A Living Frame

Tracking Changes to the National Health Interview Survey's
Address-Based Sample Frame over Time

Bryan Schar
U.S. Census Bureau

Any views expressed are those of the author and not necessarily those of the
U.S. Census Bureau or the National Center for Health Statistics

Major Goal of NHIS Frame

- Design a frame of *locatable living quarters* that are likely to contain the target population of the National Health Interview Survey (NHIS) for a *ten year sample design*.
  - Locatable Living Quarters: Any site where people live, stay, or could live that can be located for in-person interviews
  - Want to track the status of living quarters over the life of a frame that is updated semi-annually
  - Maintain the high quality of the frame over its lifetime

United States™
Census
Bureau

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

Addresses Versus Living Quarters

- Design a frame of *locatable living quarters…*
  - A frame of sites
- Addresses are characteristics of living quarters and can be reconfigured (or changed) without new living quarters being created or an old ones being deleted
- We want to add and delete living quarters over the life of the frame to reflect changes on the ground
  - Part of this involves separating address reconfigurations in updates from true adds and deletes

Example of Address Reconfigurations

- Street name changes
  - 101 Main St becomes 101 Washington St
  - 102 Main St becomes 102 Washington St
  - …
- Unit designation changes
  - 101 Main St Apt 1 becomes 101 Main St Apt A
  - 101 Main St Apt 2 becomes 101 Main St Apt B
  - …
- ZIP code or ZIP+4 code changes

Examples of Living Quarters Changes

- A single unit structure is subdivided into a duplex
  - 101 Main St becomes 101 Main St A and 101 Main St B
- Existing living quarters no longer exists, due to demolition or other such event
- New living quarters are built

United States **Census** Bureau

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

Unique ID

- Unique ID is designed to be like the Social Security Number for living quarters

- A Unique ID was assigned to every address in the first delivery of the Marketing Systems Group (MSG) data in early 2015

- With each semiannual update, new Unique IDs are created when we determine that new living quarters exists and existing Unique IDs are retired for living quarters that no longer exist

Initial Frame Updating Methodology

- MSG received a delivery of the U.S. Post Service's Computerized Delivery System and No-Stat files along with additional Address Information System products from Compact Information Systems

- They compared this new address information to the addresses on their existing database
  - Any address that was in the new information but not on the existing database was added to the database and given a new Unique ID
  - Any address that was not in the new information but existed on the database was deleted and its Unique ID was retired
  - Information from any new address that matched an existing address was used to update the existing information for that record

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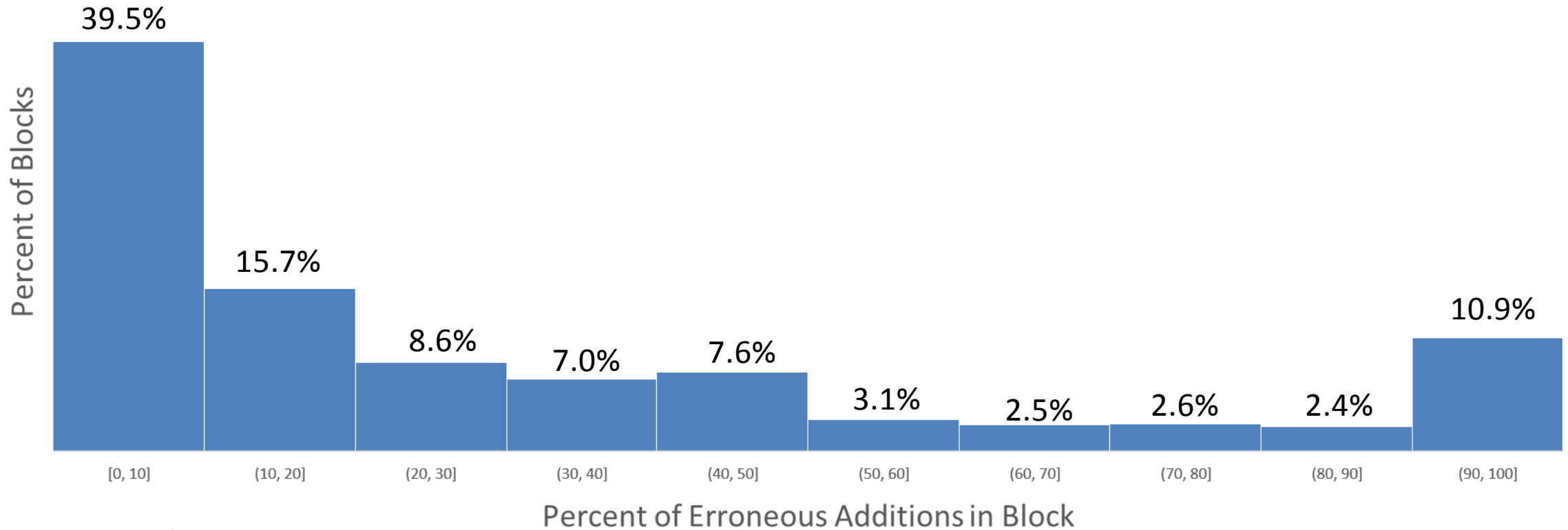U.S. CENSUS BUREAU
census.gov

Question

- What are the global and local impacts over the life of the NHIS address-based sampling frame of not correctly separating address reconfiguration from true living quarters adds and deletes?

