Accurately Identifying and Enumerating Multi-unit Housing with Remote Sensing Data for Address Frame Enhancement

FCSM 2023, Washington DC

Lee Fiorio
Ned English

NORC at the University of Chicago

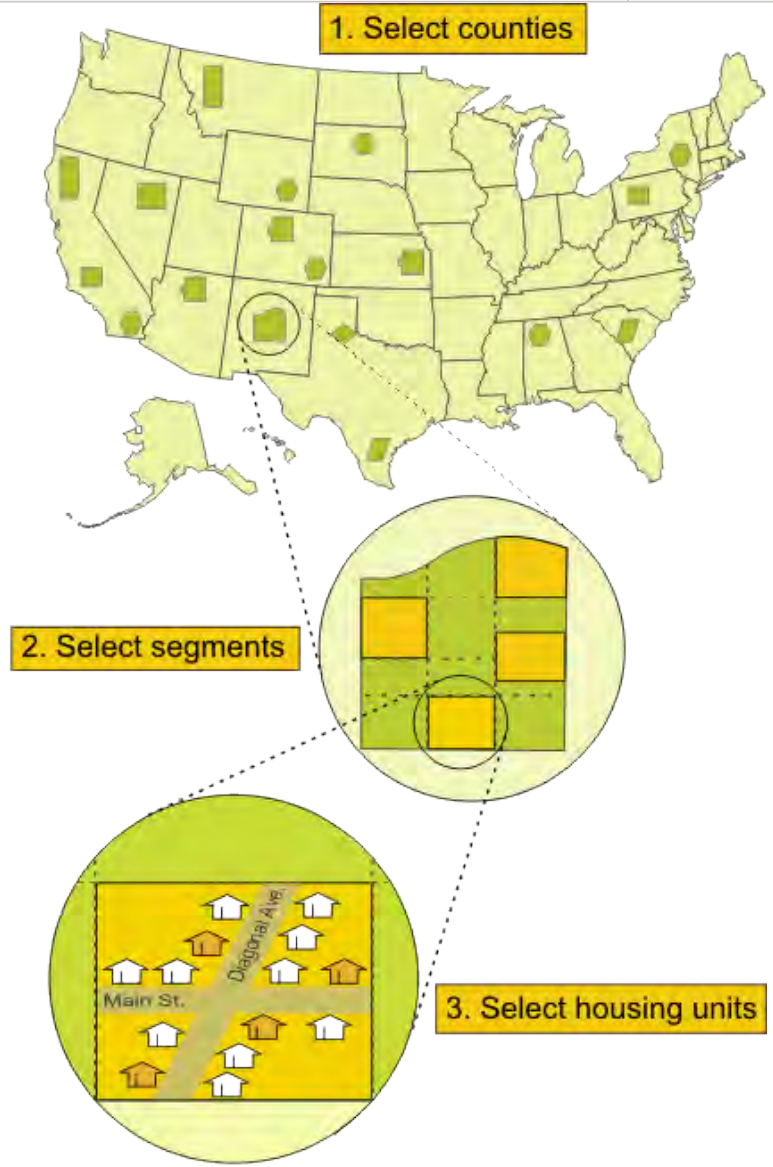
Area probability surveys typically require address lists

Main ingredient for address frame:

- Computerized Delivery Sequence file (CDS)
  - USPS address file

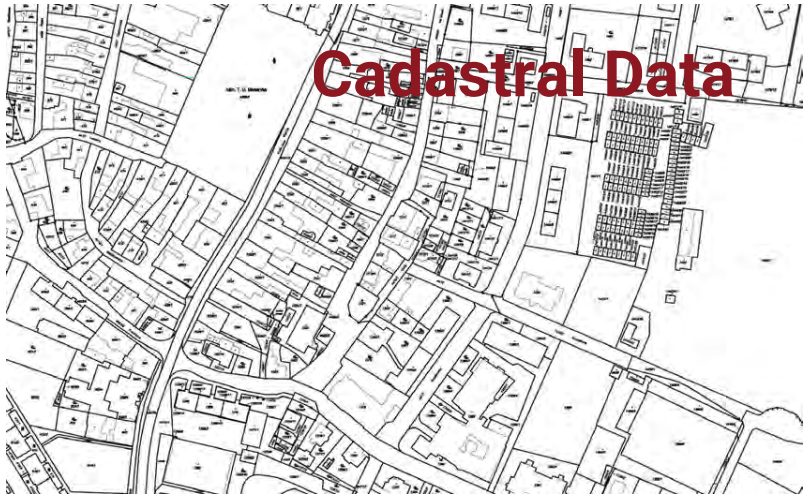
Need supplemental addresses in low-coverage areas

