown in some languages, such as Chinese. In addition, the word “both” translates to “simultaneously” in Chinese. Respondents may wonder why they were asked to answer the Hispanic origin and race items simultaneously and then were told that “Hispanic origins are not races” (Pan et al., 2009).

Race INR rates are high for Spanish returns and for the bilingual English/Spanish returns. This was expected as past research has shown that Hispanic respondents tend to view Hispanic origin and race as the same construct and thus opt to fill out the Hispanic origin item only (Humes, 2009).

Table 5: 2010 Census Item Nonresponse Rates for Self-Response Person-Level Items - National

Form Type/Operation	Percent Item Nonresponse					
	Relationship*	Sex	Age/Date of Birth	Hispanic Origin	Race	Overcount
Self-Response	1.1	1.7	0.8	4.2	3.3	1.4
MOMB	1.1	1.7	0.8	4.2	3.3	1.4
- Initial English	1.0	1.5	0.8	4.1	2.1	1.2
- Initial Bilingual**	1.6	2.8	0.9	3.8	12.2	2.6
- Replacement***	1.8	2.6	1.4	6.9	5.0	2.5
Update/Leave	1.1	1.9	1.0	4.6	2.6	1.5
Be Counted	2.0	1.9	1.1	8.5	5.8	n/a
- English	1.9	1.5	1.0	8.9	3.6	n/a
- Spanish	2.4	3.2	1.4	1.8	18.6	n/a
- Chinese	1.5	3.0	0.8	29.7	1.0	n/a
- Korean	1.1	1.7	0.7	18.8	0.4	n/a
- Russian	3.3	2.6	+	9.5	1.0	n/a
- Vietnamese	2.2	4.2	0.9	27.4	0.5	n/a
Fulfillment	1.4	2.8	1.0	3.5	10.8	4.1
- English	1.3	1.9	1.0	4.6	2.3	1.4
- Spanish	1.4	3.6	1.0	1.6	18.1	6.3
- Chinese	0.8	2.3	0.5	40.9	0.6	6.5
- Korean	0.9	1.3	0.7	18.7	0.6	4.9
- Russian	3.0	+	+	10.4	+	2.8
- Vietnamese	1.4	2.4	+	26.0	0.7	5.2

Source: Census Unedited File

n/a - item is not asked for this questionnaire type.

* Excludes Person 1 since relationship uses first person as reference.

**Note these rates differ from the INR rates in the 2010 Bilingual Assessment Report because those rates represent all bilingual returns where the rates in this report only represent returns that made it to the CUF.

*** Replacement forms were English only.

+ Cell statistic involves less than 10 cases.

Table 6 provides person-level INR rates by type of enumeration return. Rates that stand out include high INR rates for the age/date of birth item on NRFU and ETL returns. The NRFU and ETL populations are traditionally harder to enumerate, as evidenced by specialized personal visit enumeration procedures (U.S. Census Bureau, 2009c) which are, in turn, traditionally associated with lower data quality. In addition, occupied HU interviews in NRFU utilized a proxy respondent 23.8 percent of the time (Walker et al.). This may have affected INR rates as research has shown that INR is more common when using proxy respondents (Shaw et al., 2000).

Research on 2010 NRFU using behavior coding methods showed that interviewers altered the age/date of birth questions (both questions were asked together on the form) frequently. The age question was asked differently from how it was scripted 78 percent of the time for Person 1 and 33 percent of the time for Persons 2 through 5. In addition, respondents answered the age and date of birth items inadequately approximately 53 percent of the time. The inadequate responses appear to primarily be due to respondents only answering age OR date of birth (both needed to be answered to be considered adequate). This was very likely due to the interviewers changing the question to request only a single piece of information during the initial interview. The interviewers omitted either the date of birth item (accounts for 57 percent of the major changes to this item) or the age item (accounts for 25 percent of the major changes to the item). In some (but not all) cases, interviewers proceeded to collect the additional piece of information during the interview (Childs and Jurgenson, 2011).

Overall, persons with a usual-home-elsewhere that were enumerated on GQ forms have relatively high INR rates for many items except the age/date of birth item, compared to the rates for overall enumerator returns. The GQs are often enumerated from personnel at the facility or lists provided by the facility management, who may have limited knowledge of the residents' characteristics (U.S. Census Bureau, 2009a).

Table 6: 2010 Census Item Nonresponse Rates for Enumerator Return Person-Level Items - National

Form Type/Operation	Percent Item Nonresponse					
	Relationship*	Sex	Age/Date of Birth	Hispanic Origin	Race	Overcount
Enumerator Return	2.4	1.0	10.0	3.3	3.4	4.9
CFU**	0.7	0.1	0.4	0.6	1.0	n/a
- English	0.5	0.1	0.4	0.6	0.8	n/a
- Spanish	2.3	0.1	0.5	0.5	2.5	n/a
TQA	0.1	0.1	0.9	1.3	1.8	1.0
- English	0.1	0.1	0.9	1.2	1.4	0.9
- Spanish	0.1	+	0.8	1.4	4.2	1.9
NRFU and Others***	2.9	1.2	12.3	4.0	4.0	4.9
ETL	3.3	0.6	12.0	6.7	6.0	6.9
GQs on Person File****	n/a	3.1	1.9	6.5	4.5	n/a

Source: Census Unedited File

n/a - item is not asked for this questionnaire type.

* Excludes Person 1 (PNUM=1) since relationship uses first person as reference; for TQA REF_PERSON=1 is excluded.

** For CFU cases, data for all items are passed through from the original form or asked if missing, with the exception of overcount, which is not asked if missing.

*** These forms were used for NRFU, UE, RA, and RUE and did not have a Spanish version.

**** These GQ forms identified persons with a Usual-Home-Elsewhere that correctly belong in the regular housing unit universe.

+ Cell statistic involves less than 10 cases.

Table 7 provides person-level INR rates by type of GQ return. A total of 7,917,579 persons reside in GQs in the INR universe. This number is the base for the first line in Table 7. Of this number, 7,894,078 persons reside in GQs that were enumerated on Individual Census Reports (ICRs), Military Census Reports (MCRs), or Shipboard Census Reports (SCRs), as opposed to those who were originally on Be Counted forms but indicated no usual residence and thus were put in the GQ universe. Most (95.7 percent) are on ICRs, 3.7 percent are on MCRs, and 0.7 percent are on SCRs.

Looking at the unedited GQ returns by type of GQ, 49.6 percent are in institutional GQs, 50.0 percent are in noninstitutional GQs, and 0.4 percent have unknown GQ type. Note that by design, GQ type was not allowed to be changed after the Group Quarters Validation (GQV) operation.

