UNITED STATES DEPARTMENT OF COMMERCE
Economics and Statistics Administration
U.S. Census Bureau
Washington, DC 20233-0001

July 13, 2012

2010 CENSUS PLANNING MEMORANDA SERIES

No. 201 (Reissue)

MEMORANDUM FOR The Distribution List

From: Burton Reist *[signed]*
Acting Chief, Decennial Management Division

Subject: 2010 Census Non-ID Processing Assessment Report

Attached is the revised 2010 Census Non-ID Processing Assessment Report. The revised document includes corrections to Table 5.4.4 (numbers and percentages) and Table 5.4.7 (numbers only) for the Location section (Stateside/Puerto Rico). The overall totals for the tables were not affected by this correction; only the distribution of the respective Non-ID cases between Stateside and Puerto Rico were affected by the error. The totals for each table remain the same.

If you have any questions about this document, please contact Michael Niosi at (301) 763-8938.

Attachment

2010 Census Non-ID Processing Assessment Report

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

FINAL

Michael Niosi

Decennial Management Division

This document contains no Title 13 data or Personally Identifiable Information (PII)

Table of Contents

List of Tables

Executive Summary

The primary purpose of this assessment is to document and analyze what happened during the 2010 Census Non-ID Processing operation. The Non-ID Processing assessment will provide information on the background of the operation, including the planning, development, and implementation stages. The assessment will also document the final results for all aspects of the operation and provide information on the impacts to the 2010 Census address files. Overall, the assessment will produce valuable data for the next planning cycle for the 2020 Census.

Background

During the 2010 Census enumeration for stateside (the 50 states and the District of Columbia) and Puerto Rico, the majority of people were enumerated via questionnaires delivered to their living quarters by mail or hand-delivered by United States Census Bureau field staff. These questionnaires contained a census identification number (Census ID) that linked the questionnaire to the address of the living quarters. However, there were other opportunities for individuals to respond to the census outside of the two typical delivery modes, including coverage improvement programs and enumeration interviews conducted by Census Bureau field staff.

All addresses generated from coverage improvement programs and new addresses submitted by field staff originate from questionnaires that lack a Census ID, hence the term "Non-ID." These questionnaires contain a Processing ID instead that is used for tracking purposes. The address information from all eligible questionnaires is commonly referred to as "Non-ID cases." Accordingly, the process of comparing these cases to address records contained in the Master Address File/Topologically Integrated Geographic Encoding and Referencing database to determine whether or not they match existing address records and/or assigning census geographic codes to these cases is known as "Non-ID Processing." The ultimate goal of Non-ID Processing is to determine if the addresses submitted by respondents and field staff should be assigned a Census ID and added to the census universe.

The 2010 Census Non-ID Processing operation received inputs from several operations that consisted of respondent-provided address information. These operations generated questionnaires without a preassigned Census ID, and contained address information without an associated block-level geocode, referred to as Type A cases:

- Be Counted
- Telephone Questionnaire Assistance Fulfillment
- Telephone Questionnaire Assistance Interview
- Group Quarters Enumeration
- Service-Based Enumeration
- Remote Alaska
- Remote Update Enumerate
- Update Enumerate
- Nonresponse Followup

On some Be Counted forms, the respondent(s) indicated they did not live or sleep at an address on April 1, 2010. The information from these forms is referred to as Type B cases.

Type C cases were created when Census Bureau enumerators provided address information for living quarters not contained on the census address list for their assignment area during the following operations: Group Quarters Enumeration, Remote Alaska, Remote Update Enumerate, Update Enumerate, Enumeration at Transitory Locations, Update/Leave, Nonresponse Followup, and Nonresponse Followup Vacant Delete Check. These operations generated "add" questionnaires that contained the address information associated with the living quarters, a 2010 collection block geocode, and a Processing ID.

Type A and Type B Non-ID cases went through a series of steps in an attempt to match and geocode the address information because all questionnaires must be associated with a census geographic code. The initial step was an attempt to assign a case to a state and county, a subprocess known as "header coding." Type A cases required additional processing steps after being header coded. Type B cases were complete once successfully header coded and were later added to the Group Quarter universe and allocated to census counts accordingly. Type A and Type B cases not successfully header coded were sent to the clerical operation at the National Processing Center for resolution.

Type A cases successfully header coded through automated processing were then submitted to an automated process that attempted to match them to address records already in the Master Address File/Topologically Integrated Geographic Encoding and Referencing database. Cases unable to be matched were sent through automated geocoding. All unmatched cases that could not be geocoded via the automated geocoding routine were sent to the National Processing Center for a clerical matching and geocoding operation. Those cases that matched to a preexisting address record that did not have either an associated geocode or could not be geocoded through the automated process were sent to the National Processing Center for a clerical matching and geocoding operation as well.

After clerical processing, all Type A and Type B cases were returned to Census Bureau headquarters. Certain Type A cases were then sent through automated post-clerical processing. During post-clerical processing, a final attempt was made to obtain a complete match/geocode for Type A cases not fully coded during clerical processing.

Type C cases containing the required block-level geocode went through automated matching before the addresses were accepted as new address records to the Master Address File/ Topologically Integrated Geographic Encoding and Referencing database in order to avoid duplicate records. Type C cases were rejected if they did not meet certain criteria and the address information and associated response data were not included in the census universe. If a match occurred, then the existing address record was updated to reflect an additional source. If no match was found, then the address was added as a new address record.

The final step for all successfully matched and/or geocoded cases was to update the Master Address File/Topologically Integrated Geographic Encoding and Referencing database with all the matches and geocodes. Cases not successfully matched and/or geocoded were deemed uncodable and the address information and associated response data were not included in the

census universe. Once the final disposition of all the Non-ID cases was reported to the Decennial Systems and Contracts Management Office's Universe Control and Management system, further processing steps were taken to create the final census universe of housing units and persons.

Results

The following table provides summary data for the research questions in this report.

Summary Data for the Non-ID Processing Research Questions

Research Questions	Summary Data
1. How many cases were sent to automated Non-ID Processing?	Total Cases: **2,887,757** • **1,265,551** Type A Cases • **13,427** Type B Cases • **1,608,779** Type C Cases
2. What was the outcome of automated Non-ID Processing (header coding, matching, and geocoding)?	Type A Cases: • **97.36 percent** were successfully header coded • Matched and Geocoded: **656,032 (53.24 percent)** • Not Matched But Geocoded: **59,113 (4.80 percent)** • Matched But Not Geocoded: **114,672 (9.31 percent)** • Not Matched and Not Geocoded: **205,524 (16.68 percent)** • Rejected: **196,780 (15.97 percent)** Type B Cases: • **88.70 percent** were successfully header coded Type C Cases: • **98.63 percent** had the geocode provided by the enumerator accepted • Added as New Address Records: **967,842 (61.00 percent)** • Matched to Existing Address Records: **588,421 (37.08 percent)** • Rejected: **30,458 (1.92 percent)**
3. How many cases were sent to clerical Non-ID Processing?	Type A Cases: **550,406** Type B Cases: **1,517**
4. What was the outcome of clerical Non-ID Processing (header coding, matching, and geocoding)? How many cases were resolved via telephone call?	Type A Cases: • **59.19 percent** were successfully header coded • Matched to a Geocoded Record: **178,358 (33.24 percent)** • Matched During Automated Processing/Clerically Geocoded: **90,800 (16.92 percent)** • Matched During Automated Processing/Unable to Clerically Geocode: **23,870 (4.45 percent)** • Clerically Unmatched/Geocoded: **124,770 (23.25 percent)** • Clerically Unmatched/Unable to Geocode: **118,765 (22.13 percent)**

Research Questions	Summary Data
	Type B Cases: • **20.37 percent** were successfully header coded Results of Telephone Calls: A telephone call was attempted for 199,603 Type A cases during clerical processing. • Corrected Address Information and Obtained Geocode: **61,842 (30.98 percent)** • Obtained Geocode with No Additional Address Information Provided: **23,328 (11.69 percent)** • Respondent Not Reached/No Additional Address Information Provided: **114,433 (57.33 percent)**
5. What was the outcome of post-clerical Non-ID Processing (matching)?	Type A Cases: • Matched to a Geocoded Record: **33,586 (12.55 percent)** • Matched to a Ungeocoded Record: **58,240 (21.75 percent)** • Unmatched to a Record: **106,978 (39.96 percent)** • Rejected: **68,905 (25.74 percent)**
6. What were the actions taken on the records in the Master Address File/Topologically Integrated Geographic Encoding and Referencing database as a result of Non-ID Processing (automated, clerical, and post-clerical)?	Type A Cases: • **78.21 percent** were matched and geocoded; included in the census universe • **2.23 percent** were matched and ungeocoded; not included in the census universe • **9.79 percent** were unmatched and geocoded: included in the census universe • **9.77 percent** were unmatched and ungeocoded: not included in the census universe Type B Cases: • **91.04 percent** were assigned to a state and county; included in the census universe • **8.96 percent** were not assigned to a state and county; not included in the census universe Type C Cases: • **60.16 percent** were added as new address records; included in the census universe • **36.58 percent** were matched to existing address records; included in the census universe • **3.26 percent** were rejected; not included in the census universe

Research Questions	Summary Data
7. How did the actual cost of the 2010 Census Non-ID Processing operation compare to the expected budget?	Geography Division: • Total Budget: $1,580,000 • Actual Cost: $1,491,328 • Percent of Budget Spent: 94.39 percent National Processing Center: • Total Budget: $2,666,541 • Actual Cost: $2,234,227 • Percent of Budget Spent: 83.79 percent Totals for 2010 Census Non-ID Processing: • Total Budget: $4,246,541 • Actual Cost: $3,725,555 • Percent of Budget Spent: 87.73 percent

Recommendations

The following recommendations are based on the planning, development, and implementation stages of the 2010 Census Non-ID Processing operation. While not all the recommendations listed are directly linked to material already covered in the Executive Summary, they are tied to issues and challenges experienced and documented throughout the operation. Additional recommendations are included in the Recommendations section of the full assessment report.

General Planning and Development Recommendations

- The decisions regarding which software to use for the clerical operation and what staff (i.e., National Processing Center clerks, contractors, or other) will be assigned the Puerto Rico cases should be incorporated into the planning and development stages of the 2020 Census Non-ID Processing operation.

- The automated and clerical processing stages should be developed in a more parallel fashion rather than sequentially, as occurred during the planning and development stages of the 2010 Census Non-ID Processing operation.

- All stakeholders, especially staff outside of headquarters (e.g., National Processing Center, telephone centers), should be involved in the planning and development stages of the 2020 Census Non-ID Processing operation from the beginning.

Clerical Processing

- A consistent message should be delivered during all the trainings, including the goals of the training, as well as the procedures themselves. For example, a verbatim training is one possible way to maintain consistency between training sessions.

- As occurred for the 2010 Census, the National Processing Center should keep a core dedicated staff on the clerical operation from beginning to end to ensure things keep operating smoothly. The core staff would include a manager for the Geographic Technicians and Unit Supervisors that will work on Non-ID Processing exclusively through the height of the clerical operation and then part-time as needed for the remainder of the operation.

- The clerical software should be able to separate the Non-ID cases and not always keep them bundled into work units. This would allow Non-ID cases to be transferred between clerks, as well as allow cases that were resolved in error or were left unfinished by a clerk to be inserted back into production.

- The clerical software should allow the user to have multiple work units open per account. This way, if there is a problem with one work unit, the user could access another work unit without waiting for the problem with the first work unit to be fixed or having to create a second user account in order to access another work unit.

- The clerical software should have the ability to capture metrics for each clerk that can be used as part of the quality control process.

Staff Communications

- Technologies such as videoconferencing, which was used effectively for planning and implementing the 2010 Census Non-ID Processing operation, could be used again for face-to-face team and stakeholder meetings.

- On-site visits from headquarters staff to the National Processing Center prior to production are very helpful for observing the environment where the clerical operation will take place and for headquarters staff to get a better understanding of how the National Processing Center manages their staff.

This page intentionally left blank.

1. Introduction

1.1 Scope

The primary purpose of this assessment is to document and analyze what happened during the 2010 Census Non-ID Processing operation. The Non-ID Processing assessment will provide information on the background of the operation, including the planning, development, and implementation stages. The assessment will also document the final results for all aspects of the operation and provide information on the impacts to the 2010 Census address files. Overall, the assessment will produce valuable data for the next planning cycle for the 2020 Census.

1.2 Intended Audience

This document is intended for the following users:

- Address List Development Operations Integration Team (ALD OIT)
- Decennial Leadership Group
- Additional internal stakeholders, such as program managers and subject matter experts involved in the planning and implementation of the Non-ID Processing operation and the 2020 Census
- External stakeholders

2. Background

During the 2010 Census enumeration for stateside (the 50 states and the District of Columbia) and Puerto Rico, the majority of people were enumerated via questionnaires delivered to their living quarters (LQ) by mail or hand-delivered by U.S. Census Bureau field staff. These questionnaires contained a census identification number (Census ID) that linked the questionnaire to the address of the LQ. However, there were other opportunities for individuals to respond to the census outside of the two typical delivery modes, including coverage improvement programs and enumeration interviews conducted by Census Bureau field staff.

All addresses generated from coverage improvement programs and new addresses submitted by field staff originate from questionnaires that lack a Census ID, hence the term "Non-ID." These questionnaires contain a Processing ID instead that is used for tracking purposes. The address information from all eligible questionnaires is commonly referred to as "Non-ID cases." Accordingly, the process of comparing these cases to address records contained in the Master Address File (MAF)/Topologically Integrated Geographic Encoding and Referencing (TIGER) database (MTdb) to determine whether or not they match existing address records and/or assigning census geographic codes to these cases is known as "Non-ID Processing." The ultimate goal of Non-ID Processing is to determine if the addresses submitted by respondents and field staff should be assigned a Census ID and added to the census universe.

Non-ID Processing serves three main functions:

- Identifies whether an address exists in the census universe.

- Provides part of the workload for the Field Verification (FV) operation.
- Provides basic geographic information (where possible) for Be Counted (BC) responses on which the respondent has stated they did not have an address on Census Day.

2.1 Census 2000 Non-ID Processing

For Census 2000, to ensure that every known LQ either received a form or an enumeration interview, additional data collection strategies were developed to enumerate special population groups and people that believed they had not received a census questionnaire or were not included on the questionnaire for their residence. The following programs were conducted for Census 2000 to include population groups that had been undercounted in past censuses:

- BC
- Telephone Questionnaire Assistance (TQA) programs (TQA Fulfillment and TQA Interview)
- Service-Based Enumeration (SBE)

Every LQ address originating from the MAF in Census 2000 had a unique identifier, the MAF identification number (MAF ID). Accordingly, each response to the census was data-captured and linked to a LQ using the MAF ID. However, when addresses were received from such operations as BC, TQA Fulfillment, and TQA Interview, they lacked a MAF ID. This required a comparison to the address records in the MAF to determine if an address record already existed for that address. If an address matched an address record in the MAF, the address was assigned that MAF ID and the response data were associated with that MAF ID. If the address did not match an address record in the MAF, an attempt was made to obtain a census block geocode for the address from the TIGER database. If a census block geocode was obtained, it was provisionally added to the MAF and assigned a MAF ID. These addresses then were verified to exist before being permanently added to the MAF.

There were three categories established for the responses that did not have a MAF ID:

- Type A - case in which a respondent provided an address from BC and TQA (Fulfillment and Interview), plus "usual home elsewhere" (UHE) responses[1] from SBE, Special Place Group Quarters Enumeration (GQE), Military/Maritime Crews-of-Vessels Enumeration, Military Unit Enumeration, and Nonresponse Followup (NRFU).

- Type B - case in which the respondent indicated they had no address/UHE on April 1, 2000. Type B addresses originated exclusively from BC questionnaires.

[1] UHE responses are generated when a LQ is occupied by one or more people who have a usual residence at another location. This unit is classified as vacant and the residents are counted at their usual residence, for which the address is sent to Non-ID Processing to be verified. The address information from In-Mover responses, which occur when a person moves into a LQ after Census Day, is also sent to Non-ID Processing. For the purposes of this assessment, both UHE and In-Mover responses will be collectively referred to as 'UHE responses.'

- Type C - case in which an enumerator added an address during the Update/Leave (U/L), Urban U/L, NRFU, Coverage Improvement Followup, GQE, Update Enumerate (UE), List/Enumerate, and Remote Alaska (RA) operations.

The Geography Division (GEO) worked with a contractor to develop a system for the clerical geocoding of addresses that did not match the MAF and/or geocode to the TIGER database. The system was called the Interactive Mapping and Geocoding System (IMAGS). IMAGS provided a geography assignment component, a mapping component, a call center application, and a status tracking and reporting component. IMAGS was used by the staff at the National Processing Center (NPC) as part of a clerical operation.

GEO also contracted with a private firm to perform the matching and geocoding of addresses in Puerto Rico. Any fallout, after the contractor cleaned up the address information and the cases were processed through the automated system, was sent to NPC for resolution.

There were 4,189,815[2] Type A and Type B Non-ID cases processed from April 6, 2000 through August 17, 2000. Of these, 467,596 cases were completed by the 2000 FV cutoff of June 14, 2000 and eligible to be sent to the 2000 FV operation; over 65 percent of those were found to be correctly geocoded[3].

GEO provided the results of Non-ID Processing to the Decennial Systems and Contracts Management Office (DSCMO), who stored the information in the Decennial Master Address File (DMAF). The DMAF was the Census 2000-specific housing unit (HU) inventory, containing only those HU addresses initially identified as being eligible for the census, as well as those added to the inventory during the course of the census.

2.2 2008 Census Dress Rehearsal Non-ID Processing

There was no clerical stage for the 2008 Census Dress Rehearsal (DR) Non-ID Processing operation due to a lack of clerical software; only the automated matching/geocoding phase occurred. The automated operation was largely the same as it was for Census 2000. However, the systems used to implement the operation changed. The MAF and TIGER became one consolidated system (i.e., the MTdb) and the equivalent to the DMAF for this census cycle was the Universe Control and Management system (UCM). DSCMO UCM staff delivered Non-ID cases to GEO weekly during the 2008 Census DR and GEO accumulated them in a holding area. Ultimately, GEO processed these cases in two separate groups: an initial batch and a final batch.

After some operations were eliminated for the 2008 Census DR, TQA Fulfillment (Type A cases) remained the sole source of input. A total of 438 Type A cases were received for Non-ID Processing from May 22, 2008 through August 8, 2008. The results for all the cases were reported to the DSCMO UCM system via the Non-ID Feedback Table (NIFT). The results of the 2008 Census DR Non-ID Processing were also delivered to the Decennial Management Division

[2] There were approximately 2.3 million Non-ID cases erroneously included in the 2000 Non-ID Processing operation from the GQE operation. Further information is provided in Section 4 – Limitations.

[3] The percentage of Non-ID cases sent to the 2000 FV operation that were correctly geocoded was calculated by adding the number of addresses found in the block to which it was geocoded to the number of addresses found to be a duplicate of an address in the block to which it was geocoded, and dividing by the total number of Non-ID cases sent to 2000 FV.

(DMD) for a one-time Cost and Progress (C&P) report and to the Decennial Statistical Studies Division (DSSD) for assessment purposes. There was no delivery made to 2008 FV because there was no 2008 FV operation conducted for the 2008 Census DR. Reasons for the reduced scope and canceled operations are detailed in 2008 Census Dress Rehearsal Memorandum No. 50, *Reduced Scope of the 2008 Census Dress Rehearsal and a One-Month Delay of Census Day.*

2.3 2010 Census Non-ID Processing

Types of Non-ID Cases

The 2010 Census Non-ID Processing operation received inputs from several operations that consisted of respondent-provided address information. These operations generated questionnaires without a preassigned Census ID, and contained address information without an associated block-level geocode:

- BC
- TQA Fulfillment
- TQA Interview
- GQE
- SBE
- RA
- Remote UE (RUE)
- UE
- NRFU

In addition, Census Bureau enumerators provided address information for LQs not contained on the census address list for their assignment area during GQE, RA, RUE, UE, Enumeration at Transitory Locations (ETL), U/L, NRFU, and NRFU Vacant Delete Check (VDC). These operations generated "add" questionnaires that were expected to contain the address information associated with the LQs, a 2010 collection block geocode, and a Processing ID.

For the 2010 Census, Non-ID cases were classified as follows:

1. **Type A** cases were ungeocoded and had a Processing ID. Type A cases consisted of the following:

    - Respondent-provided addresses from BC, TQA Fulfillment, and TQA Interview
    - UHE addresses from GQE[4], RA, UE, RUE, and NRFU[5]. UHE addresses from GQE/SBE Individual Census Reports (ICR) and Shipboard Census Reports (SCR) were sent to Non-ID Processing only for eligible group quarters (GQs). Eligibility for Non-ID Processing was determined by GQ type code. Addresses

[4] GQ UHE addresses were also collected during SBE, but they were assigned to GQE when delivered for Non-ID Processing.
[5] There were a number of UHE addresses that originated from VDC. The Non-ID Processing subteam was not expecting UHEs from VDC, so there was no requirement for DSCMO to deliver them to GEO for processing. However, when these addresses came into HQ processing, DSCMO assigned the NRFU source code to them and sent them on to GEO.

from UHE responses with an invalid or unknown GQ type code were not sent to Non-ID Processing and were enumerated at the GQ instead.

For all Type A cases, DSCMO considered the presence of a person record in the persons table for the ID and a value of 1 in the Priority Name Capture column, in addition to the other requirements in the specification for Address Update (ADDUP) file composition. No cases should have been delivered to GEO without the presence of a person record.

Type A cases went through the full automated and clerical processes described later in this section.

2. **Type B** cases were ungeocoded and had a Processing ID. These cases came from BC forms in which the respondent(s) indicated they had no address on April 1, 2010. Type B cases were derived only from BC forms.

   Type B cases were only assigned to higher-level geographic units (i.e., state and county) based on any location information provided by the respondent on the form. These cases were later added to the GQ universe, where they were randomly allocated to a shelter in the assigned county once it was determined that the respondent was not already being counted at a GQ in that county.

3. **Type C** cases were geocoded by field staff to the 2010 Census collection block level and had a Processing ID, although a small percentage were ungeocoded (i.e., had a missing or invalid geocode) and had a Processing ID. Type C cases consisted of enumerator-generated address adds from ETL, GQE[6], U/L, RA, UE, RUE, NRFU[7], and VDC.

   Type C cases went through an automated process that attempted to match the addresses to the MTdb first. If no match was found, then the addresses were added as new MTdb records.

Exception Check-In Records

When a form was encountered during data capture that had a missing or unreadable barcode and associated eye-readable ID, an exception check-in record was generated. The Decennial Response Integration System (DRIS) depended on each form having a unique identifier, which is the Census ID or Processing ID. If an exception check-in record was encountered, an attempt was made to find the preprinted address in the universe of census-eligible addresses and generate the associated Census ID for that record. If the address did not get a direct one-to-one match to an existing address in the universe, DRIS generated a new Processing ID based on the form type.

Any paper form processed by DRIS had the potential to be an exception check-in record. This meant exception check-in records could have originated from a mail return or an enumerator

[6] GQ adds also originated from SBE, but they were assigned to GQE when delivered for Non-ID Processing.
[7] There were a number of Type C cases that originated from NRFU Reinterview (RI). The Non-ID Processing subteam was not expecting adds from NRFU RI, so there was no requirement for DSCMO to deliver them to GEO for processing. However, when these addresses came into HQ processing, DSCMO assigned them the NRFU source code and sent them on to GEO.

return from any of the previously mentioned operations, except GQE. However, exception check-in records from D-1 Mailout/Mailback forms and U/L addressed forms (form types 111 and 112) were not sent to GEO for Non-ID Processing because addresses from the regular versions of these forms were not involved in Non-ID Processing and were covered by NRFU if the respondents did not return a form or if their form was not able to be processed. For all other exception check-in records, the Processing ID and form type was used to classify the exception check-in records sent to GEO as Type A, Type B, or Type C, and to assign the appropriate source code.

Processing Flow

Type A and Type B Cases

The initial step in the processing of both Type A and Type B cases was an attempt to assign a case to a state and county, a subprocess known as "header coding." Type A cases required additional processing steps after being header coded. Type B cases were complete once successfully header coded and later allocated to census counts accordingly. Type A and Type B cases not successfully header coded were sent to the clerical operation at NPC for resolution.

Type A cases successfully header coded were submitted to an automated process in an attempt to match them to address records already in the MTdb. Cases that did not match a MTdb record were sent through automated geocoding. All unmatched cases that could not be geocoded via the automated geocoding routine were sent to NPC for a clerical matching and geocoding operation. Those cases that matched to a MTdb record that did not have either an associated geocode or could not be geocoded through the automated process were sent to NPC for clerical processing as well.

NPC clerical staff used an interactive software application to look for errors in the address information and attempted to match and/or geocode them. The initial attempt to match and geocode cases was conducted by staff referred to as Front Line clerks. Cases not resolved by the Front Line clerks were referred to other staff members known as Referral clerks. If the assigned Referral clerk was unable to resolve a case with the information available, they attempted to call the respondents[8] to obtain better address or location information. Using the respondent-provided information, the clerk made another attempt to match the address; otherwise, the clerk tried to geocode the address by referring to an electronic map and asking the respondent to provide directions from a known intersection or landmark.

After clerical processing, the cases were returned to Census Bureau headquarters (HQ). Type A cases that still required a match, geocode, or both went through automated post-clerical processing. During post-clerical processing a final attempt was made to match and/or geocode these cases.

The final step for the Type A and Type B cases was to update the MTdb with all the matches and geocodes for all successfully matched and geocoded cases. Cases that were unmatched/

[8] The name and, when available, telephone number for "Person 1" on the census forms was supplied to the clerical staff along with the address information for the Type A Non-ID cases.

ungeocoded or matched/ungeocoded were deemed uncodable, and the address information and associated response data were not included in the census universe.

Certain Type A cases were delivered to 2010 FV for validation before being included in the census universe. Type A cases completed before the 2010 FV cutoff date were eligible for 2010 FV if they fell into one of the following categories:

- Cases matched to a MTdb record without a preexisting geocode and geocoded via the automated or clerical process
- Cases not matched but geocoded via the automated or clerical process.

Once the 2010 FV universe was set, only Type A cases matched to a MTdb record with a preexisting geocode resulted in the addresses being included in the census universe.

GEO reported to the DSCMO UCM system via a NIFT the disposition of every Type A and Type B case delivered for Non-ID Processing. The results of Non-ID Processing were also delivered to DMD for C&P reports and to DSSD for assessment purposes.

Type C Cases

Type C cases containing the required block-level geocode went through automated matching before the addresses were accepted as new adds to the MTdb in order to avoid duplicate records. There was no header coding or automated geocoding phase for the Type C cases because the cases were supposed to be delivered with a geocode[9]. Type C cases were rejected if they did not meet certain criteria and the address information and associated response data were not included in the census universe. If a match occurred, then the MTdb record was updated to reflect an additional source. If no match was found, then the address was added as a new MTdb record.

GEO reported to the DSCMO UCM system via a NIFT the disposition of every Type C case delivered for Non-ID Processing. The results of Non-ID Processing were also delivered to DSSD for assessment purposes.