- 3,096 segments in NORC National Frame
- 5% needed enhancement
- *Important for reducing coverage bias*



Source: U.S. Energy Information Administration, 2009 Residential Energy Consumption Survey

Address listing remotely? Integrating CDS with spatial data



Source: Zamarova (2020)

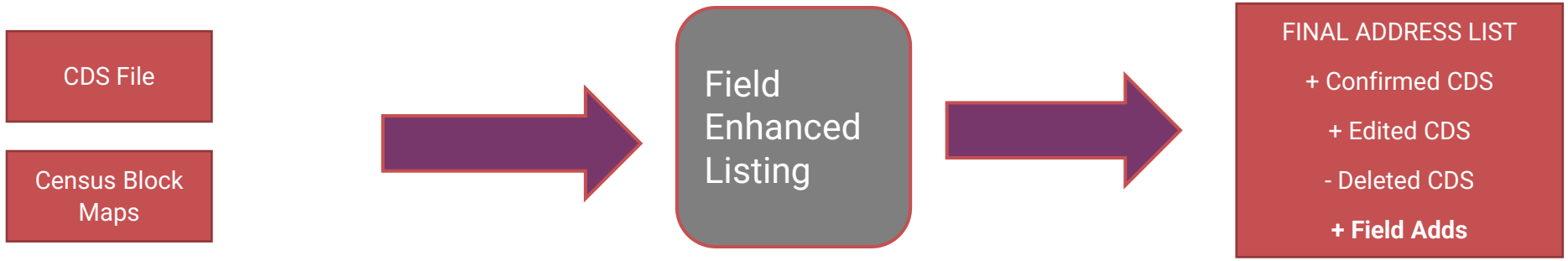
Source: OpenStreetMaps

Source: WikiCommons

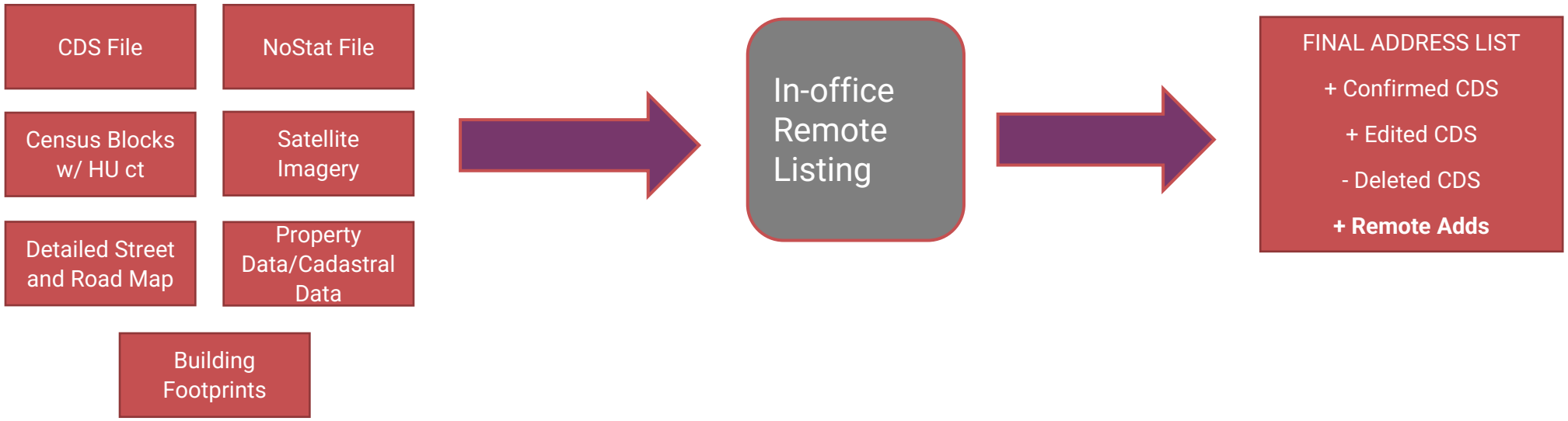
Source: OpenStreetMaps

Conceptualization for remote listing

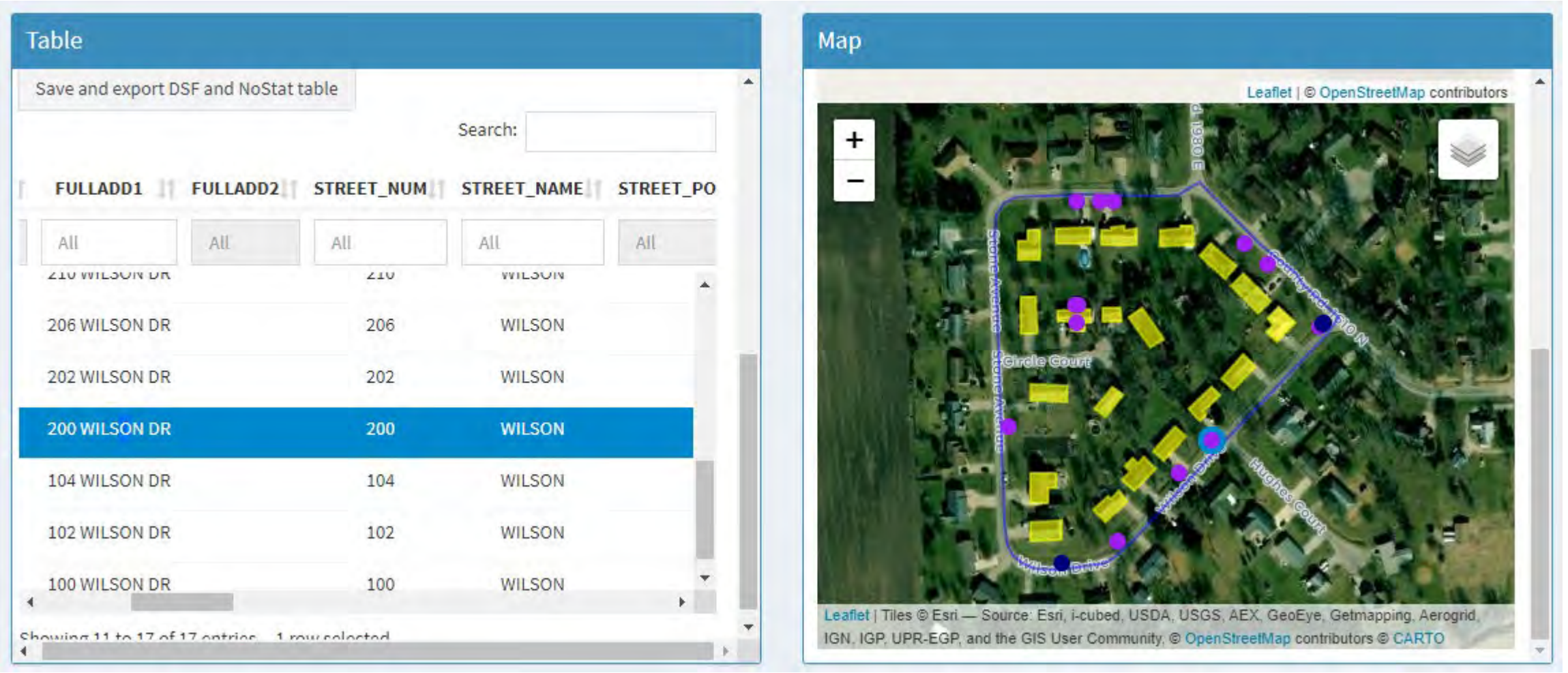
Enhanced Listing



In-office Remote Listing



NORC Interactive Mapping Address Listing Tool



Address lists produced with the remote listing tool were validated in a sample of blocks

Total Remote Listed

61 Segments

3,245 Blocks

~33,500 anticipated HUs (2020 Census)

~13,500 CDS addresses (40% coverage)

+30,500 Remote Adds

Total coverage: **130%**

Sample

8 Segments

74 Blocks

1,185 anticipated HUs (2020 Census)

415 CDS addresses (35% coverage)

+951 Remote Adds

Total coverage **115%**

Round 1 Validation: Summary Findings

Initial Address List

415 CDS lines

951 Remote Adds

1,366 Total addresses

Field-Validated Address List

389 validated CDS lines (accuracy 94%)

889 validated Remote Adds (accuracy 93%)

+100 Field additions (7% miss rate)

1,378 Total addresses

- In aggregate, CDS and Remote Adds were validated at a similar rate
- Similar number of addresses added by field as deleted (i.e. marked not valid)

What is our accuracy rate with multi-unit housing?