This meant that for any GQ that was vacant during GQV and for which we could not establish a type, the GQ retained the unknown GQ type at the end of the census.

Race and Hispanic origin INR rates are relatively high across all form types and all GQ types. This is possibly caused by:

- difficulty in contacting respondents, especially in noninstitutional GQs,
- the limited knowledge of the personnel who might be providing the enumeration data, or
- limited information available from a list provided by the property management (U.S. Census Bureau, 2009c).

Institutional GQs have lower INR rates than noninstitutional GQs for every item except Hispanic origin, where they are slightly higher.

Table 7: 2010 Census Item Nonresponse Rates for GQ Person-Level Items on GQ Forms by Form Type and by GQ Type (when GQ type is known) - National

Form Type/Operation	Percent Item Nonresponse			
	Sex	Age/Date of Birth	Hispanic Origin	Race
GQs on GQ File	3.0	6.5	25.0	18.1
ICR	2.9	6.4	24.7	17.6
- ICR English	2.9	6.2	24.6	17.3
- ICR Spanish	7.4	16.0	30.1	31.5
MCR	2.0	9.7	35.5	32.3
SCR	13.3	12.5	15.9	15.1
Institutional GQs	2.9	5.2	25.4	14.7
- Correctional Facilities for Adults	3.3	7.0	27.8	14.7
- Juvenile Facilities	2.5	3.2	18.9	13.6
- Nursing Facilities	2.4	2.7	22.2	14.7
- Other Institutional Facilities	2.7	3.1	27.8	17.7
Noninstitutional GQs	3.1	7.9	24.7	21.5
- College/University Student Housing	2.7	5.8	26.8	24.1
- Military Quarters	3.4	9.9	32.3	29.5
- Other Noninst. Facilities	3.6	12.0	17.7	13.0

Source: Census Unedited File

4.2 Imputation Rates

Housing unit data in the imputation universe include all occupied HUs in the final census, excluding housing units from 2010 Census experiments and substituted HUs. Substituted housing units/persons were also excluded from the 2008 Census Dress Rehearsal and the Census 2000 imputation reports. The HU imputation universe is 113,757,437. Note that this is lower than the INR HU universe because the imputation analysis excludes HUs that contain substituted persons (the difference is 1,898,558 HUs).

Person-level items include all non-substituted persons on primary returns who were not part of 2010 Census experiments. Substituted persons are not included in the data for item imputation (i.e., assignment and allocation) workload since whole persons are substituted from existing person-level items. Totally allocated persons are included. This constitutes a person universe of 293,761,205 persons. This number is higher than the INR person universe because it includes totally allocated persons.

There are 7,987,323 persons in the GQ imputation universe. Of these, 7,894,078 GQ persons were enumerated on an ICR, MCR, or SCR, which is equivalent to the GQ INR universe. The remaining persons in the GQ imputation universe were originally listed on Be Counted forms, but the respondents indicated no usual residence and thus were put in the GQ universe.

For a detailed description of how “as reported,” assignment, and allocation rates are computed, refer to Section 2.2.2.

4.2.1 Assignment and Allocation rates

Tables 8 through 11 present “as reported,” assignment, and allocation rates for household-level and person-level items. In Tables 8 through 11, “as reported” rates are based on valid responses, whereas INR rates in Tables 1 through 7 were based on the presence of a response, regardless of validity. Thus, the imputation rates in these tables cannot be directly compared to the previous INR rates.

Table 8 presents imputation rates for each of the six data items. Age/date of birth, Hispanic origin, and race have the highest imputation rates. Similar to INR rates, the enumerator returns seem to be the reason for the high age/date of birth and tenure imputation rates.

Keeping in mind the limitations described in Section 2.2.4, the overall imputation rates in Table 8 are in line with what was seen in Census 2000 (Appendices C and D). Self-response rates in the 2010 Census are similar to short-form self-response imputation rates in Census 2000.

Table 8: Overall “As Reported” and Imputation Rates (in percent) - National

Mode	Person-Level Items					Household-Level Item
	Relationship	Sex	Age/Date of Birth	Hispanic Origin	Race	Tenure*
Overall**						
As Reported	97.9	98.4	95.0	95.5	95.9	96.5
Imputed:						
Assigned	2.1	1.6	5.1	4.5	4.1	3.5
Allocated	0.5	1.3	1.5	1.7	1.2	*
	1.7	0.3	3.6	2.8	2.9	3.5
Self-Response						
As Reported	98.3	98.3	98.1	95.2	95.8	97.4
Imputed:						
Assigned	1.7	1.8	1.9	4.8	4.2	2.6
Allocated	0.3	1.6	1.1	2.0	1.3	*
	1.4	0.2	0.7	2.8	2.9	2.6
Enumerator Return						
As Reported	97.1	99.0	87.7	96.7	96.3	94.1
Imputed:						
Assigned	2.9	1.0	12.3	3.3	3.7	5.9
Allocated	0.9	0.7	2.4	0.9	1.0	*
	2.0	0.3	9.9	2.4	2.7	5.9
GQ Enumeration						
As Reported	n/a	96.2	91.4	74.3	81.1	n/a
Imputed:						
Assigned	n/a	3.8	8.6	25.7	19.0	n/a
Allocated	n.a	2.4	1.2	13.7	9.8	n/a
	n/a	1.4	7.5	11.9	9.2	n/a

Source: Census Edited File

n/a - item is not asked for this questionnaire type.

*Tenure could not be assigned in the 2010 Census in the absence of long form data.

** The Overall category includes 223,206 persons on “other returns” (besides self-response and enumerator returns), which represent HU or person records that were created during Response Processing System activities as a result of various situations/processes in which an original return/person was unavailable, unusable, or insufficient. For example, person records were created to match population counts derived when not all person data items were received. These persons are categorized as totally allocated persons.

Note: Imputation rates include assignment and allocation rates.

Table 9 presents the imputation rates for self-response returns by form type. Similar to the INR rate trend in Tables 3 and 5, replacement forms have higher imputation rates for all items, compared to the initial MOMB English forms. This is possibly due to population differences – people who respond to the initial MOMB form are likely different than those who respond to the replacement forms. As expected, Be Counted forms have relatively high imputation rates for relationship, Hispanic origin, and race, since the population enumerated on these forms is

traditionally hard to enumerate, as evidenced by the use of specialized self-response enumeration procedures (U.S. Census Bureau, 2009d).