2.3.1 Planning and Development

The planning and development cycle for the automated 2010 Census Non-ID Processing operation began with the preparations for the 2008 Census DR Non-ID Processing operation. Representatives from DMD, GEO, DSSD, DSCMO, and NPC formed a subteam in August 2006 that reported to ALD OIT. The subteam developed many documents for the 2008 Census DR Non-ID Processing operation that were subsequently used to develop the plans for the 2010 operation, including, but not limited to, the schedule, operational plan, risk register, C&P reports, deliverables list, workflow diagrams, and customer requirements documents (CRDs). The

[9] Almost all ungeocoded Type C cases were automatically rejected and progressed no further in the operation. However, due to the large number (roughly 105,000) of NRFU and VDC Type C cases rejected for a missing or incomplete block geocode and the fact that this was the last chance for these addresses to get into the census universe, an exception was made. These cases were sent through the automated geocoding process, and if unsuccessful, to the clerical operation at NPC in an attempt to obtain a geocode. More detail regarding these cases is provided in the text following Table 5.2.14.

development of some of these documents as they relate to 2010 Census Non-ID Processing will be discussed further in the sections below, along with the planning and development of the budget, automated and clerical processes, and staffing.

The planning and development for the processing of Type C cases was not in scope for the Non-ID Processing subteam; all specifications regarding the processing of the Type C cases were provided by the respective enumeration operation subteams. Therefore, all information in Section 2.3.1 relates to Type A and Type B cases only, unless stated otherwise.

Schedule Development

The Non-ID Processing subteam started with the activities and durations from the 2008 Census DR schedule and met over a period of several months to further develop and refine the 2010 Census Non-ID Processing schedule. This work included adding the activities that did not exist in the 2008 Census DR, such as the clerical operation. The schedule development involved a coordinated group effort with members from DMD, GEO, DSSD, and DSCMO responsible for specific tasks covering the scope of the operation.

Once the activities were agreed upon, the Non-ID Processing subteam reviewed the activity linkages (predecessors/successors), logic/relationships (e.g., Start to Start, Finish to Finish), and activity durations. The subteam returned the modified schedule to DMD Management Information Systems (MIS) staff for analysis. This was an iterative process that continued until the schedule was ready to be baselined.

The baselining process began when the detailed integrated schedule was considered final by ALD OIT and the MAF/TIGER Integrated System Team and then sent to the Census Integration Group (CIG). The CIG accepted the 2010 Census Non-ID Processing schedule, along with the schedule for all decennial operations, once they reached a consensus that the schedule was final and there were only minimal issues to be handled through later schedule analysis and the change control management process. The 2010 Census baseline schedule was established on May 22, 2008. A global schedule change request (CR) containing changes for all operations was submitted in December 2008 before the formal change control management process was put into effect.

The activities related to the clerical process had to be added to the 2010 Census schedule because there was not a clerical component of the 2008 Census DR Non-ID Processing operation. The activities added to the schedule included those related to the software development and testing of the clerical application to be used at NPC.

The original version of the schedule had sections titled “Automated Non-ID Processing” and “Clerical Non-ID Processing,” with the activities within those sections using the same terminology. It was difficult to track the development of the clerical application with activities that simply used the term “clerical processing.” To distinguish between the development of the clerical application and the implementation of the clerical process, the decision was made to combine the activities related to automated and clerical processing in one section and move the activities related to the clerical application to a separate section. Therefore, a schedule CR was

submitted and enacted in December 2009, resulting in an "Automated and Clerical Non-ID Processing" section and a "Non-ID Processing Clerical Application" section; the activities in those sections were renamed accordingly.

Other modifications to the schedule that occurred after the baselining of the 2010 Census schedule included changes to the C&P schedule. GEO determined they would not be able to deliver files for C&P until April 26, 2010, instead of the April 19, 2010 date originally put into the schedule. Thus the availability of the Non-ID Processing C&P system and the delivery of the MTdb output files for C&P had their start dates changed to April 26, 2010. Moving the date back allowed GEO to deliver a full set of data for the C&P reports including all the information needed regarding the clerical operation. GEO agreed to send DMD via email the number of records received and header coded for the first three weeks of production because the automated and clerical operations would be starting prior to April 26, 2010.

There were parts of the Non-ID Processing operation that were not reflected as activities in the official 2010 Census schedule. There were no schedule activities related to the clerical training and the development of the clerical procedures. For the CRDs, only the Prepare/Deliver and Receive lines were entered into the schedule; there were no separate lines for tracking the review and comment periods. GEO developed and maintained the schedule for clerical training, including the creation of the training materials. This schedule was modified close to the start of the clerical operation due to changes in the training and materials necessitated by requirements from NPC. Generally, NPC requires draft copies of training materials be delivered at least three weeks in advance of the first training date to allow for revisions, and final versions be delivered to allow adequate time for printing. The tracking of the review and comment periods for the CRDs was handled by a deliverable list, which was a Non-ID Processing subteam product used to track the dates for all deliverables.

Although the 2010 FV cutoff was a key activity for the 2010 Census Non-ID Processing operation, it was not specifically mentioned in the 2010 Census schedule. The activities related to the automated and clerical operations spanned the whole length of the Non-ID Processing operation without any indication of a 2010 FV cutoff milestone. There was no 2010 FV cutoff activity in the 2010 Census Non-ID Processing schedule because there was the possibility that Census Bureau management would decide to extend the cutoff in order to include more addresses for 2010 FV. The Non-ID Processing subteam felt that we should not include a date that could potentially move multiple times in the official schedule. Indeed, this turned out to be the case. The original 2010 FV cutoff date for clerical processing was May 18, 2010, while the original 2010 FV cutoff date for automated processing was May 25, 2010. At the time these dates were selected, it was believed that GEO required a week to process the clerical updates, leading to a week between the two 2010 FV cutoff dates. The 2010 FV cutoff date for both the automated and clerical operations was ultimately May 28, 2010. The events leading to the final 2010 FV cutoff date are further discussed later in this section.

Budget Development

Initial Budget Formulation

The funding for the 2010 Census Non-ID Processing operation began with fiscal year (FY) 2009 and continued through FY10. In FY09, NPC was not initially funded for the project, but GEO's initial allocation was 1.5 million dollars in non-salary object funds in anticipation of hiring a contractor to design the clerical software equivalent to the IMAGS system used in Census 2000 Non-ID Processing. Salaries for GEO staff working on planning and development activities for the project were allocated in a different project code and were not tracked separately from other hours charged by other projects to that project code. In FY10, both GEO and NPC had initial allocations.

The original FY10 allocation for GEO ($490,000) was based on the $240,000 budgeted for the Census Bureau Telecommunications Office to provide services for the clerical operation and $250,000 set aside for the Non-ID Processing share of the cost for FastData, an application that allows a user to research public information databases for a fee. The telephone portion of the clerical operation required the use of FastData in instances where a clerk needed to obtain a telephone number because the respondent did not provide one. The estimates for these two services were based on actual cost data from the Census 2000 Non-ID Processing operation. GEO's initial budget plan for the allocation anticipated that spending would not start until March 2010. The working assumption was that funds would not be needed until the services were utilized, which would be when training began for the clerical operation.

For NPC, the FY10 budget was originally $4,966,541. This total was based on the cost of the clerical operation for Census 2000, plus extra funds believed to be needed for a contractor to process the Puerto Rico addresses at NPC. Several years prior to the 2010 Census, automated Non-ID Processing workloads were estimated by adding 20 percent to the Census 2000 overall workload to match the projected national HU increase from 2000 to 2010. The potential workload estimate for the clerical operation used a 72 percent resolution rate for the automated process. This resolution rate was used because it was the rate for the resolution of BC cases for Census 2000 and BC was anticipated to be the largest source of inputs into Non-ID Processing due to publicity events during the 2010 Census. This resulted in estimated clerical workloads of 713,898 cases for stateside and 48,847 cases for Puerto Rico.

However, as the start of the 2010 Census Non-ID Processing operation drew closer, there were concerns from senior decennial census managers that the workload may be larger than anticipated. This led to the belief that insufficient money was allocated to the project to cover all the costs of the clerical process. In response to these concerns, DMD staff generated contingency estimates assuming a 75 percent increase in cases for stateside and 50 percent increase for Puerto Rico; the latter figure was based on the logic that U/L would be leaving forms directly on the door of every housing unit in Puerto Rico, minimizing the need for BC and TQA because people would be less likely to think they had been missed. Additionally, a 50 percent referral rate (rate at which Front Line clerks would pass cases to Referral clerks) was assumed for the stateside contingencies. There was no need to add a referral rate to the Puerto Rico contingencies because all Puerto Rico cases were to be treated as referral cases (i.e., there

would be no Front Line clerks for Puerto Rico Non-ID Processing). Finally, the requirement for hosting the Puerto Rico Non-ID contractor staff in NPC was dropped from the scope of NPC's budget.

NPC supplied cost estimates to DMD based on both workload scenarios, and DMD management decided to base the NPC budget on the contingency estimate of a 75 percent increase in stateside cases. As a result, the NPC cost estimate was $4,732,095 for the clerical work that would be required to meet that contingency, an amount close to the existing FY10 NPC allotment. Table 2.3.1 and Table 2.3.2 display the figures for the modified stateside and Puerto Rico workload estimates used for budget development. The tables include the estimated clerical production rates.

Table 2.3.1 Stateside Clerical Workload Estimates Based on 75 Percent Case Increase

Initial 2010 FV Cutoff for Clerical Processing (May 18, 2010)							
Clerical Staff	**Individual Production Rate (cases/hr)**	**Hours/ day**	**Days Worked**	**Cumulative Cases Worked**	**Clerks/ day**	**Total Cases/ hour**	**Total Cases/ day**
Front Line	10	6.5	24	547,176	350.8	1754	11,400
Referral	5	6.5	24	227,990	292.3	1461	9,500
Remainder of Clerical Operation							
Clerical Staff	**Individual Production Rate (cases/hr)**	**Hours/ day**	**Days Worked**	**Cumulative Cases Worked**	**Clerks/ day**	**Total Cases/ hour**	**Total Cases/ day**
Front Line	10	6.5	61	199,255	50.3	251	1,633
Referral	5	6.5	61	45,710	23.1	115	749

Source: DMD.

Table 2.3.2 Puerto Rico Clerical Workload Estimates Based on 75 Percent Case Increase

Initial 2010 FV Cutoff for Clerical Processing (May 18, 2010)							
	Individual Production Rate (cases/hr)	**Hours/ day**	**Days Worked**	**Cumulative Cases Worked**	**Clerks/ day**	**Total Cases/ hour**	**Total Cases/ day**
Staff	5	8	28	24,133	21.5	108	862
Remainder of Clerical Operation							
	Individual Production Rate (cases/hr)	**Hours/ day**	**Days Worked**	**Cumulative Cases Worked**	**Clerks/ day**	**Total Cases/ hour**	**Total Cases/ day**
Staff	5	8	61	33,062	13.6	68	542

Source: DMD.

Additional Funding

There were several occasions on which additional funding was sought for the 2010 Census Non-ID Processing operation.

In September 2008, it was determined that new hardware and software would be required to host the Non-ID Processing software in NPC (see *Clerical Software* subsection for the details regarding the decisions involving the clerical software). DMD staff worked with the Census Bureau Information Systems Support and Review Office staff to assemble an Information Technology Purchase Plan and cost estimate for the equipment. Based on the initial estimate, an unfunded requirement (UFR) for $600,000 for NPC was submitted and approved for FY08. In April 2009, a UFR for $90,000 was submitted to cover the cost of a test of the clerical software in NPC. No additional funding was sought for NPC for FY10.

GEO received no additional funding in FY08 or FY09, but did so on three different occasions in FY10. In February 2010, once the decision was made to utilize a private firm to conduct clerical processing on the addresses from Puerto Rico, 1.6 million dollars was added to GEO's FY10 allocation to fund the cost of the contract. Additionally, in March 2010 it was determined that GEO staff would have a significant presence in NPC for the duration of the Non-ID clerical operation. As a result, $60,000 was added to GEO's FY10 allocation to fund travel of GEO staff between HQ and NPC throughout the summer. Finally, once clerical production was underway and the amount of overtime required for GEO staff at NPC was realized, $200,000 was added to the GEO FY10 allotment to fund this additional cost. The FY10 GEO Non-ID Processing budget eventually reached $2,350,000.

Risk Mitigation

The Non-ID Processing subteam developed a risk register to identify and manage risks to the Non-ID Processing operation. The subteam assigned each risk a likelihood rating and an impact rating, each on a scale from one to three, with one being the lowest rating. All of the risks had a

low to medium overall rating, which was calculated by multiplying the likelihood rating by the impact ratings. The subteam also devised a mitigation plan and a contingency plan for each risk. The risk register was reviewed and updated monthly. The risk register was posted to the 2010 Census Operations Center at the end of each month.

There were nine risks identified by the Non-ID Processing subteam. The highest rated risk involved the proper data capture of the address information from the various input operations. The mitigation plan for this risk involved the subteam meeting with the respective planning teams for each operation (e.g., BC and TQA teams) to discuss the issue of address data capture, which was done throughout 2008. The contingency plan involved rejecting the addresses that did not conform to the proper format during automated processing and sending them to the clerical operation for resolution.

For more information on all nine risks identified by the Non-ID Processing subteam, see Appendix A.

Development of Automated Processing

Address Update File Composition

The development of the automated process began with updating all the documents from the 2008 Census DR Non-ID Processing operation to reflect the greater scope of the 2010 Census operation. One of the documents requiring updating was the specification to DSCMO for the ADDUP file composition. Due to the reduced scope of the 2008 Census DR, the only input to Non-ID Processing was the address information from TQA, whereas the 2010 Census Non-ID Processing operation had to account for inputs from eight operations. The 2010 Census version of the specification described, by operation and type of Non-ID case, the specific criteria address records had to meet for inclusion in the delivery from DSCMO to GEO for Non-ID Processing. The criteria included whether the address was geocoded or ungeocoded, presence of a Processing ID, form type(s), form number(s), and any additional criteria unique to that form(s). The list of eligible and ineligible GQs that could generate a UHE response eligible for Non-ID Processing was provided by the Data Processing Techniques and Requirements Branch within DSSD.

The specification also noted that DSCMO would not deliver adds from U/L because GEO would get those records from the results of the keying of the address registers at NPC. Likewise, the GQE/SBE adds would not be delivered in the same manner as the other Type C cases. The Local Census Offices (LCOs) would key the adds from GQE/SBE into the GQE Operation Control System (OCS). The paper-based OCS would create GQ add records and generate a Universe Enumeration Control Table and transmit it via Product Services to Headquarters Processing (HQP). HQP would convert the GQ add records into ADDUP table records and transmit them to GEO for processing.

As a supplement to the ADDUP file specification, GEO developed crosswalk creation spreadsheets that provided instructions to DSCMO for mapping from the Universal Response database Schema (URdbS) to the ADDUP files. There were three ADDUP file crosswalks: a combined one for Type A and Type B cases, one for Type C cases, and one for GQ records.

The specification also provided the details to DSCMO on the start and finish dates for the ADDUP file deliveries, as well as on the frequency of deliveries. The ADDUP file deliveries were scheduled to begin on March 30, 2010 and end on August 2, 2010, with the deliveries being on a weekly basis. The schedule for the delivery of the U/L and GQE records was not covered in the specification due to those records not being delivered via the usual ADDUP delivery system.

Automated Non-ID Processing Requirements

For the 2008 Census DR, DSSD wrote one CRD containing the requirements for automated processing, C&P data, the NIFT, and assessment files. For the 2010 Census, separate CRDs containing the detailed requirements were written for each.

DSSD wrote the CRD for the processing of Type A and Type B cases. This document established the rules for preprocessing, header coding, automated matching, automated geocoding, and updating the MTdb, including when to add a new record and when to update an existing record with a Non-ID source. Table 2.3.3 contains a sample of the requirements.

Table 2.3.3 Sample Requirements for Automated Processing

Processing Stage	Sample Requirements
Preprocessing	Establish whether a case was Type A or Type B. Standardize the address information.
Header Coding	Obtain a corrected and/or updated ZIP code for all addresses. Reject a case if it is header coded to a foreign country.
Automated Matching	Attempt to only match to HU, GQ, and 2010 Transient Unit (TU) MTdb records. If a Type A case successfully matches a MTdb record with a geocode, the software shall: 1. Assign the MAFID of the matching MTdb record. 2. Assign the geocode of the matching MTdb record. 3. Update the MAF source of the MTdb record. 4. Assign a Non-ID action code to the MTdb record.
Automated Geocoding	If the Type A case matches a MTdb record without a geocode and is able to be geocoded to a block, the software shall: 1. Assign the MAFID of the matching MTdb record. 2. Assign the geocode obtained through automated geocoding. 3. Update the MAF source of the MTdb record. 4. Update the MTdb record with the state, county, and block information. 5. Assign a Non-ID action code to the MTdb record. 6. Exclude the case from further processing until it is outputted to the NIFT.
Updating MTdb	The software shall update the MTdb record with a Non-ID source for a Type A case that matches to a HU, GQ, or TU record if the matched record is "in census."

Source: *Customer Requirements for 2010 Census Non-ID Processing*.

The document also described the characteristics of the Non-ID cases that were eligible for clerical processing and post-clerical processing (see Table 2.3.4). Requirements to track the results of automated processing in a NIFT and an assessment file (Clerical Review and Data Assessment (CRDA) file) were also included in this specification. The CRDA is a transaction file that recorded the results for all the Non-ID cases, plus had additional fields appended to it for assessment purposes.

Table 2.3.4 Criteria for Clerical and Post-Clerical Processing

Processing Stage	Criteria for Eligibility
Clerical Processing	Type A cases not matched to a geocoded MTdb record and not geocoded in the automated geocoding process. Type A cases that do not have complete location or mailing address information that can be used for matching and/or geocoding are excluded from the automated process. Type A and Type B cases not header coded in the automated header coding process.
Post-Clerical Processing	Type A cases that have not matched to the MTdb after having gone through clerical processing. Type A cases that have not matched to the MTdb but were geocoded by the automated geocoding process.

Source: *Customer Requirements for 2010 Census Non-ID Processing*.

Additional matching rules were developed due to concern about addresses deleted during Address Canvassing appearing during Non-ID Processing. If a Type A case matched a record deleted during Address Canvassing, the original geocode from the MTdb record was not assigned to the Non-ID Processing address. The situation was treated like a "matched – no geocode" case. If a new geocode was not obtained during automated processing, the case was sent to clerical processing. If a geocode (whether a new collection block or the original collection block) was not found during clerical processing, then the original geocode was assigned during post-clerical processing and the address was sent to 2010 FV.

The requirements for the automated processing of the Type C cases were written in operation-specific CRDs. In general, the CRDs covered the rules for how to update the MTdb when an address matches a MTdb record and when it does not match, including the circumstances for when to update the existing MTdb record, when to create a new MTdb record, and when to reject the address.

Testing and Quality Control

The automated process developed for the 2010 Census required considerable testing to ensure it would work as needed because the automated process for the 2008 Census DR had such a limited scope. A test plan was developed that included an overview of the testing GEO designed for the systems involved in automated processing. The major activities included:

- unit tests
- developer integration tests
- system tests
- data quality edits
- user acceptance tests (UAT)
- product quality control (QC)

Developers performed unit and developer integration testing in the development environment. System testing was conducted by GEO's Independent Test and Verification Team in a separate test environment. Data quality edits, UATs, and product QC were conducted by GEO in both the testing and production environments.

Automated Processing Workflow

Appendix B contains the detailed workflow for the automated processing of Type A and Type B cases, as well as a detailed workflow for the Type C cases.

Development of Clerical Processing

Clerical Software

The initial step in the development of the clerical process was to decide on the software that would be used by the clerical staff. In early 2008 the Non-ID Processing subteam, along with GEO and DMD management, had to decide whether to work with a contractor to develop software, as was done for Census 2000, or to modify existing internal software. The in-house software options included the American Community Survey (ACS) clerical geocoding software designed and developed by the Geocoding Software Branch (GSB) within GEO. A decision was made in April 2008 to go with the ACS clerical software because it was a proven system, and even with modifications and added functionality it would be less costly and risk-prone than working with a contractor to develop a whole new system. It was also believed that any improvements to the ACS clerical software could potentially be utilized in the future by other Census Bureau operations with similar needs.

The Non-ID Processing subteam decided to rename the adaptation of the ACS clerical software to reflect the modifications and additional functionality of the application. The name that was chosen was the Matching and Geocoding Interactive Clerical (MAGIC) software. Some of the main requirements for MAGIC included:

- Compatibility with data from HQ systems
- The data included a MTdb extract that was used to build the address and address range reference files
- The data included a universe file containing the addresses requiring clerical resolution
- Generate work units from the input universe based on the initial level of resolution required and geographic areas (coding areas plus adjacent counties)
- Allow a clerk to skip a case and proceed to the next case; ensure all cases are resolved prior to the work unit being closed
- Allow a clerk to retrieve and review the status of all Non-ID cases within their assigned work unit

- Perform a search of all streets or MAF units within specific state/county/tract/block geography

- Allow a clerk to correct respondent-provided address information

The customer requirements for MAGIC were finalized in November 2008. The Address Programs Management Branch (APMB) within GEO subsequently began working on the software requirements specification (SRS) for MAGIC. The SRS was baselined in June 2009 and was updated in an iterative process as planning and software development went on.

In addition to the address matching and geocoding interface provided by MAGIC, the clerical staff needed mapping software for use as a spatial reference during the telephone portion of the clerical operation. GEO recommended a spatial application they developed in-house that was being used for other programs called the Search Utility for Mapping Objects (SUMO). SUMO would be used by the clerks to locate streets while speaking to the respondent on the telephone, and then to capture the geographic codes by selecting the area on the map where the LQ is located. The geographic codes would be transferred to MAGIC to aid the clerk in their search for an address match or geocode.

Shortly after the completion of the MAGIC SRS, NPC staff came to HQ to discuss the logistics for the clerical operation and for a clerical test at NPC in mid-July 2009. In addition, the first functional versions of MAGIC and SUMO were demonstrated.

The July 2009 clerical test at NPC consisted of a functionality test and a load test. The functionality test for MAGIC and SUMO involved approximately 40 clerks who were monitored by HQ staff who noted any problems the clerks encountered while using the software. Two versions of SUMO were tested during the clerical test at NPC. One was the MapViewer version and the other was the Esri version. The functionality test also determined if additional upgrades were required and ensured the critical requirements of the software were met, such as the match and search capabilities.

The load test was performed on MAGIC and SUMO to simulate the estimated user activity during the clerical operation and determine if the system resources could adequately support the needs of these applications. In addition to the live users, the load test simulated about 200 MAGIC users and about 200 SUMO users.

The feedback from the July 2009 tests was used by HQ staff to modify and retest both MAGIC and SUMO. Due to the available technical support and some unique functionality only available in the MapViewer version of SUMO (e.g., ability to search by intersection), that version of SUMO was selected for use in August 2009.

Once the clerical workflow (i.e., the steps a clerk should take to resolve a Non-ID case) was finalized, MAGIC underwent modifications to provide the functionality necessary to complete each step in the workflow. The clerical workflow involved dividing the staff into Front Line clerks and Referral clerks. The Front Line clerks would work cases that required only the use of MAGIC and send the rest to the Referral staff, who utilized additional sources such as SUMO,

information from FastData, and information provided by respondents via the telephone. MAGIC and SUMO were continually tested and modified as a result of the training exercises and dry runs until the start of clerical training in late March 2010.

Clerical Processing Workflow

See Appendix B for the detailed workflow for the clerical process.

QC Method for Clerical Processing

MAGIC utilized the preexisting QC method from the ACS clerical software for Front Line cases. The QC method consisted of a "triplicate method" in which an automated process built into MAGIC selected every tenth case from four separate work units (a set of cases assigned to a specific clerk) to obtain a ten percent sample. If the number of cases in a work unit was below ten (a threshold identified by the Non-ID Processing subteam), the work unit would not have any QC cases pulled from it. The sample cases were worked by three clerks and the results were compared using a "majority rules" algorithm. The case passed QC if all three cases matched and/or geocoded the same. The case also passed QC if two of the three cases matched and/or geocoded the same. If all three cases were matched and/or geocoded differently, the case failed QC and was submitted to the Referral staff to resolve. If a work unit had QC differences of 25 percent or greater, then the work unit failed QC. Work units that failed QC were reworked by a clerk who was not one of the original clerks who reviewed the work units. The QC method was applied again until all work units passed, up to a maximum of three cycles.

The triplicate method was designed to identify consistency, not verify accuracy like the standard QC method for most operations. Due to technical limitations that prevented a work unit from being reopened once it was completed, there was no opportunity to correct cases, unless errors were discovered while the clerk was still active in the work unit. For these instances, the work unit could be reassigned with the resolutions blanked out from the first clerk. If errors were discovered after the work unit was completed, the clerk was shown their errors so they could avoid making the same mistakes on future cases. DSSD quality assurance subject matter experts concurred with the decision to retain this QC methodology.

Calling Component

Early into the development process the Non-ID Processing subteam concluded that little from the Census 2000 operation could be utilized for the 2010 Census. For example, the Census 2000 operation used a call scheduler as part of the IMAGS software. This was deemed ineffective by the Census 2000 Non-ID Processing planning staff, plus there was no possibility of developing a similar component for MAGIC. Specifically, due to limitations in the core MAGIC software, there was not the flexibility to transfer cases or work units between clerks. This meant it was impossible to have one clerk attempt to call a respondent on one shift and then have another clerk make another attempt on another shift.

The concern about work units being left unfinished because a clerk needed to contact a respondent to resolve a number of cases led to the initial plan to limit the calls per case to just

one. To be consistent with other operations with a calling component and to ensure the public was given a reasonable chance to get their form accepted for the 2010 Census, the decision was eventually made to cap the number of calls per case at three.

It was decided to group the referral cases by the county to which they were header coded, to allow the timing of the calls to be based on time zone. The clerks would not be allowed to make calls prior to 8 a.m. nor after 9 p.m. in any time zone. NPC is located in the Eastern time zone, so that meant calls could start as early as 8 a.m. Eastern Standard Time (EST) for Eastern time zone cases and end as late as midnight EST (9 p.m. Pacific Standard Time) for Pacific time zone cases.

NPC planned to handle foreign language calls that required a translator by using staff within NPC's Geography Branch who spoke the required language, staff from the Jeffersonville Telephone Center (JTC), or staff from the Tucson Telephone Center (TTC). If there was no one within the NPC Geography Branch who could assist with the call, then the Referral staff could request the assistance of JTC/TTC. The plan was for the Referral staff to send JTC/TTC information of how many cases needed a translator and which languages needed to be translated, and then to conduct a conference call between the Referral clerk, the translator, and the respondent for each case.

It was known that the NPC clerical staff would need a tool to look up a respondent's telephone number for those cases where it was not provided, but the decisions on how and when to use FastData were made once the HQ staff attended a training session in late 2009 conducted by the vendor explaining the capabilities of the database. The Non-ID Processing subteam agreed that FastData would only be used by Referral clerks to obtain a telephone number when there was enough address information for the case but the respondent's telephone number had not been provided[10]. No address information from FastData would be used to make corrections to the address information in MAGIC. The reason for the limited use of FastData was because the address information provided could not be verified due to the database containing multiple forms of the same address and multiple addresses for the same person; therefore, only the respondent could confirm the address information via telephone.

