Update on Administrative Record Usage

Thomas Mule
Special Assistant to the Chief
Decennial Statistical Studies Division

Presentation to the Census Scientific Advisory Meeting

September 17, 2020

Disclaimer: The information provided in presentation materials is for informational purposes only and may not represent the official position of the Census Bureau or the Department of Commerce. Statements made by individual presenters may not represent the agency's final position on any matter.

Shape
your future
START HERE >

United States®
Census
2020

Administrative Record Usage in the 2020 Census

Major initiative area of research and development this decade included
- Administrative record modeling to support the reduction of contacts in the Nonresponse Followup Operation
- Assigning a census ID to self response without one (NonID operation)
- Self-Response and Nonresponse Followup Quality Assurance

This presentation will focus on the committees interest in how the administrative record modeling was modified based on the delayed start in the Nonresponse Followup Operation and the extension of the Internal Revenue Service tax filing deadline from April 15th to July 15th

This presentation will also include an update on the off-campus student record initiative

This presentation will not focus on the Citizen Voting Age Population and the Presidential Memorandum usages

Shape
your future
START HERE >

United States®
Census
2020

Outline

Original 2020 Census plans
- Planned 2020 administrative record modeling
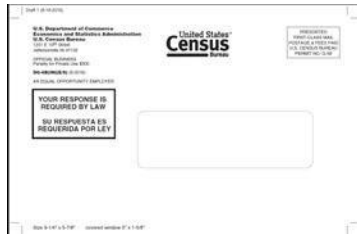- Highlight changes from the 2018 End-to-End Test for vacant and delete addresses

Changes to the methods and processing to compensate for the delay
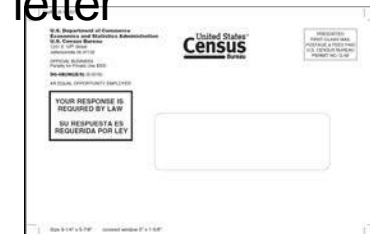
Off-campus student administrative record initiative

Summary

Shape
your future
START HERE >

United States®
Census
2020

Original 2020 Census Plan:
2020 Census Self-Response Contact Strategy

#1 Initial letter

#2 Reminder letter

#3 Reminder postcard

#4 Questionnaire

#5 Not too late postcard

This example is for Self Response (TEA1).

- Part of TEA1 received a paper questionnaire on the first mailing.

For Update/Leave TEA, #1 to #5 are replaced with enumerator leaving questionnaire packet at door and two reminder mailings in April

Shape your future START HERE >

United States® Census 2020

Original 2020 Census Plan:
Identifying Vacant and Nonexistent Addresses
Can we determine if address is vacant or does not meet our definition of a housing unit?

Example sources for AR

•United States Postal Service information
  • USPS Undeliverable-as-Addressed (UAA) reasons for census mailings made around April 1
  • Delivery Sequence File information