Table 9: 2010 Census “As Reported” and Imputation Rates (in percent) for Self-Response - National

Form Type		Person-Level Items					Household-Level Item
		Relationship	Sex	Age/Date of Birth	Hispanic Origin	Race	Tenure*
MOMB	Overall:						
	As Reported	98.4	98.3	98.2	95.2	95.9	97.4
	Assigned/allocated	1.6	1.7	1.9	4.8	4.2	2.6
	Initial English:						
	As Reported	98.6	98.4	98.3	95.4	97.0	97.7
	Assigned/Allocated	1.4	1.6	1.7	4.6	3.0	2.4
	Initial Bilingual:						
	As Reported	97.2	97.2	97.5	95.1	86.7	95.0
	Assigned/Allocated	2.8	2.8	2.5	4.9	13.3	5.1
	Replacement:						
As Reported	96.9	97.3	96.9	91.7	93.3	96.0	
Assigned/Allocated	3.1	2.7	3.1	8.3	6.7	4.0	
UL	As Reported	97.7	98.1	97.9	94.0	95.9	96.5
	Assigned/Allocated	2.3	1.9	2.1	6.0	4.1	3.6
Be Counted	Overall:						
	As Reported	75.5	98.1	96.6	84.5	86.8	97.3
	Assigned/Allocated	24.6	1.9	3.4	15.5	13.2	2.7
	English:						
	As Reported	78.3	98.5	96.9	85.8	90.6	97.6
	Assigned/Allocated	21.7	1.6	3.2	14.2	9.4	2.4
	Spanish:						
	As Reported	64.7	96.7	95.3	82.4	68.5	93.6
	Assigned/Allocated	35.3	3.3	4.7	17.6	31.5	6.4
	Chinese:						
	As Reported	62.7	97.0	96.5	65.6	92.3	97.2
	Assigned/Allocated	37.3	3.0	3.5	34.5	7.7	2.8
	Korean:						
	As Reported	72.4	98.2	96.8	78.3	96.2	96.2
Assigned/Allocated	27.6	1.8	3.2	21.7	3.8	3.8	
Russian:							
As Reported	69.9	97.4	98.1	85.2	93.0	93.5	
Assigned/Allocated	30.1	2.6	1.9	14.8	7.1	6.5	

Fulfillment	Vietnamese:							
	As Reported	67.8	95.8	96.6	63.7	87.5	93.2	
	Assigned/Allocated	32.2	4.2	3.4	36.3	12.5	6.8	
	Overall:							
	As Reported	96.3	97.1	97.4	93.8	86.7	95.3	
	Assigned/Allocated	3.8	2.9	2.6	6.2	13.3	4.7	
	English:							
	As Reported	98.0	98.1	98.1	94.5	96.2	97.0	
	Assigned/Allocated	2.0	1.9	1.9	5.5	3.8	3.1	
	Spanish:							
	As Reported	94.9	96.3	96.9	94.1	78.9	93.0	
	Assigned/Allocated	5.1	3.7	3.1	5.9	21.1	7.0	
	Chinese:							
	As Reported	96.5	97.7	97.2	57.9	97.5	95.2	
	Assigned/Allocated	3.5	2.3	2.8	42.1	2.5	4.8	
	Korean:							
	As Reported	98.3	98.7	97.0	81.3	99.4	96.8	
	Assigned/Allocated	1.7	1.4	3.0	18.7	0.6	3.2	
	Russian:							
As Reported	97.9			89.2		93.9		
Assigned/Allocated	2.1	+	+	10.8	+	6.2		
Vietnamese:								
As Reported	95.4	97.6	98.0	71.8	96.8	89.8		
Assigned/Allocated	4.7	2.4	2.0	28.2	3.3	10.2		

Source: Census Edited File

*Tenure could not be assigned in the 2010 Census in the absence of long form data.

+ Cell statistic involves less than 10 cases.

Table 10 presents the imputation rates for enumerator returns by form type. Similar to the INR trend shown in Table 6, imputation rates are substantially higher for the age/date of birth item on NRFU and ETL returns.

Table 10: “As Reported” and Imputation Rates (in percent) for Enumerator Returns - National

Form Type	Person-Level Items					Household-Level Item
	Relationship	Sex	Age/Date of Birth	Hispanic Origin	Race	Tenure*
CFU						
As Reported	97.6	99.9	97.5	99.2	98.6	99.3
Imputed:	2.4	0.1	2.5	0.8	1.4	0.7
Assigned	1.0	0.1	2.0	0.2	0.3	*
Allocated	1.4	<0.1	0.5	0.6	1.1	0.7
TQA						
As Reported	91.6	99.9	98.9	98.8	97.8	99.1
Imputed:	8.4	0.1	1.1	1.3	2.2	0.9
Assigned	0.9	0.1	0.1	0.4	0.6	*
Allocated	7.5	<0.1	1.0	0.9	1.6	0.9
NRFU and Others**						
As Reported	97.0	98.8	85.2	96.0	95.7	93.3
Imputed:	3.1	1.2	14.8	4.0	4.3	6.8
Assigned	0.9	0.8	2.5	1.1	1.2	*
Allocated	2.2	0.4	12.3	2.9	3.1	6.8
ETL						
As Reported	98.2	99.4	85.7	93.3	93.6	94.6
Imputed:	1.9	0.7	14.3	6.7	6.4	5.4
Assigned	0.5	0.6	2.2	0.8	0.8	*
Allocated	1.3	0.1	12.1	6.0	5.6	5.4
GQ on Person File***						
As Reported	n/a	96.9	95.1	93.4	95.3	n/a
Imputed:	n/a	3.2	5.0	6.6	4.7	n/a
Assigned	n/a	2.9	3.1	+	+	n/a
Allocated	n/a	0.2	1.8	6.6	4.7	n/a

Source: Census Edited File

n/a - item is not asked for this questionnaire type.

*Tenure could not be assigned in the 2010 Census in the absence of long form data.

** These forms were used for NRFU, UE, RA, and RUE and did not have a Spanish version.

*** These GQ forms identified persons with a Usual-Home-Elsewhere that correctly belong in the regular housing unit universe.

+ Cell statistic involves less than 10 cases.

Note: Imputation rates include assignment and allocation rates.

Table 11 shows the imputation rates for GQs. The first set of rows in Table 11 represents 7,987,323 persons living in GQs in the imputation universe. Of these, 7,894,078 GQ persons were enumerated on an ICR, MCR, or SCR as opposed to those originally on Be Counted forms that indicated no usual residence and thus were put in the GQ universe.

The remaining rows of Table 11 include only those GQ persons enumerated on one of the GQ forms. About half of the persons residing in these GQs are in institutional GQs. The other half are in noninstitutional GQs.

Similar to the INR rate trend in Table 7, imputation rates are relatively high for Hispanic origin and race across all GQ types.

