



Give a Second Thought to the Secondary City: New applications of the USPS City State file

Derick Brown, RTI International
Joe McMichael, RTI International

AAPOR 2018 Denver, Colorado

Let me tell you a story...

- A boy from a small town in Maryland wants to look up his home town on the vendor provided Computerized Delivery Sequence File (CDS): Eldersburg

- A frequency of cities in Maryland on CDS the beginning with an “E”

City	Count
EARLEVILLE	2,656
EAST NEW MARKET	1,507
EASTON	15,723
ECKHART MINES	33
EDEN	1,679
EDGEWATER	10,454
EDGEWOOD	10,893
ELK MILLS	252
ELKRIDGE	14,288
ELKTON	22,064
ELLERSLIE	441
ELLCOTT CITY	37,761
EMMITSBURG	2,831
ESSEX	20,223
EWELL	296



What happened to Eldersburg?

- Very strange: Eldersburg definitely still exists, right?
- What about looking for Eldersburg's ZIP code? 21784

ZIP code	Count
21778	557
21779	523
21780	883
21782	2,308
21783	4,164
21784	17,012
21787	4,955
21788	5,393
21790	142
21791	2,513
21793	4,478
21794	1,141
21795	4,974

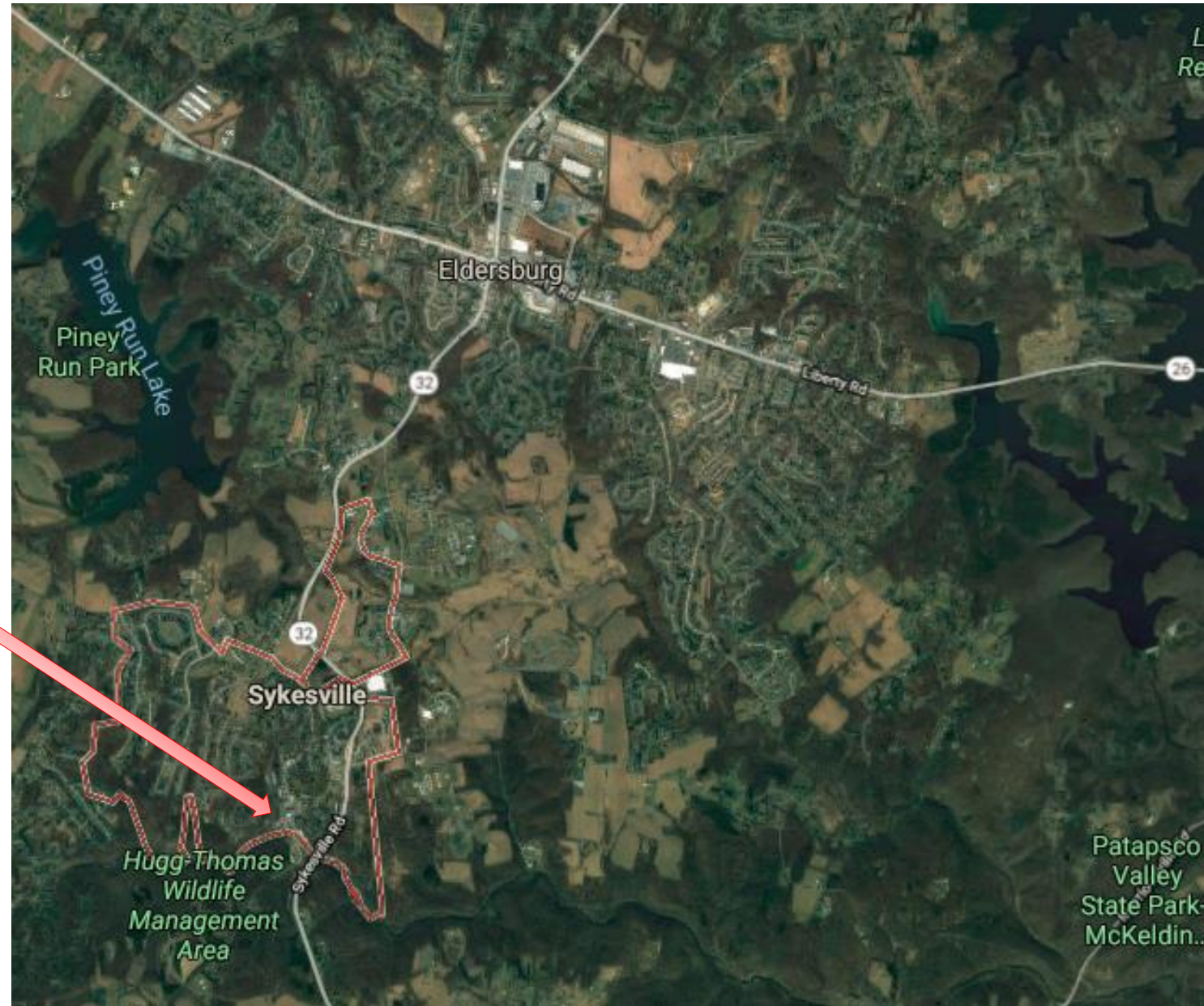
What happened to Eldersburg?