Internal Revenue Service (IRS) 1040 filings

IRS 1099 information returns

Centers for Medicare and Medicaid Services Medicare Enrollment database

Indian Health Service Patient database

Veterans Service Group of Illinois (VSGI) third-party national files

Census Bureau Master Address File

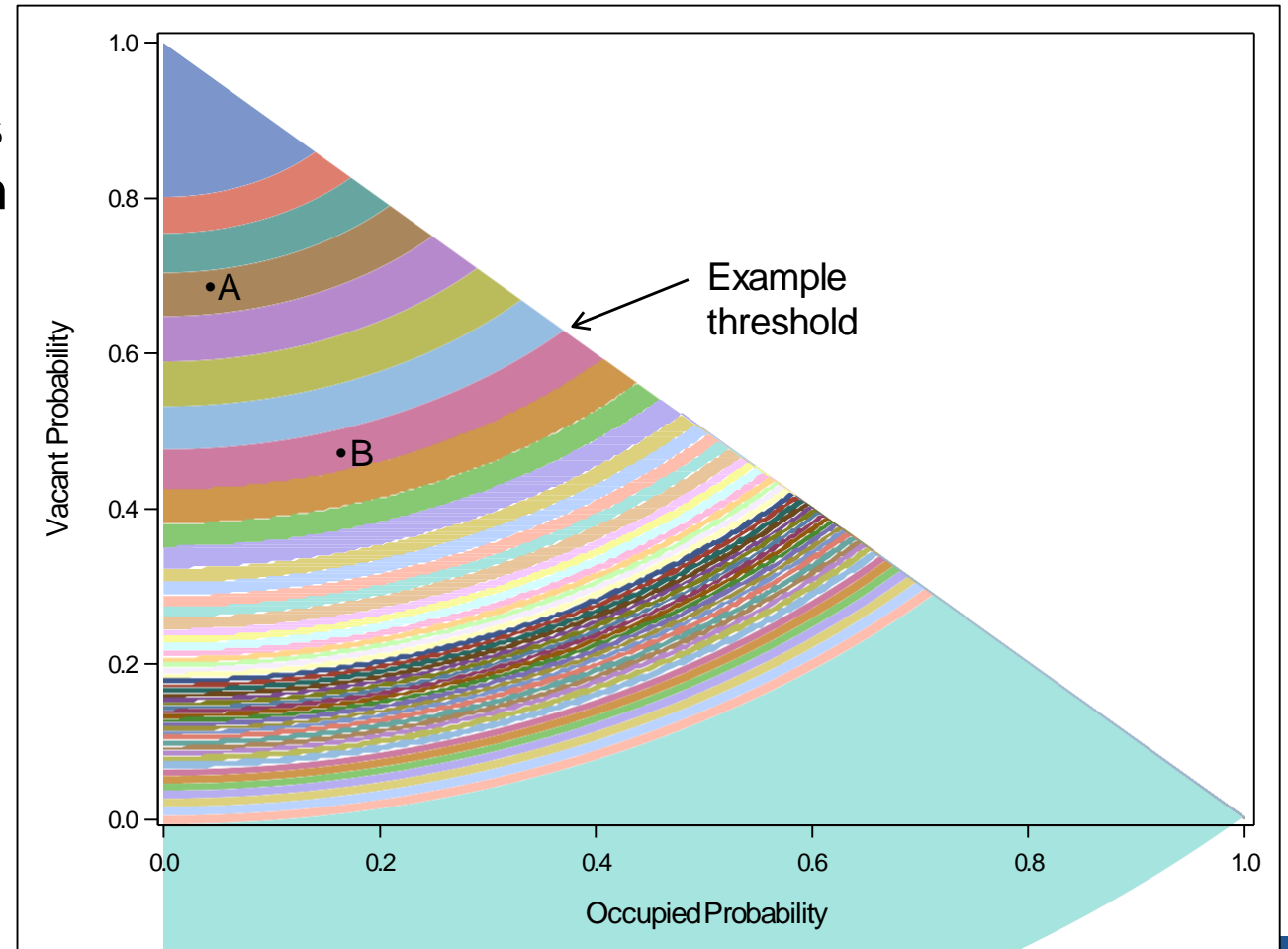ACS Area-level estimates: % vacancy, % poverty, % Hispanic, etc.

Shape
your future
START HERE >

United States®
Census
2020

Original 2020 Census Plan:
Identifying Vacant and Nonexistent Addresses
Distance Function

The distance functioncan be visualized as successive bands of cases emanating from the point (0,1) in the top left corner

Each successive band represents an additional amount of the NRFU workload

In this example, unit A is identified as initial AR vacant while unit B is not

Similar approach implemented for Non-Existent or addresses that need to be deleted

March 2018

April 2018

May 2018

Use administrative records to determine possible vacant and nonexistent address

Address has to have at least one UAA in TEA1 or TEA 6 mailings

Send mailing to address about 6 weeks after Census Day

Address receives one field visit

Mail undelivered and no sign of occupancy

Field work resolution or self-response
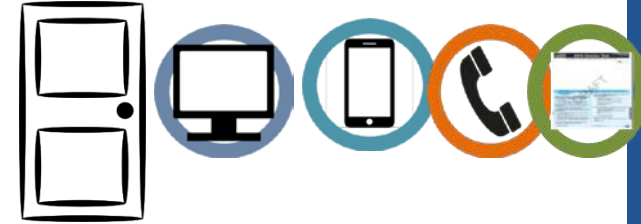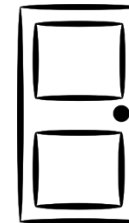
Administrative record vacant

Administrative record nonexistent address

VACANT

Address receives one field visit

Mail delivered or Mail undelivered but sign of occupancy

Address has opportunity to self-respond

Address receives full NRFU contact strategy

Shape your future START HERE >

United States® Census 2020

Original 2020 Census Plan:
Using Administrative Records to Enumerate NRFU Housing Units

Can we reduce the number of contact for 101 Main Street, Anytown USA?

1. Build a roster from most recent administrative record sources

TY 2019 Internal Revenue Service Individual Tax Returns 1040 TY 2019 Internal Revenue Service Informational Returns

Centers for Medicare and Medicaid Services Medicare Enrollment database Indian Health Service Patient Database

Census Bureau Household Composition Key File

2. Check that multiple sources indicate the household lives at an address

This uses wider set of sources than those listed above.