Table 11: “As Reported” and Imputation Rates (in percent) for Group Quarters - National

Form Type	Person-Level Items			
	Sex	Age/Date of Birth	Hispanic Origin	Race
GQ Overall				
As Reported	96.2	91.4	74.3	81.1
Imputed:	3.8	8.6	25.7	19.0
Assigned	2.4	1.2	13.7	9.8
Allocated	1.4	7.5	11.9	9.2
Institutional GQs				
As Reported	97.1	93.6	74.6	85.1
Imputed:	2.9	6.4	25.4	14.9
Assigned	2.6	1.2	13.7	7.4
Allocated	0.3	5.3	11.7	7.5
- Correctional Facilities for Adults				
As Reported	96.7	92.5	72.2	85.1
Imputed:	3.3	7.5	27.8	14.9
Assigned	3.0	0.9	13.1	5.4
Allocated	0.3	6.6	14.8	9.5
- Juvenile Facilities				
As Reported	97.4	90.9	81.8	86.4
Imputed:	2.6	9.1	18.2	13.6
Assigned	1.8	0.9	7.2	4.7
Allocated	0.8	8.3	11.0	9.0
- Nursing Facilities				
As Reported	97.6	95.4	77.9	85.3
Imputed:	2.4	4.6	22.2	14.7
Assigned	2.2	1.6	15.5	10.5
Allocated	0.2	3.0	6.7	4.2
- Other Institutional Facilities				
As Reported	97.3	95.0	70.5	81.1
Imputed:	2.7	5.0	29.5	18.9
Assigned	1.9	1.2	12.2	8.0
Allocated	0.9	3.7	17.4	10.9

NonInstitutional Facilities					
	As Reported	96.9	90.8	75.3	78.4
	Imputed:	3.1	9.3	24.7	21.6
	Assigned	2.3	1.2	14.1	12.4
	Allocated	0.8	8.1	10.7	9.3
<hr/>					
-	College/University Housing				
	As Reported	97.3	93.3	73.2	75.7
	Imputed:	2.8	6.7	26.8	24.3
	Assigned	2.3	0.8	17.1	15.3
	Allocated	0.4	5.9	9.8	9.0
-	Military Quarters				
	As Reported	96.6	88.9	67.7	70.3
	Imputed:	3.5	11.2	32.4	29.7
	Assigned	1.8	1.2	18.4	17.3
	Allocated	1.7	9.9	13.9	12.4
-	Other NonInstitutional Facilities				
	As Reported	96.4	85.5	82.3	86.9
	Imputed:	3.6	14.5	17.7	13.1
	Assigned	2.2	2.1	5.9	4.1
	Allocated	1.4	12.4	11.7	9.0

Source: Census Edited File

Note: Imputation rates include assignment and allocation rates.

4.2.2 Previous Census Response

As stated earlier, a new approach to assigning race and Hispanic origin responses was implemented during the 2010 Census. When race or Hispanic origin was left blank or was inconsistent with other responses and a value could not be determined based on information provided for that same person or from within the household, a value was assigned based on matched Census 2000 responses or previous ACS responses, if possible. Note that the responses used from Census 2000 and ACS were “as reported” (i.e., not imputed). If a value could not be determined based on matched Census 2000 or ACS responses, then a value from a person in a nearby HU was allocated.

Approximately 36.6 percent of all imputed Hispanic origin responses and 28.6 percent of all imputed race responses used information from Census 2000 or a previous ACS response. This new approach for assigning race and Hispanic origin appears to be beneficial, in terms of the quantity of usable data. (See Cardella, 2010 for details of the use of Census 2000 and ACS response data)

4.2.3 Data Completeness Statistic

Table 12 provides the number and percent of persons with non-imputed responses to the person-level items (relationship, sex, age/date of birth, Hispanic origin, race). This is a distribution of the level of INR by person (i.e., number of persons with zero to five person-level item responses). Note that results for Table 12 include all non-substituted persons on primary returns in occupied HUs.

On every form type, the majority of persons responded to all five of the person-level data items (four for GQs). Over 97 percent of all persons in this universe answered at least four of five data items.

Table 12: Data Completeness Statistic

Form Type	Number of Person-Level Characteristics Reported						
	Total	5 of 5	4 of 5	3 of 5	2 of 5	1 of 5	0 of 5
Overall*	292,742,798 (100.0)	255,451,086 (87.3)	30,054,023 (10.3)	4,466,459 (1.5)	2,144,788 (0.7)	380,052 (0.1)	246,390 (0.1)
Self-Response	205,159,791 (100.0)	182,775,485 (89.1)	19,005,583 (9.3)	2,614,197 (1.3)	625,642 (0.3)	126,819 (0.1)	12,065 (<0.1)
- MOMB	203,745,574 (100.0)	181,645,463 (89.2)	18,771,689 (9.2)	2,577,154 (1.3)	614,568 (0.3)	124,780 (0.1)	11,920 (<0.1)
• Initial English	174,233,941 (100.0)	157,995,168 (90.7)	13,772,149 (7.9)	1,886,918 (1.1)	471,479 (0.3)	98,324 (0.1)	9,903 (<0.1)
• Bilingual	20,123,323 (100.0)	15,851,233 (78.8)	3,692,603 (18.4)	476,764 (2.4)	86,303 (0.4)	15,339 (0.1)	1,081 (<0.1)
• Replacement	9,388,310 (100.0)	7,799,062 (83.1)	1,306,937 (13.9)	213,472 (2.3)	56,786 (0.6)	11,117 (0.1)	936 (<0.1)
- Update/ Leave	535,317 (100.0)	477,611 (89.2)	48,153 (9.0)	7,167 (1.3)	1,966 (0.4)	385 (0.1)	35 (<0.1)
- Be Counted	573,513 (100.0)	406,520 (70.9)	135,498 (23.6)	22,612 (3.9)	7,444 (1.3)	1,355 (0.2)	84 (<0.1)
- Fulfillment	305,387 (100.0)	245,891 (80.5)	50,243 (16.5)	7,264 (2.4)	1,664 (0.5)	299 (0.1)	26 (<0.1)
Enumerator Return	87,359,801 (100.0)	72,675,601 (83.2)	11,048,440 (12.7)	1,852,262 (2.1)	1,519,146 (1.7)	253,233 (0.3)	11,119 (<0.1)
- CFU	17,326,236 (100.0)	16,251,587 (93.8)	959,442 (5.5)	81,122 (0.5)	27,136 (0.2)	6,571 (<0.1)	378 (<0.1)
- TQA	148,919 (100.0)	132,134 (88.7)	15,100 (10.1)	998 (0.7)	510 (0.3)	142 (0.1)	35 (<0.1)
- NRFU	69,680,985 (100.0)	56,137,248 (80.6)	10,039,315 (14.4)	1,763,426 (2.5)	1,484,419 (2.1)	245,929 (0.4)	10,648 (<0.1)
- ETL	186,946 (100.0)	151,378 (81.0)	23,223 (12.4)	4,943 (2.6)	6,824 (3.7)	565 (0.3)	13 (<0.1)
- GQ on Person File**	16,715 (100.0)	3,254 (19.5)	11,360 (68.0)	1,773 (10.6)	257 (1.5)	26 (0.2)	45 (0.3)
GQ Enum.	7,987,323 (100.0)	n/a	5,275,862 (66.1)	1,296,791 (16.2)	1,143,327 (14.3)	108,547 (1.4)	162,796 (2.0)