- What city corresponds to 21784 in the CDS?
- Sykesville:
 - Small town
 - Constantly mispronounced
 - The only city name associated with 21784 on the CDS

ZIP Code	City	Count
21778	ROCKY RIDGE	557
21779	ROHRERSVILLE	523
21780	SABILLASVILLE	883
21782	SHARPSBURG	2,308
21783	SMITHSBURG	4,164
21784	SYKESVILLE	17,012
21787	TANEYTOWN	4,955
21788	THURMONT	5,393
21790	TUSCARORA	142
21791	UNION BRIDGE	2,513
21793	WALKERSVILLE	4,478
21794	WEST FRIENDSHIP	1,141
21795	WILLIAMSPORT	4,974

Eldersburg vs. Sykesville: On a map

- Eldersburg
 - North of Sykesville
 - Mostly consists of the area surrounding the intersection of route 26 and route 32
 - Suburbs
- Sykesville
 - Primarily consists of main street and surrounding residential areas



Eldersburg is a secondary city

USPS City State Detail Record Data File – all records with ZIP code 21784

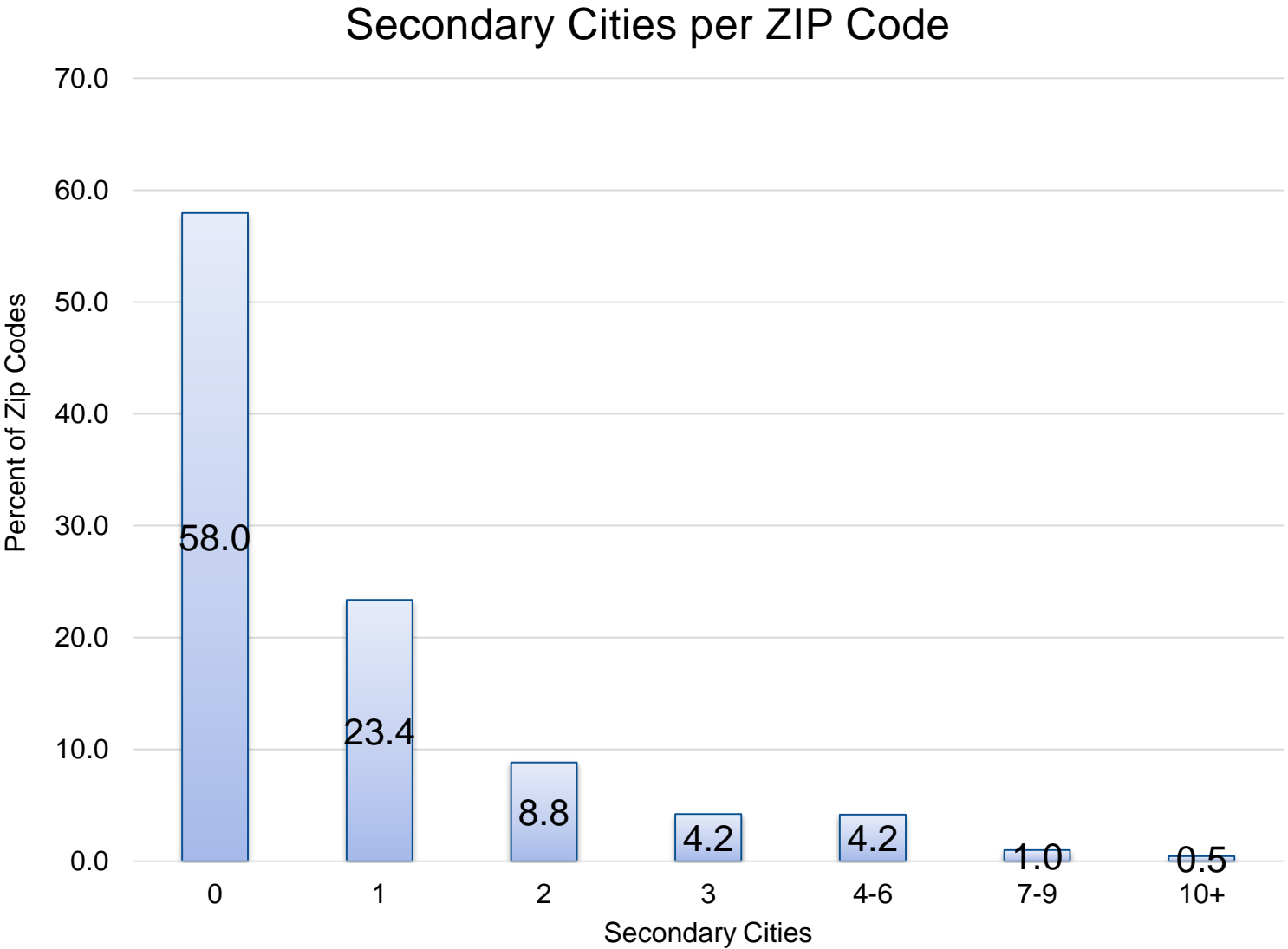
ZIP code	Preferred Last Line (PLL) City Name	Secondary City Name
21784	SYKESVILLE	ELDERSBURG
21784	SYKESVILLE	GAITHER
21784	SYKESVILLE	SYKESVILLE