3. Evaluate the roster

How likely is it that we are counting all of the people rostered in the right place?

How likely is it that the household composition of the rostered family matches the Census?

4. Decision whether to use administrative record data for 101 Main Street

Shape
your future
START HERE >

United States®
Census
2020

Original 2020 Census Plan:
Identifying Occupied Addresses
Distance Function

The distance function can be visualized as successive bands of cases emanating from the point (1,1) in the top right corner

Each successive band represents additional addresses of the NRFU workload that could reduce contacts

In this example, unit A is identified as AR occupied while unit B is not
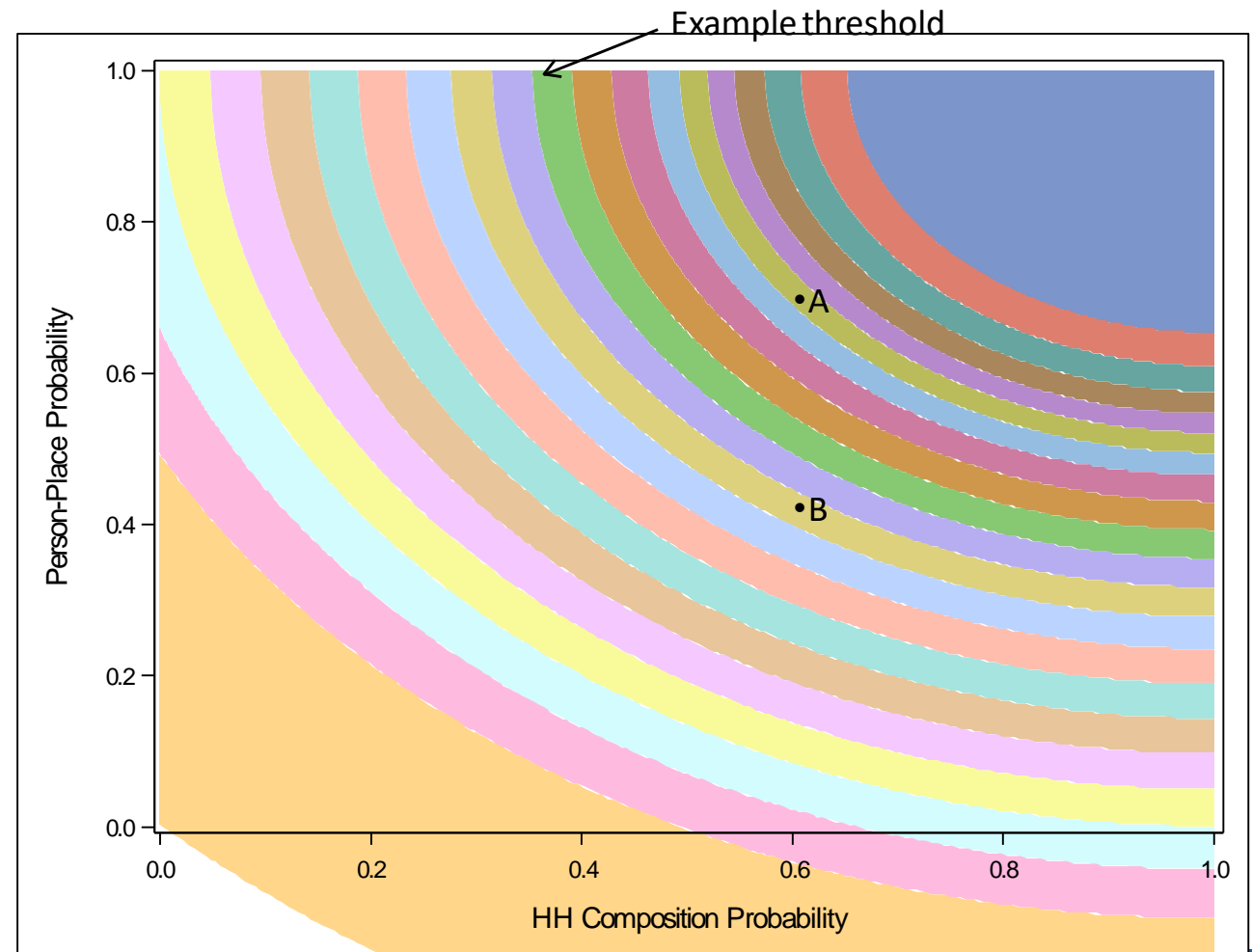
Original 2020 Census Plan:
Using Administrative Records to Enumerate NRFU Housing Units