Remote listing requires reconciling three distinct forms of data

USPS Addresses (CDS)

- Mailing addresses
- Problems: low coverage; not necessarily linked to a physical address

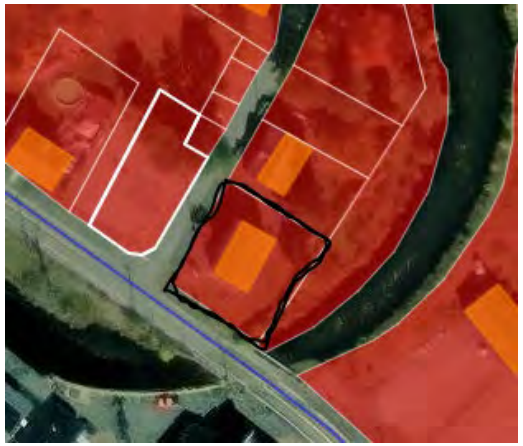
Satellite Data/Building Footprints

- Visible structures
- Problems: out buildings; multi-unit buildings; business vs. residential; no address info

Tax Parcels

- Properties
- Problems: multiple HUs on a property

Multi-unit problem: There is not a one-to-one relationship between addresses, visible structures, housing units, or properties



NORC's solution: detailed protocol for handling different multi-unit scenarios

Scenarios

- Apartment buildings, duplexes, row homes
- Manufactured home communities
- Multi-HU parcels: e.g. trailer behind house; multiple single-family homes on a lot

Manual review

Contextual data

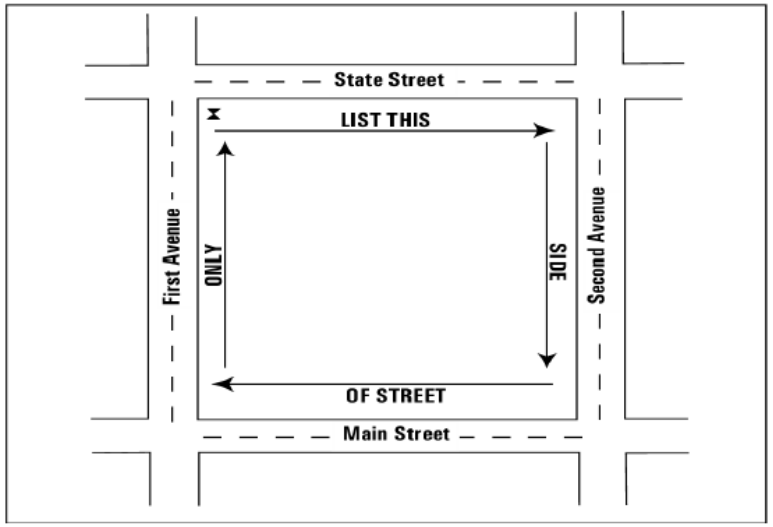
- Google StreetView, real estate listings, tax records

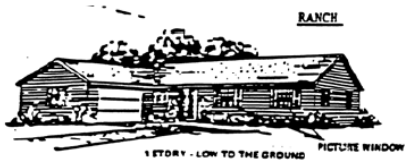
For the 2020 National Frame, 61 segments were listed with the Remote Listing Tool

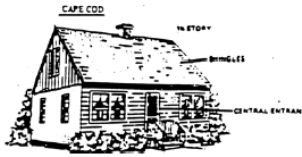
1. What percentage of addresses added with remote listing tool are multi-unit?
2. How well does the remote listing protocol capture multi-unit scenarios?
    a. Are the multi-unit adds accurate?
    b. How many multi-units are missed?
    c. Is there anything systematic about the inaccuracy?

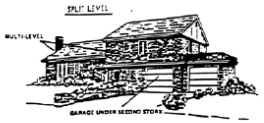
NORC Field Staff used standard NORC 'Enhanced Listing' procedure to validate 6 remote listed segments in-person

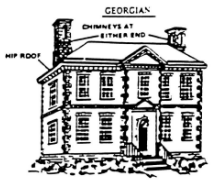
6 segments in Massachusetts, Pennsylvania, Indiana, Louisiana, and California

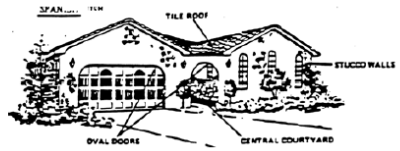












Two FIs reviewed lists block by block

- Starting in NW corner
- Traveling in clockwise direction
- **Confirming** existing addresses
- **Deleting** non-existent or non-valid addresses
- **Adding** missing addresses

RQ1: What percentage of addresses added with remote listing tool are multi-unit?