Toll-free Telephone Number

One of the final things that needed to be set up before production started was a toll-free telephone number that would show on respondents' Caller ID when a clerk attempted to make contact via telephone. The toll-free number allowed respondents to hear a prerecorded message confirming that the Census Bureau attempted to reach them for the purposes of verifying their information for the 2010 Census, and that they may be contacted again.

The purpose for the toll-free number was two-fold. First it served to provide respondents with an additional sense of validity for the operation, and second it alleviated concern about call backs. If a respondent requested proof that the clerk really worked for the Census Bureau, the clerk was

[10] For numerous BC Non-ID cases, the respondent's telephone number was not in the expected response field but was instead entered in the "House Number" field. If the "House Number" field contained a valid phone number, the clerk did not utilize FastData.

to direct the respondent to the toll-free number where they could obtain more information. Additionally, if a respondent did not answer, but attempted to return the call, they would be using the same toll-free number. These procedures reflected the Non-ID Processing subteam's decision that respondents should not have the ability to call clerks directly at their desks because the clerk probably would not be able to retrieve the specific case that the respondent was calling about given the limitation within MAGIC to retrieve cases from closed work units.

Puerto Rico Contracting Staff

In developing the components of the 2010 Census Non-ID Processing operation, the Non-ID Processing subteam decided to use a contractor to perform the clerical matching and geocoding of all the cases originating from Puerto Rico. The decision was based on the complexity of Puerto Rico addresses and the potential communication requirements involved in the calling component of the clerical operation. It was believed that the requirements for qualified individuals who had a comprehensive knowledge of the addressing system in Puerto Rico and the ability to communicate with the respondents in Puerto Rico could only be met by a contractor. The contract was awarded in February 2010 and the contractor staff was given office space near Census HQ.

Unlike stateside clerical processing, there were no Front Line clerks for Puerto Rico clerical processing. Given the smaller expected workload and staff, the contracting staff would act as the equivalent of the Referral clerks at NPC, meaning they would use MAGIC, SUMO, information from FastData, and information provided by respondents via the telephone to resolve cases. GSB developed a modified version of MAGIC that incorporated the additional Puerto Rico address elements. For the most part, the versions of MAGIC used for stateside and Puerto Rico had the same functionality. HQ staff began testing the prototype version of the Puerto Rico clerical coding system during late summer/early autumn 2009.

2.3.2 Implementation

This section describes the activities involved in the execution of the 2010 Census Non-ID operation. It includes a discussion of the automated process, the staffing and training process for the clerical operation, and the clerical process. The activities described in this section cover both stateside and Puerto Rico, identified where appropriate.

Automated Process

ADDUP File Deliveries

DSCMO delivered the first Non-ID Type A/B ADDUP file on March 30, 2010, and GEO started running these cases through the automated process on April 1, 2010. The ADDUP file deliveries were planned for every Tuesday throughout the duration of the operation, but starting with the sixth delivery on May 4, 2010, the frequency increased in order to process more cases, which meant continuous work could be provided for NPC clerical staff. This allowed more cases to be processed before the 2010 FV cutoff date of May 28, 2010.

Ten Type A/B ADDUP files were delivered from May 4 through May 27, 2010, resulting in a total of 15 ADDUP files delivered before the 2010 FV cutoff date. All Type A and Type B cases GEO received from DSCMO by the May 28 cutoff date were processed and included in the 2010 FV universe, if eligible. The number of addresses sent to 2010 FV from the Non-ID cases completed by the May 28 cutoff date was 295,371.

The operational plan for 2010 Census Non-ID Processing included an assumption that no Non-ID inputs would be provided to 2010 FV beyond the planned May 2010 cutoff date. However, in July 2010, Census Bureau management decided more Non-ID cases should be given the opportunity to be part of the 2010 FV workload. As a result, 11 Type A/B ADDUP files were delivered and processed by the 2010 FV supplemental cutoff date of August 4, 2010, generating an additional 34,154 addresses for the 2010 FV workload. A total of 329,525 addresses were sent to 2010 FV from Non-ID Processing.

After the 2010 FV initial cutoff date, the Type A/B ADDUP file deliveries resumed a weekly schedule until the last delivery on August 9, 2010, except for the week of July 26, 2010, when DSCMO provided two deliveries. There were a total of 27 Type A/B ADDUP file deliveries for the operation containing a total of 1,278,978 cases.

The Type C ADDUP files were delivered weekly starting March 30, 2010 and ending August 16, 2010. The ADDUP files for ETL were the first to arrive, but eventually the files containing the Type C cases from the other operations began arriving. A total of 1,608,779 Type C cases were delivered for Non-ID Processing. The processing of the Type C cases occurred on a flow basis with the other records from the respective operations, with the U/L records being the first to be processed starting on April 1, 2010, and the VDC records the last to be processed on August 23, 2010.

Resolution Rate

For the 2010 Census, the resolution rate for the automated process prior to clerical processing was expected to equal the Census 2000 Non-ID Processing automated resolution rate used for the budget estimates (72 percent). However, the percent of cases resolved per ADDUP file delivery ranged from 41.9 percent to 69.1 percent, with an average of 56.9 percent (see Table 2.3.5). The lowest resolution rate (41.9 percent) can be attributed to the approximately 90,000 TQA cases delivered that week with the street name omitted due to a data mapping issue between the UCM and the ADDUP file. Without street names, these cases were unmatched and ungeocoded in the automated process and sent to the clerical process for resolution. The issue was quickly identified and resolved, and the street names were delivered to GEO, who then supplied them to NPC staff, allowing them to work the cases.

Table 2.3.5 Automated Non-ID Processing Totals

ADDUP File Delivery Date	Total Number of Type A and Type B Cases Received by GEO from DSCMO	Total Number of Type A and Type B Cases Resolved During Automated Processing	Percent Resolved During Automated Processing
3/30/2010	24,943	15,394	61.7%
4/6/2010	41,210	25,596	62.1%
4/13/2010	216,771	134,278	61.9%
4/20/2010	222,464	135,385	60.9%
4/27/2010	312,695	130,883	41.9%
5/4/2010	174,022	99,841	57.4%
5/6/2010	45,549	26,482	58.1%
5/11/2010	21,837	13,375	61.2%
5/13/2010	13,457	8,388	62.3%
5/18/2010	21,429	13,697	63.9%
5/20/2010	9,512	6,066	63.8%
5/24/2010	9,757	6,297	64.5%
5/25/2010	3,238	2,192	67.7%
5/26/2010	3,737	2,449	65.5%
5/27/2010	3,342	2,146	64.2%
6/1/2010	4,578	3,010	65.7%
6/8/2010	6,470	4,253	65.7%
6/15/2010	5,624	3,644	64.8%
6/22/2010	7,309	4,748	65.0%
6/29/2010	11,313	7,518	66.5%
7/6/2010	35,400	24,459	69.1%
7/13/2010	75,482	52,007	68.9%
7/20/2010	2,585	1,526	59.0%
7/26/2010	1,906	1,152	60.4%
7/28/2010	1,015	643	63.3%
8/2/2010	2,468	1,396	56.6%
8/9/2010	865	570	65.9%
Total	**1,278,978**	**727,395**	**56.9%**

Source: GEO/CRDA file.

Another reason for the lower than estimated resolution rate was the approximately 20,600[11] noncity-style addresses sent to Non-ID Processing. For the 2010 Census, exact matching was only performed on rural route addresses. In contrast, during Census 2000 Non-ID Processing,

[11] See Section 4 – Limitations for explanation regarding how the number of noncity-style addresses delivered to 2010 Non-ID Processing was determined.

attempts were made to match all noncity-style addresses in the automated process if the respondent's information contained a name and/or a telephone number. The matching software used a combination of match keys that included the noncity-style address, ZIP code, occupant name, and telephone number. The cases matched through this method increased the Census 2000 automated resolution rate. For the 2010 Census, the method used for Census 2000 could not be used because names and telephone numbers are no longer stored in the MTdb; therefore, these cases could not be resolved during the automated process and were sent to the clerical process. Some other reasons that led to a number of unresolved cases during automated processing were:

- Misfielded data items that were due to the forms being completed incorrectly, either by the public or Census Bureau field staff
- Inconsistent form design concerning respondent address fields, which may have led to data capture and processing problems
- New HUs that had no MTdb record to match
- No program implemented for the latter part of the past decade to continuously update address ranges interactively in TIGER led to a large number of ungeocoded MTdb records

Staffing and Training

Staffing

When it came time for NPC to hire staff for the clerical portion of Non-ID Processing, they experienced hiring issues and delays due to circumstances outside their control. Even though the response to the vacancy announcements in early autumn 2009 was overwhelming with about ten people applying for every one vacancy, the pool of potential hires was much smaller. Roughly half the people who accepted positions showed up and some applicants said they were not interested in the position after starting the hiring process, possibly due to a 26-week extension of unemployment benefits that was authorized by the U.S. Congress in March, 2010 as part of their efforts to address the financial hardships confronting many of the nation's citizens at that time.

Of particular significance was the impact the hiring issues had on recruiting supervisory staff for the clerks. At the start of production, there were not enough supervisors to cover all the units working on Non-ID Processing due to the inability to fill all the positions. HQ staff from the Non-ID Processing subteam and NPC Geographic Technicians assisted by handling non-supervisory duties for both shifts, especially during training. NPC eventually filled all supervisor positions.

The bulk of the clerical staff was needed for the first five weeks of production to complete as many cases as possible by the 2010 FV cutoff date, but the actual number of NPC staff working on production changed as the workload fluctuated. Staff worked on other projects during lulls that were caused either by low workload or technical issues.

The original plan was for a 50-50 split between Front Line and Referral units, but when the work started coming in, it became apparent that adjustments needed to be made. The processing of cases by the Front Line clerks occurred more quickly than projected, resulting in the need for fewer Front Line units and more Referral units. Accordingly, only four Front Line units were created: two on the day shift and two on the night shift. During the first week of May 2010, the night shift units were converted to Referral units and only the two day shift Front Line units remained. Starting May 24, 2010, only two Referral units were assigned to the clerical process, with the actual number of clerks working on the project varying depending on the workload. Table 2.3.6 shows how many hours per week were devoted to Non-ID Processing, allocated by training, production, leave, and overtime.

Table 2.3.6 Hours Devoted to Non-ID Processing at NPC

Week of:	Training	Production	Leave (Annual, Sick, AWOP*)	Total Hours	Overtime (Additional Hours Worked Beyond the 40 hr Work Week)
4/11/2010	4,297	2,177	374	6,847	143
4/18/2010	3,744	5,605	424	9,773	77
4/25/2010	5,379	7,519	1,009	13,907	443
5/2/2010	1,161	12,532	1,214	14,907	716
5/9/2010	519	13,855	1,309	15,683	1,576
5/16/2010	71	9,145	969	10,185	490
5/23/2010	736	1,514	286	2,535	9
5/30/2010	283	550	432	1,265	0
6/6/2010	0	656	92	747	79
6/13/2010	0	376	31	407	1
6/20/2010	2	459	42	503	23
6/27/2010	2	1,069	219	1,289	85
7/4/2010	0	499	75	574	19
7/11/2010	0	1,809	99	1,908	70
7/18/2010	0	2,145	279	2,424	342
7/25/2010	0	656	192	848	49
8/1/2010[1]	162	530	124	816	0
8/8/2010	23	199	11	233	0
Total	**16,378**	**61,294**	**7,180**	**84,851**	**4,121**

Source: NPC, Financial Services Staff. Note: Hours are rounded to nearest whole number. *Absent Without Pay.
[1]Starting 8/3/2010, two clerical units at NPC were trained and did production work on Type C NRFU add cases and their hours were charged to Non-ID Processing.

Initially, the two groups of clerks were hired with different responsibilities and as different pay grades; Front Line clerks were Grade 4 employees and Referral clerks were Grade 4/5 (Grade 4 with the potential to be promoted to Grade 5) employees. However, during the implementation of the operation, when more Referral clerks and fewer Front Line clerks were needed, NPC

reexamined the performance plans for each grade to determine how best to shift staff. It was determined that Grade 4 clerks could perform more tasks, such as making telephone calls, and were able to become Referral clerks. The Grade 4 Referral clerks, referred to as First Pass Referral clerks, were able to make telephone calls while working on referrals, but they were not allowed to use FastData. The FastData searches were performed by the Grade 5 Referral clerks, Supervisors, Geographic Technicians, and HQ staff monitoring the operation on site.

Table 2.3.7 shows information from MAGIC that estimates the number of clerks working on each type of work unit per week. In this table, a clerk working on header coding or block coding cases is counted only once per week per work unit type. In either header coding or block coding, a clerk is counted for an individual week if a work unit they were assigned was completed. However, a Referral clerk could be counted up to three times each week because there are three types of referral work units (block coding referral, three-way tie from QC, and automated match with no geocode). Referral clerks also worked on header coding cases, so they could be counted up to a total of four times per week.

Table 2.3.7 Estimated Number of Clerks by Work Unit Type

Week of:	Number of Clerks Working on Header Coding Cases	Number of Clerks Working on Block Coding Cases	Number of Clerks Working on Referral Cases
4/11/2010	45	83	19
4/18/2010	61	101	102
4/25/2010	41	102	163
5/2/2010	28	61	282
5/9/2010	16	101	328
5/16/2010	18	155	314
5/23/2010	20	2	85
5/30/2010	5	0	45
6/6/2010	4	0	50
6/13/2010	6	0	55
6/20/2010	10	0	57
6/27/2010	0	0	4
7/4/2010	8	0	52
7/11/2010	5	0	57
7/18/2010	18	0	85
7/25/2010	1	0	77
8/1/2010	20	0	64
8/8/2010	17	0	48

Source: GEO/MAGIC.

Training

Training at NPC began on March 29, 2010 with the "Train the Trainers" sessions. These sessions, conducted for both shifts, served the purpose of training some of the Supervisors, Geographic Technicians, and Lead Clerks on the facets of Non-ID Processing, focusing on MAGIC and SUMO. These "Train the Trainers" sessions lasted seven workdays.

The next training sessions started two days later and marked the beginning of the Front Line and Referral clerical training. The Front Line and Referral clerks were trained separately. These training sessions were also conducted for both shifts and initially lasted six workdays. Clerical training sessions were conducted on a continuous basis through the end of May 2010 as new staff was brought onto the project. As the management staff became more familiar with the program and the training, they were able to get clerks into production more quickly.

The original plan called for a test to be administered to each clerk as they completed their training to evaluate their understanding of the process and to determine their readiness for production. Those that failed the test would undergo additional training and would retake the test until they successfully completed it. However, in an effort to get clerks into production more quickly, each trainee was given a practice work unit instead of a test, and received feedback and the opportunity to rework cases from a Geographic Technician or one of the on-site HQ staff before moving onto production.

A two-day supervisor training was also conducted in mid-April 2010. This training focused on teaching the supervisors how to utilize the Supervisor section of MAGIC, including using reports and assigning work units.

Toward the end of clerical processing, a training session occurred associated with some unplanned work. It was discovered in late July 2010 that a large number of NRFU and VDC add records were being rejected during processing due to missing or incomplete block geocodes. Due to the volume of rejected NRFU and VDC Type C cases and the fact that this was the last chance for these addresses to get into the census universe, these cases were sent through the automated geocoding software. If a geocode still could not be obtained, the cases were sent to the clerical operation at NPC in an attempt to obtain a geocode. Two clerical units were trained for this work on August 3, 2010.

Puerto Rico

The initial workload estimates for Puerto Rico cases based on a 20 percent increase from Census 2000 called for a staff of 12 contractors. The revised estimates that used a 50 percent workload increase determined the need for a staff of 21 contractors. The contractor managed to fulfill the staffing requirements even with the request for additional staff coming a few weeks before training started.

The contractor staff originally consisted of 18 Junior Analysts, two Senior Analysts, and one Project Manager. The number of staff was down to seven by the end of May 2010. In early June 2010, the Junior Analysts started working part-time. The workload was so small by the end of

June that the contractor was no longer needed. Table 2.3.8 shows the number of contractor staff per week, by position.

Table 2.3.8 Puerto Rico Non-ID Processing Contractor Staff

Week of:	Junior Analysts	Senior Analysts	Project Manager	Total Contractor Staff
4/4/2010	11	2	1	14
4/11/2010	11	2	1	14
4/18/2010	18	2	1	21
4/25/2010	19	1	1	21
5/2/2010	18	1	1	20
5/9/2010	18	1	1	20
5/16/2010	18	1	1	20
5/23/2010	18	0	1	19
5/30/2010	6	0	1	7
6/6/2010*	4	0	1	5
6/13/2010	4	0	1	5
6/20/2010	4	0	1	5
6/27/2010	3	0	1	4

Source: GEO/Geographic Contracts Management Branch.
*Starting the week of 6/6/2010, the Junior Analysts worked part-time.

There were two training sessions conducted by GEO/APMB for the Puerto Rico contracting staff. The first training, which involved all the staff, began April 7, 2010 and lasted six workdays, with a supervisor training for the Senior Analysts and Project Manager occurring on the fifth day. The second training started April 19, 2010 and lasted five workdays. The second training was conducted for the group of analysts hired to fill the request for additional staff for the estimated workload increase. There was one fewer training day for the second group because they were able to sit with the first group who had started production that week and learn more about the process. Sample cases using Puerto Rico addresses were used for the trainings.

Clerical Process

Clerical Deliveries

The first clerical delivery to NPC was on April 12, 2010. The plan was for the clerical deliveries to occur once a week. Similar to what occurred with the flow of ADDUP files to GEO, the deliveries sometimes occurred more than once a week to ensure a steady flow of work for the clerical staff. Sometimes clerical deliveries consisted of just header coding cases because this was the first automated process to complete and generate cases for the clerical process. The Puerto Rico cases were included in the same clerical deliveries to NPC as the stateside cases and were accessed by the contractor staff through MAGIC.

Each night the back-end processing ran on the clerical cases. This processing applied the QC method and moved cases over to referral. There were occasions when this process was performed in-between shifts to make cases available for the night shift. Completed cases were sent to HQ on a flow basis for post-clerical processing and MTdb updating.

The total number of Type A/B cases delivered to the clerical process was 551,583. The first five weeks of production were when the majority of cases were delivered. This was expected because the vast majority of BC and TQA cases were anticipated to come in during this timeframe. The distribution of each clerical delivery by case type is shown in Appendix C. The stateside numbers (Header Coding, Block Cases, and Referral Block Cases) represent the final totals per category for each clerical delivery, which are larger than the initial totals delivered. Each time a case moved to another category, it was counted as a separate case, allowing the opportunity for cases to be counted two or three times.

The clerical delivery on May 14, 2010 contained header coding cases that were deemed uncodable when the cases were originally worked. The cases were reworked because GEO/APMB felt there was additional information for these cases that could be used to possibly header code the cases and move them to block coding. The clerks were unable to see this information in MAGIC due to data mapping issues. GEO/APMB staff worked these cases themselves. After the May 14 clerical delivery, the data mapping issue was corrected and no further cases were reworked.

There were three clerical deliveries for the NRFU/VDC Type C cases. These cases were originally rejected for a missing or incomplete block geocode and were sent through the automated geocoding process to obtain a complete geocode, but were unsuccessful. The total number of NRFU/VDC Type C cases delivered was 35,213. The distribution of each clerical delivery by case type is shown below in Table 2.3.9. Similar to the numbers for the Type A and Type B cases in Appendix C, the numbers in Table 2.3.9 represent the final totals per category for each clerical delivery, which are larger than the initial totals delivered. Each time a case was successfully header coded and became either a referral block case or a Puerto Rico case, it was counted as a separate case.

Table 2.3.9 Clerical Deliveries – NRFU/VDC Type C Cases

Delivery Date	Header Coding Cases	Referral Block Cases	Puerto Rico Cases	Total
8/2/2010	10,661	12,944	86	23,691
8/5/2010	2,708	1,148	0	3,856
8/10/2010	5,727	1,962	32	7,721
Totals	**19,096**	**16,054**	**118**	**35,268**

Source: GEO/MAGIC reports. Note: A NRFU/VDC Type C case can possibly be worked twice and is counted as a separate case each time. For example, a case can be successfully header coded and then resolved as a referral block case.

Production

Production began as scheduled on April 15, 2010. The clerical process continued until work on the final delivery was completed on August 12, 2010, four days earlier than scheduled. Work was performed on a number of cases from the first clerical delivery prior to the official April 15, 2010 start of production. Some of the cases were worked by the clerical staff as part of a load test of MAGIC conducted on April 13, 2010. Other cases were worked by GEO/APMB to help ensure the whole system was working correctly. All the work done prior to April 15, 2010 counted as production and the results were retained.

Table 2.3.10 presents the production rate for each type of work unit. The total cases per week represent how many cases of that type were completed during each work week, which may include weekends. The production time data represent how many hours each week it took to complete all the cases. The cases per hour columns show the average number of the cases completed per hour. The numbers for the referral cases represent the combined totals for the three types of referral cases (block coding referral, three-way tie from QC, and automated match with no geocode).

The production rates for the Front Line and Referral clerks were dramatically higher than estimated. It was estimated the Front Line clerks would average ten cases per hour, and the actual rate was 48 cases per hour. The Referral clerks were estimated to have a production rate of five cases per hour, but the actual rate was 15 cases per hour.

Table 2.3.10 Production Rate (Cases Per Hour)

Week of:	Total Header Coding Cases Per Week	Production Time (Hours)	Header Coding Cases Per Hour	Total Block Coding Cases Per Week	Production Time (Hours)	Block Coding Cases Per Hour	Total Referral Cases Per Week	Production Time (Hours)	Referral Cases Per Hour
4/11/2010	9,255	102	91	27,626	882	31	270	43	6
4/18/2010	6,514	67	97	118,041	3,076	38	14,878	1,653	9
4/25/2010	7,468	55	135	116,954	2,061	57	37,907	4,331	9
5/2/2010	4,780	37	128	107,649	2,040	53	71,569	7,670	9
5/9/2010	4,521	33	136	123,773	1,854	67	123,129	10,338	12
5/16/2010	740	10	75	19,050	391	49	60,730	5,368	11
5/23/2010	426	5	82	42	1	42	6,599	475	14
5/30/2010	87	1	102	0	0	0	1,292	80	16
6/6/2010	158	1	155	0	0	0	2,348	235	10
6/13/2010	161	1	161	0	0	0	1,696	169	10
6/20/2010	227	3	80	0	0	0	2,433	214	11
6/27/2010	0	0	0	0	0	0	68	4	16
7/4/2010	438	2	219	0	0	0	3,498	340	10
7/11/2010	2,347	10	239	0	0	0	9,752	884	11
7/18/2010	662	9	70	0	0	0	20,482	1,626	13
7/25/2010	52	0	173	0	0	0	3,015	293	10
8/1/2010	13,456	30	449	0	0	0	15,133	293	52
8/8/2010	5,766	13	457	0	0	0	2,225	59	38
Average Per Hour			**158**			**48**			**15**

Source: GEO/MAGIC. Note: The numbers in the Production Time (Hours), Header Coding Cases Per Hour, Block Coding Cases Per Hour, and Referral Cases Per Hour columns are rounded to the nearest whole number, with all production time of less than an hour rounded up to one hour.
[1]Starting the week of 5/30/2010, all Block Coding cases became Referral Cases.

There were some cases that contained unexpected information in the address fields that led to changes in procedures. For example, cases that contained wording in the address information indicating homelessness appeared in the workload. The expectation was respondents without an address on April 1, 2010 would have checked the appropriate box on the BC form, allowing these cases to be designated as Type B cases and processed accordingly. Some respondents though did not check the box and wrote the word "Homeless" or some variation thereof in the address fields on the BC form. As a result, these cases met the criteria to be classified as Type A cases because the box was not checked, and were sent through the full automated process. Predictably, these cases were not resolved during automated processing due to the nonstandard information in the address fields, and thus were sent to clerical processing for resolution. As with automated processing, these cases were considered uncodable due to the lack of address information. However, in order to give them the opportunity to be allocated to the GQ population, the cases were changed to Type B cases at HQ after the clerical staff marked the cases as uncodable.

In addition, a small number of cases contained an address that was associated with a private mailbox, indicated by a "PMB ###" at the end of the address. These addresses were discovered to be for businesses that provide private mailboxes and were not the location address for the respondent. The Front Line clerks were instructed to refer these cases if there was a name or telephone number provided with the address information. Otherwise, these cases were classified as uncodable.

Other kinds of cases that led to revised procedures were hyphenated addresses in Queens, New York and in Hawaii, as well as unusual living situations such as boats. In the case of the hyphenated address, the clerks were allowed to make exact matches between the unhyphenated address (e.g., 10001 Hill St) from the respondent and the hyphenated address (e.g., 100-01 Hill St) in the MAF extract. For respondents living on a boat, the clerks were instructed to geocode the address to the nearest marina or land block where they were docked. For other unusual living situations, such as extended stay motels or campgrounds, the clerks attempted to only match to the MAF reference file, and failing that, they marked the record as uncodable.

MAGIC and SUMO

Some changes related to MAGIC were made early in production. First, "House Number" was added as a field in the Search window, allowing the clerks to perform searches using the house number listed in the respondent's address information in conjunction with the street name and ZIP code. Second, the City, County, and State fields were removed from the Corrected Information tab. Initially, the clerks were instructed to fill out the Corrected Information tab regardless of the outcome of the case. Early into production, GEO realized there was no need for the clerks to enter corrected information when the address matched a record in the reference file. The automated process did not require the city/county/state information to be updated, as the geocode that gets associated with the address determines the geography and will overwrite any incorrect header coding information.

Overall, the performance of MAGIC met the requirements of the program. The times it was unavailable to users were usually related to the distribution of an updated version. This was

generally handled between shifts so as to minimize work disruptions. The fact that the software was hosted on several servers meant, if it was necessary to push out the new version during a shift, the clerks could log into one server while the new version was distributed on another server, and then switch back after the distribution was completed.

There were some issues with SUMO that arose during production. The load tests performed on SUMO prior to production indicated that the software could handle 200 clerks simultaneously during production. However, at the beginning of production, the software performed very slowly and would stop working for periods of time. Approximately a month into production, the server configuration files required adjustment due to the high number of runaway sessions generated by the SUMO application. These additional processes overutilized the server's processing and memory resources, degrading SUMO's performance and causing error messages. After updates were applied to the server configuration files, the performance of SUMO improved.

Another software issue that developed early into production was that the previous cookie (a collection of information stored on a computer and used to relate one computer transaction to a later one) left by SUMO was not being overwritten by the current cookie, therefore populating MAGIC with the incorrect block information. This problem was quickly identified and corrected through the use of a cookie clean-up application installed by NPC. The incorrect block information though could not be fixed for those cases that had already been completed and gone through MTdb updating.

Many clerks would use Google Maps as a substitute for SUMO while on the telephone with a respondent due to the slow performance issues or because SUMO was down. They would locate the address in Google Maps using street intersections and landmarks, and after the telephone call had concluded, they would locate the same area in SUMO to get the geocodes.