#1 Initial letter
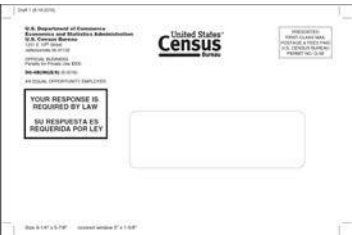
#2 Reminder letter
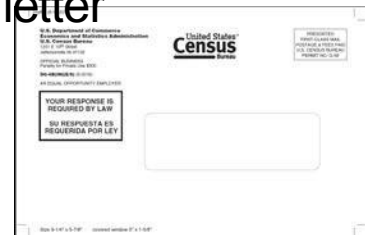
#3 Reminder postcard

#4 Questionnaire

#5 Not too late postcard

#6 First visit by enumerator and notice of visit
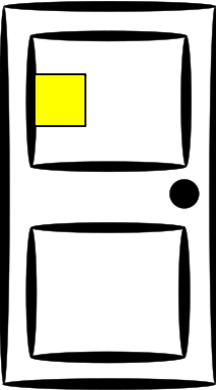
#7 Final postcard about one week after visit

Receive a self-response return – we use the respondent provided data

No return received –

we use Administrative Records Data

This example is for TEA1. For Update/Leave TEA, #1 to #5 are replaced with enumerator leaving questionnaire packet at door and two reminder mailings.

Shape
your future
START HERE >

United States®
Census
2020

Original 2020 Census Plan:
NRFU Contact Strategy

Phase 1 – Full Optimization

- Use of administrative record modeling to reduce to one day of contacts for selected occupied with rosters, vacant and addresses needed to be deleted
- Change from 2018, all vacant or delete addresses received at least one visit
- Major factor in addresses receiving one or six visits was UAA information from the May vacant/delete mailing
  - Operational control system implementation based on addresses being in the vacant/delete mailing workload

Phase 2 - Permanent Assignment

Closeout Phase – Get to Done

- One change from 2018 was to determine additional occupied, vacant and delete addresses.
- For this we utilized our post-processing distances from the 2018 End-to-End Tests
- Earliest areas eligible for closeout was June 23rd so we could include these determinations in our June update delivery.

Shape
your future
START HERE >

United States®
Census
2020

Changes to the IRS Tax Filing Deadline and NRFU Operation

Internal Revenue Service (IRS)
- On March 21, 2020, IRS announced that the tax filing deadline was extended from April 15, 2020 to July 15, 2020
  - While there was a delay in the deadline, the Census Bureau continued to receive monthly deliveries of processed 1040 records.

2020 Census
- Start of the NRFU operation delayed until August 9, 2020
  - Soft-launch of NRFU operation started on July 16, 2020 in selected ACOs

Shape
your future
START HERE >

United States®
Census
2020

Revised Administrative Record Modeling

May Vacant and Delete Determinations

Determined vacant and delete addresses at end of May with post card in-home arrival of June 12th

Update/Leave addresses had only one mailing in early April

- Modified our vacant and delete modeling to use only one UAA determination

Update/Leave mailing in April included P.O. Box Only addresses

- Implemented rule to not allow determination if address was in P.O. Box only zip code

- 430,000 vacant and 20,000 delete addresses

Concern about identifying vacant or delete addresses in zip codes with high concentrations of UAAs

- Implemented rule to not allow determinations for the top 5th percentile of zip codes had a with rate of UAAs on the first mailing   (57 percent or higher)

- 1,000,000 vacant and 20,000 delete addresses

Shape
your future
START HERE >

United States®
Census
2020

Revised Administrative Record Modeling

June Administrative Record Update

Identified initial administrative record occupied cases

- Count agreement comparisons Adrec occupied determinations to where we had self-responses

- Preliminary analysis was showing 81 percent agreement and similar amounts of availability of age, sex, race and Hispanic origin as during tests.

Changing May vacant or delete addresses to full contacts

- May vacant or delete determinations to addresses in Early NRFU areas were changed to full contacts
    - 440,000 vacant and 200,000 delete addresses changed

- Distance function value in June modeling run was now outside the distance cutoff
    - 100,000 vacant and 8,000 delete addresses changed

Shape your future START HERE >

United States® Census 2020

Revised Administrative Record Modeling

Remaining Administrative Record Update

Early July Processing
- Additional one-visit occupied cases
- First batch of closeout occupied cases if a 1040 was returned for the address
- No closeout vacant or delete addresses in this run since we did not have the IRS delivery that included the July 15th submissions

Early August Processing
- Additional one-visit occupied and closeout occupied cases including additional sources and years used to support matching
- Additional one-visit vacant and delete cases that had over 90 percent probability of being unoccupied
- First batch of closeout vacant and delete addresses