Source: Census Edited File

Note: Only persons who were requested to respond to full person-level data are included in this table. On the initial English questionnaire, replacement questionnaire, update/leave questionnaire, and fulfillment questionnaire, full person-level data were collected for Persons 1 through 6. On the bilingual questionnaire, full person-level data were collected for Persons 1 through 8. On the NRFU, Be Counted, and ETL questionnaires, full person-level data were collected for Persons 1 through 5. Also, Person 1 is included in the relationship analysis on the CEF as “Householder”, so Person 1s always have at least one item filled with the exception of totally allocated persons.

* The “Overall” category includes 223,206 persons on “other returns” (besides self-response and enumerator returns) which represent HU or person records that were created during Response Processing System activities as a result of various situations/processes in which an original return/person was unavailable, unusable, or insufficient. For example, person records were created to match population counts derived when not all person-level items were received. These persons are categorized as totally allocated persons.

** These GQ forms identified persons with a Usual-Home-Elsewhere that correctly belong in the regular housing unit universe. n/a – GQ forms do not collect information for relationship so they can have a maximum of only four responses per person.

Table 13 provides the number and percent of non-imputed responses to person-level items by person number. Results are relatively consistent across persons. Full person-level responses (5 of 5) decreases for each additional person listed on the questionnaire. This is possibly due to householders (Person 1) who may not be able to answer all the data items for other household members, especially non-family members.

Table 13: Data Completeness Statistic - Number and Percent of Persons in Housing Units by Sum of Non-Imputed Responses to Person-Level Items

Person Number*	Number of Characteristics Reported						
	Total	5 of 5	4 of 5	3 of 5	2 of 5	1 of 5	0 of 5
Overall	292,742,798 (100.0)	255,451,086 (87.3)	30,054,023 (10.3)	4,466,459 (1.5)	2,144,788 (0.7)	380,052 (0.1)	246,39 0 (0.1)
Person 1	113,757,356 (100.0)	100,929,366 (88.7)	10,402,259 (9.1)	1,396,258 (1.2)	912,679 (0.8)	103,221 (0.1)	13,573 (<0.1)
Person 2	83,617,509 (100.0)	73,471,384 (87.9)	8,133,058 (9.7)	1,188,570 (1.4)	564,509 (0.7)	144,905 (0.2)	115,08 3 (0.1)
Person 3	46,207,808 (100.0)	39,985,780 (86.5)	5,049,865 (10.9)	779,339 (1.7)	289,051 (0.6)	61,410 (0.1)	42,363 (0.1)
Person 4	27,877,847 (100.0)	23,682,816 (85.0)	3,380,248 (12.1)	563,708 (2.0)	191,085 (0.7)	36,525 (0.1)	23,465 (0.1)
Person 5	12,598,175 (100.0)	10,403,537 (82.6)	1,755,660 (13.9)	312,420 (2.5)	99,492 (0.8)	18,071 (0.1)	8,995 (0.1)
Person 6**	5,179,139 (100.0)	4,212,725 (81.3)	755,446 (14.6)	133,689 (2.6)	48,253 (0.9)	8,766 (0.2)	20,260 (0.4)

Source: Census Edited File

Note: Person 1 is included in the relationship analysis on the CEF as “Householder”, so Person 1s always have at least one item filled with the exception of totally allocated persons.

*There are 3,504,964 persons who were enumerated as Person 7 or higher in this universe. Since most form types collect all five characteristics for Persons 1 through 6 but only partial information for Persons 7 and higher, these persons are not comparable and thus excluded from this table .

** Be Counted and ETL Person 6 and higher are not included in this analysis because those forms collected full person-level data for Persons 1 through 5 only.

4.2.4 Totally Allocated/Assigned Persons

When each characteristic for a person record requires editing, allocation, or assignment, the person record is considered totally allocated/assigned (as long as one person within the household has reported data). By this definition, the last column in Table 12 (0 out of 5 responses) is comprised of totally allocated/assigned persons.

Table 14 shows the number and percent of totally allocated/assigned persons from Table 12, as well as totally allocated/assigned persons excluded in the Table 12 analysis. The percent of totally allocated persons across form types is also provided in Table 14. These results include all non-substituted persons in occupied, primary HUs.

Overall, totally allocated persons make up about 0.1 percent of our universe. Most of the totally allocated persons have form type values indicating the person records were created during processing. The remaining totally allocated persons are split between self-response forms (primarily MOMB) and enumerator returns (primarily NRFU).

Table 14: Number and Percent of Totally Allocated Persons by Form Type

	Number	Percent Within Form Type	Percent Across Form Type
Overall	247,904	0.1	100.0
Self-Response	13,579	<0.1	5.5
• All MOMB	13,234	<0.1	5.3
• Update/Leave	48	<0.1	<0.1
• Be Counted	264	<0.1	0.1
• Fulfillment	33	<0.1	<0.1
Enumerator Return	11,119	<0.1	4.5
• CFU	378	<0.1	0.2
• TQA	35	<0.1	<0.1
• NRFU	10,648	<0.1	4.3
• ETL	13	<0.1	<0.1
• GQ on Person File*	45	0.3	<0.1
Other Returns**	223,206	100.0	90.0

Source: Census Edited File

* These GQ forms identified persons with a Usual-Home-Elsewhere that correctly belong in the regular housing unit universe.

** The Other Returns category includes 223,206 persons on “other returns” (besides self-response and enumerator returns) which represent HU or person records that were created during Response Processing activities as a result of various situations/processes in which an original return/person was unavailable, unusable, or insufficient. For example, person records were created to match population counts derived when not all person-level items were received. These persons are categorized as totally allocated persons.

4.2.5 Substitution Rates

A whole-household substitution is used when all the person-level characteristics for every person in the HU are missing. A nearby HU of up to six people with complete data is selected to represent the missing data items for up to the first six persons in the HU needing substitution.¹⁰ For HUs with more than six people that require substitution, all characteristic data for the remaining persons are allocated.