Calling Component

In order to increase the number of cases resolved per day, the calling component of the clerical operation underwent some changes a few weeks into production. Data from a sample of clerks showed that the majority of cases that were resolved with a telephone call were resolved after two calls. The decision was made on May 6, 2010 to change the maximum number of calls per case from three to two.

In early May 2010, another alteration of the calling component occurred when the concept of the "phone helper" was introduced. A phone helper was a clerk who was paired up with a Referral clerk(s) and assisted them by making additional telephone calls for cases where the Referral clerk was unable to contact the respondent during the first call. The Referral clerk provided the phone helper with a screenshot of MAGIC containing the respondent's address, name, and telephone number for each case. While the Referral clerk continued working the rest of the cases in the work unit, or switched to another work unit using an alternate account, the phone helper made the additional calls to the respondent. The phone helper recorded on the MAGIC screenshot any additional address information they received, and if they were able to get location information from the respondent, they made a screenshot from Google Maps displaying the longitude and latitude of the area near the location and used an "X" to mark the location of the address. The Referral clerk then used the additional address information to attempt an address or

address range match, or entered the location information in SUMO to get a block geocode. Two of the day shift Referral units utilized phone helpers from the night shift on an as-needed basis.

A change to the clerical procedures occurred due to the clerks making contact with GQs during their telephone calls. The procedures did not cover what the clerks should do when the telephone numbers provided by the respondents led to a GQ being contacted. There was a concern that matching to a GQ, especially a sensitive GQ (e.g., domestic violence shelter), would add a HU MAF Unit that would be eligible for the 2010 FV universe. The procedures were changed so that the Referral clerks were instructed to only attempt to match these cases to the MAF reference file once the address was identified as belonging to a GQ. If no match could be made, then the Referral clerks were instructed not to attempt an address range search using MAGIC or SUMO, and to just make the case uncodable. There was no mechanism to track how many times this situation occurred, but that is functionality to be considered for future versions of the clerical application.

The procedures for handling foreign language cases that required a translator evolved as production continued. An attempt was made to identify foreign language cases in MAGIC by using the form type that was sent in as an indication of the language the respondent spoke. The Referral clerk worked the case first without making a telephone call. Once the call was made and the Referral clerk identified that a respondent spoke a language other than English, they communicated as best they could that they would call back shortly with a translator. The Referral clerk then filled out a sign-up sheet noting what language they needed translated. A NPC Geography staff member was utilized as a translator if there was someone who spoke the necessary language or a NPC Supervisor or Geographic Technician called JTC or TTC from the Referral clerk's workstation and requested a translator. When a translator was secured, the Referral clerk provided the pertinent information on the case to the translator and then called the respondent back. The initial call where the respondent was identified as speaking a language other than English did not count towards the maximum number of calls allowed by the Referral clerk.

Puerto Rico

The contractor staff began production on April 15, 2010, the same day as NPC staff. The contractor staff worked until June 30, 2010. Staff from GEO/APMB worked the remaining cases until they were completed on August 12, 2010.

The updates made to the stateside version of MAGIC were also simultaneously made to the version used for the Puerto Rico cases. The performance of MAGIC was the same for the contractor as it was for NPC.

The same stability issues regarding SUMO that existed for NPC were also experienced by the contractor staff. The analysts supplemented SUMO with a website created by the government of Puerto Rico that included boundaries not contained within SUMO. When this website demonstrated performance issues, the staff was permitted to use Google Maps.

The calling portion of the operation was critical for the Puerto Rico cases. SUMO contained a large number of unnamed roads, which limited the search ability. Without calling the respondent and gathering the necessary information, the analysts would not have been able to geocode many of the cases. The contractors had the time and the resources to allow for up to three telephone calls per case for the entire time they were working on the project, so the maximum number of telephone calls was never limited to two as was done for stateside calls.

The Puerto Rico Non-ID analysts provided the telephone number of the JTC to respondents who were hesitant to cooperate and wanted to verify the Non-ID Processing operation. The JTC staff was provided the names of the analysts so that if a respondent called in response to receiving a call, the JTC representative could verify the legitimacy of the operation and confirm the analyst worked for the Census Bureau. The JTC was willing to do this because of the relatively small number of analysts working on Puerto Rico clerical processing.

3. Methodology

The assessment questions for the 2010 Census Non-ID Processing operation were answered by analyzing data from the CRDA file for Type A and Type B cases and various assessment files for the Type C cases. Additional information was gathered from debriefings, observations, and lessons learned exercises. The results are presented in the following categories when appropriate: case type, source, location (i.e., stateside, Puerto Rico, and foreign country), milestone totals, and cumulative totals at the end of the operation.

4. Limitations

Comparison of Results to Census 2000

A direct comparison between Non-ID Processing during Census 2000 and the 2010 Census is difficult due to the differences between each version of the operation, including:

- Different matching rules used in automated processing (e.g., matching of noncity-style addresses in 2000)
- Different clerical applications
- Changes to the MTdb throughout the past decade

The number of UHE Type A cases cannot be compared between Census 2000 and the 2010 Census because the numbers of UHE cases from Census 2000 was tallied by enumeration form type (e.g., ICR, MCR) and not by operation. Another limitation is that in Census 2000, there were 2,281,712 UHE addresses from GQ questionnaires that should have been excluded from the process. GQ questionnaires with a UHE address were to be screened for exclusion based on GQ type and the outcome of the screening questions; this was done to prevent people in certain UHE-ineligible types of GQs (e.g., prisons) from being improperly enumerated at a residence other than their GQ. The screening questions were intended to ensure that if persons whose primary residence was a GQ also provided a UHE address, they would not be enumerated elsewhere on that basis. Both filters were not applied during production, but the screening by GQ type was done after the Non-ID Processing clerical operation was completed, prior to the

addresses going to 2000 FV. Thirty-seven percent of the GQ questionnaires resolved during Non-ID Processing were in that process inappropriately because the filtering of GQ questionnaires was not done correctly.

Data issues

Some variable values were not populated or were incorrectly populated in the CRDA file used to generate the numbers used in Section 5. A relatively small number of cells were affected, but this may have caused certain numbers used to answer the assessment questions to be off from the true number by no more than a few cases.

Another situation that may have affected the results is the 603 cases that were originally Type A cases but were converted to Type B cases after post-clerical processing because the address fields contained information suggesting the respondent was experiencing homelessness (see Section 2.3.2 for the complete explanation). These cases are classified as Type B cases in the CRDA file; however, because they were treated as Type A cases throughout automated, clerical, and post-clerical processing, they were treated as Type A cases when calculating the results for assessment questions one through five.

The cases from RA, RUE, and UE were combined into one source (RA/RUE/UE) for the CRDA file and MTdb updates. This occurred because the three operations used the same form to collect their UHEs, resulting in all the forms having the same form type, which is included in the Processing ID. This made it impossible to distinguish between the three operations when the forms arrived for data capture. Information on the add records from each of the operations can be found in the 2010 Update Enumerate Operations Assessment.

The exact number of Non-ID cases containing noncity-style addresses is difficult to determine because of all the variations of noncity-style addresses delivered for processing. Also, the noncity-style addresses are not all located in one address field, making it difficult to account for all of them. As a result, only an estimated number of noncity-style addresses delivered to Non-ID Processing can be given. The estimate was derived by searching in the CRDA for common incarnations of noncity-style addresses such as Post Office (P.O.) Box, Rural Route, and Highway Contract Route addresses.

The number of Non-ID cases presented for each assessment question may not match the number of records reported in the assessments for each respective operation. The numbers presented by each assessment may represent a different stage in HQ processing, therefore leading to different numbers being reported in the assessments. In a similar manner, all results involving cases originating from GQE are preliminary because the 2010 GQE Assessment has not been finalized and the results have not been formally presented.

5. Results

This section presents the results of the 2010 Census Non-ID Processing operation. The results are presented as the answers to the seven research questions:

1. How many cases were sent to automated Non-ID Processing?

2. What was the outcome of automated Non-ID Processing (header coding, matching, and geocoding)?

3. How many cases were sent to clerical Non-ID Processing?

4. What was the outcome of clerical Non-ID Processing (header coding, matching, and geocoding)? How many cases were resolved via telephone call?

5. What was the outcome of post-clerical Non-ID Processing (matching)?

6. What were the actions taken on the records in the MTdb as a result of Non-ID Processing (automated, clerical, and post-clerical)?

7. How did the actual cost of the 2010 Census Non-ID Processing operation compare to the expected budget?

The results will be presented by case type and by the following categories when appropriate:

a. Source
b. Location
c. Milestone and cumulative totals at the end of the operation

5.1 How many cases were sent to automated Non-ID Processing?

For the 2010 Census, there were 1,265,551 Type A cases, 13,427 Type B cases, and 1,608,779 Type C cases delivered to GEO for automated Non-ID Processing. Each case type is subject to different processing workflows; therefore the results are presented in the appropriate tables for each type.

Type A Cases

There were 1,265,551 Type A cases delivered by DSCMO to GEO for automated Non-ID Processing. Of these, an estimated 20,600 cases contained a noncity-style address, including P.O. Box, Rural Route, and Highway Contract Route address information.

Table 5.1.1 shows the distribution of Type A cases by source. BC contributed the vast majority of cases (60.64 percent), while TQA Fulfillment and NRFU contributed over 10 percent each.

Table 5.1.1 Type A Cases Sent to Automated Non-ID Processing by Source

Source	Number of Cases	Percent of Total
Be Counted	767,487	60.64
TQA Fulfillment	200,704	15.86
TQA Interview	112,662	8.90
GQE[12]	51,197	4.05
NRFU	131,517	10.39
RA/RUE/UE	1,984	0.16
Total Cases	**1,265,551**	**100.00**

Source: CRDA file.

Table 5.1.2 presents a comparison of the number of cases from Census 2000 and the 2010 Census for BC and the TQA programs.

Table 5.1.2 BC and TQA Type A Cases - Comparing Census 2000 to the 2010 Census

Source	2000 Total	2010 Total	Percent Change
Be Counted	551,596	767,487	39.14
TQA Fulfillment	197,171	200,704	1.79
TQA Interview	510,996	112,662	-77.95
Total Cases	**1,259,763**	**1,080,853**	**-14.20**

Source: 2000 totals from Census 2000 Non-ID Processing Assessment Report.

The Non-ID Processing subteam initially estimated an increase of Non-ID inputs consistent with the 20 percent increase in the national HU inventory over the course of the preceding decade, and later adjusted the estimate to a 50 percent increase, which was appropriate in the case of BC. However, workload from TQA Fulfillment was relatively stable with an increase of less than two percent, and cases from TQA Interview decreased considerably; there were almost 78 percent less than in the Census 2000. The net difference of Type A cases from these three sources was a decrease of about 14 percent from Census 2000 figures.

The number of UHE Type A cases cannot be compared between Census 2000 and the 2010 Census because the number of cases from Census 2000 was tallied by enumeration form type (e.g., ICR, MCR) and not by operation. Also, during Census 2000 there was an error in the process that selected records from operations in which a UHE response was permitted, and therefore figures from Census 2000 Non-ID Processing appear grossly disproportionate with the 2010 figures, and comparisons would be misleading.

[12] It was discovered in November 2010 that, due to response data being mapped incorrectly during HQ processing, 57,426 UHE addresses from Shipboard Census Reports were not delivered to Non-ID Processing. The numbers presented for GQE in the 2010 Non-ID Processing Assessment reflect only the addresses that were actually delivered to GEO. Refer to the 2010 Census Shipboard Enumeration Operation Assessment for further information on this situation.

Table 5.1.3 displays the number of Type A cases by the location from which they originated. The location of a case is determined by the state and county assigned to a case, which could have occurred in either the automated or clerical process. Cases that have an address from outside the 50 states, the District of Columbia, and Puerto Rico were assigned a foreign country code of 99.

Table 5.1.3 Type A Cases Sent to Automated Non-ID Processing by Location

Location*	Number of Cases	Percent of Total
Stateside	1,239,424	97.94
Puerto Rico	12,963	1.02
Foreign Country	418	0.03
Not Associated with a State/County	12,746	1.01
Total Cases	**1,265,551**	**100.00**

Source: CRDA file. *The location where a case originated from may have been determined in either automated or clerical processing.

Table 5.1.4 displays the distribution of Type A cases delivered to automated Non-ID Processing by source and location.

Table 5.1.4 Type A Cases by Source and Location

	Stateside		Puerto Rico		Foreign Country		Not Associated with a State/County	
Source	Number of Cases	Percent of Total	Number of Cases	Percent of Total	Number of Cases	Percent of Total	Number of Cases	Percent of Total
Be Counted	748,082	60.36	10,192	78.62	43	10.93	9,170	71.69
TQA Fulfillment	199,922	16.13	755	5.82	1	0.24	26	0.20
TQA Interview	112,304	9.06	298	2.30	4	0.95	56	0.44
GQE	49,968	4.03	635	4.90	307	72.92	287	2.24
NRFU	127,581	10.29	1,083	8.35	62	14.73	2,791	21.82
RA/RUE/UE	1,522	0.12	0	0.00	1	0.24	461	3.60
Total Cases	**1,239,379**	**100.00**	**12,963**	**100.00**	**418**	**100.00**	**12,791**	**100.00**

Source: CRDA file. Note: Percentages may not sum to 100.00 due to rounding.

BC cases made up the majority of cases from stateside and Puerto Rico. BC cases also accounted for the largest group of cases not associated with a state/county. The proportion of cases coming from TQA Fulfillment and TQA Interview was significantly larger stateside than it was for the other areas. Of the 418 cases that were associated with a foreign country, nearly three-fourths (72.92 percent) originated from GQE.

Table 5.1.5 displays the distribution of Type A cases by milestone date. The three milestones were the 2010 FV initial cutoff, the 2010 FV supplemental cutoff, and the time between the 2010 FV supplemental cutoff and the end of the Non-ID Processing operation. It should be noted that each case delivered was fully processed by the close-of-business of the cutoff date. The vast

majority (87.76 percent) of Type A cases was matched and geocoded by the initial 2010 FV cutoff date.

The total number of Type A cases sent to 2010 FV from the 2010 FV initial cutoff was 295,371 (26.59 percent of the cases for that milestone), and the number of Type A cases sent from the 2010 FV supplemental cutoff was 34,154 (22.18 percent of the cases delivered for that milestone).

Table 5.1.5 Type A Cases Sent to Automated Non-ID Processing by Milestone

Milestone	Number of Cases	Percent of Total	Number of Cases Sent to 2010 FV	Percent of Total
Cases Received Prior to 2010 FV Initial Cutoff (May 28, 2010)	1,110,692	87.76	295,371	89.64
Cases Received After 2010 FV Initial Cutoff and Prior to 2010 FV Supplemental Cutoff (August 4, 2010)	154,007	12.17	34,154	10.36
Cases Received After 2010 FV Cutoff (Between August 4 and August 9, 2010)	852	0.07	0	0.00
Total Cases	**1,265,551**	**100.00**	**329,525**	**100.00**

Source: CRDA file.

Type B Cases

There were 13,427 Type B cases delivered by DSCMO to GEO for automated Non-ID Processing. As in Census 2000, BC was the only source for Type B cases. Additionally, there were 603 cases that were originally Type A cases but were converted to Type B cases after post-clerical processing because the address fields contained information suggesting the respondent was experiencing homelessness. These cases are classified as Type B cases in the CRDA file; however, because they were treated as Type A cases throughout automated, clerical, and post-clerical processing, they were treated as Type A cases when calculating the results for assessment questions one through five.

Table 5.1.6 presents a comparison of the number of Type B cases from Census 2000 and the 2010 Census.

Table 5.1.6 Type B Cases - Comparing Census 2000 to the 2010 Census

Source	2000 Total	2010 Total	Percent Change
Be Counted	17,106	13,427	-21.51

Source: 2000 totals from Census 2000 Non-ID Processing Assessment Report.

The decrease in 2010 Census of over 21 percent of Type B cases from BC could possibly be attributed to more people who did not have an address on April 1 being enumerated by ETL and SBE than in Census 2000.

Table 5.1.7 displays the number of Type B cases by the location from which they originated. The location of a case is determined by the state and county assigned to a case, which could have occurred in either the automated or clerical process. Cases that have an address from outside the 50 states, the District of Columbia, and Puerto Rico were assigned a foreign country code of 99.

Table 5.1.7 Type B Cases Sent to Automated Non-ID Processing by Location

Location*	Number of Cases	Percent of Total
Stateside	12,102	90.13
Puerto Rico	145	1.08
Foreign Country	3	0.02
Not Associated with a State/County	1,177	8.77
Total Cases	**13,427**	**100.00**

Source: CRDA file. *The location where a case originated from may have been determined in either automated or clerical processing.

The vast majority (90.13 percent) of Type B cases originated from stateside and just over one percent originated from Puerto Rico. The combined total number of cases that were coded to a foreign country or not associated with a state/county at all equals less than nine percent.

Table 5.1.8 displays the distribution of Type B cases by milestone date. Even though Type B cases were not eligible for 2010 FV, the 2010 FV cutoff represents a significant date that the Non-ID Processing subteam used as a marker for keeping track of workloads and staffing needs. The three milestones were the 2010 FV initial cutoff, the 2010 FV supplemental cutoff, and the time between the 2010 FV supplemental cutoff and the end of the operation. It should be noted that each case delivered was fully processed by the close-of-business of the cutoff date.

Table 5.1.8 Type B Cases Sent to Automated Non-ID Processing by Milestone

Milestone	Number of Cases	Percent of Total
Cases Received Prior to 2010 FV Initial Cutoff (May 28, 2010)	13,271	98.84
Cases Received After 2010 FV Initial Cutoff and Prior to 2010 FV Supplemental Cutoff (August 4, 2010)	143	1.06
Cases Received After 2010 FV Cutoff (Between August 4 and August 9, 2010)	13	0.10
Total Cases	**13,427**	**100.00**

Source: CRDA file.

Nearly all the Type B cases were delivered to Non-ID Processing by the 2010 FV initial cutoff date of May 28, 2010. Just over one percent of the remaining cases arrived after that date.

Type C Cases

There were 1,608,779 Type C cases delivered to GEO for automated Non-ID Processing. Table 5.1.9 shows the distribution of Type C cases by source. NRFU and U/L contributed the majority

of cases, with over 39 percent and over 33 percent respectively. While GQE and RA/RUE/UE contributed the lowest number of cases among the source operations, their respective areas of coverage were relatively smaller, so it follows that fewer addresses were added via these two operations.

Table 5.1.9 Type C Cases Sent to Automated Non-ID Processing by Source

Source	Number of Cases	Percent of Total
ETL	121,290	7.54
GQE	18,818	1.17
NRFU	630,175	39.17
VDC	230,619	14.34
RA/RUE/UE	73,039	4.54
U/L	534,838	33.24
Total Cases	**1,608,779**	**100.00**

Source: Assessment files provided by GEO for each operation.

Table 5.1.10 displays the number of Type C cases by the location from which they originated. The location of a case is determined by the state and county assigned to a case by the Census Bureau field staff, except for the group of NRFU cases that were sent to the clerical process because of a missing/incomplete geocode. Cases from any of the operations that had missing or incomplete geocodes may have been due to an omission by an enumerator, a keying error, or other factors. Cases in which the state could not be determined were designated as 'Not Associated with a State/County'; cases that have an address from outside the 50 states, the District of Columbia, and Puerto Rico were also put into this category.

Table 5.1.10 Type C Cases Sent to Automated Non-ID Processing by Source and Location

	Stateside		Puerto Rico		Not Associated with a State/County	
Source	Number of Cases	Percent of Total	Number of Cases	Percent of Total	Number of Cases	Percent of Total
ETL	120,698	8.32	141	0.10	451	2.04
GQE	18,626	1.28	192	0.14	0	0.00
NRFU	602,044	41.51	18,402	13.48	9,729	44.11
VDC	214,985	14.82	4,307	3.15	11,327	51.35
RA/RUE/UE	72,488	5.00	0	0.00	551	2.50
U/L	421,348	29.05	113,490	83.12	0	0.00
Total Cases	**1,450,189**	**100.00**	**136,532**	**100.00**	**22,058**	**100.00**

Source: Assessment files provided by GEO for each operation. Note: Percentages may not sum to 100.00 due to rounding.

The largest group of stateside cases originated from NRFU (41.51 percent), while U/L generated the most cases from Puerto Rico (83.12 percent). This makes sense because the entire island was covered during U/L, which was conducted earlier than NRFU, providing a greater chance for adds. Of particular note, NRFU and VDC cases, when combined, make up over 95 percent of the cases that could not be assigned to a state/county.

5.2 What was the outcome of automated Non-ID Processing (header coding, matching, and geocoding)?

Type A Cases

Of the 1,265,551 Type A cases sent through the automated process, 1,232,121 (97.36 percent) cases were successfully header coded to a state and county during the automated process. The detailed results for the cases header coded through automated processing are displayed in Table 5.2.1.

Table 5.2.1 Type A Cases Header Coded Through Automated Processing

Category		Number of Cases per Category	Percent of Total per Category
Source	Be Counted	743,423	60.34
	TQA Fulfillment	198,843	16.14
	TQA Interview	110,525	8.97
	GQE	50,032	4.06
	NRFU	127,787	10.37
	RA/RUE/UE	1,511	0.12
Location	Stateside	1,221,494	99.14
	Puerto Rico	10,627	0.86
Milestone	Initial 2010 FV Cutoff	1,081,595	87.78
	2010 FV Supplemental Cutoff	149,712	12.15
	After 2010 FV Cutoff	814	0.07
Total Cases		**1,232,121**	**100.00**

Source: CRDA file.

The number of Type A cases that were not successfully header coded to a state and county during the automated process was 33,430 (2.64 percent). This total includes 33,090 cases not header coded at all and 340 cases header coded to a foreign country. The 33,090 cases that did not receive a header code were sent to clerical processing for resolution.

Table 5.2.2 presents the results of automated processing for the cases that were successfully header coded.

Table 5.2.2 Automated Processing Results for Type A Cases

Category	Number of Cases	Percent of Total
Matched and Geocoded	656,032	53.24
Not Matched But Geocoded	59,113	4.80
Matched But Not Geocoded	114,672	9.31
Not Matched and Not Geocoded	205,524	16.68
Rejected	196,780	15.97
Total Cases	**1,232,121**	**100.00**

Source: CRDA file.

Table 5.2.3 through Table 5.2.7 present the detailed results for each category listed in Table 5.2.2. The detailed results for the Type A cases that matched and geocoded during automated processing are shown below in Table 5.2.3. A case was considered matched and geocoded if it either matched to a geocoded MTdb record (642,154 cases (97.88 percent)) or matched to an ungeocoded MTdb record but obtained a geocode via the automated geocoding routine (13,878 cases (2.12 percent)). Of significance, the 414,108 BC cases that were matched and geocoded during automated processing represented nearly 54 percent of the total BC cases delivered to Non-ID Processing.

Table 5.2.3 Type A Cases Matched and Geocoded Through Automated Processing

Category		Number of Cases per Category	Percent of Total per Category
Source	Be Counted	414,108	63.12
	TQA Fulfillment	54,955	8.38
	TQA Interview	74,088	11.29
	GQE	29,140	4.44
	NRFU	83,062	12.66
	RA/RUE/UE	679	0.10
Location	Stateside	656,029	99.99
	Puerto Rico	3	0.00
Milestone	Initial 2010 FV Cutoff	560,258	85.40
	2010 FV Supplemental Cutoff	95,255	14.52
	After 2010 FV Cutoff	519	0.08
Total Cases		**656,032**	**100.00**

Source: CRDA file. Note: Percentages may not sum to 100.00 due to rounding.

Table 5.2.4 below shows the detailed results for the Type A cases that were unable to be matched during automated processing but were able to be geocoded via the automated geocoding routine. The reason these cases were only able to be geocoded during automated processing may be a result of updates to TIGER that did not get reflected in the MAF, but it warrants further investigation. Any information discovered could assist in improving the matching routines.

Table 5.2.4 Type A Cases Not Matched But Geocoded Through Automated Processing

Category		Number of Cases per Category	Percent of Total per Category
Source	Be Counted	41,706	70.55
	TQA Fulfillment	2,148	3.63
	TQA Interview	4,202	7.11
	GQE	3,090	5.23
	NRFU	7,877	13.33
	RA/RUE/UE	90	0.15
Location	Stateside	59,107	99.99
	Puerto Rico	6	0.01
Milestone	Initial 2010 FV Cutoff	50,149	84.84
	2010 FV Supplemental Cutoff	8,925	15.10
	After 2010 FV Cutoff	39	0.07
Total Cases		**59,113**	**100.00**

Source: CRDA file. Note: Percentages may not sum to 100.00 due to rounding.

The detailed results for the Type A cases that matched to an ungeocoded MTdb record are shown below in Table 5.2.5. The majority of ungeocoded addresses in the MTdb originated from the Delivery Sequence File (DSF) from the United States Postal Service. The DSF is delivered twice a year and used to update the Census Bureau's address inventory.

Table 5.2.5 Type A Cases Matched But Not Geocoded Through Automated Processing

Category		Number of Cases per Category	Percent of Total per Category
Source	Be Counted	71,929	62.73
	TQA Fulfillment	20,825	18.16
	TQA Interview	15,362	13.40
	GQE	773	0.67
	NRFU	5,676	4.95
	RA/RUE/UE	107	0.09
Location	Stateside	114,672	100.00
	Puerto Rico	0	0.00
Milestone	Initial 2010 FV Cutoff	105,169	91.71
	2010 FV Supplemental Cutoff	9,440	8.23
	After 2010 FV Cutoff	63	0.05
Total Cases		**114,672**	**100.00**

Source: CRDA file. Note: Percentages may not sum to 100.00 due to rounding.

The cases that matched to an MTdb record and could not be geocoded were sent to clerical processing for another attempt at being geocoded.

The detailed results for the Type A cases that were neither matched nor geocoded during automated processing are shown below in Table 5.2.6.

Table 5.2.6 Type A Cases Not Matched and Not Geocoded Through Automated Processing

Category		Number of Cases per Category	Percent of Total per Category
Source	Be Counted	146,026	71.05
	TQA Fulfillment	6,007	2.92
	TQA Interview	13,880	6.75
	GQE	11,944	5.81
	NRFU	27,120	13.20
	RA/RUE/UE	547	0.27
Location	Stateside	200,115	97.37
	Puerto Rico	5,409	2.63
Milestone	Initial 2010 FV Cutoff	174,647	84.98
	2010 FV Supplemental Cutoff	30,714	14.94
	After 2010 FV Cutoff	163	0.08
Total Cases		**205,524**	**100.00**

Source: CRDA file.

The cases that were unmatched and ungeocoded as a result of automated processing were sent to clerical processing for further resolution.

Table 5.2.7 below shows the detailed results for the Type A cases rejected during the automated processing. A Type A case was rejected for one of the following reasons:

- Incomplete address (194,741 cases (98.96 percent))
- Illegal or missing value in the CRDA (2,033 cases (1.03 percent))
- Illegal ZIP code for Puerto Rico (6 cases (0.00 percent))

Table 5.2.7 Type A Cases Rejected During Automated Processing

Category		Number of Cases per Category	Percent of Total per Category
Source	Be Counted	69,654	35.40
	TQA Fulfillment	114,908	58.39
	TQA Interview	2,993	1.52
	GQE	5,085	2.58
	NRFU	4,052	2.06
	RA/RUE/UE	88	0.04
Location	Stateside	191,571	97.35
	Puerto Rico	5,209	2.65
Milestone	Initial 2010 FV Cutoff	191,372	97.25
	2010 FV Supplemental Cutoff	5,378	2.73
	After 2010 FV Cutoff	30	0.02
Total Cases		**196,780**	**100.00**

Source: CRDA file.

