Geographical Information Improvements

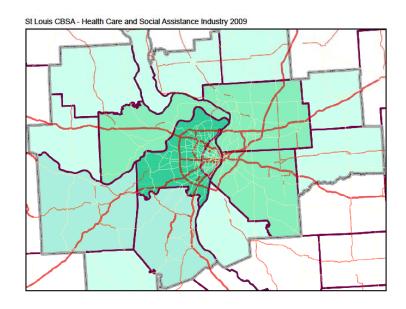
Carl Anderson, GISP

LED Partnership Workshop Arlington, VA March 9-10, 2011

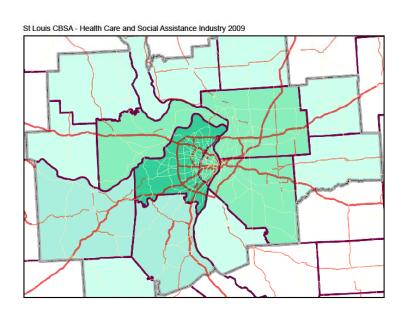
Administrative record-based systems like OnTheMap use the universe of outcomes of interest measured directly at the geographical level of interest.*

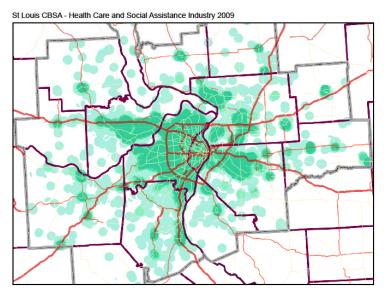
^{*} John M Abowd - APDU Newsletter Vol 33 No 2

County Level Detail



More detailed Geography, Allows for much more useful detail





Local Employment Dynamics

Local

Geography

Addresses

Geography

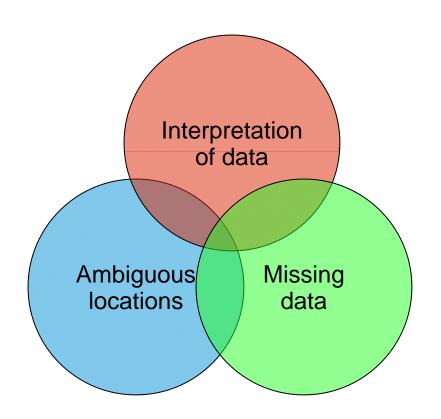
Worker/ Establishment

Geography

Challenges in Addressing

- Jurisdictions that have no addresses
- Unusual addressing schemes
- Establishments located by ZIP code only
- Establishments located without ZIP code
- Changes in ZIP code vs the longitudinal data
- Timing of Changes How and when do changes appear in administrative records

Causes of Addressing Challenges



Missing data

Longitudinal nature of the records

2006Q2	ABC Ltd	451 Maple St	Anytown		20789
2006Q3	ABC Ltd	451 Maple St			20789
2006Q4	ABC Ltd	451 Maple St	Anytowwn	ОН	20789
2007Q1	ABC Ltd	451 Maple St	Anytown	ОН	20789
2007Q2	ABC Ltd	451 Maple St	Anitown		20789
2007Q3	ABC Ltd	451 Maple St	Anytown	ОН	
2007Q4	ABC Ltd	St Rte 65 & 451 Maple St	Anytown	MD	21789
2008Q1	ABC Ltd	451 Maple St		MD	21789
2008Q2	ABC Ltd	451 Maple St	Anytown	MD	21789

^{*} Construed data, for use as an example only

Interpretation of Addresses

- Postal vs Colloquial Addressing
- Develop strategies for specific unusual Address forms

Example:

40 Miles N of Sand Creek, Barrow Ak US 4 & Rt 401 / 407 Westhill Road

Interpretation of Addresses Cont'd

Requires Developing Careful Strategies

Example:

403-22A Main Street NW

West 403 Main Street North Apt 22A

N403W22 Main Street Apt A

Ambiguous Addresses

- Location may be indistinct unless all address elements are perfect
- Uncertain geography

Example:

```
2<sup>nd</sup> Ave SE, Atlanta GA, 30317

2<sup>nd</sup> Ave NE, Atlanta GA, 30030

2<sup>nd</sup> Ave, Atlanta GA, 30032

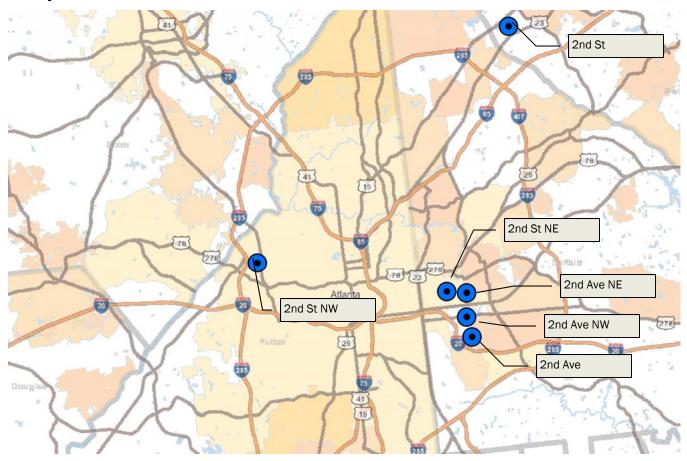
2<sup>nd</sup> St, Atlanta GA, 30360

2<sup>nd</sup> St NE, Atlanta GA, 30317

2<sup>nd</sup> St NW, Atlanta GA, 30318
```