Sample

6 Segments

40 Blocks

808 anticipated HUs (2020 Census)

306 CDS addresses (38% coverage)

+632 Remote Adds

Total coverage **116%**

Remote Adds

422 single HU (67%)

210 multi-unit HU (33%)

Multi-unit scenarios made up one third of all addresses added

RQ2: How well does the remote listing protocol capture multi-unit scenarios?

How accurate were the adds?

92% single HU remote adds confirmed in the field

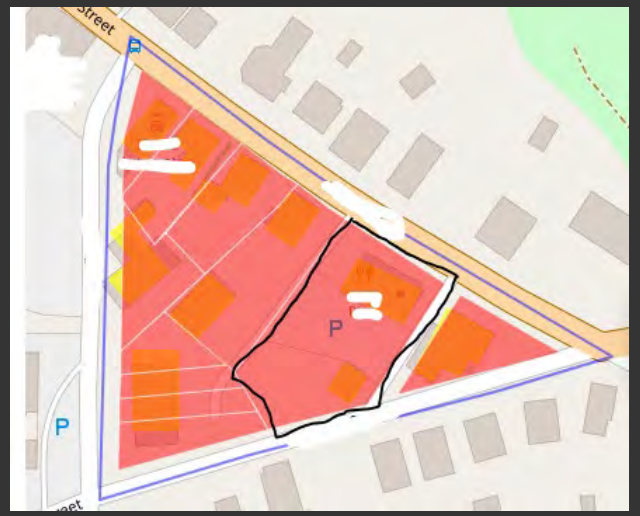
79% multi-unit HU remote adds confirmed by field

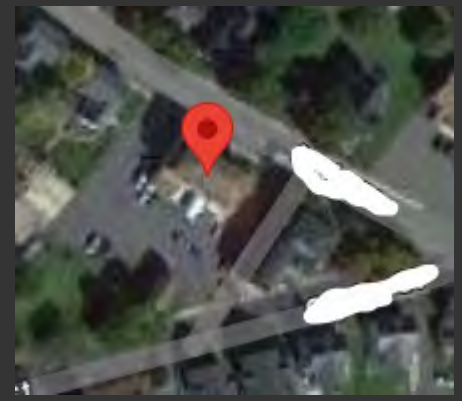
How many multi-units were among the misses?

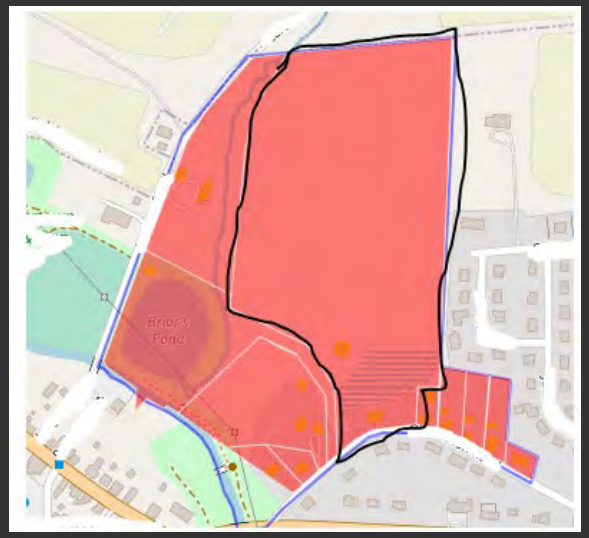
Multi-unit HUs comprised 30% (28 of 95) of the addresses missed

Multi-unit HUs confirmed at a relatively high rate (but lower than single HUs)

Multi-unit HUs do not make up a disproportionate number of the HUs missed


Street
P
P




Bride's
Pond

Multi-unit HUs pose a serious challenge to remote listing…

… but this challenge is (somewhat) mitigated by having a detailed protocol for handling wide variety of multi-unit scenarios

In a sample of 6 segments enhanced using remote listing tool

- 33% of all remote adds were generated from a multi-HU scenario
- Multi-unit remote adds were validated at a rate of 79%
- Missed addresses were not disproportionately multi-unit

Benefits of Remote Listing

Considerably more cost-effective and less time-consuming than field listing

It is easier to refresh remote lists (i.e. field lists are ‘frozen’ for the duration of the frame)