National Impact

- Based on the methods we currently have developed, we identified approximately **0.01 percent** of the records eligible for sample are address reconfigurations that were treated as new living quarters (erroneous additions)
  - These records were given new Unique IDs but should have been recorded as changes to existing records
  - Almost all of these erroneous additions (**99.95 percent**) were from the first update to the frame because we addressed this issue early on in the frame development
- About **0.2 percent** of the blocks eligible for sample contained erroneous additions

United States™
Census
Bureau

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

Block Level Impact

Blocks by Percent of Erroneous Additions to Total Units

Percent of Blocks

39.5%

15.7%

8.6%

7.0%

7.6%

3.1%

2.5%

2.6%

2.4%

10.9%

[0, 10]   (10, 20]   (20, 30]   (30, 40]   (40, 50]   (50, 60]   (60, 70]   (70, 80]   (80, 90]   (90, 100]

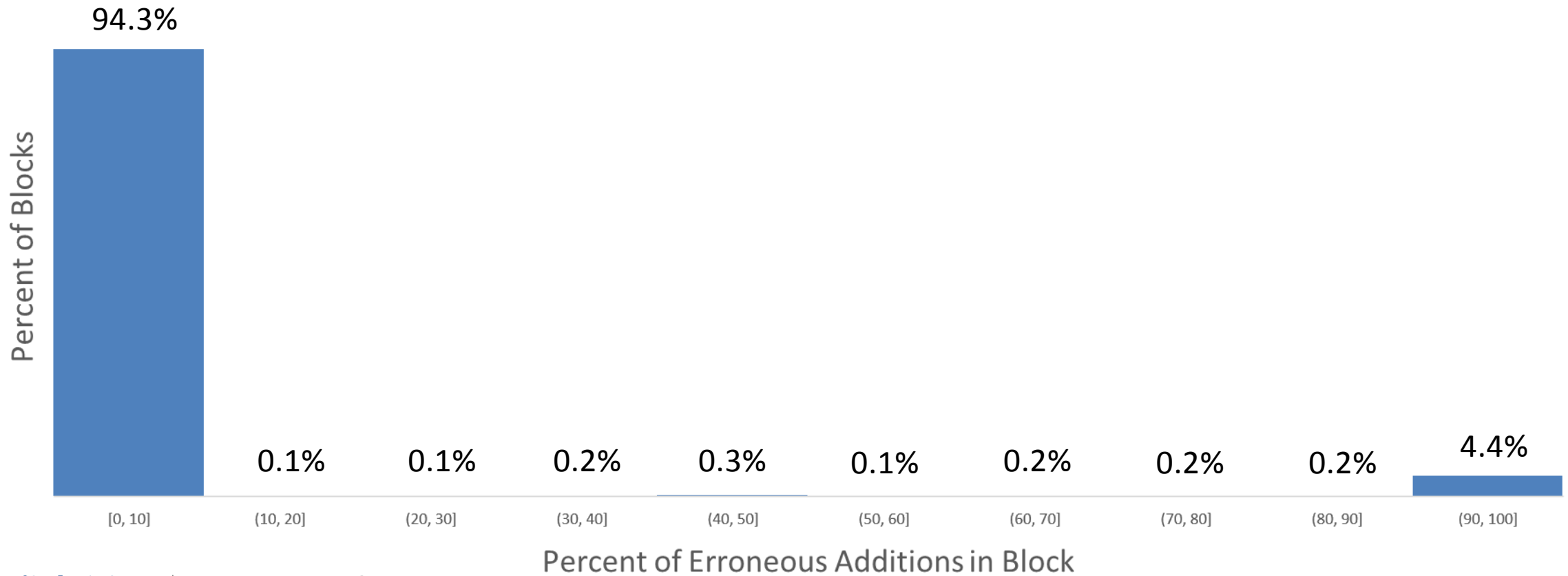Percent of Erroneous Additions in Block

Prevalence in Updates

- Almost **16 percent** of the adds that were eligible for sample in the first semiannual update were actually address reconfigurations that were erroneously treated as new units

- The erroneous additions were highly concentrated to specific blocks
  - **0.38 percent** of the updated blocks contained **50 percent** of the erroneous additions
  - All the erroneous additions were contained in **5.8 percent** of the updated blocks

Block Level Impact

Blocks by Percent of Erroneous Additions in First Update

94.3%   0.1%   0.1%   0.2%   0.3%   0.1%   0.2%   0.2%   0.2%   4.4%

[0, 10]   (10, 20]   (20, 30]   (30, 40]   (40, 50]   (50, 60]   (60, 70]   (70, 80]   (80, 90]   (90, 100]

Percent of Blocks

Percent of Erroneous Additions in Block

United States™
Census
━ Bureau ━

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

Conclusions

- The issues of address reconfigurations being treated as erroneous additions seems to have an overall limited impact on the quality of the frame
  - However, the local impact can be significant even if rare
- There appears to be a non-negligible number of erroneous additions in the updates when appropriate methods are not used to mitigate the issue
  - These erroneous additions are concentrated to specific blocks
- There are methods that can be used to mitigate this issue as MSG will discuss in their presentation
- Limitations
  - Have limited data to make this assessment
  - Our measures of erroneous additions are evolving

Future Work

- Measure the number of erroneous additions that would have occurred had they not been identified in our later updates
  - That is, had we continued to use the original update methodology what would have been the impact on the frame and on each update?
- Continue to improve our ability to identify address reconfigurations that result in erroneous additions
  - This includes investigating if we can better use U.S. Postal Service's products to identify address reconfigurations

Contact Information

- Bryan Schar
- bryan.g.schar@census.gov
- 301-763-8498

United States Census Bureau

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