The high number of rejected cases for TQA Fulfillment can be attributed to the cases delivered to automated processing without the street name due to a data mapping issue between UCM and the ADDUP file (as described in Section 2.3.2).

The cases rejected during automated processing were sent to clerical processing for further resolution.

Table 5.2.8 presents the results of the automated matching and geocoding process for Type A cases, summarized by source operation.

Table 5.2.8 Automated Processing Results for Type A Cases by Source

Source	Number of Cases	Matched and Geocoded	Percent of Total	Unmatched and Geocoded	Percent of Total	Matched and Ungeocoded	Percent of Total	Unmatched and Ungeocoded	Percent of Total	Rejected	Percent of Total
Be Counted	**743,423**	414,108	55.70	41,706	5.61	71,929	9.68	146,026	19.64	69,654	9.37
TQA Fulfillment	**198,843**	54,955	27.64	2,148	1.08	20,825	10.47	6,007	3.02	114,908	57.79
TQA Interview	**110,525**	74,088	67.03	4,202	3.80	15,362	13.90	13,880	12.56	2,993	2.71
GQE	**50,032**	29,140	58.24	3,090	6.18	773	1.55	11,944	23.87	5,085	10.16
NRFU	**127,787**	83,062	65.00	7,877	6.16	5,676	4.44	27,120	21.22	4,052	3.17
RA/RUE/UE	**1,511**	679	44.94	90	5.96	107	7.08	547	36.20	88	5.82
Total Cases	**1,232,121**	**656,032**	**53.24**	**59,113**	**4.80**	**114,672**	**9.31**	**205,524**	**16.68**	**196,780**	**15.97**

Source: CRDA file. Note: Percentages may not sum to 100.00 due to rounding.

Of the 1,232,121 Type A cases that were successfully header coded and proceeded to automated matching and geocoding, there were 516,976 cases not resolved and sent to clerical processing.

The 59,113 cases that were unmatched and geocoded were sent to 2010 FV, along with 13,878 cases that matched to an ungeocoded MTdb record but obtained a geocode via the automated geocoding routine.

Type B Cases

Of the 13,427 Type B cases sent through the automated process, 11,910 (88.70 percent) were successfully header coded; which means those cases were complete and later added to the GQ universe and allocated to census counts accordingly. The remaining 1,517 (11.30 percent) cases not successfully header coded were sent to clerical processing for resolution.

The detailed results for the cases header coded through automated processing are displayed in Table 5.2.9.

Table 5.2.9 Type B Cases Header Coded Through Automated Processing

Category		Number of Cases per Category	Percent of Total per Category
Source	Be Counted	11,910	100.00
Location	Stateside	11,794	99.03
	Puerto Rico	116	0.97
Milestone	Initial FV Cutoff	11,768	98.81
	FV Supplemental Cutoff	131	1.10
	After FV Cutoff	11	0.09
Total Cases		**11,910**	**100.00**

Source: CRDA file.

Type C Cases

Of the 1,608,779 Type C cases sent through the automated process, 1,586,721 (98.63 percent) cases had the geocode supplied by the enumerator accepted and were successfully associated with a state and county. Table 5.2.10 presents the detailed results of automated processing for those cases associated with a state.

Table 5.2.10 Type C Cases Associated With a State/County Through Automated Processing

Category		Number of Cases per Category	Percent of Total per Category
Source	ETL	120,839	7.62
	GQE	18,818	1.19
	NRFU	620,446	39.10
	VDC	219,292	13.82
	RA/RUE/UE	72,488	4.57
	U/L	534,838	33.71
Location	Stateside	1,450,189	91.40
	Puerto Rico	136,532	8.60
Total Cases		**1,586,721**	**100.00**

Source: Assessment files provided by GEO for each operation. Note: Percentages may not sum to 100 due to rounding.

The number of Type C cases not successfully associated with a state and county during the automated process was 22,058 (1.37 percent).

Table 5.2.11 presents the results of automated processing for the 1,586,721 Type C cases successfully associated with a state. At the end of automated processing, a case was added to the MTdb as a new record, matched to an existing MTdb record, or rejected from processing.

Table 5.2.11 Automated Processing Results for Type C Cases

Category	Number of Cases	Percent of Total
Added	967,842	61.00
Matched	588,421	37.08
Rejected	30,458	1.92
Total Cases	**1,586,721**	**100.00**

Source: Assessment files provided by GEO for each operation.

Table 5.2.12 through Table 5.2.14 present the detailed results for each category listed in Table 5.2.11. For each table, the results for the NRFU and VDC cases reflect the final results after additional processing was completed.

The detailed results for the Type C cases added as new MTdb records during automated processing are shown below in Table 5.2.12. These cases represent new addresses added to the census universe.

Table 5.2.12 Type C Cases Added as New MTdb Records During Automated Processing

Category		Number of Cases per Category	Percent of Total per Category
Source	ETL	92,781	9.59
	GQE	17,171	1.77
	NRFU	280,263	28.96
	VDC	77,242	7.98
	RA/RUE/UE	66,880	6.91
	U/L	433,505	44.79
Location	Stateside	841,433	86.94
	Puerto Rico	126,409	13.06
Total Cases		**967,842**	**100.00**

Source: Assessment files provided by GEO for each operation.

The detailed results for the Type C cases matched to an existing MTdb record during automated processing are shown below in Table 5.2.13.

Table 5.2.13 Type C Cases Matched During Automated Processing

Category		Number of Cases per Category	Percent of Total per Category
Source	ETL	27,231	4.63
	GQE	1,625	0.28
	NRFU	328,393	55.81
	VDC	139,697	23.74
	RA/RUE/UE	5,256	0.89
	U/L	86,219	14.65
Location	Stateside	588,150	99.95
	Puerto Rico	271	0.05
Total Cases		**588,421**	**100.00**

Source: Assessment files provided by GEO for each operation.

The detailed results for the Type C cases rejected during automated processing are shown below in Table 5.2.14. A Type C case was rejected for one of the following reasons:

- Illegal or missing value (12,696 cases (41.68 percent))
- Illegal block geocode (4,536 cases (14.89 percent))

- Duplicate Customer ID[13] (9,955 cases (32.68 percent))
- Action code of 'Z'[14] (3,271 cases (10.74 percent))

Table 5.2.14 Type C Cases Rejected During Automated Processing

Category		Number of Cases per Category	Percent of Total per Category
Source	ETL	827	2.43
	GQE	22	0.04
	NRFU	11,790	40.98
	VDC	2,353	26.05
	RA/RUE/UE	352	1.72
	U/L	15,114	28.78
Location	Stateside	20,606	67.65
	Puerto Rico	9,852	32.35
Total Cases		**30,458**	**100.00**

Source: Assessment files provided by GEO for each operation.

There were 57,342 NRFU Type C cases and 47,802 VDC Type C cases rejected for incomplete or missing block geocode during their initial run through the automated process. Due to the large number of rejects and the fact that this was the last chance for these addresses to get into the census universe, these cases were sent through the automated geocoding software. The cases that were still unable to be geocoded after going through the automated process were sent to the clerical operation at NPC in a final attempt to obtain a geocode. The results of the automated geocoding cannot be determined because the assessment tally file does not distinguish the results between the automated process and the clerical process. The combined results are presented in section 5.3.

5.3 How many cases were sent to clerical Non-ID Processing?

Type A Cases

A total of 550,066 Type A cases were not resolved through the automated process and were sent to the clerical process. The categories from automated processing that the clerical cases originated from are shown in Table 5.3.1. The cases in the not matched and not geocoded through automated processing category contributed the most (37.34 percent), but the cases contributed by the rejected during automated processing category were a close second with 35.75 percent.

[13] A Customer ID was a number assigned sequentially to every record for tracking purposes in the files delivered to GEO for Non-ID Processing.

[14] A 'Z' action code was assigned to Other Living Quarters and preprinted on the listing pages for enumerators to assist in the identification of HUs and were not action codes returned by the enumerators. GEO was instructed to reject cases with a 'Z' action code during MTdb updating.

Table 5.3.1 Type A Cases Requiring Clerical Processing

Category	Number of Cases	Percent of Total
Cases Unable to be Header Coded Through Automated Processing	33,090	6.07
Cases Matched But Not Geocoded Through Automated Processing	114,672	20.83
Cases Not Matched and Not Geocoded Through Automated Processing	205,524	37.34
Cases Rejected During Automated Processing	196,780	35.75
Total Cases	**550,066**	**100.00**

Source: CRDA file. Note: Percentages may not sum to 100.00 due to rounding.

The details of the Type A cases delivered to clerical processing are shown below in Table 5.3.2. The ranking of source operations based on the number of cases delivered to clerical processing is the same for the clerical process as it was for the automated process. Most cases resulted from BC (56.66 percent), followed by TQA Fulfillment, NFRU, TQA Interview, GQE, and RA/RUE/UE.

Table 5.3.2 Type A Cases Delivered to Clerical Processing

Category		Number of Cases per Category	Percent of Total per Category
Source	Be Counted	311,642	56.66
	TQA Fulfillment	143,600	26.11
	TQA Interview	34,368	6.25
	GQE	18,688	3.40
	NRFU	40,554	7.37
	RA/RUE/UE	1,214	0.22
Location*	Stateside	524,243	95.31
	Puerto Rico	12,954	2.35
Milestone	Initial 2010 FV Cutoff	499,991	90.90
	2010 FV Supplemental Cutoff	49,782	9.05
	After 2010 FV Cutoff	293	0.05
Total Cases		**550,066**	**100.00**

Source: CRDA file. Note: Percentages may not sum to 100.00 due to rounding.
*The stateside and Puerto Rico cases do not add up to 550,066 nor 100.00 percent because the remaining cases were either coded to a Foreign Country (78 cases) or were not assigned to a State/County (13,425 cases) after they went through clerical processing.

Type B Cases

A total of 1,517 Type B cases were not resolved through the automated process and were sent to the clerical process. The details of the Type B cases delivered to clerical processing are shown below in Table 5.3.3.

Table 5.3.3 Type B Cases Delivered to Clerical Processing

Category		Number of Cases per Category	Percent of Total per Category
Source	Be Counted	1,517	100.00
Location*	Stateside	835	55.04
	Puerto Rico	29	1.91
Milestone	Initial FV Cutoff	1,503	99.08
	FV Supplemental Cutoff	12	0.79
	After FV Cutoff	2	0.13
Total Cases		**1,517**	**100.00**

Source: CRDA file. *The stateside and Puerto Rico cases do not add up to 1,517 nor 100.00 percent because the remaining cases were not assigned to a State/County (1,208 cases) after they went through clerical processing.

Type C Cases

Normally, all ungeocoded Type C cases are automatically rejected and go no further in the operation because the cases are supposed to be delivered with a complete geocode. The exception for the 2010 Census was the rejected NRFU and VDC cases. Due to the large number of rejected NRFU/VDC Type C cases and the fact that this was the last chance for these addresses to get into the census universe, these cases were sent through the automated geocoding software. There were 35,213 cases sent to the clerical operation at NPC in a final attempt to obtain a geocode because the automated geocoding process was not successful.

5.4 What was the outcome of clerical Non-ID Processing (header coding, matching, and geocoding)? How many cases were successfully resolved via telephone call?

Type A Cases

Of the 33,090 Type A cases not assigned to a state and county during automated processing, 19,587 (59.19 percent) cases were successfully header coded during the clerical process and proceeded to clerical matching and geocoding. Table 5.4.1 presents the detailed results of clerical header coding.

Table 5.4.1 Type A Cases Header Coded Through Clerical Processing

Category		Number of Cases per Category	Percent of Total per Category
Source	Be Counted	14,336	73.19
	TQA Fulfillment	1,834	9.36
	TQA Interview	2,078	10.61
	GQE	567	2.89
	NRFU	762	3.89
	RA/RUE/UE	10	0.05
Location	Stateside	17,556	89.63
	Puerto Rico	2,031	10.37
Milestone	Initial 2010 FV Cutoff	18,394	93.91
	2010 FV Supplemental Cutoff	1175	6.00
	After 2010 FV Cutoff	18	0.09
Total Cases		**19,587**	**100.00**

Source: CRDA file.

The number of cases not successfully header coded during the clerical process and deemed uncodable was 13,503 (40.81 percent). This total includes 13,425 cases not assigned to a state and county and 78 cases assigned to a foreign country.

The 19,587 Type A cases that were header coded proceeded to clerical matching and geocoding with the 516,976 cases not resolved during automated processing, for a total of 536,563 cases. Table 5.4.2 presents the results of the Type A cases sent to clerical processing.

Table 5.4.2 Clerical Processing Results for Type A Cases

Category	Number of Cases	Percent of Total
Matched to a Geocoded MTdb Record	178,358	33.24
Matched During Automated Processing/ Clerically Geocoded	90,800	16.92
Matched During Automated Processing/ Unable to Clerically Geocode	23,870	4.45
Clerically Unmatched/Geocoded	124,770	23.25
Clerically Unmatched/Unable to Geocode	118,765	22.13
Total Cases	**536,563**	**100.00**

Source: CRDA file. Note: Percentages may not sum to 100.00 due to rounding.

Table 5.4.3 through Table 5.4.7 present the detailed results for each category listed in Table 5.4.2.

The detailed results for the Type A cases matched to a geocoded MTdb record during clerical processing are shown below in Table 5.4.3. All addresses in the MTdb extract used during clerical processing were geocoded; therefore any case that resulted in a match was also automatically geocoded.

Table 5.4.3 Type A Cases Matched to a Geocoded MTdb Record During Clerical Processing

Category		Number of Cases per Category	Percent of Total per Category
Source	Be Counted	95,650	53.63
	TQA Fulfillment	53,400	29.94
	TQA Interview	8,215	4.61
	GQE	6,266	3.51
	NRFU	14,699	8.24
	RA/RUE/UE	128	0.07
Location	Stateside	173,719	97.40
	Puerto Rico	4,639	2.60
Milestone	Initial 2010 FV Cutoff	161,723	90.67
	2010 FV Supplemental Cutoff	16,568	9.29
	After 2010 FV Cutoff	67	0.04
Total Cases		**178,358**	**100.00**

Source: CRDA file.

The detailed results for the Type A cases matched during automated processing and clerically geocoded are shown below in Table 5.4.4. These cases were matched to an ungeocoded MTdb record during the automated process and were sent to clerical processing to be geocoded.

Table 5.4.4 Type A Cases Matched During Automated Processing/Clerically Geocoded

Category		Number of Cases per Category	Percent of Total per Category
Source	Be Counted	57,163	62.95
	TQA Fulfillment	16,370	18.03
	TQA Interview	12,152	13.38
	GQE	535	0.59
	NRFU	4,495	4.95
	RA/RUE/UE	85	0.09
Location	Stateside	90,800	100.00
	Puerto Rico	0	0.00
Milestone	Initial 2010 FV Cutoff	83,283	91.72
	2010 FV Supplemental Cutoff	7,468	8.22
	After 2010 FV Cutoff	49	0.05
Total Cases		**90,800**	**100.00**

Source: CRDA file. Note: Percentages may not sum to 100.00 due to rounding.

The detailed results for the Type A cases matched during automated processing and clerically geocoded are shown below in Table 5.4.5. These cases were matched to an ungeocoded MTdb record during the automated process and were sent to clerical processing to be geocoded, but a geocode was unable to be obtained.

Table 5.4.5 Type A Cases Matched During Automated Processing/ Unable to Clerically Geocode

Category		Number of Cases per Category	Percent of Total per Category
Source	Be Counted	14,765	61.86
	TQA Fulfillment	4,455	18.66
	TQA Interview	3,210	13.45
	GQE	237	0.99
	NRFU	1,181	4.95
	RA/RUE/UE	22	0.09
Location	Stateside	23,870	100.00
	Puerto Rico	0	0.00
Milestone	Initial 2010 FV Cutoff	21,884	91.68
	2010 FV Supplemental Cutoff	1,972	8.26
	After 2010 FV Cutoff	14	0.06
Total Cases		**23,870**	**100.00**

Source: CRDA file.

The detailed results for the Type A cases clerically unmatched but geocoded are shown below in Table 5.4.6. These cases were not able to be matched in either the automated process or the clerical process, but a geocode was obtained during clerical processing.

Table 5.4.6 Type A Cases Clerically Unmatched/Geocoded

Category		Number of Cases per Category	Percent of Total per Category
Source	Be Counted	70,434	56.45
	TQA Fulfillment	39,268	31.47
	TQA Interview	5,766	4.62
	GQE	2,260	1.81
	NRFU	6,769	5.43
	RA/RUE/UE	273	0.22
Location	Stateside	120,332	96.44
	Puerto Rico	4,438	3.56
Milestone	Initial 2010 FV Cutoff	116,353	93.25
	2010 FV Supplemental Cutoff	8,368	6.71
	After 2010 FV Cutoff	49	0.04
Total Cases		**124,770**	**100.00**

Source: CRDA file.

The detailed results for the Type A cases clerically unmatched and ungeocoded are shown below in Table 5.4.7. Neither a match nor a geocode could be obtained for these cases during clerical processing.

Table 5.4.7 Type A Cases Clerically Unmatched/Unable to Geocode

Category		Number of Cases per Category	Percent of Total per Category
Source	Be Counted	63,933	53.83
	TQA Fulfillment	30,081	25.33
	TQA Interview	4,970	4.18
	GQE	9,072	7.64
	NRFU	10,465	8.81
	RA/RUE/UE	244	0.21
Location	Stateside	115,193	96.99
	Puerto Rico	3,572	3.01
Milestone	Initial 2010 FV Cutoff	106,298	89.50
	2010 FV Supplemental Cutoff	12,372	10.42
	After 2010 FV Cutoff	95	0.08
Total Cases		**118,765**	**100.00**

Source: CRDA file.

A telephone call was attempted for 199,603 cases during clerical processing. Of the 360,969 stateside referral cases, telephone calls were attempted for 194,807 (53.97 percent) cases. Of the 12,983 Puerto Rico clerical cases, telephone calls were attempted for 4,796 (36.94 percent) cases. Table 5.4.8 displays the characteristics of the cases where a telephone call was attempted. The results of the telephone calls are displayed in Table 5.4.9.

Table 5.4.8 Type A Cases Where a Telephone Call Was Attempted

Category		Number of Cases per Category	Percent of Total per Category
Source	Be Counted	113,610	56.92
	TQA Fulfillment	54,156	27.13
	TQA Interview	13,843	6.94
	GQE	1,848	0.93
	NRFU	15,759	7.90
	RA/RUE/UE	387	0.19
Location	Stateside	194,807	97.60
	Puerto Rico	4,796	2.40
Milestone	Initial 2010 FV Cutoff	180,879	90.62
	2010 FV Supplemental Cutoff	18,597	9.32
	After 2010 FV Cutoff	127	0.06
Total Cases		**199,603**	**100.00**

Source: CRDA file.

Table 5.4.9 Results of Telephone Calls

Category	Number of Cases	Percent of Total
Corrected Address Information and Obtained Geocode	61,842	30.98
Obtained Geocode with No Additional Address Information	23,328	11.69
No Additional Address Information	6,532	3.27
Hang Up/Hostile/Refusal	13,776	6.90
Bad Connection	637	0.32
No Answer	17,786	8.91
Invalid Phone Number	14,205	7.12
Busy Signal	2,430	1.22
Answering Machine	59,067	29.59
Total Cases	**199,603**	**100.00**

Source: CRDA file.

Of the 61,842 calls that resulted in corrected information and a geocode, 27,172 cases were matched to a geocoded MTdb record and 34,286 cases were unmatched and geocoded. For the 23,328 calls that resulted in a geocode and obtained no additional address information, 2,038 matched to a geocoded MTdb record and 21,124 cases were unmatched and geocoded. The

remaining 384 cases flagged as “Corrected Address Information and Obtained Geocode” and the remaining 166 cases flagged as “Obtained Geocode with No Additional Address Information” were not clerically geocoded and were likely assigned the incorrect telephone resolution type by the clerk. For the remaining 114,433 calls that were attempted, either the respondent was not reached or provided no information to help resolve the case.

Table 5.4.10 displays the summarized results of clerical processing for Type A cases by source operation.

Table 5.4.10 Clerical Processing Results for Type A Cases by Source

Source	Number of Cases	Matched to Geocoded MTdb Record	Percent of Total	Matched During Automated Processing/ Clerically Geocoded	Percent of Total	Matched During Automated Processing/ Unable to Clerically Geocode	Percent of Total	Clerically Unmatched/ Geocoded	Percent of Total	Clerically Unmatched/ Unable to Geocode	Percent of Total
Be Counted	**301,945**	95,650	31.68	57,163	18.93	14,765	4.89	70,434	23.33	63,933	21.17
TQA Fulfillment	**143,574**	53,400	37.19	16,370	11.40	4,455	3.10	39,268	27.35	30,081	20.95
TQA Interview	**34,313**	8,215	23.94	12,152	35.42	3,210	9.36	5,766	16.80	4,970	14.48
GQE	**18,370**	6,266	34.11	535	2.91	237	1.29	2,260	12.30	9,072	49.38
NRFU	**37,609**	14,699	39.08	4,495	11.95	1,181	3.14	6,769	18.00	10,465	27.83
RA/RUE/UE	**752**	128	17.02	85	11.30	22	2.93	273	36.30	244	32.45
Total Cases	**536,563**	**178,358**	**33.24**	**90,800**	**16.92**	**23,870**	**4.45**	**124,770**	**23.25**	**118,765**	**22.13**

Source: CRDA file. Note: Percentages may not sum to 100 due to rounding.

The 90,800 cases matched during automated processing and clerically geocoded were sent to 2010 FV. The 124,770 cases clerically unmatched but geocoded were also sent to 2010 FV.

At the end of clerical processing, any Type A case that was successfully header coded but still required a match, geocode, or both proceeded to post-clerical processing.

Type B Cases

Of the 1,517 Type B cases not assigned to a state and county during automated processing, 309 (20.37 percent) cases were successfully header coded during the clerical process and were considered complete. Table 5.4.11 presents the detailed results of clerical header coding.

Table 5.4.11 Type B Cases Header Coded Through Clerical Processing

Category		Number of Cases per Category	Percent of Total per Category
Source	Be Counted	309	100.00
Location	Stateside	285	92.23
	Puerto Rico	24	7.77
Milestone	Initial 2010 FV Cutoff	306	99.03
	2010 FV Supplemental Cutoff	3	0.97
	After 2010 FV Cutoff	0	0.00
Total Cases		**309**	**100.00**

Source: CRDA file.

The remaining 1,208 (79.63 percent) cases not header coded during clerical processing were considered uncodable and did not advance further in the process.

Type C Cases

A combined 105,144 rejected NRFU/VDC Type C cases were sent through automated and clerical processing in an attempt to obtain a geocode. There were 69,931 (66.51 percent) cases resolved in the automated process, leaving 35,213 (33.49 percent) cases sent to the clerical process.

The tally files for the NRFU and VDC Type C cases contain a combination of the results of the automated process and the clerical process and do not distinguish between the two processes. Therefore, Table 5.4.12 displays the combined results of automated and clerical processing for the NRFU/VDC Type C cases initially rejected during processing. At the end of automated and clerical processing, a case was added to the MTdb as a new record, matched to an existing MTdb record, or still classified as a reject due to a missing or incomplete block geocode.

Table 5.4.12 Automated and Clerical Processing Results for NRFU/ VDC Type C Cases by Source

Source	Number of Cases	Added	Percent of Total	Matched	Percent of Total	Rejected	Percent of Total
NRFU	57,342	14,458	25.21	25,757	44.92	17,127	29.87
VDC	47,802	8,934	18.69	25,591	53.54	13,277	27.77
Total Cases	105,144	23,392	22.25	51,348	48.84	30,404	28.92

Source: Assessment files provided by GEO for each operation.

For both NRFU and VDC, the majority of the rejected cases matched to a preexisting MTdb record. After all the automated and clerical processing had completed, an additional 74,740 addresses (added or matched) had been accepted into the census universe. The remaining 30,404 cases were still rejected during processing due to an incomplete or missing block geocode.

5.5 What was the outcome of post-clerical Non-ID Processing (matching)?

Post-clerical processing was a stage of Non-ID Processing only for Type A cases that, upon completion of the clerical process, still required a match, geocode, or both. The cases also had to have been previously header coded to a state and county in order to qualify for post-clerical processing.

A total of 267,709 Type A cases still not fully matched and geocoded were sent to post-clerical processing. There were three categories of cases requiring a match, a geocode, or both:

- Automated unmatched/ungeocoded, clerical unmatched/ungeocoded (119,069 cases, 44.48 percent)
- Automated matched/ungeocoded, clerical unmatched/ungeocoded (23,870 cases, 8.92 percent)
- Automated unmatched/ungeocoded, clerical unmatched/geocoded (124,770 cases, 46.61 percent).

Table 5.5.1 presents the results of post-clerical processing for Type A cases.

Table 5.5.1 Post-Clerical Processing Results for Type A Cases

Category	Number of Cases	Percent of Total
Matched to a Geocoded MTdb Record	33,586	12.55
Matched to a Ungeocoded MTdb Record	58,240	21.75
Unmatched to a MTdb Record	106,978	39.96
Rejected	68,905	25.74
Total Cases	**267,709**	**100.00**

Source: CRDA File.

Table 5.5.2 through Table 5.5.5 present the detailed results for each category listed in Table 5.5.1.

The detailed results for the Type A cases matched to a geocoded MTdb record during post-clerical processing are shown below in Table 5.5.2. These cases may have matched to new addresses from a DSF update or new addresses added from operations that occurred after the MTdb extract was created for clerical processing. If a case entered post-clerical processing with a geocode and matched a geocoded MTdb record, the geocode already associated with the case was retained.

Table 5.5.2 Type A Cases Matched to a Geocoded MTdb Record During Post-Clerical Processing

Category		Number of Cases per Category	Percent of Total per Category
Automated Unmatched/Ungeocoded, Clerical Unmatched/Ungeocoded		3,464	10.31
Automated Matched/Ungeocoded, Clerical Unmatched/Ungeocoded		664	1.98
Automated Unmatched/Ungeocoded, Clerical Unmatched/Geocoded		29,458	87.71
Source	Be Counted	16,033	47.74
	TQA Fulfillment	14,286	42.54
	TQA Interview	1,208	3.60
	GQE	732	2.18
	NRFU	1,312	3.91
	RA/RUE/UE	15	0.04
Location	Stateside	33,570	99.95
	Puerto Rico	16	0.05
Milestone	Initial 2010 FV Cutoff	31,905	94.99
	2010 FV Supplemental Cutoff	1,668	4.97
	After 2010 FV Cutoff	13	0.04
Total Cases		**33,586**	**100.00**

Source: CRDA File. Note: Percentages may not sum to 100.00 due to rounding.