- Sykesville is the **PLL city** for ZIP code 21784
 - Preferred last line on a piece of mail
 - There is only one PLL city per ZIP code
- Eldersburg and Gaither are **secondary cities** for ZIP code 21784
 - There can be multiple secondary cities per ZIP code

City State Product

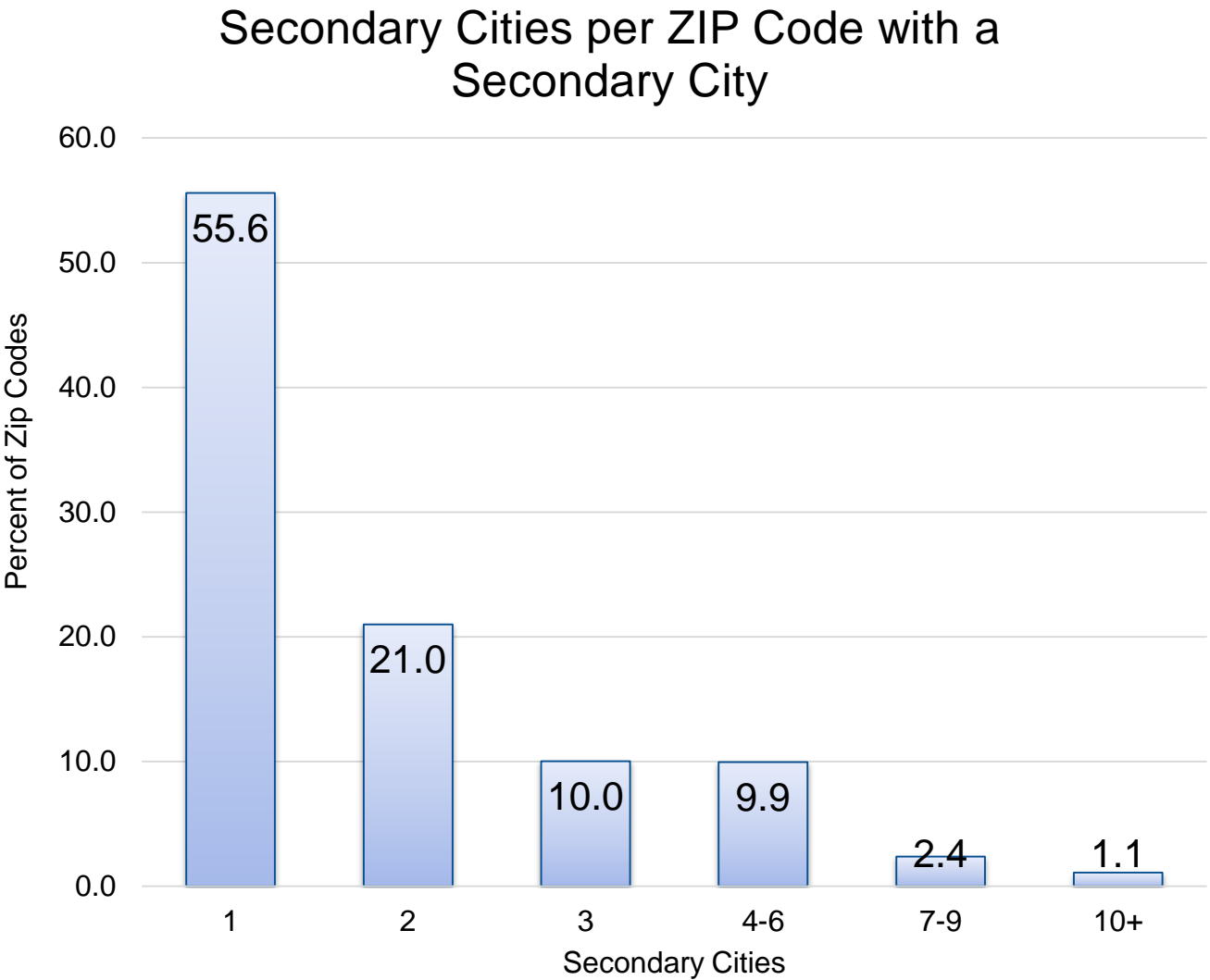
- Provided by the USPS
- Assists with mailing standardization
- Six different data files encapsulated within the City State Product
- "Detail Record" contains one record for each secondary city and ZIP code combination in the United States
 - Corresponding ZIP code
 - Corresponding PLL city
- There is no one-to-one mapping from CDS address to secondary city