¹⁰ If the HU to be substituted is a mail return, then the selected nearby HU is a mail return. Otherwise the selected nearby HU is a nonmail return. This stratification of return type only occurs for substitution, not for allocation.

There are 5,770,791 persons substituted in the 2010 Census. This represents 1.9 percent of all persons. All of the substituted persons are on form types created during Response Processing System activities as a result of various situations/processes in which an original return/person was unavailable, unusable, or insufficient. For example, person records were created to match population counts derived when not all person-level items were received.

5. Summary and Recommendations for Future Research

Table 15 provides an overall summary of INR and imputation in one table. Note that the INR rates cannot be directly compared to the imputation rates because the imputation rates are based on valid responses, whereas INR rates are based on the presence of a response, regardless of validity.

Table 15. Overall Item Nonresponse and Imputation Rates

	Person-Level Items					Household
	Relationship	Sex	Age/Date of Birth	Hispanic Origin	Race	-Level Item Tenure
Item Nonresponse	1.5	1.5	3.5	3.9	3.3	4.5
Total	100.0	100.0	100.0	100.0	100.0	100.0
“As Reported”	97.9	98.4	95.0	95.5	95.9	96.5
Imputed	2.1	1.6	5.1	4.5	4.1	3.5
Assigned	0.5	1.3	1.5	1.7	1.2	n/a
Allocated	1.7	0.3	3.6	2.8	2.9	3.5
Substituted	1.9 percent of all persons					

Suggestions for Future Research:

- Age/date of birth and tenure INR and imputation rates are high for enumerator NRFU and ETL returns; research on source of error should be focused on this area. Since the tenure issue is likely related to the placement of the question on the questionnaire, an automated instrument for the followup process might resolve this issue.
- Research to reduce Hispanic origin and race item nonresponse is recommended.
- Using previous Census and ACS response data for allocation of missing Hispanic origin and race proved to be beneficial. Expanding this allocation by using these sources, as well as other administrative records (including potentially non-primary returns), for all missing items is recommended research.
- It is also recommended that a quality check of the previous response data be examined by comparing the previous Census and ACS responses for Hispanic origin and race to the “as reported” Hispanic origin and race responses in the 2010 Census.
- Data for roughly 5.7 million people (1.9 percent) were acquired through substitution in the

2010 Census. Investigation into the operational source of substitutions is recommended. This includes operations producing cases needing substitution and operations contributing data for substitution.

6. References

- Alberti, Nicholas (2008), *Specifications for the 2010 Census Data Capture Audit and Resolution*, DSSD Decennial Census 2010 Memorandum Series #C-01, December 18, 2008.
- Cardella, Michael (2010), *Transmittal of Baseline Requirements for Editing 100% Items – 2010 Census*, Internal Memo from Lamas to Stoudt, January 4, 2010. Requirements, v1.2, December 28, 2010.
- Childs, Jennifer Hunter and Nathan Jurgenson (2011), *Behavior Coding of the 2010 Nonresponse Followup (NRFU) Interviews Report*, 2010 Census Planning Memoranda Series No. 150, August 30, 2011.
- Coon, David and Neil Osborne (expected, 2012), *2010 Census Decennial Response Integration System Paper Questionnaire Data Capture Assessment Report*, 2010 Decennial Memorandum Series, (Forthcoming Report).
- Hill, Joan, Frederic Lestina, Jason Machowski, Cynthia Rothhaas, and Kimberly Roye (2010), *Study of Respondents Who List Themselves as Person 1*, DSSD 2010 Memorandum Series # G-09, October 1, 2010.
- Humes, Karen (2009), *2010 Census Alternative Questionnaire Experiment: Race and Hispanic Origin Treatments, Information for Census Advisory Committees on Purpose and Rationale of 2010 AQE Treatments*, June 15, 2009.
- Norris, Sherri (2003), *Analysis of Item Nonresponse Rates for the 100 Percent Housing and Population Items from Census 2000*, Census 2000 Evaluation B.1.b, September 23, 2003.
- Pan, Yuling, M. Mandy Sha, Hyunjoo Park, Alisu Schoua-Glusberg (2009), *2010 Census Language Program: Pretesting of Census 2010 Questionnaire in Five Languages*, Research Report Series (Survey Methodology #2009-01), February 10, 2009.
- Pan, Yuling, Barbara Craig, and Suzanne Scollon (2005), *Results from Chinese Cognitive Interviews on the Census 2000 Long Form: Language, Literacy, and Cultural Issues*, Research Report Series (Survey Methodology #2005-09), September 6, 2005.

- Reiser, Courtney, Samantha Stokes, Elizabeth Compton, Michael Bentley (expected, 2012), *2010 Census Alternative Questionnaire Experiment (AQE): Census 2000 Form Replication Panel*, 2010 Decennial Memorandum Series, (Forthcoming Report).
- Rothhaas, Cynthia, Michael Bentley, Joan M. Hill, and Frederic Lestina (2011), *2010 Census: Bilingual Questionnaire Assessment Report*, 2010 Census Planning Memorandum Series No. 156, November 4, 2011.
- Shaw, Chris, Elaine McColl, and Senga Bond (2000), *Functional Abilities and Continence: The Use of Proxy Respondents in Research Involving Older People*, Vol. 9, No. 10, pp. 1117-1126, December, 2000.
- U.S. Census Bureau (2011), *2010 Census Operational Plan and Accompanying Operational Requirements Document*, 2010 Census Information Memorandum Series, No. 2 (Reissue 3), June 29, 2011.
- U.S. Census Bureau (2010), *Requirements for Editing 100% Items for 2010 Dress Rehearsal*, Version 1.2.1, November 17, 2010.
- U.S. Census Bureau (2009a), *2010 Census Detailed Operational Plan for Group Quarters Operation*, 2010 Census Information Memorandum Series No. 44, December 10, 2009.
- U.S. Census Bureau (2009b), *2010 Census Detailed Operational Plan for Telephone Questionnaire Assistance and Fulfillment Operations*, 2010 Census Informational Memorandum Series No. 48, November 30, 2009.
- U.S. Census Bureau (2009c), *2010 Census Detailed Operational Plan for Enumeration at Transitory Locations Field Operations*, 2010 Census Information Memorandum Series No. 28, September 23, 2009.
- U.S. Census Bureau (2009d), *2010 Census Detailed Operations Plan for Be Counted/Questionnaire Assistance Center Operations*, 2010 Census Information Memorandum Series No. 33, September 17, 2009.
- Walker, Shelley, Susanna Winder, Geoff Jackson, and Sarah Heimel (expected, 2012), *2010 Census Nonresponse Followup Operations (NRO) Assessment*, 2010 Decennial Memorandum Series, (Forthcoming Report).
- Zajac, Kevin (2003), *Analysis of Imputation Rates for the 100 Percent Person and Housing Unit Data Items from Census 2000*, Census 2000 Evaluation B.1.a, September 25, 2003.