Ambiguous Addresses

Any one of 3 Counties and 4 Cities



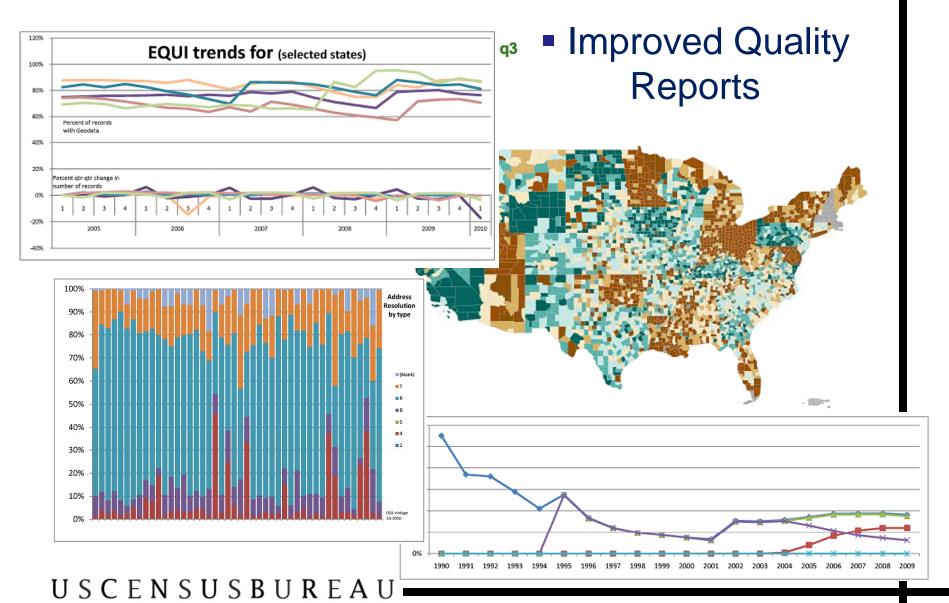
What is LEHD doing to improve Geography & Addressing?

- 1. Redesigning Address Geocoding Process:
- Improve interpretation of varying address styles (formats, special cases etc.)
- Improve handling of missing data
- Improve spatial accuracy
- Utilize longitudinal address review
- Utilize supplemental reference address data
- 2. Building improved tools to monitor the quality of Addresses & Geography within LED

Geocoding Process Improvements

Existing Redesigned **Data Sources** ES-202 + BR + Lat/ **ACS POW** AHS ES-202 SSEL MAF **ACS POW** AHS SSEL MAF CPR Lat/Long Long Apply cross-walk Collect Addresses connecting Data Sources By entity Key Mix Address Data from All Sources to the entity Key Data Preparation and Mix Address Data from Combination Sources with no Build longitudinal "Best Observation" connection to the entity table for entities data Key Longitudinal Edits To infill missing observations and partial observations Match. Standardize. Match, Standardize, Geocoding and Geocode, and De-Duplicate Geocode, and De-Duplicate Matching Linking and Link Datasets using Matched Addresses Update entity Keys Link Datasets using Matched Addresses Preservation of Keys Add Preliminary Add Census Geographic Codes Build Add Census Geographic Codes Census Geographic Cross-Walk based on links into Census Bureau based on links into Census Bureau Build Codes Master Address File **Tables** Master Address File Cross-Walk Tables Validate assigned Census Geographic Codes using information in Data Sources Augment Data, Adding Missing Approximate Census Geographic Codes Approximate Census Geographic Codes using range searches into Census Geographic using range searches into Geocoded Geocoded Addresses Codes Extrapolate Census Geographic Codes Extrapolate Census Geographic Codes From other data (Tiger, TFN, NSDI) From other data (Tiger) Finalize Data Finalize Geographic codes with additional Finalize Geographic codes with additional attributes attributes

Quality Improvements & Visualization Cont'd



How Can State LMI offices Help?

- Assist in understanding History of data submissions
- Participate in Data Quality Feedback Loops
- Act as Liaisons to other State offices who collect and record address data

Questions?

CES.Local.Employment.Dynamics@census.gov