Secondary Cities per ZIP Code



- .87 secondary cities per ZIP code
- 58% of ZIP codes have no secondary city
- 23% of ZIP codes have a single secondary city

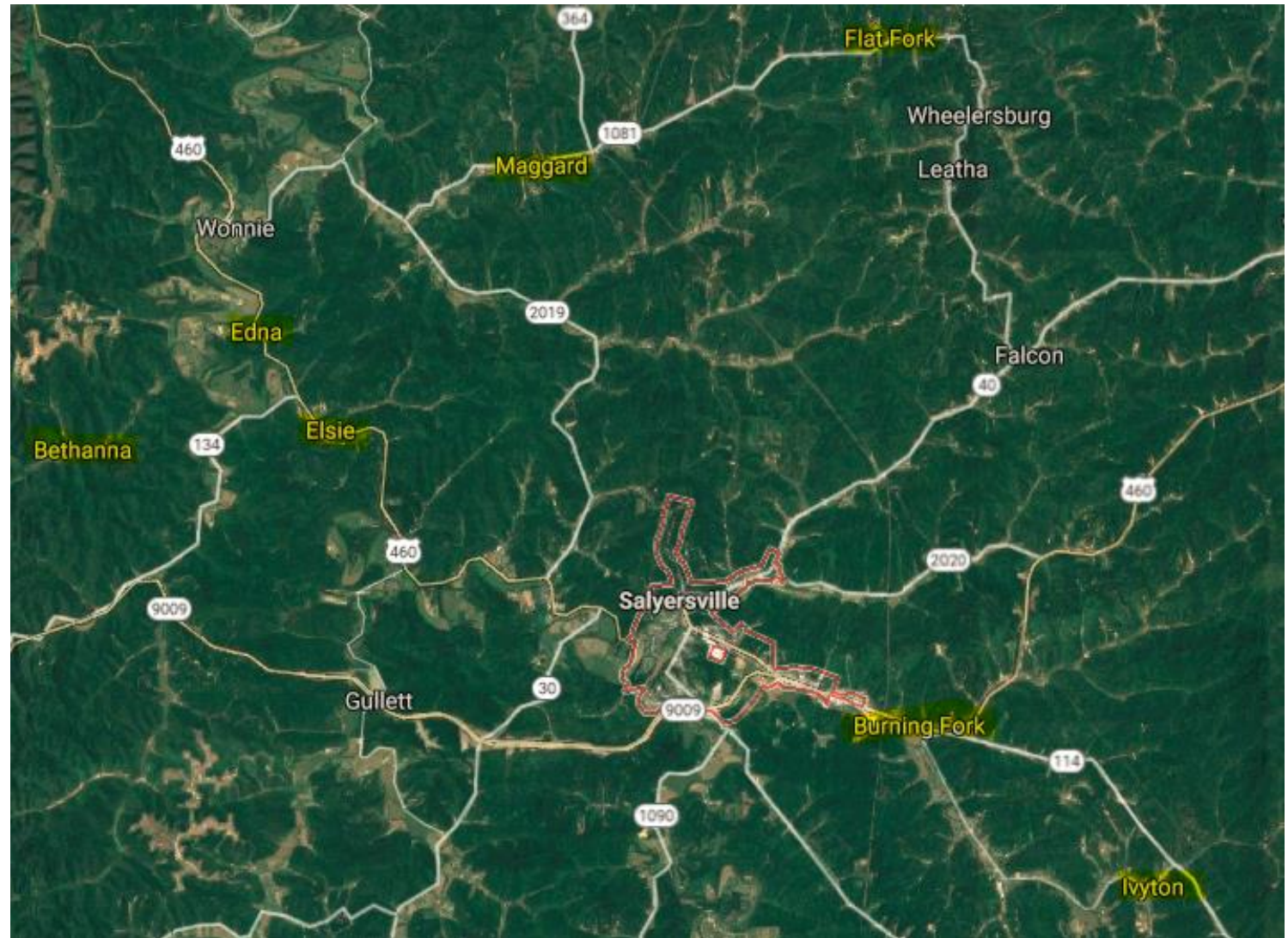
Secondary Cities per ZIP Code with a Secondary City