Appendix A: Classification of Edit/Allocation Flag Variables Into Categories from the Census Edited File

Item	Variable	Category	Values
Tenure*	FTEN	Non-imputed values	0 = As Reported
		Assigned values	1 = Assigned by Consistency Check
		Allocated values	4 = Allocated due to no response 5 = Allocated because reported owned with and without a mortgage 6 = Allocated because reported rented and without payment of rent 7 = Allocated because reported owned and rented
Age/Date of Birth	FAGE	Non-imputed values	0 = Consistent as reported 1 = Age only (DOB not fully reported) 2 = Date of birth only (age is not reported)
		Assigned values	3 = Inconsistent age and date of birth (more than one year) 8 = Age of householder or spouse adjusted to be consistent with number of children
		Allocated values	4 = Allocated from hot deck
		Substituted values	7 - Substituted
Hispanic Origin	FSPAN	Non-imputed values	00 = As reported 01 = Mixed: single Hispanic code retained 02 = Mixed: single Hispanic code randomly selected from reported responses 03 = Multiple Hispanic response: single Hispanic code randomly selected from reported responses 04 = Multiple Non-Hispanic response: single non-Hispanic code randomly selected from reported responses
		Assigned values	05 = Assigned Hispanic from race code 07 = Assigned from previous census response
		Allocated values	06 = Allocated from within household 08 = Allocated from hot deck (Spanish surname used) 09 = Allocated from hot deck (Non-Spanish surname used) 10 = Allocated from hot deck (surname not used)
		Substituted values	11 = Substituted

Race	FRACE	Non-imputed values	0 = As reported
		Assigned values	1 = Code changed through consistency edit 3 = Assigned race from response in Hispanic question 9 = Assigned from previous census response
		Allocated values	4 = Allocated from within household 5 = Allocated from hot deck
		Substituted values	7 = Substituted
Relationship	FREL	Non-imputed values	0 = As reported
		Assigned values	1= Householder/spouse report same sex; relationship of spouse changed to unmarried partner. 2 = Value edited for household consistency 3 = Reported value changed for person in a GQ or else housing unit person is on a GQ form
		Allocated values	4 = Allocated from hot deck 5 = Allocated due to consistency check 8 = Allocated due to a response for extended roster person
		Substituted values	7 = Substituted
Sex	FSEX	Non-imputed values	0 = As reported
		Assigned values	1 = From first name 2 = Value edited for household consistency
		Allocated values	4 = Allocated from hot deck 5 = Allocated from consistency check
		Substituted values	7 = Substituted

*Tenure could not be assigned in the 2010 Census in the absence of long form data.

The Census 2000 INR rates are listed below. The self-response mode in the Census 2000 report included United States Postal Service delivery and Local Census Office delivery for Undeliverable As Addressed for Mailout/Mailback categories, Update Leave, Urban Update Leave, Internet, and Be Counted. Enumerator Return mode in the Census 2000 report included Coverage Followup, Telephone Questionnaire Assistance, Nonresponse Followup, Coverage Improvement Followup, Update Enumerate, List/Enumerate, and Remote Alaska.

Appendix B: Census 2000 Item Nonresponse Rates by Form Type and Response Mode (Norris, 2003)

		Percent Item Nonresponse					
		Relationship	Sex	Age	Hispanic Origin	Race	Tenure
Overall		1.3	1.1	3.7	3.1	2.9	4.1
Short Form		1.3	1.2	3.7	3.1	3.0	3.1
Long Form		1.7	1.0	4.1	3.2	2.6	9.6
Self-Response		1.1	0.9	1.9	3.2	2.8	2.6
Enumerator Response		2.1	1.7	8.8	2.7	3.2	8.8
Short Form	Self-Response	1.0	1.0	1.9	3.3	3.0	2.1
	Enumerator Return	2.1	1.7	8.9	2.6	3.1	6.3
Long Form	Self-Response	1.4	0.5	1.9	3.1	2.3	5.7
	Enumerator Return	2.2	1.9	8.4	3.2	3.4	17.8

Data source: HCUF

Note: Table includes only data-defined persons

Appendix C: Census 2000 Imputation Rates* by Form Type and Response Mode (Zajac, 2003)

		Percent Imputation					Tenure
		Relationship	Sex	Age	Hispanic Origin	Race	
Overall		2.6	2.0	5.1	4.4	4.0	5.5
Short Form		2.5	2.0	5.0	4.3	4.0	3.7
Long Form		3.2	2.1	5.7	4.6	3.8	13.2
Self-Response		2.0	1.7	3.0	4.5	3.8	3.0
Enumerator Response		4.1	2.8	10.9	4.1	4.3	12.2
Short Form	Self-Response	1.9	1.7	2.9	4.5	3.9	2.1
	Enumerator Return	4.1	2.7	11.0	3.9	4.2	9.0
Long Form	Self-Response	2.5	1.5	3.1	4.5	3.4	8.0
	Enumerator Return	4.4	3.2	10.6	4.7	4.7	23.6

Data source: HCEF_D', HCUF

* Imputation rates include assignment and allocation rates.

Appendix D: Census 2000 Imputation Rates* by Form Source and Response Mode (Zajac, 2003)

		Percent Imputation					
Form Source		Relationship	Sex	Age	Hispanic Origin	Race	Tenure
Self-Response	Overall Self	2.0	1.7	3.0	4.5	3.8	3.0
	• USPS Delivery	2.0	1.7	3.0	4.4	4.1	2.8
	• LCO Delivery-UAA	2.5	2.1	3.8	5.5	4.2	4.1
	• Update/Leave	1.7	1.6	2.7	4.6	2.6	3.8
	• Urban Update/Leave	2.3	2.2	3.3	5.4	4.5	3.4
	• Internet	1.3	1.2	1.5	3.0	3.3	1.4
	• Be Counted	18.0	7.8	9.2	12.5	11.7	3.6
Enumerator Response	Overall Enumerator	4.1	2.8	10.9	4.1	4.3	12.2
	• Coverage Edit Followup	4.3	3.4	9.7	6.8	10.9	6.9
	• Telephone Questionnaire Assistance	2.9	0.5	2.9	1.7	2.2	0.8
	• Nonresponse Followup	3.9	2.6	10.4	3.6	3.4	11.4
	• Coverage Improvement Followup	6.3	4.3	21.6	6.7	6.3	22.8
	• Update/Enumerate	3.9	3.1	6.8	3.6	3.2	10.1
	• List/Enumerate	4.5	3.8	7.3	4.4	4.3	9.9
	• Remote Alaska	12.6	12.2	15.3	12.7	12.0	3.7
	• Other: T-Night, Orphans	10.1	2.8	11.6	8.9	5.5	46.0

Data source: HCEF_D', HCUF

* Imputation rates include assignment and allocation rates.