The detailed results for the Type A cases matched to an ungeocoded MTdb record during post-clerical processing are shown below in Table 5.5.3. Although having a match without a geocode is insufficient for the address to be included in the census universe, the match proves the existence of the address and that it may be considered for future census operations. The unmatched and clerically geocoded cases that matched to an ungeocoded MTdb record during post-clerical processing were then considered matched and geocoded.

Table 5.5.3 Type A Cases Matched to an Ungeocoded MTdb Record During Post-Clerical Processing

Category		Number of Cases per Category	Percent of Total per Category
Automated Unmatched/Ungeocoded, Clerical Unmatched/Ungeocoded		5,037	8.65
Automated Matched/Ungeocoded, Clerical Unmatched/Ungeocoded		22,685	38.95
Automated Unmatched/Ungeocoded, Clerical Unmatched/Geocoded		30,518	52.40
Source	Be Counted	24,757	42.51
	TQA Fulfillment	27,854	47.83
	TQA Interview	4,006	6.88
	GQE	200	0.34
	NRFU	1,401	2.41
	RA/RUE/UE	22	0.04
Location	Stateside	58,234	99.99
	Puerto Rico	6	0.01
Milestone	Initial 2010 FV Cutoff	55,838	95.88
	2010 FV Supplemental Cutoff	2,378	4.08
	After 2010 FV Cutoff	24	0.04
Total Cases		**58,240**	**100.00**

Source: CRDA File. Note: Percentages may not sum to 100.00 due to rounding.

The detailed results for the Type A cases not matched to a MTdb record during post-clerical processing are shown below in Table 5.5.4.

Table 5.5.4 Type A Cases Unmatched to a MTdb Record During Post-Clerical Processing

Category		Number of Cases per Category	Percent of Total per Category
Automated Unmatched/Ungeocoded, Clerical Unmatched/Ungeocoded		51,395	48.04
Automated Matched/Ungeocoded, Clerical Unmatched/Ungeocoded		105	0.10
Automated Unmatched/Ungeocoded, Clerical Unmatched/Geocoded		55,478	51.86
Source	Be Counted	72,778	68.03
	TQA Fulfillment	7,751	7.25
	TQA Interview	6,577	6.15
	GQE	6,362	5.95
	NRFU	13,091	12.24
	RA/RUE/UE	419	0.39
Location	Stateside	103,948	97.17
	Puerto Rico	3,030	2.83
Milestone	Initial 2010 FV Cutoff	91,738	85.75
	2010 FV Supplemental Cutoff	15,138	14.15
	After 2010 FV Cutoff	102	0.10
Total Cases		**106,978**	**100.00**

Source: CRDA File. Note: Percentages may not sum to 100.00 due to rounding.

The detailed results for the Type A cases rejected during post-clerical processing are shown in Table 5.5.5. A Type A case was rejected for one of the following reasons:

- Incomplete address (59,221 cases (85.95 percent))
- Not valid for matching (2,007 cases (2.91 percent))
- Illegal or missing value in the CRDA (7,677 cases (11.14 percent))
- Illegal block (0 cases (0.00 percent))

Table 5.5.5 Type A Cases Rejected During Post-Clerical Processing

Category		Number of Cases per Category	Percent of Total per Category
Automated Unmatched/Ungeocoded, Clerical Unmatched/Ungeocoded		59,173	85.88
Automated Matched/Ungeocoded, Clerical Unmatched/Ungeocoded		416	0.60
Automated Unmatched/Ungeocoded, Clerical Unmatched/Geocoded		9,316	13.52
Source	Be Counted	35,831	52.00
	TQA Fulfillment	23,913	34.70
	TQA Interview	2,155	3.13
	GQE	4,275	6.20
	NRFU	2,648	3.84
	RA/RUE/UE	83	0.12
Location	Stateside	63,642	92.36
	Puerto Rico	5,263	7.64
Milestone	Initial 2010 FV Cutoff	65,356	94.85
	2010 FV Supplemental Cutoff	3,530	5.12
	After 2010 FV Cutoff	19	0.03
Total Cases		**68,905**	**100.00**

Source: CRDA File. Note: Percentages may not sum to 100.00 due to rounding.

Table 5.5.6 displays the summarized results of post-clerical processing for Type A cases by source operation.

Table 5.5.6 Post-Clerical Processing Results for Type A Cases by Source

Source	Number of Cases	Matched to Geocoded MTdb Record	Percent of Total	Matched to Ungeocoded MTdb Record	Percent of Total	Unmatched to a MTdb Record	Percent of Total	Rejected	Percent of Total
Be Counted	**149,399**	16,033	10.73	24,757	16.57	72,778	48.71	35,831	23.98
TQA Fulfillment	**73,804**	14,286	19.36	27,854	37.74	7,751	10.50	23,913	32.40
TQA Interview	**13,946**	1,208	8.66	4,006	28.73	6,577	47.16	2,155	15.45
GQE	**11,569**	732	6.33	200	1.73	6,362	54.99	4,275	36.95
NRFU	**18,452**	1,312	7.11	1,401	7.59	13,091	70.95	2,648	14.35
RA/RUE/UE	**539**	15	2.78	22	4.08	419	77.74	83	15.40
Total Cases	**267,709**	**33,586**	**12.55**	**58,240**	**21.75**	**106,978**	**39.96**	**68,905**	**25.74**

Source: CRDA file. Note: Percentages may not sum to 100.00 due to rounding.

Another 33,586 cases were matched and geocoded during post-clerical processing, allowing those addresses to be accepted into the census universe.

5.6 What were the actions taken on the records in the MTdb as a result of Non-ID Processing (automated, clerical, and post-clerical)?

Type A Cases

Table 5.6.1 displays the cumulative results of Non-ID Processing for the Type A cases.

Table 5.6.1 Cumulative Results for Type A Cases by Source

Source	Number of Cases Matched and Geocoded	Percent of Total	Number of Cases Matched and Ungeocoded	Percent of Total	Number of Cases Unmatched and Geocoded	Percent of Total	Number of Cases Unmatched and Ungeocoded*	Percent of Total	Total Cases (By Source)
Be Counted	592,696	77.25	15,273	1.99	88,010	11.47	71,225	9.28	**767,204**
TQA Fulfillment	158,676	79.06	8,351	4.16	9,426	4.70	24,251	12.08	**200,704**
TQA Interview	96,454	85.62	3,274	2.91	8,098	7.19	4,832	4.29	**112,658**
GQE	36,713	72.15	162	0.32	4,748	9.33	9,258	18.20	**50,881**
NRFU	103,834	78.95	1,160	0.88	13,275	10.09	13,248	10.07	**131,517**
RA/RUE/UE	911	45.92	20	1.01	258	13.00	795	40.07	**1,984**
Total Cases (By Category)	**989,284**	**78.21**	**28,240**	**2.23**	**123,815**	**9.79**	**123,609**	**9.77**	**1,264,948**

Source: CRDA File. Note: Percentages may not sum to 100.00 due to rounding. *The "Number of Cases Unmatched and Ungeocoded" category includes cases that did not header code or were header coded to a foreign country.

The vast majority of Type A cases (78.21 percent) was matched and geocoded, which indicates that most of the addresses received for Non-ID Processing were already in the MTdb. At the source level, less than half of the RA/RUE/UE cases were matched and geocoded. The percentage of unmatched and ungeocoded RA/RUE/UE cases was nearly as high as the percentage of matched and geocoded. The 123,815 cases that were unmatched and geocoded had new MTdb records created and were included in the census

universe if they were validated in 2010 FV. The number of cases matched and geocoded combined with the number of cases unmatched and geocoded means over 1.1 million cases were successfully resolved and had their addresses included in the census universe.

The matched and ungeocoded cases were not included in the census universe, but the match proves the existence of the address and future operations may obtain a geocode. Less than 10 percent of all Type A cases were unmatched and ungeocoded, meaning the addresses and associated response data were not included in the census universe.

Type B Cases

Table 5.6.2 displays the cumulative results of Non-ID Processing for the Type B cases. The Type A cases switched to Type B cases are included in this table. Although BC is supposed to be the only source of Type B cases, because of the switch from Type A to Type B, some of the Type B cases originated from TQA Interview and GQE.

Table 5.6.2 Cumulative Results for Type B Cases by Source

Source	Number of Cases Header Coded	Percent of Total	Number of Cases Not Header Coded	Percent of Total	Total Cases (By Source)
Be Counted	12,459	90.88	1,251	9.12	13,710
TQA Interview	4	100.00	0	0.00	4
GQE	310	98.10	6	1.90	316
Total Cases (By Category)	**12,773**	**91.04**	**1,257**	**8.96**	**14,030**

Source: CRDA File. *The "Number of Cases Not Header Coded" category includes cases that did not header code or were header coded to a foreign country.

Over 91 percent of the Type B cases were successfully header coded, meaning the cases were assigned to a state and county. These cases were later added to the GQ universe and allocated to census counts accordingly.

Type C Cases

Table 5.6.3 displays the cumulative results of Non-ID Processing for the Type C cases.

Table 5.6.3 Cumulative Results for Type C Cases by Source

Source	Number of Cases Added	Percent of Total	Number of Cases Matched	Percent of Total	Number of Cases Rejected*	Percent of Total	Total Cases (By Source)
ETL	92,781	76.50	27,231	22.45	1,278	1.05	121,290
GQE	17,171	91.25	1,625	8.64	22	0.12	18,818
NRFU	280,263	44.47	328,393	52.11	21,519	3.41	630,175
VDC	77,242	33.49	139,697	60.57	13,680	5.93	230,619
RA/RUE/UE	66,880	91.57	5,256	7.20	903	1.24	73,039
U/L	433,505	81.05	86,219	16.12	15,114	2.83	534,838
Total Cases (By Category)	**967,842**	**60.16**	**588,421**	**36.58**	**52,516**	**3.26**	**1,608,779**

Source: Assessment tally files provided by GEO for each operation. Note: Percentages may not sum to 100.00 due to rounding.
*The "Number of Cases Rejected" category includes cases that were not associated with a state.

Overall, the majority of Type C cases (60.16 percent) was added as new MTdb records. At the source level, the majority of cases for NRFU and VDC were matched to preexisting MTdb records, indicating that most of the 'adds' were addresses already in the MTdb. The number of cases added to the MTdb as new records combined with the number of cases matched to preexisting MTdb records means more than 1.5 million addresses were successfully processed and were included in the census universe.

5.7 How did the actual cost of the 2010 Census Non-ID Processing operation compare to the expected budget?

In FY08, the expenditure for computer software and hardware for the Non-ID Processing servers at NPC was approximately $588,000, or about 98 percent of the funds added after the unfunded requirement was approved. GEO's entire FY09 budget of 1.5 million dollars was expended on contract funding, and NPC's only allocation for FY09 involved the testing conducted in July 2009. The project code used in 2009 for this work was also used for a significant amount of other work, so there is no way to extract the actual cost for the clerical software testing. The combined FY10 cost for GEO and NPC was only about half of their collective budget; details for each division's expenditures in FY10 are provided in Table 5.7.1.

GEO spent approximately two-thirds of its allotment for FY10. GEO's cost for FastData was $102,486, about 40 percent of the allotted funds. In addition, the funds set aside by GEO for telecommunication services for the clerical operation were not required as the telephones and lines were supplied by NPC with separate funding. The actual cost for the contract for the Puerto Rico clerical processing was nearly $604,000, roughly 38 percent of the $1.6 million allotted. This was due to a much smaller than expected workload for Puerto Rico. The number of overtime hours worked by GEO staff at NPC was less than anticipated as well; only about $130,000 was needed. As a result GEO spent 1.49 million dollars of the 2.3 million dollar FY10 budget, or approximately 63.5 percent.

NPC FY10 costs were less than half of their allotment. The charges were primarily from salaries and overheads; approximately 1.6 million dollars. NPC also spent approximately $600,000 on

non-salary objects, such as computer equipment for NPC clerical staff and equipment maintenance costs. Ultimately, the final cost for clerical work in NPC was just over 2.2 million dollars, or about 45 percent of their FY10 budget, which turned out to be consistent with the original estimated 20 percent increase over Census 2000 workloads and not the contingency estimates actually used to calculate the budget for the 2010 Census.

Table 5.7.1 shows the planned and actual costs for both divisions, as well as the overall totals for the operation.

Table 5.7.1 Initial Budget and Actual Cost

Division	Total Initial Budget	Actual Cost	Percent of Total Budget Spent
GEO	$2,350,000	$1,491,328	63.46
NPC	$4,966,541	$2,234,227	44.99
TOTAL	**$7,316,541**	**$3,725,555**	**50.92**

Source: Financial Management Reports.

After the height of the operation was over in late May 2010 and it became evident that spending would be less than anticipated, both GEO and NPC identified surpluses in their respective allocations and the funds were reprogrammed. Table 5.7.2 indicates the revised budget plans for both divisions and the actual costs, as well as overall totals for the operation.

Table 5.7.2 Revised Budget and Actual Cost

Division	Total Revised Budget	Actual Cost	Percent of Revised Budget Spent
GEO	$1,580,000	$1,491,328	94.39
NPC	$2,666,541	$2,234,227	83.79
TOTAL	**$4,246,541**	**$3,725,555**	**87.73**

Source: Financial Management Reports.

6. Related Evaluations, Experiments, and/or Assessments

The assessments for the following 2010 Census operations provide additional information on the Non-ID cases originating from their operation:

Be Counted and Questionnaire Assistance Centers Assessment
Enumeration at Transitory Locations Assessment
Group Quarters Enumeration Assessment
Nonresponse Followup Operations Assessment

Shipboard Enumeration Operation Assessment
Telephone Questionnaire Assistance Assessment
Update Enumerate Operations Assessment
Update/Leave Assessment

The 2010 Census Field Verification Operation Assessment provides information on the final determination of the addresses sent to 2010 FV from 2010 Census Non-ID Processing.

The 2010 Census Universe Control and Management/Response Processing System Assessment documents the processes involved in the 2010 UCM and provides information on the addresses processed through the 2010 UCM.

Staff in DMD conducted research on a small sample of Non-ID cases deleted during 2010 FV. The purpose of the research was to determine the causes for the deletions and offer some recommendations for future Non-ID Processing and FV operations. The results of the research are documented in a memorandum for the record titled *Initial Research Regarding Addresses from 2010 Non-ID Processing Subsequently Deleted during 2010 Field Verification*.

7. Conclusions and Recommendations

The 2010 Census Non-ID Processing stakeholders dealt with challenges throughout the planning/ development and implementation stages of the operation. The knowledge and experience gathered during the operation will assist in more effective and efficient planning and development of future census operations. This section provides conclusions drawn from this assessment and identifies key recommendations to consider for implementation in the 2020 Census.

7.1 Conclusions

The key points of the Non-ID Processing assessment are summarized in this section.

- The operation finished on time, under budget, and without any serious issues affecting production.

- Over 1.2 million Type A and Type B cases were processed, with over 1.1 million cases positively resolved (match and/or geocode found), allowing more addresses to be included in the 2010 Census and improving census coverage.

- Over 1.6 million Type C cases were processed, with over 1.5 million cases positively resolved (added or matched), allowing more addresses to be included in the 2010 Census and improving census coverage.

- DSCMO was able to deliver the Type A/B ADDUP files more frequently than planned and GEO was able to process a larger volume of Non-ID cases through automated processing at a faster rate than expected.

- MAGIC was successfully adapted from the ACS geocoding software and was well tested and ready for production in advance of the start of the clerical operation in April 2010. Also, MAGIC was designed to handle multiple types of inputs without requiring major software changes.

- A large number of clerical staff was trained over a short period of time in order to get the clerical operation underway.

7.2 Recommendations

This section covers recommendations for future Non-ID Processing operations for the following subject matter areas:

- General planning and development
- Clerical processing (staffing and training, production and clerical software)
- Staff communications

General Planning and Development

- The decisions regarding which software to use for the clerical operation and what staff (i.e., NPC clerks, contractors, or other) will be assigned the Puerto Rico cases should be incorporated into the planning and development stage of 2020 Census Non-ID Processing from the beginning. If the decision is made to use a contractor, sufficient lead time will be needed to obtain the services of a contractor.

- If Non-ID Processing and FV are conducted in the same manner for the 2020 Census, the FV cutoff can be scheduled much later into production, possibly near the end of production. This recommendation is based upon the assumption that processing will run as quickly as it did during the 2010 Census and that the FV operation will not require their universe of addresses until early to mid-August 2020. If the FV stakeholders' require preliminary workload figures from Non-ID Processing in June 2020 for planning purposes, this could be achieved by an interim product that contains counts of eligible cases by collection block within each LCO. This will allow each LCO to determine their workload distribution and plan staff accordingly.

- If Non-ID Processing and FV are conducted differently for the 2020 Census, then the delivery of addresses to FV can possibly occur on a flow basis, especially if FV begins much earlier than it did for the 2010 Census. It may also be possible to have addresses not located during FV to be sent back to Non-ID Processing for review, which could result in corrections to the address information or a different geocode being assigned the case.

- The official schedules for future Non-ID Processing operations should be more inclusive. For example, there should be activities in the schedule that track the clerical training and the development of the clerical procedures.

- The automated and clerical processing stages should be developed in a more parallel fashion, rather than sequentially, as occurred during the planning and development stage of the 2010 Census Non-ID Processing operation. If automated processing is developed first and takes longer than anticipated, then there will be less time to devote to clerical processing than needed.

- NPC stakeholders should be involved in the planning and development stage of the 2020 Census Non-ID Processing operation from the beginning. NPC staff can provide information on the clerical staff, including any restrictions on what tasks the clerical staff can perform.

- The JTC/TTC staff should be included in the planning and development stage of the 2020 Census Non-ID Processing operation so that they are fully aware what is needed from them regarding foreign language calls and can plan accordingly.

- The roles and responsibilities of all the stakeholders should be made clear during the planning and development stage of the 2020 Census Non-ID Processing operation.

- Requirements for the assessment data should include information on the operation(s) that updated an address prior to Non-ID Processing to allow additional analysis. It would be useful to know the distribution of matched cases among the various operations, as well as if Non-ID Processing reinstated any deletes from other operations, which was not possible to track during 2010 Census Non-ID Processing.

- The possibility should be considered of submitting all Type C cases that get rejected due to an incomplete or missing block geocode through the Type A/B automated and clerical header coding and geocoding processes in order to give these cases another chance at being positively resolved and included in the census.

- Any case that includes information in the address fields that suggests the address is nonresidential should be flagged during the automated process and sent to the clerical process so that the respondent can be contacted for their residential address. For example, address information containing "PMB" within it should be flagged because the address belongs to a business that rents private mailboxes and should not be included in the census universe. If the clerical staff is unable to acquire a residential address from the respondent, then the case should be made uncodable. If the address turns out to be a residence and not commercial, or is mixed-use, then the case can proceed through the process and have its status updated accordingly.

- Any Type A case that includes information in the address fields that suggests that the respondent(s) are experiencing homelessness should be converted into a Type B case during the automated process. Converting as many as possible during the automated process will reduce the effort needed by the clerical staff to identify these cases. Also, the risk of a case being deemed uncodable during clerical processing would be lower because only a state and county code are required to resolve Type B cases.

Clerical Processing - Staffing and Training

- The clerical trainings at NPC could be more productive if held in a training room with the necessary equipment away from the production area. The trainings could also be conducted for smaller groups so that participants can see and hear more clearly.

- A consistent message should be delivered during all the trainings, including the goals of the training, as well as the procedures themselves. A verbatim training is one possible way to ensure that the same training is being given to all clerks. Another way to help ensure consistency in the training sessions might be to have a GEO staff member run the clerical trainings with a NPC staff member supporting them. Finally, using computer-based training may be a way to help the trainers maintain consistency between training sessions.

- The "Train the Trainers" sessions should ensure the staff learns all they can about the operation prior to the start of clerical training, and should be scheduled with sufficient time to achieve that goal. It would be beneficial for the "Train the Trainer" sessions to include numerous training cases and possibly "live" cases as the initial production workload comes in.

- A broader set of examples should be used in the clerical trainings for 2020 Census Non-ID Processing. The complexities of the Non-ID Processing operation are difficult to account for during the training. All the variations of address information that could be received by the input operations could never be covered by the examples used in the training. The unusual Non-ID cases that were encountered during 2010 Census could be used as training cases for 2020 Census Non-ID Processing. Given a more comprehensive set of examples, the clerical staff will be better prepared to handle the majority of cases and learn how to be flexible when it comes to the challenging cases.

- All stakeholders must recognize that the workload for the clerical operation fluctuates on a weekly basis, even on a daily basis, and that the staffing levels will need to adjust accordingly. This means it may be difficult for future Non-ID Processing clerical operations to avoid moving staff around, but it is recommended that NPC continue to keep a core dedicated staff on the operation from beginning to end to ensure things keep operating smoothly. The core staff would include a manager for the Geographic Technicians and Unit Supervisors that will work on Non-ID Processing exclusively through the height of the clerical operation (i.e., the first six weeks) and then part-time as needed for the remainder of the operation. A possible way to maintain a core staff would be to conduct similar clerical processing work for other customers, such as ACS and the current surveys.

- A turnover rate should continue to be factored in when estimating the staff levels needed during clerical production.

- The feasibility of having the calling component of the clerical process be conducted during a split shift should be investigated. The split shift would start in the middle of

the day shift and end in the middle of the night shift. This may mitigate concerns about calling respondents too early or too late in the day, as well as ensure the clerks would be calling at times when people were more likely to be home and willing to talk with the clerks.

Clerical Processing – Production and Clerical Software

- The foreign language Non-ID cases should be sorted into separate work units so NPC staff can more easily determine the number of foreign language calls that potentially need to be made and can schedule the necessary translators at the beginning of the shift rather than having to react to the cases as they appear randomly within regular work units.

- The clerical software should be able to separate the Non-ID cases and not always keep them bundled into work units. The ability to separate cases could allow the transfer of cases between clerks, another functionality that should be built into the clerical software. This could permit the cases that require additional telephone calls on one shift to be removed from the original clerk's work unit and combined into a new work unit for a different clerk on the other shift. This ability would also allow cases that were resolved in error or were left unfinished by a clerk to be inserted back into production.

- The clerical software should allow the user to have multiple work units active per account. This way, if there is a problem with one work unit, the user could access another work unit without waiting for the problem with the first work unit to be fixed or having to create a second user account in order to access another work unit.

- The clerical software should have the ability to capture metrics for each clerk that can be used as part of the QC process.

- A sample of the cases worked by each clerk should be part of the QC process.

- NPC clerical unit supervisors should monitor the status of the work units using the reports in MAGIC on a daily basis. It would be helpful for each clerk to be able to check on the status of their own work units as well. This reduces the chance of work units getting left "hanging" in MAGIC because either the clerk moved on to another work unit without completing the first one or the clerk was moved onto another project with completing the work unit.

Staff Communications

- An official communication plan should be developed for the 2020 Census Non-ID Processing operation. The plan should include the following, at a minimum:

    - Contact information for all stakeholders as well as explanations of when specific staff should be contacted.

  - A description of meetings, including the timing, frequency, and purpose of any regularly scheduled meetings. As an example, during the early planning stages of the operation, quarterly face-to-face meetings could be set up for HQ and NPC staff, which would become more frequent when training begins, and perhaps even more frequent during production.

  - An explanation of the formal mechanism for requests, software related or otherwise. A possible mechanism is a request form that is emailed to the subteam to review. This would be very helpful during the gathering of initial requirements during the planning and development stage.

  - Details on creating an email distribution list to ensure that all relevant stakeholders receive pertinent emails. An email distribution list was created for 2010 Census Non-ID Processing and worked well when utilized.

- All stakeholders need to be informed of any updates to the operation. All subject matter areas involved in the operation should have a representative at the subteam meetings and should be included in the distribution of the meeting notes. One area should be responsible for updating and distributing updates to the procedures to ensure consistency.

- Regarding the clerical staff, the methods used during 2010 Census Non-ID Processing for communicating to a large and diffuse workforce could be enhanced to ensure everyone is current on procedures. Some of the practices used for the 2010 Census that could be used more are regular staff briefings and handouts given to each clerk explaining the changes to the operation. A recommended new method for use during the 2020 Census is a website that the staff can use to access all Non-ID Processing documentation, including the latest procedural updates.

- Technologies such as videoconferencing, which was used effectively for planning and implementing the 2010 Census Non-ID Processing operation, could be used again for face-to-face team and stakeholder meetings. For the 2010 Census, videoconferences generally were more productive than teleconferences for subteam meetings. The visual aspect of the videoconferences allowed for team members who could not attend the meeting in person to feel like they were fully contributing to the meeting. Videoconferences also allowed for team members who do not regularly see each other to get to know each other better.

- On-site visits from HQ staff to NPC prior to production are very helpful for observing the environment where the clerical operation will be taking place and for HQ staff to get a better understanding of how NPC manages their staff. As occurred for the 2010 Census, this will allow the subteam to recognize any necessary adjustments to the training plan.

8. Acknowledgements

ALD OIT would like to sincerely thank the following divisions and groups for their valuable contributions to this assessment:

Decennial Management Division
Decennial Statistical Studies Division
Decennial Systems and Contract Management Office
Federal Working Group
Geography Division
National Processing Center

9. References

Address List Development Operations Integration Team (2012), "2010 Census Field Verification Operational Assessment," U.S. Census Bureau, 2012.

Alberti, Nicholas and Robin Pennington (2010), "Customer Requirements for 2010 Census Non-ID Processing", DSSD Decennial Census Memorandum No. A-04R1, U.S. Census Bureau, May 18, 2010.

Barrett, Diane, Andrew Ciango, Brian De Vos, and Andre Williams (2012), "2010 Census Shipboard Enumeration Operation Assessment Report," U.S. Census Bureau, March 7, 2012.

Jonas, Kimball (2003), "Census 2000 Evaluation E.5, Revision 1 – Group Quarters Enumeration," U.S. Census Bureau, August 6, 2003.

Kebede, Henock (2009), "Test Plan for the MAF/TIGER System Support of the 2010 Non-ID Processing Operation," U.S. Census Bureau, October 28, 2009.

Medina, Karen (2003), "Assessment Report for Census 2000 Non-ID Questionnaire Processing (including BCF/TQA Field Verification)," Census 2000 Information Memorandum No. 141, U.S. Census Bureau, September 25, 2003.

Niosi, Michael (2008), "Customer Requirements for the 2010 Non-ID Processing Clerical Application," U.S. Census Bureau, December 19, 2008.

Niosi, Michael (2010a), "2010 Census Operational Plan for the Non-ID Processing Operation," 2010 Census Information Memorandum No. 10, U.S. Census Bureau, March 9, 2010.

Niosi, Michael (2010b), "2010 Non-ID Processing Assessment Study Plan," 2010 Census Planning Memorandum No. 74, U.S. Census Bureau, May 12, 2010.

Niosi, Michael (2010c), "Specification for 2010 Non-ID Processing ADDUP File Composition *Revision 2*," U.S. Census Bureau, June 2, 2010.

Pike, A. Edward (2001), "Program Master Plan: Census 2000 Matching/Geocoding Non-Master Address File (MAF) Identification (ID) Questionnaires," Census 2000 Information Memorandum No. 98, U.S. Census Bureau, April 12, 2001.

Treat, James (2009), "2010 Census Schedule Management Plan," 2010 Census Information Memorandum No. 34, U.S. Census Bureau, July 1, 2009.