- 4.0 secondary cities per ZIP code
- 56% of ZIP codes have a single secondary city
- The most secondary cities in a single ZIP code is 31!

Salyersville, Kentucky 41465

- Bethanna, Burning Fork, Carver, Cisco, Conley, Cutuno, Cyrus, Duco, Edna, Elsie, Ever, Flat Fork, Foraker, Fredville, Fritz, Gapville, Gifford, Hager, Harper, Hendricks, Ivyton, Lickburg, Logville, Maggard, Marshallville, Mashfork, Salyersville, Seitz, Stella, Sublett, Swampton
- 31 secondary cities
- A small rural mountain city in eastern Kentucky



Secondary Cities and ABS

- Potential applications of secondary cities for ABS frames and studies
 1. Geocoding
 2. Address Matching
 3. Field work

Question #1: Can secondary cities be used to improve geocoding?

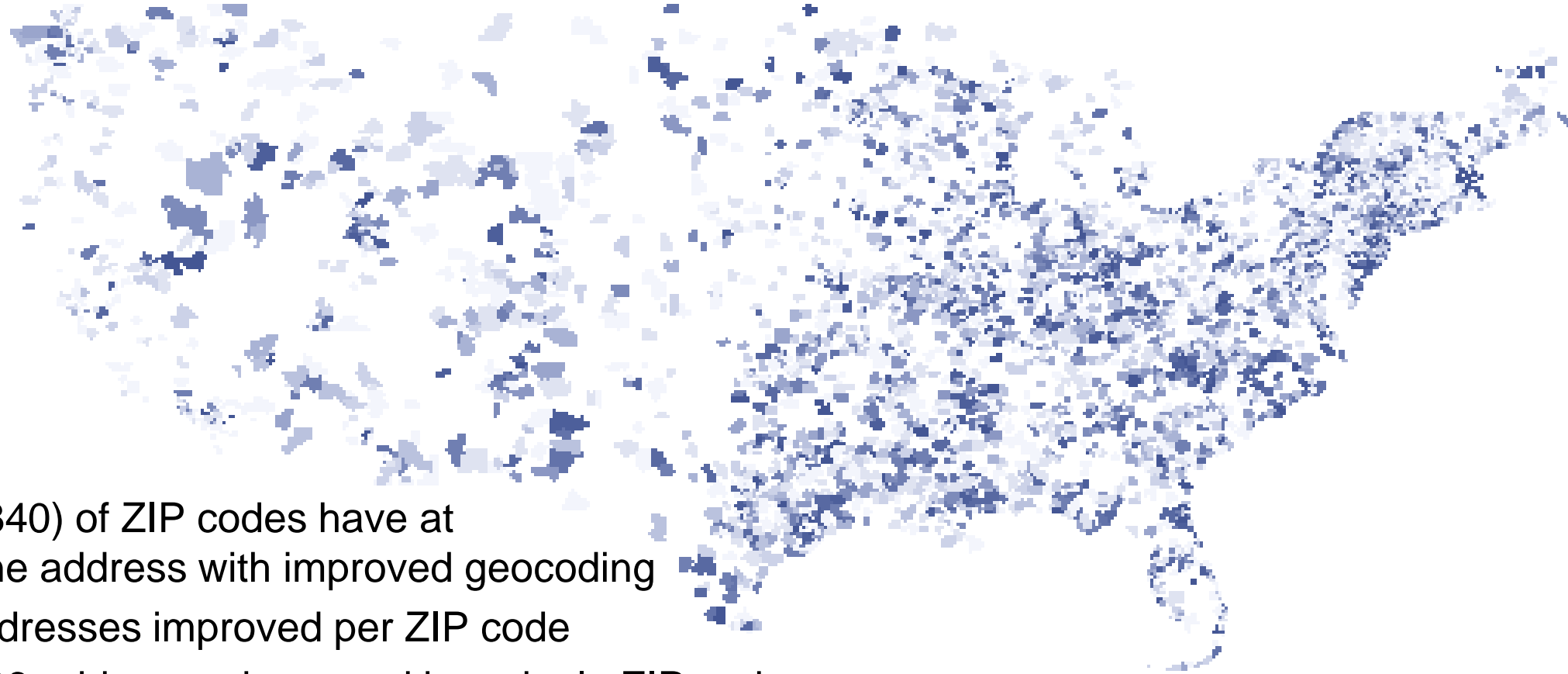
- We geocoded the July 2017 residential CDS
- 31.1% (n=44.4 million) of addresses coded to something less precise than a point address
- 21.7 million addresses of the 44.4 million have at least one secondary city associated with the address's ZIP code
- Geocoded these 21.7 million addresses using secondary cities
 - 57.4 million records attempted for secondary city geocoding (2.65 records per an address)

Question #1: Results

- .15% precision improved, 93.83% stayed the same, and 6.02% precision decreased
 - Considering the secondary city record(s) with the most precise geocode level
- The number of addresses on the residential CDS for which this improves geocoding is very small, .022% (n=31,766)
- There still is utility for these results
 - Addresses with improvement are clustered geographically

Addresses with improved geocoding precision by zip code

Density map of ZIP code tabulation areas (ZCTAs) and number addresses with improved geocoding