The table above was not separated out by form type (short or long) in the Census 2000 report.

Note that in the Census 2000 report and for the 2010 Census report, this analysis is done by the type of form returned. Because form types changed between Census 2000 and the 2010 Census, the categories are not equivalent. For example, the areas where the Remote Alaska form was used in Census 2000 was replaced by the Update/Enumerate form in 2010. Also note that there was no List Enumerate operation for the 2010 Census.

Table A1: 2010 Census Item Nonresponse Rates for Household-Level Items - Puerto Rico

Form Type/Operation	Percent Item Nonresponse			
	Household Count	Undercount	Tenure	Phone Number
Overall	1.3	6.9	3.3	7.6
Self-Response	2.0	8.6	3.0	7.2
Enumerator Return	<0.1	3.6	3.8	8.3

Source: Census Unedited File

Table A2: 2010 Census Item Nonresponse Rates for Person-Level Items - Puerto Rico

Form Type/Operation	Percent Item Nonresponse					
	Relationship*	Sex	Age/Date of Birth	Hispanic Origin	Race	Overcount
Overall	0.7	2.0	2.3	0.7	2.8	2.6
Self-Response	0.7	2.6	0.6	0.6	3.7	3.4
Enumerator Return	0.8	1.0	5.1	1.0	1.3	1.2
GQ Enumeration	n/a	2.1	19.9	6.4	31.2	n/a

Source: Census Unedited File

n/a - item is not asked for this questionnaire type.

* Excludes reference person since relationship uses first person as reference.

Table A3: 2010 Census Item Nonresponse Rates for Self-Response Household-Level Items - Puerto Rico

Form Type/Operation	Percent Item Nonresponse			
	Household Count	Undercount	Tenure	Phone Number
Self-Response	2.0	8.6	3.0	7.2
Update/Leave	2.0	8.6	3.0	7.3
Be Counted	n/a	n/a	6.3	2.7
Fulfillment	2.5	11.5	3.0	3.4

Source: Census Unedited File

n/a - item is not asked for this questionnaire type.

Table A4: 2010 Census Item Nonresponse Rates for Enumerator Return Household-Level Items - Puerto Rico

Form Type/Operation	Percent Item Nonresponse			
	Household Count	Undercount	Tenure	Phone Number
Enumerator Return	<0.1	3.6	3.8	8.3
CFU	**	***	0.1	*
TQA	**	+	0.0	0.0
NRFU	<0.1	3.6	4.1	8.3
ETL	0.0	n/a	+	11.4

Source: Census Unedited File

n/a - item is not asked for this questionnaire type.

* Telephone number is required for the CFU operation.

** Household Count is automatically set in CFU and TQA based on number of person records and thus these operations are excluded from the household count item.

*** Undercount is not asked in CFU but is passed through from the original form (including self-response forms).

+ Cell statistic involves less than 10 cases but not zero.

Table A5: 2010 Census Item Nonresponse Rates for Self-Response Person-Level Items - Puerto Rico

Form Type/Operation	Percent Item Nonresponse					
	Relationship*	Sex	Age/Date of Birth	Hispanic Origin	Race	Overcount
Self-Response	0.7	2.6	0.6	0.6	3.7	3.4
Update/Leave	0.7	2.6	0.6	0.6	3.7	3.4
Be Counted	0.8	1.9	0.6	0.5	3.3	n/a
Fulfillment	+	2.1	+	1.5	3.6	3.0

Source: Census Unedited File

n/a - item is not asked for this questionnaire type.

* Excludes reference person since relationship uses first person as reference.

+ Cell statistic involves less than 10 cases, but not zero.

Table A6: 2010 Census Item Nonresponse Rates for Enumerator Return Person-Level Items - Puerto Rico

Form Type/Operation	Percent Item Nonresponse					
	Relationship*	Sex	Age/Date of Birth	Hispanic Origin	Race	Overcount
Enumerator Return	0.8	1.0	5.1	1.0	1.3	1.2
CFU	0.4	0.1	0.4	0.4	0.9	n/a
TQA	0.0	0.0	+	+	+	+
NRFU	0.8	1.1	5.6	1.1	1.4	1.2
ETL	+	0.0	9.5	+	+	0.0
GQ Enumeration	n/a	2.1	19.9	6.4	31.2	n/a

Source: Census Unedited File

n/a - item is not asked for this questionnaire type.

* Excludes reference person (PNUM=1) since relationship uses first person as reference, except for TQA which uses REF_PERSON=1.

+ Cell statistic involves less than 10 cases, but not zero.

Table A7: Overall “As Reported” and Imputation Rates (as a percentage) – Puerto Rico

Form Type/Operation	Person-Level Items					Household-Level Item
	Relationship	Sex	Age/Date of Birth	Hispanic Origin	Race	Tenure
Overall						
As Reported	98.3	97.8	95.1	98.6	96.5	96.9
Imputed:	1.7	2.3	4.9	1.4	3.5	3.1
Assigned	0.4	1.9	2.5	0.1	0.7	*
Allocated	1.3	0.4	2.4	1.3	2.8	3.1
Self-Response						
As Reported	98.5	97.3	97.5	98.7	95.7	96.5
Imputed:	1.5	2.7	2.5	1.3	4.3	3.5
Assigned	0.3	2.4	1.9	0.1	0.9	*
Allocated	1.3	0.3	0.5	1.2	3.4	3.5
Enumerator Return						
As Reported	98.6	99.0	91.6	98.9	98.5	97.7
Imputed:	1.4	1.0	8.4	1.1	1.5	2.3
Assigned	0.6	0.9	3.4	0.1	0.2	*
Allocated	0.9	0.2	5.1	1.0	1.3	2.3
GQ Enumeration						
As Reported	n/a	97.2	76.5	92.9	68.1	n/a
Imputed:	n/a	2.9	23.5	7.1	31.9	n/a
Assigned	n/a	1.9	2.6	0.7	7.4	n/a
Allocated	n/a	1.0	21.0	6.5	24.5	n/a

Source: Census Edited File

*Tenure could not be assigned in the 2010 Census in the absence of long form data.

n/a - item is not asked for this questionnaire type.

Note: Imputation rates include assignment and allocation rates.