U.S. Census Bureau (2010), "ALD OIT Operational Status Report," September 14, 2010.

Vitrano, Frank A. (2007), "Reduced Scope of the 2008 Census Dress Rehearsal and a One-Month Delay of Census Day," 2008 Census Dress Rehearsal Memoranda Series No. 50, Decennial Management Division, U.S. Census Bureau, November 19, 2007.

Whittington, Charles (2009), "Non-ID Load Test Summary July 27 – July 31, 2009," U.S. Census Bureau, August 13, 2009.

Appendix A: Summary of 2010 Census Non-ID Processing Risk Assessment

Brief Risk Description	Mitigation Strategy	Contingency Plan	Outcome
If UCM cannot accurately distinguish between input sources, GEO will be unable to process the records correctly.	DMD and GEO worked with DSCMO to determine best methodology to ensure accurate identification of input sources, including use of operation code, processing ID, and form type.	GEO would have to assign a "generic" source code and process the records all in the same way if the individual sources cannot be identified.	No problems in assigning the proper source to all the cases during the course of the operation.
A larger workload than anticipated could add significant cost to Non-ID Processing.	Census 2000 workloads provide an adequate baseline from which to estimate. Also, current equipment and processes are both faster and more efficient than in 2000.	Clerical processing would require additional staff hours, which could be accommodated by overtime before the day shift and after the night shift Monday-Friday, and staff could also work weekends.	Workload was actually less than anticipated, even during the peak of operations (April and May).
If there are problems with C&P receiving updates from GEO the ability to monitor progress would be hampered.	DMD delivered requirements for C&P data well in advance of production and also conducted User Acceptance Testing of C&P.	Progress data from the clerical work could be monitored via reports generated from the MAGIC software. However, details on the results of automated processing would be unavailable unless GEO could provide ad hoc reports.	For the first three weeks of production, GEO provided progress data via email. Afterward, they were able to deliver the C&P data for Type A and Type B cases on a weekly basis.

Brief Risk Description	Mitigation Strategy	Contingency Plan	Outcome
If the address data from the various operational inputs are not data-captured correctly, the address records could be incorrectly categorized and processed incorrectly, leading to rejected or uncodable records.	The Non-ID Processing subteam met with other operational teams (e.g., Be Counted, TQA) to inform them about the requirements for address data to be successfully processed.	The clerical operation provides an opportunity to resolve a case without adequate address information. However, it requires a contact name from the response form, which is not always supplied by a respondent or data-captured correctly.	There were no significant issues regarding quality of address data from any one source.
If the address data for a case are mapped to the wrong fields, the difficulty for matching and/or geocoding the record will be significantly increased.	DMD and GEO provided detailed specifications to DSCMO to ensure the data were mapped accurately from URdbS to UCM and then to the ADDUP files generated for GEO.	If GEO determined data were incorrect or missing, DSCMO would be asked to redeliver the data from UCM.	Due to an error in the crosswalk from UCM regarding data from TQA Fulfillment, street names were omitted from the initial delivery to GEO, which severely impacted the match rate during Non-ID Processing. DSCMO redelivered the file and corrected their software for future deliveries.
If clerical production rate is lower than anticipated, there will be a subsequent impact on the Field Verification workload, which impacts census coverage.	Two weekday shifts for stateside clerical processing, and use of overtime as needed for both stateside and Puerto Rico clerical Non-ID Processing.	Phone calls could be eliminated from the clerical process to expedite resolution of cases – “low hanging fruit” rationale.	Production rate was faster than anticipated for both Front Line and Referral clerks.

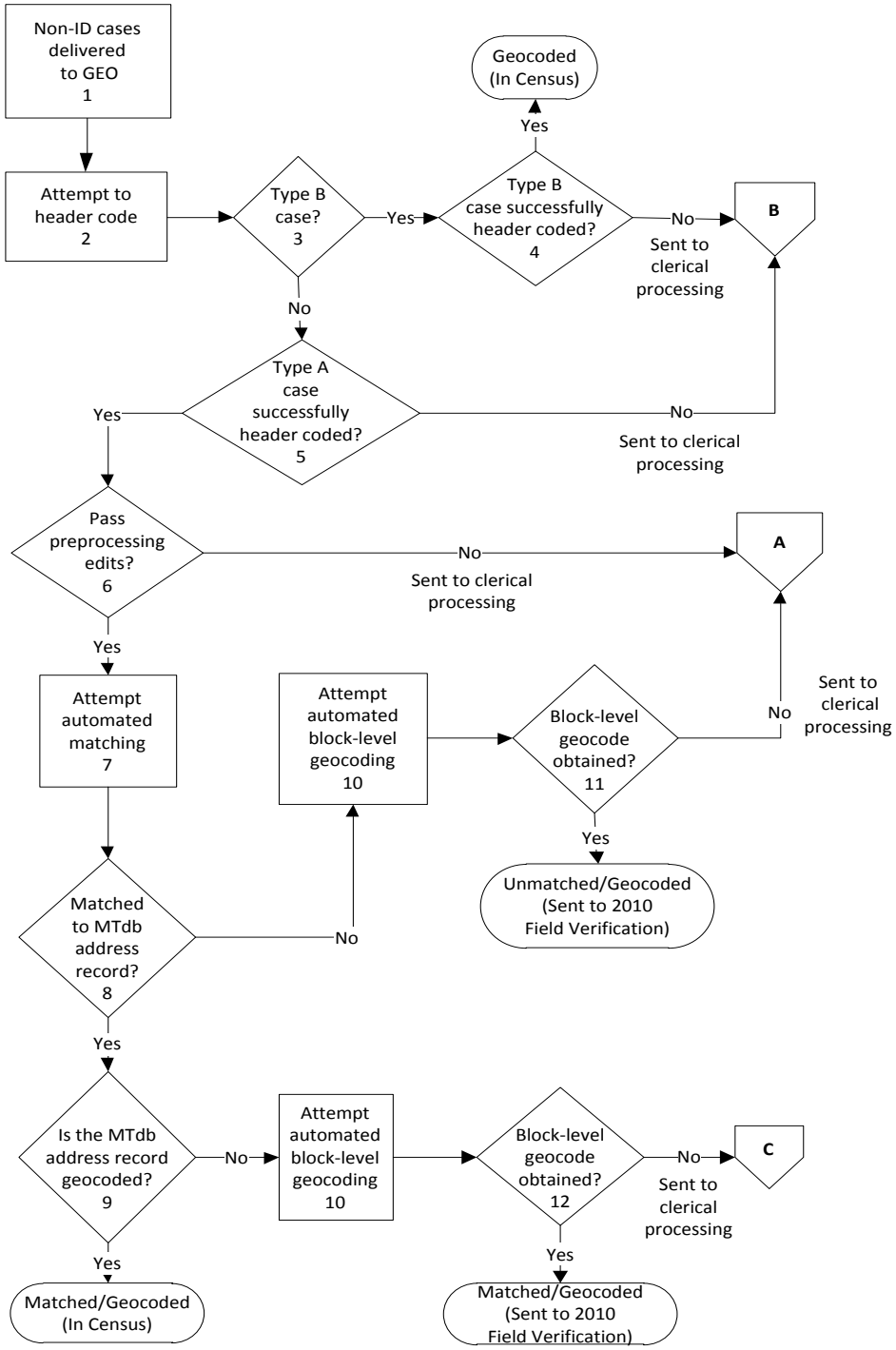
Brief Risk Description	Mitigation Strategy	Contingency Plan	Outcome
Difficulty obtaining cooperation from respondents during telephone calls could reduce success rate.	Training materials provide a script for clerks to follow, and some opportunities to practice dealing with difficult respondents.	If a high average rate of refusals is observed, monitor individual performance of clerks to detect low success rates; retrain or reassign clerks not suited to making telephone calls.	Rate of refusals was low (6.90 percent of total calls attempted).
If records are received too late to be processed they will miss the opportunity to be included in the census.	DSCMO initially delivered records weekly, but then changed to semi-weekly for the bulk of the operation. GEO also made frequent deliveries of records to clerical processing.	Census Bureau management would have to decide if processing should extend beyond current schedule in order to accommodate late records.	There were 865 Type A and Type B cases received after the 2010 FV supplemental cutoff. These cases had the opportunity to be included in the census if they were matched to a geocoded MTdb record.
If the automated matching has significant difficulty matching Puerto Rico addresses, the clerical staff will have to resolve most of them, which impacts the ability to resolve them by the 2010 FV cutoff date.	Hire contractor staff knowledgeable about PR addresses to assure high production/success rate. Staff the PR clerical operation adequately to deal with the estimated 60,000 records.	Use overtime to extend work hours. Also, ensure that adequate equipment, other infrastructure is in place for additional staff if feasible to bring on (depending on timing of decision and proximity of the 2010 FV cutoff date).	The approximately 13,000 records for PR was less than one-quarter of the estimate.

Appendix B: 2010 Census Non-ID Processing Workflow

2010 Automated and Clerical Non-ID Processing
Stateside and Puerto Rico
Type A and Type B Cases
(Pre-2010 Field Verification Cutoff)

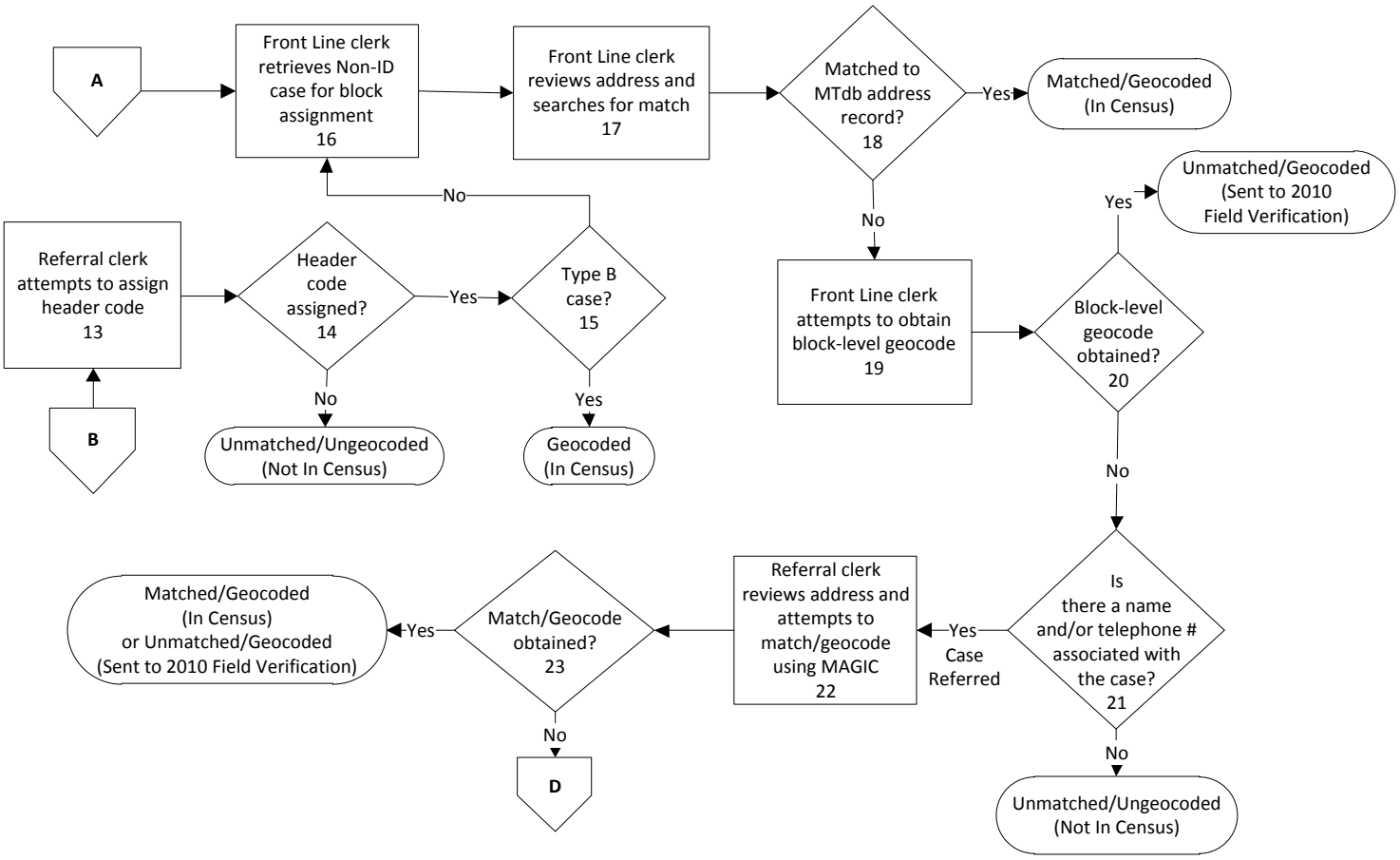
(Part 1 of 3)

Narrative begins on page B-4.

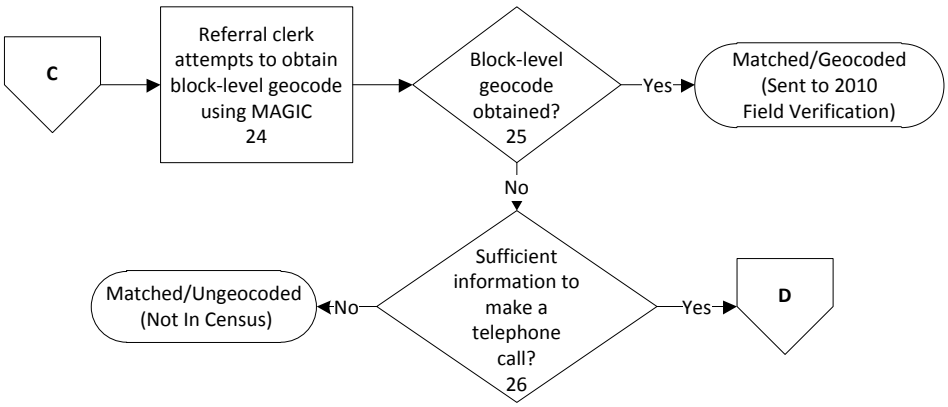


2010 Automated and Clerical Non-ID Processing
Stateside and Puerto Rico
Type A and Type B Cases
(Pre-2010 Field Verification Cutoff)

(Part 2 of 3)

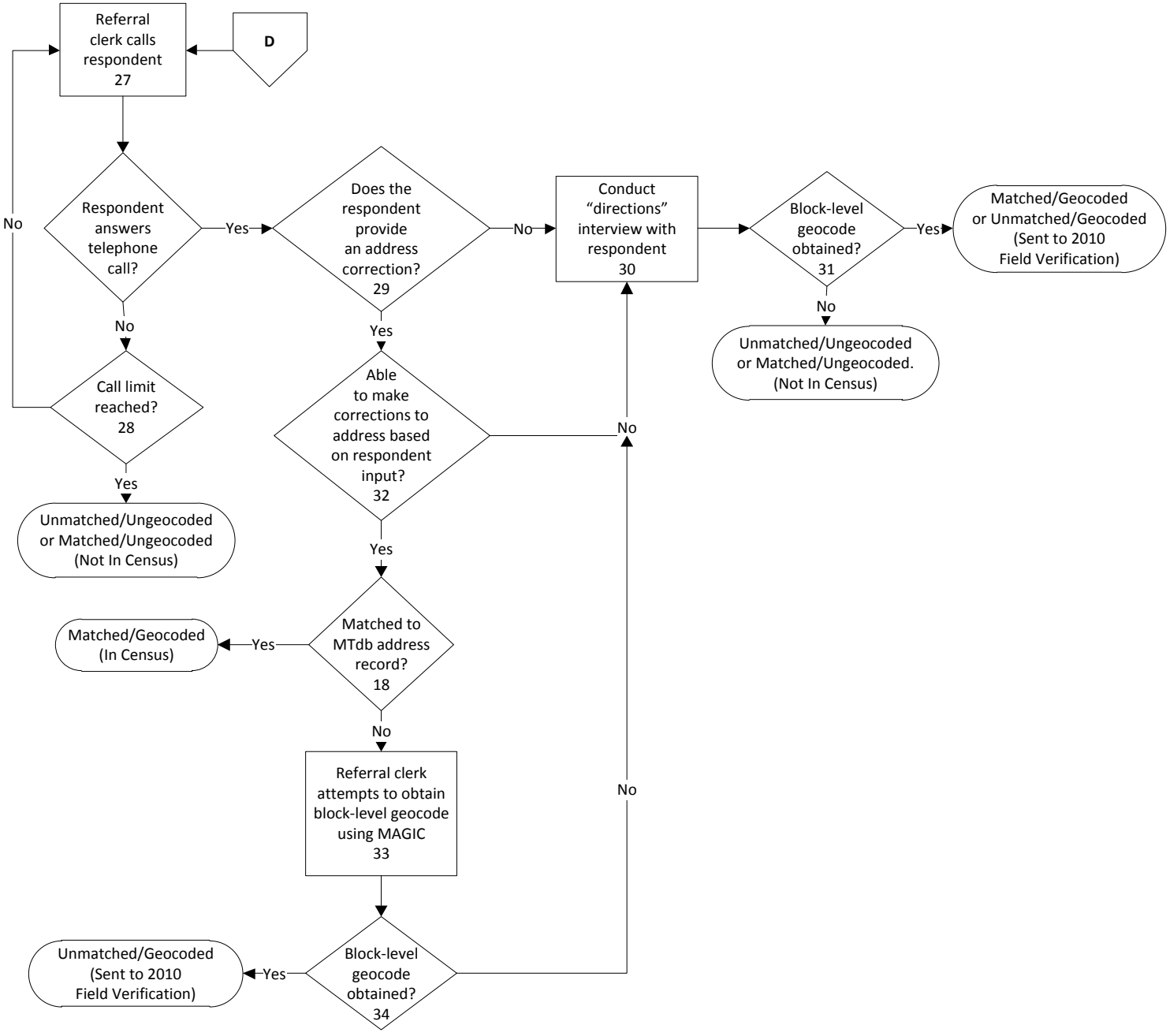


Workflow for cases that go directly to Referral Clerks



2010 Automated and Clerical Non-ID Processing
Stateside and Puerto Rico
Type A and Type B Cases
(Pre-2010 Field Verification Cutoff)

(Part 3 of 3)



Narrative for Automated and Clerical Non-ID Processing Workflow
Stateside and Puerto Rico
Type A and Type B Cases
(Pre-2010 Field Verification Cutoff)

Box	Box Title	Description
1	Non-ID cases delivered to GEO	Type A and Type B Non-ID cases are delivered to GEO in ADDUP files for Non-ID Processing.
2	Attempt to header code	The process of using the respondent's address information and assigning state and county codes to the Non-ID case.
3	Type B case?	The automated process checks for Type B cases, which are cases that have a "No Address" flag. If the case contains this flag, then the case requires only header coding. If the case does not contain the flag, then it is a Type A case and an attempt to match and geocode the case will be made after it is header coded.
4	Type B case successfully header coded?	If the case is Type B, and it is successfully header coded, then the case goes no further in the Non-ID process. The case is header coded and passed on to DSCMO to be added to the GQ universe and allocated to census counts accordingly. If the Type B case fails to header code, then it is passed on to clerical processing (Referral clerk).
5	Type A case successfully header coded?	If a type A case is successfully header coded, it moves on to preprocessing edits. If it is not successfully header coded, it is sent to clerical processing (Referral clerk).
6	Pass preprocessing edits?	Edits are run on the address information checking for completeness, legal values, and rejection criteria. If the address information passes, then it moves on to the automated matching process. If it fails, it is sent to clerical processing (Front Line clerk).
7	Attempt automated matching	The process by which addresses go through Census Bureau software in an attempt to match the address to an existing MTdb address record. Automated matching includes city-style (house number and street name) addresses and Rural Route addresses (exact matches only for Rural Route addresses).
8	Matched to MTdb address record?	If the address does match a MTdb address record, then the system checks to see if the MTdb address record is already geocoded. If the address does not match a MTdb address record, then an attempt is made to obtain a block-level geocode for the Non-ID case through automated geocoding.
9	Is the MTdb address record geocoded?	If the MTdb address record that was matched to is already geocoded, then the Non-ID case is done processing because it has been matched and geocoded, and will be included in the census counts. If the MTdb address record does not have a geocode, then an attempt is made to obtain a block-level geocode for the Non-ID case through automated geocoding.
10	Attempt automated block-level geocoding	The process of assigning a Non-ID case to a census block using an address range in the MTdb.
11	Block-level geocode obtained?	If a block-level geocode is obtained within the MTdb, then the Non-ID case is considered unmatched and geocoded, and it is then added to the MTdb to be sent to 2010 FV. If a block-level geocode is not obtained within the MTdb, then the address is passed on to clerical processing (Front Line clerk).
12	Block-level geocode obtained?	If a block-level geocode is obtained through automated geocoding for the Non-ID address, then the Non-ID case is assigned a provisional geocode and sent to 2010 FV. If a block-level geocode is not obtained within the MTdb, then the address is passed on to clerical processing (Referral clerk).
13	Referral clerk attempts to assign header code	The Referral clerk brings up the Non-ID case in the MAGIC software in order to attempt to header code the address. The Referral clerk reviews the address information and searches for the coding area to which that address belongs.

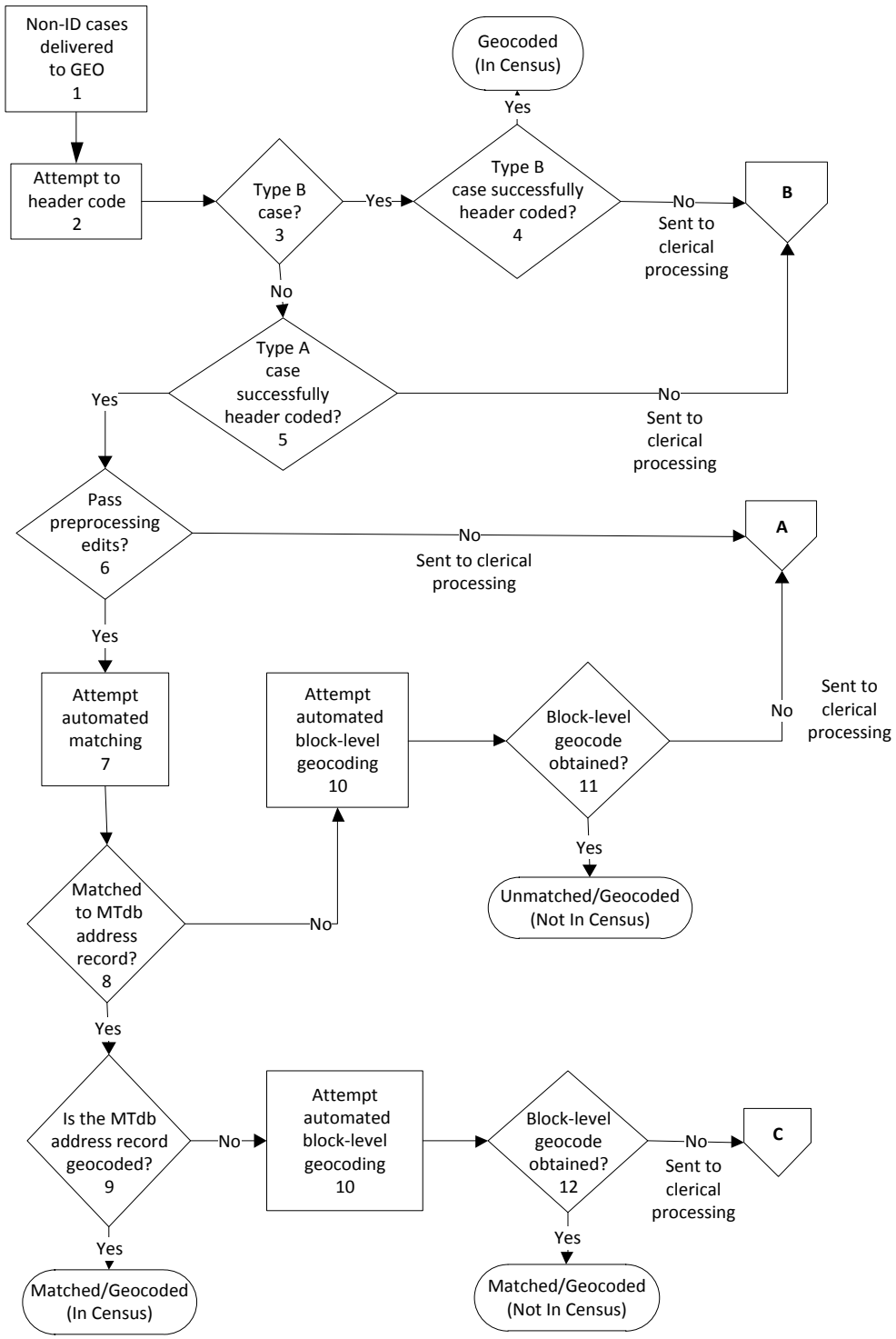
Box	Box Title	Description
14	Header code assigned?	If the Non-ID case is not able to be header coded, then the Non-ID case is unmatched and ungeocoded, and it proceeds no further. If the Non-ID case is successfully header coded, then the software checks to see if it is a Type B case.
15	Type B case?	If the address record is a Type B case (see Box 3 for more information), then the address record is sent to DSCMO to be added to the GQ universe and allocated to census counts accordingly. If it is a Type A case, then the Non-ID case proceeds on to block assignment.
16	Front Line clerk retrieves Non-ID case for block assignment	The Front Line clerk brings up the Non-ID case in the MAGIC software in order to attempt to assign a block-level geocode.
17	Front Line clerk reviews address and searches for match	The Front Line clerk reviews the address information and checks for obvious errors. Using the Search Window, the clerk enters a corrected version of the address and searches for a match in the MTdb.
18	Matched to MTdb address record?	If the address does match a MTdb address record, the Non-ID case will automatically be geocoded because the reference file that the clerks match against only contains MTdb address records with geocodes. If the address does not match a MTdb address record, then an attempt is made to obtain a block-level geocode for the Non-ID case.
19	Front Line clerk attempts to obtain block-level geocode	The Front Line clerk attempts an address range match in order to geocode the address at the block-level.
20	Block-level geocode obtained?	If a block-level geocode is obtained from within the MTdb, then the Non-ID case is assigned a provisional geocode and sent to 2010 FV. If a block-level geocode is not obtained within MTdb, then the Front Line clerk must look for additional information in order to determine if the Non-ID case can be referred.
21	Is there a name and/or telephone number associated with the case?	If no match is made to the MTdb, and the address does not geocode to the MTdb, then the clerk will look at the Additional Info tab in the MAGIC software to see if a name and/or telephone number is listed. If a name and/or telephone number is found, then the case can be referred. If the information is not there, then the case will be unmatched and ungeocoded.
22	Referral clerk reviews address and attempts to match/geocode using MAGIC	The Referral clerk reviews the case and attempts to obtain a match and/or geocode using only the MAGIC software.
23	Match/Geocode obtained?	If the address does match a MTdb address record, the Non-ID case will automatically be geocoded because the reference file that the clerks match against only contains MTdb address records with geocodes. If a block-level geocode is obtained from within the MTdb, then the Non-ID case is assigned a provisional geocode and sent to 2010 FV. If no match or geocode is obtained within the MTdb, then the next step is to attempt to contact the respondent by making a telephone call.
24	Referral clerk attempts to obtain block-level geocode using MAGIC	The Referral clerk reviews the case and attempts to obtain a geocode using only the MAGIC software. The case is already matched from the automated process.
25	Block-level geocode obtained?	If a block-level geocode is obtained from within the MTdb, then the Non-ID case is assigned a provisional geocode and sent to 2010 FV. If the address does not geocode to the MTdb, then the Referral clerk has to determine if there is sufficient information to make a telephone call.