- 11.8% (n=4,840) of ZIP codes have at least one address with improved geocoding
- Median: 3 addresses improved per ZIP code
- Maximum: 426 addresses improved in a single ZIP code
- 55% of addresses with improved geocodes occur in 10% of ZIP codes with improved geocoding
 - Improvements clustered by ZIP code

Question #2: Can Secondary Cities and PLL Cities assist in address matching?

- Problem: Non-standardized addresses to be merged to a standardized frame
- Example non-standardized addresses:
 - Field enumerated addresses
 - Survey responses
 - Consumer databases
 - Registries
- Solution: Convert the city on non-standardized address to the PLL city

Question #2: Example

- Merging a national listing of 10,445 field enumerated (FE) addresses against the CDS file
- Initially, 43.5% (n= 4,547) of FE addresses matched to the CDS
- Converted the FE address city to PLL city
- 54 additional FE addresses matched to the CDS
 - 1.2% of all FE addresses that matched

Question #3: Can secondary cities assist with locating addresses in-person?

- Example: secondary cities as an obstacle to field staff
- Provide field staff each secondary city associated with the ZIP code
- Currently implementing this on a project
 - Too early to provide results
 - Difficult to measure this method's efficacy

Conclusions

- Secondary cities are an interesting thing to explore
 - The rabbit hole is seemingly endless
- Secondary cities can be utilized to improve geocoding
 - Improvements are clustered by ZIP codes
- Matching addresses
 - Utilizing the PLL city can result in improved address match rates
- Locating addresses
- Other ABS applications?



delivering **the promise of science**
for global good



Derick Brown
919.541.7420
derickbrown@rti.org