Integrating address lists with other kinds of data, like property records and aerial imagery, has other useful applications

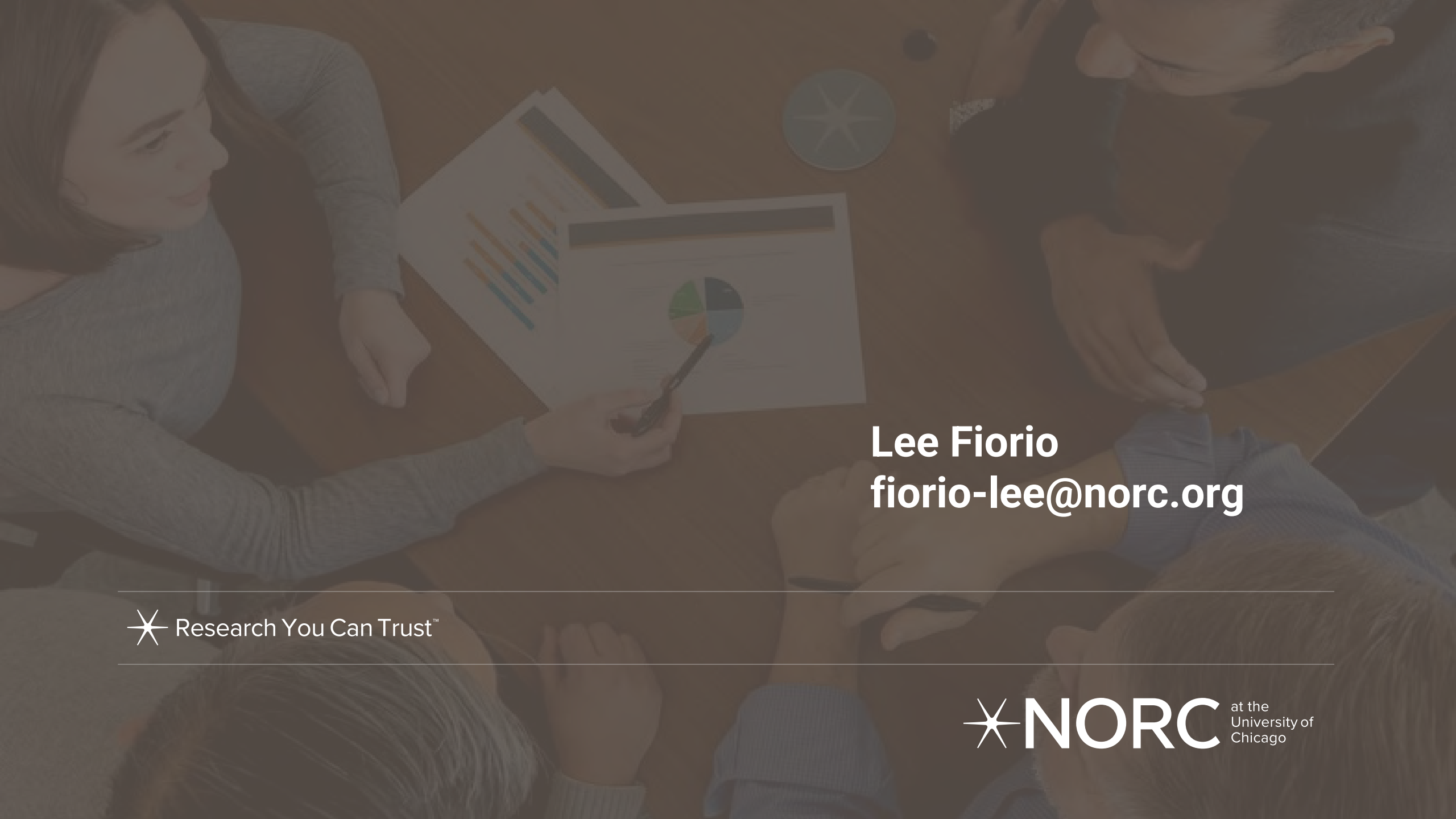
Limits of Remote Listing and Next Steps

Quality is highly dependent on accuracy of administrative data

- Does not work well at all in places where administrative data contain limited or inaccurate address information

Identify covariates associated with accuracy of remote lists

- Using segment level data from ACS on demographics and socio-economic characteristics to model list accuracy (if sample size not too small)



Lee Fiorio
fiorio-lee@norc.org

Research You Can Trust™

NORC at the University of Chicago