Box	Box Title	Description
26	Sufficient information to make a telephone call?	The Referral clerk reviews the information in the Additional Info tab in the MAGIC software, and if there is no telephone number listed there, then the Referral clerk uses a commercial database to obtain a telephone number. If a telephone number is obtained then the Referral clerk attempts to contact the respondent. If no telephone number is obtained then the case is matched and ungeocoded and is not added to the census.
27	Referral clerk calls respondent	The Referral clerk attempts to call the respondent to obtain better address information or get directions to the LQ in order to geocode the address.
28	Call limit reached?	The Referral clerk will attempt to contact the respondent until the call limit has been reached. If after the final attempt is made and there was no success in reaching the respondent, then the Non-ID case is closed, and the address will either be unmatched and ungeocoded, or matched and ungeocoded.
29	Does the respondent provide an address correction?	If the respondent provides an address correction, then the Referral clerk will attempt to update the address. If the respondent does not provide an address correction, then the Referral clerk will move on to the "directions" interview. If the address is already matched, then the Referral clerk will proceed directly to the "directions" interview since there is no need to correct the address.
30	Conduct "directions" interview with respondent	During the "directions" interview, the Referral clerk asks the respondent to guide them to their house while the Referral clerk follows along with the SUMO application.
31	Block-level geocode obtained?	If the address is already matched, and the address is able to be geocoded using the SUMO application, then the address will be matched and provisionally geocoded and sent to 2010 FV. If the address is unmatched, and the address is able to be geocoded using the SUMO application, then the address will be unmatched and provisionally geocoded, and it is added to the MTdb to be sent to 2010 FV. If the address is matched and a geocode cannot be obtained, then the address is matched and ungeocoded. If the address is unmatched and a geocode cannot be obtained, then the address record is unmatched and ungeocoded, and it is not added to the MTdb or the census counts.
32	Able to make corrections to address based on respondent input?	If the Referral clerk makes an address correction based on the respondent's input, then the Referral clerk attempts to match and geocode the address. If the Referral clerk is unable to make any corrections based on the respondent's input, then the Referral clerk will proceed directly to the "directions" interview.
33	Referral clerk attempts to obtain block-level geocode using MAGIC	The Referral clerk attempts an address range match using the MAGIC software in order to geocode the address at the block-level.
34	Block-level geocode obtained?	If a block-level geocode is obtained within the MTdb for the corrected address, then the Non-ID case is considered unmatched and geocoded, and it is then added to the MTdb to be sent to 2010 FV. If a block-level geocode is not obtained within MTdb, then the then the Referral clerk will move on to the "directions" interview.

2010 Automated and Clerical Non-ID Processing
Stateside and Puerto Rico
Type A and Type B Cases
(Post-2010 Field Verification Cutoff)

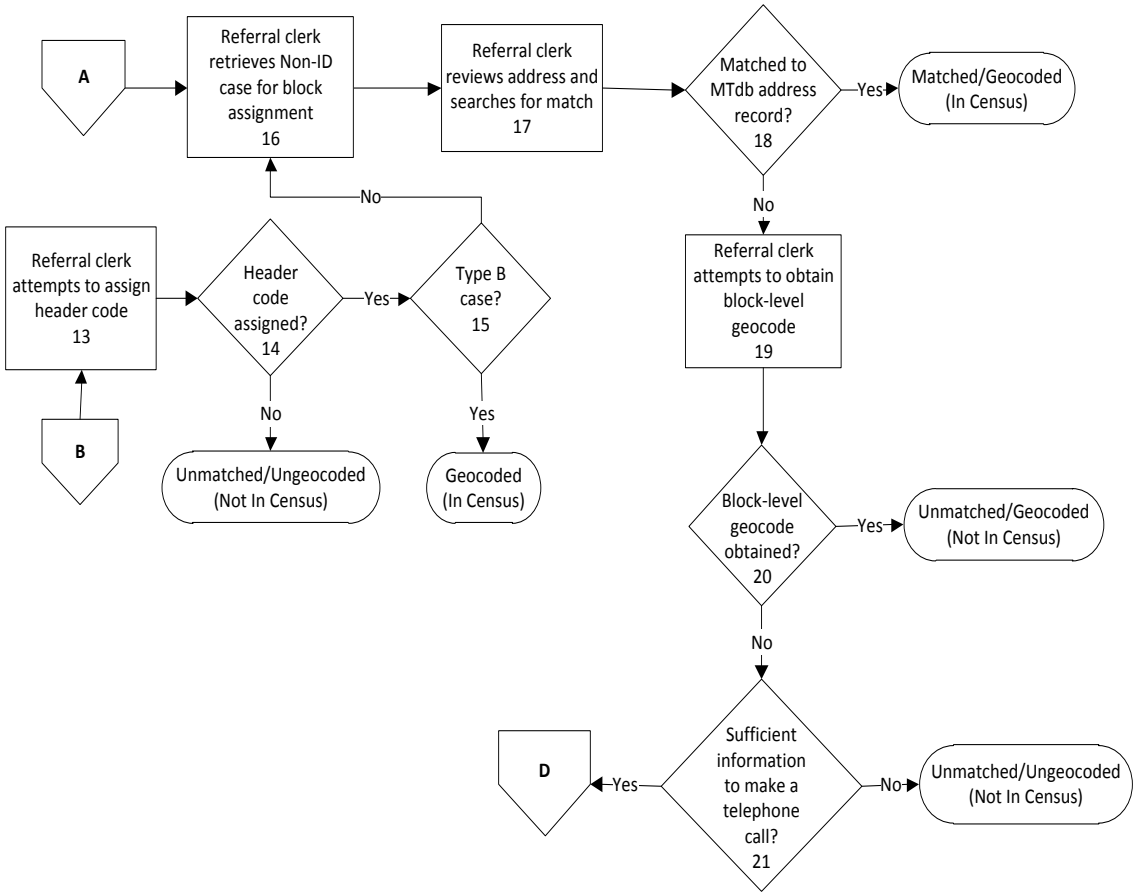
(Part 1 of 3)

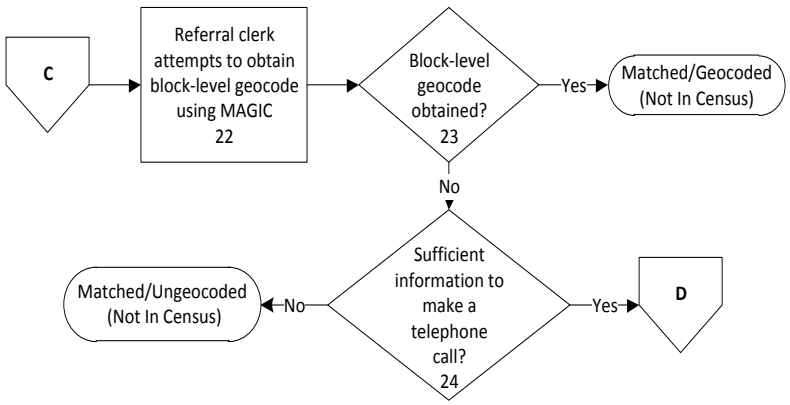
Narrative begins on page B-10.



2010 Automated and Clerical Non-ID Processing
Stateside and Puerto Rico
Type A and Type B Cases
(Post-2010 Field Verification Cutoff)

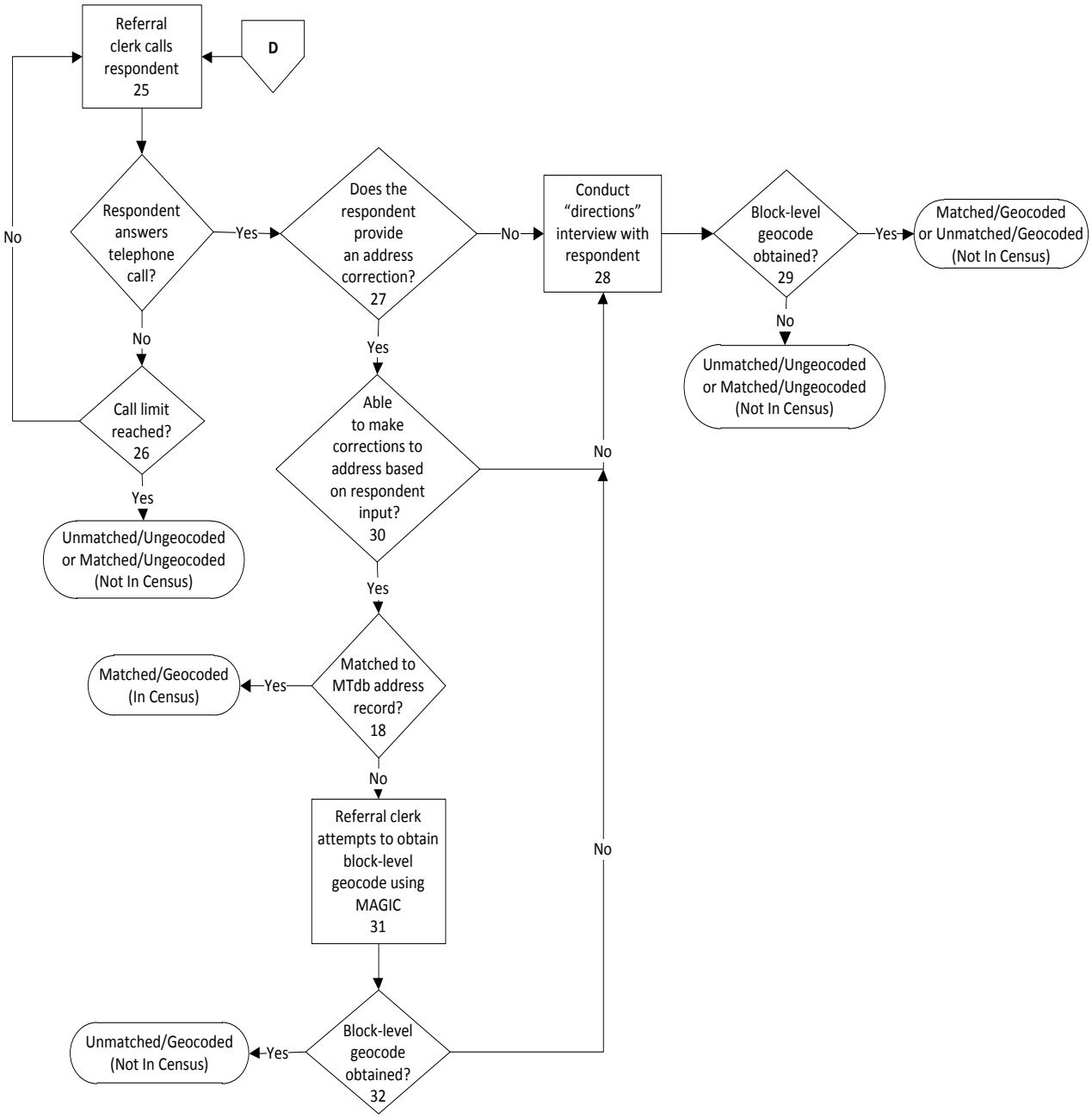
(Part 2 of 3)





2010 Automated and Clerical Non-ID Processing
Stateside and Puerto Rico
Type A and Type B Cases
(Post-2010 Field Verification Cutoff)

(Part 3 of 3)



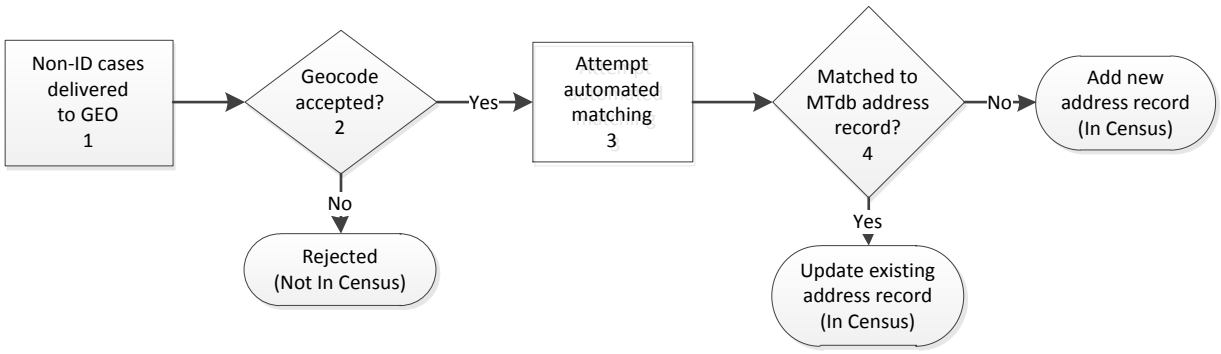
Narrative for Automated and Clerical Non-ID Processing
Stateside and Puerto Rico
Type A and Type B Cases
(Post-2010 Field Verification Cutoff)

Box	Box Title	Description
1	Non-ID cases delivered to GEO	Type A and Type B Non-ID cases are delivered to GEO in ADDUP files for Non-ID Processing.
2	Attempt to header code	The process of using the respondent's address information and assigning state and county codes to the Non-ID case.
3	Type B case?	The automated process checks for Type B cases, which are cases that have a "No Address" flag. If the case contains this flag, then the case requires only header coding. If the case does not contain the flag, then it is a Type A case and an attempt to match and geocode the case will be made after it is header coded.
4	Type B case successfully header coded?	If the case is Type B, and it is successfully header coded, then the case goes no further in the Non-ID process. The case is header coded and passed on to DSCMO to be added to the GQ universe and allocated to census counts accordingly. If the Type B case fails to header code, then it is passed on to clerical processing (Referral clerk).
5	Type A case successfully header coded?	If a Type A case is successfully header coded, it moves on to preprocessing edits. If it is not successfully header coded, it is sent to clerical processing (Referral clerk).
6	Pass preprocessing edits?	Edits are run on the address information checking for completeness, legal values, and rejection criteria. If the address information passes, then it moves on to the automated matching process. If it fails, it is sent to clerical processing (Referral clerk).
7	Attempt automated matching	The process by which addresses go through Census Bureau software in an attempt to match the address to an existing MTdb address record. Automated matching handles city-style (house number and street name) addresses and Rural Route addresses (exact match only for Rural Route addresses).
8	Matched to MTdb address record?	If the address does match a MTdb address record, then the system checks to see if the MTdb address record is already geocoded. If the address does not match a MTdb address record, then an attempt is made to obtain a block-level geocode for the Non-ID case through automated geocoding.
9	Is the MTdb address record geocoded?	If the MTdb address record that was matched to is already geocoded, then the Non-ID case is done processing because it has been matched and geocoded, and will be included in the census counts. If the MTdb address record does not have a geocode, then an attempt is made to obtain a block-level geocode for the Non-ID case through automated geocoding.
10	Attempt automated block-level geocoding	The process of assigning a Non-ID case to a census block using an address range in the MTdb.
11	Block-level geocode obtained?	If a block-level geocode is obtained within the MTdb for the Non-ID case, then the Non-ID case is considered unmatched and provisionally geocoded, and it is not added to the census counts. If a block-level geocode is not obtained within the MTdb, then the address is passed on to clerical processing (Referral clerk).
12	Block-level geocode obtained?	If a block-level geocode is obtained through automated geocoding for the Non-ID case, then the Non-ID case is matched and provisionally geocoded, and it is not added to the census counts. If a block-level geocode is not obtained within the MTdb, then the address is passed on to clerical processing (Referral clerk).
13	Referral clerk attempts to assign header code	The Referral clerk brings up the Non-ID case in the MAGIC software in order to attempt to header code the Non-ID case. The Referral clerk reviews the address information and searches for the coding area to which that address belongs.

Box	Box Title	Description
14	Header code assigned?	If the Non-ID case is not able to be header coded, then the Non-ID case is unmatched and ungeocoded, and it proceeds no further. If the Non-ID case is successfully header coded, then the software checks to see if it is a Type B case.
15	Type B case?	If the address record is a Type B case (see Box 3 for more information), then the address record is sent to DSCMO to be added to the GQ universe and allocated to census counts accordingly. If it is a Type A case, then the Non-ID case proceeds on to block assignment.
16	Referral clerk retrieves Non-ID case for block assignment	The Referral clerk brings up the Non-ID case in the MAGIC software in order to attempt to assign a block-level geocode.
17	Referral clerk reviews address and searches for match	The Referral clerk reviews the address information and checks for obvious errors. Using the Search Window, the clerk enters a corrected version of the address and searches for a match in the MTdb.
18	Matched to MTdb address record?	If the address does match a MTdb address record, the Non-ID case will automatically be geocoded because the reference file that the clerks match against only contains MTdb address records with geocodes. The address will then be included in the census. If the address does not match a MTdb address record, then an attempt is made to obtain a block-level geocode for the Non-ID case.
19	Referral clerk attempts to obtain block-level geocode	The Referral clerk attempts an address range match in order to geocode the address at the block-level.
20	Block-level geocode obtained?	If a block-level geocode is obtained from within the MTdb, then the Non-ID case is assigned a provisional geocode and not added to the census. If a block-level geocode is not obtained within MTdb, then the clerk must look for additional information in order to determine if the Non-ID case can proceed to the telephone stage.
21	Sufficient information to make a telephone call?	If no match is made to the MTdb, and the address does not geocode to the MTdb, then the clerk must locate the information necessary to make a telephone call, either in the Additional Info tab in the MAGIC software or in the commercial database. If the necessary information can be found, then the case can proceed to the telephone stage. If the information is not obtained, then the case will be unmatched and ungeocoded.
22	Referral clerk attempts to obtain block-level geocode using MAGIC	The Referral clerk reviews the case and attempts to obtain a geocode using only the MAGIC software. The case is already matched from the automated process.
23	Block-level geocode obtained?	If a block-level geocode is obtained from within the MTdb, then the Non-ID case is matched and provisionally geocoded, and it is not added to the census counts. If the address does not geocode to the MTdb, then the Referral clerk has to determine if there is sufficient information to make a telephone call.
24	Sufficient information to make a telephone call?	The Referral clerk attempt to locate the necessary information to make a telephone call, either in the Additional Info tab in the MAGIC software or in the commercial database. If the necessary information is located, then the Referral clerk attempts to contact the respondent. If the information is not obtained then the case is matched and ungeocoded, and is not added to the census.
25	Referral clerk calls respondent	The Referral clerk attempts to call the respondent to obtain better address information or get directions to their LQ in order to geocode the address.
26	Call limit reached?	The Referral clerk will attempt to contact the respondent until the call limit has been reached. If after the final attempt is made and there was no success in reaching the respondent, then the Non-ID case is closed, and the address will either be unmatched and ungeocoded, or matched and ungeocoded. In both cases, the address will not be added to the census.

Box	Box Title	Description
27	Does the respondent provide an address correction?	If the respondent provides an address correction, then the Referral clerk will attempt to update the address. If the respondent does not provide an address correction, then the Referral clerk will move on to the "directions" interview. If the address is already matched, then the Referral clerk will proceed directly to the "directions" interview since there is no need to correct the address.
28	Conduct "directions" interview with respondent	During the "directions" interview, the Referral clerk asks the respondent to guide them to their house while the Referral clerk follows along with the SUMO application.
29	Block-level geocode obtained?	If the address is already matched, and the address is able to be geocoded using the SUMO application, then the address will be matched and provisionally geocoded and not added to the census. If the address is unmatched, and the address is able to be geocoded using the SUMO application, then the address will be unmatched and provisionally geocoded, and it is not added to the census. If the address is matched and a geocode cannot be obtained, then the address record is matched and ungeocoded. If the address is unmatched and a geocode cannot be obtained, then the address record is unmatched and ungeocoded. In both cases, the address is not added to the census.
30	Able to make corrections to address based on respondent input?	If the Referral clerk makes an address correction based on the respondent's input, then the clerks attempts to match and geocode the address. If the clerk is unable to make any corrections based on the respondent's input, then the clerk will proceed directly to the "directions" interview.
31	Referral clerk attempts to obtain block-level geocode using MAGIC	The Referral clerk attempts an address range match using the MAGIC software in order to geocode the address at the block-level.
32	Block-level geocode obtained?	If a block-level geocode is obtained within the MTdb for the corrected address, then the Non-ID case is considered unmatched and geocoded, and it is not added to the census. If a block-level geocode is not obtained within MTdb, then the then the Referral clerk will move on to the "directions" interview.

2010 Automated Non-ID Processing
Stateside and Puerto Rico
Type C Cases



Note: Normally, Type C cases with a missing or invalid geocode are rejected. However for the 2010 Census, there were a large number of rejected cases for this reason from NRFU and VDC and, due to the fact that this was the last chance for these addresses to get into the census universe, these cases were sent first through the automated geocoding software, and if that was unsuccessful, then the cases were sent to the clerical process.

Narrative for Automated and Non-ID Processing
Type C Cases

Box	Box Title	Description
1	Non-ID cases delivered to GEO	Type C Non-ID cases are delivered to GEO in ADDUPs for Non-ID Processing.
2	Geocode Accepted?	The geocode provided by the enumerator is checked for completeness and validity. If the geocode is accepted, the case proceeds to automated matching. If the geocode is not accepted, the case is rejected and proceeds no further in the process.
3	Attempt Automated Matching	The process by which addresses go through Census Bureau software in an attempt to match the address to an existing MTdb address record.
4	Matched to MTdb address record?	If the address does match an existing MTdb address record, then the MTdb address record is updated to reflect an additional source. If the address does not match an existing MTdb address record, then the address becomes a new MTdb address record. In both cases, the address is included in the census universe.

Appendix C: Clerical Deliveries – Type A and Type B Cases

Delivery Date	Header Coding Cases	Block Cases	Referral Block Cases	Puerto Rico Cases	Total
4/12/2010	2,747	9,413	5,185	450	17,795
4/14/2010	6,568	16,908	9,141	789	33,406
4/17/2010	0	67,201	48,024	1,553	116,778
4/22/2010	6,460	3,677	1,455	547	12,139
4/22/2010	0	65,372	51,454	1,913	118,739
4/28/2010	7,470	40,555	32,469	835	81,329
4/29/2010	0	122,489	59,504	3,127	185,120
5/5/2010	0	19,487	11,913	0	31,400
5/6/2010	3,869	51,941	55,327	222	111,359
5/7/2010	912	5,102	5,185	1,045	12,244
5/8/2010	0	10,835	11,481	0	22,316
5/12/2010	541	3,294	3,338	243	7,416
5/13/2010	0	2,616	2,576	0	5,192
5/14/2010	3,745	1,436	950	172	6,303
5/15/2010	327	3,541	3,236	207	7,311
5/18/2010	0	520	423	0	943
5/19/2010	429	3,271	2,428	16	6,144
5/19/2010	0	3,143	2,383	232	5,758
5/21/2010	210	2,752	2,221	94	5,277
5/21/2010	0	14	17	0	31
5/24/2010	189	0	60	3	252
5/24/2010	0	0	735	0	735
5/25/2010	0	0	2,317	215	2,532
5/26/2010	67	0	971	49	1,087
5/27/2010	78	0	1,223	27	1,328
5/28/2010	92	0	1,124	21	1,237
6/2/2010	87	0	45	3	135
6/3/2010	0	0	1,317	0	1,317
6/5/2010	0	0	143	20	163
6/9/2010	158	0	1,983	58	2,199
6/10/2010	0	0	93	0	93
6/16/2010	161	0	1,607	49	1,817
6/17/2010	0	0	47	0	47
6/18/2010	0	0	42	0	42
6/19/2010	0	0	98	0	98

Delivery Date	Header Coding Cases	Block Cases	Referral Block Cases	Puerto Rico Cases	Total
6/22/2010	0	0	39	0	39
6/23/2010	227	0	2,223	9	2,459
6/24/2010	0	0	68	37	105
6/26/2010	0	0	5	0	5
6/29/2010	0	0	68	0	68
7/6/2010	348	0	3,472	57	3,877
7/10/2010	1,137	0	9,804	246	11,187
7/17/2010	1,883	0	21,350	688	23,921
7/22/2010	79	0	1,020	12	1,111
7/27/2010	34	0	737	2	773
7/29/2010	18	0	256	0	274
7/30/2010	0	0	108	1	109
8/3/2010	87	0	1,041	2	1,130
8/11/2010	39	0	263	11	313
Total	**37,962**	**433,567**	**360,969**	**12,955**	**845,453**

Source: GEO/MAGIC reports. Note: A stateside case can be worked up to three times and is counted as a separate case each time. For example, a case can be successfully header coded, referred as a block case, and then resolved as a referral block case.
*The initial delivery on 5/24/10 marked the beginning of the post-2010 FV period where there were no longer any block cases.

Appendix D: Non-ID Processing Assessment Acronyms and Initialisms

ACS	American Community Survey
ADDUP	Address Update file
ALD OIT	Address List Development Operations Integration Team
APMB	Address Programs Management Branch (GEO)
BC	Be Counted
C&P	Cost and Progress system
Census ID	Census Identification number
CIG	Census Integration Group
CR	Change Request
CRD	Customer Requirements Document
CRDA	Clerical Review and Data Assessment file
DMAF	Decennial Master Address File
DMD	Decennial Management Division
DR	Dress Rehearsal
DRIS	Decennial Response Integration System
DSCMO	Decennial Systems and Contracts Management Office
DSF	Delivery Sequence File
DSSD	Decennial Statistical Studies Division
ETL	Enumeration at Transitory Locations
EST	Eastern Standard Time
FV	Field Verification
FY	Fiscal Year
GEO	Geography Division
GSB	Geocoding Software Branch (GEO)
GPB	Geographic Programs Branch (DMD)
GQE	Group Quarters Enumeration
HQ	Headquarters
ICD	Interface Control Document
ICR	Individual Census Report
IMAGS	Interactive Matching and Geocoding System
JTC	Jeffersonville Telephone Center
LQ	Living Quarters
MAF	Master Address File
MAFID	Master Address File Identification number
MAGIC	Matching and Geocoding Interactive Clerical software
MAS	Master Activity Schedule
MTdb	MAF/TIGER database
NIFT	Non-ID Feedback Table
NPC	National Processing Center
NRFU	Nonresponse Followup
QC	Quality Control
RA	Remote Alaska
RI	Reinterview
RUE	Remote Update Enumerate

SBE	Service-Based Enumeration
SCR	Shipboard Census Report
SRS	Software Requirements Specification
SUMO	Search Utility Mapping Object
TIGER	Topologically Integrated Geographic Encoding and Referencing system
TQA	Telephone Questionnaire Assistance
TTC	Tucson Telephone Center
TU	Transitory Unit
UAT	User Acceptance Test
UCM	Universe Control and Management
UE	Update Enumerate
UFR	Unfunded Request
UHE	Usual Home Elsewhere
U/L	Update/Leave
URdbS	Universal Response database Schema
VDC	Vacant Delete Check