Early September processing
- Additional one-visit occupied addresses
- Additonal closeout occupied, vacant and delete addresses

Post Processing
- Introducing a model to determine administrative record occupied addresses on American Indian Reservations

Shape
your future
START HERE >

United States®
Census
2020

Revised Administrative Record Modeling

AR Occupied modeling for American Indian Reservation

Original plan was for no administrative record modeling on reservations

To mitigate potential undercoverage, researched a similar approach to what we do for the rest of the country to determine occupied with a roster if unresolved after data collection is completed for self-resonse (TEA1) and update/leave (TEA6)

Developed a training modeling on 2010 American Indian Reservation addresses to apply to 2020 AIR address

Evaluation Statistics Using 2010 Census Data

Difference	AIR Models on AIR	NRFU Model on all NRFU*
AR count lower	14%	16%
Same count	67%	62%
AR count higher	19%	22%

*Results from Administrative Records Modeling Team CSAC March 2017 paper

Proportion of People in AIR AR Occupied MAFIDs with AR Demographic Information

	Age	Sex	Race/Hispanic Origin
AIR Models on AIR	100%	100%	94%
NRFU Model on all NRFU*	100%	100%	90%

*Results from Administrative Records Modeling Team CSAC March 2017 paper

Shape your future START HERE >

United States® Census 2020

Revised Administrative Record Modeling
Administrative Record Modeling results (as of August)

AR Determination	Count
Occupied June	8,035,000
Occupied July	356,000
Occupied August	697,000
Vacant Prior to August	8,503,000
Vacant August	619,000
Delete Prior to August	2,889,000
Delete August	521,000
Closeout Occupied July	770,000
Closeout Occupied August	826,000
Closeout Vacant August	909,000
Closeout Delete August	160,000

Note: Vacant and delete prior to August reflect the removal of the addresses changed in May and June processing
These address counts are based on a list of unresolved addresses for modeling on July 30, 2020.
The June mailing results is part of determination whether Vacant Prior to August and Delete Prior to August receive one visit

Shape
your future
START HERE >

United States®
Census
2020

Administrative Record Enumeration

If an address does not self-respond by the end of data collection, we need to make response records. If it is occupied, we need to enumerate the roster and their available characteristics.

Original specification used roster rules that reflected May and June IRS 1040

Concern because we started identifying occupied addresses in June but 1040 returns could be received in August and September after the July 15th filing deadline

Modified our roster formation for enumeration

- Started by using the roster developed for the housing unit when the AR occupied status was determined (An address identified in June would start with the June roster, July would use July,…)

- If an occupied address did not have a 1040 return when determination was made then we would look at the September 1040 returns the address. If all of the original people match to the 1040 return then we would add the additional 1040 people to the enumeration roster.

Shape your future START HERE >

United States® Census 2020

Off Campus Student Records

In May, the Census Bureau decided to contact universities and colleges to see if they could provide us students who were living off-campus in the spring 2020 semester
- Name, Date of Birth and Age
- Local off-campus address
- Alternative permanent address

Between June 18th and August 14th, Census Bureau contacted over 1,300 schools to see if they could participate.

Over 600 schools provided us with the information

- Assigning our Census Master Address File identifier (MAFID) to local off-campus addresses

Our research is leading us to determine how we can use this information in our post-processing operation to enumerate households that may be vacant, delete or unresolved after data collection is done.

- Develop rosters based on off-campus and the AR modeling roster sources that can be used for enumeration

- Designate addresses as "Occupied but population count unknown" if the off campus roster is incomplete.

Shape
your future
START HERE >

United States®
Census
2020

Summary

- Changes to the administrative record modeling due to changes to the IRS tax filing deadline to July 15, 2020 and the delayed start of the 2020 Census NRFU operation

- To mitigate potential coverage error, changes in May and June modeling resulted in 2.2 million addresses receiving full contacts that would have otherwise been classified as AR vacant and AR delete

- To mitigate potential undercoverage, analysis and changes to one-visit and closeout occupied determinations were implemented in June and July to compensate for change in tax filing deadline

- Introduced a model for American Indian Reservations to use after data collection is completed

- Roster determination for AR enumeration if a household does not self-respond were adjusted to allow household members included on later filed tax returns to be included

- Implementation of the off-campus student initiative to help improve the census results of students whose local school address may have been determined to have been unoccupied.

Shape
your future
START HERE >

United States®
Census
2020

Questions

What is your reaction to the administrative record usage for NRFU based on the NRFU operation changes and the the delay of receiving the IRS 1040 tax information?

2020CENSUS.GOV

Shape
your future
START HERE >

United States®
**Census
2020**