

Comparing Person Records on the Demographic Frame and LEHD's Residence Candidate File

Deliverance Bougie

AAPOR | May 16, 2024



Gap Analysis for the LEHD

- The Demographic Frame (Demo Frame) is expected to provide information about the residential location of jobholders for the Longitudinal Employer-Household Dynamics (LEHD) program
 - Residence Candidate File (RCF)
 - Residence for workers derived from administrative records
- The Gap Analysis evaluated how well the Demo Frame overlaps with the RCF through PIK and MAFID-level comparisons
 - Do we find the same people and do we put them at the same address?*
 - If not at the same address, do they match as some level of geography?*
- Can the Demographic Frame feasibly replace the RCF?
- What improvements are needed to meet the needs of the LEHD program?

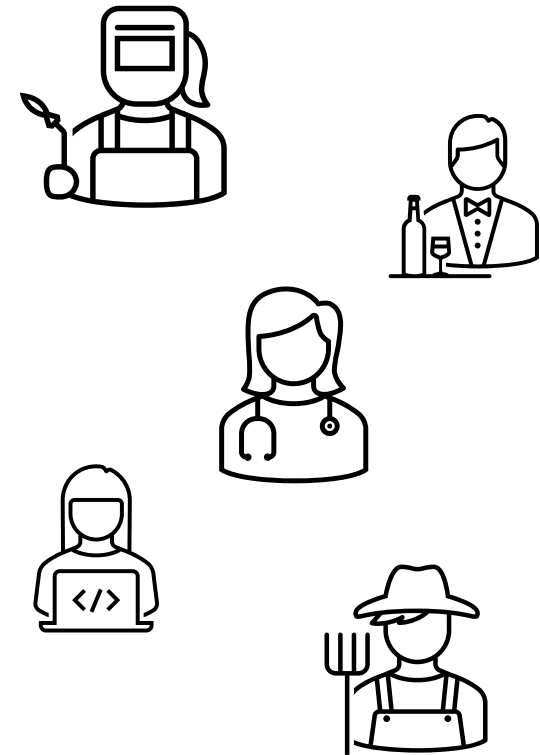
Data used

- Data sources
 - 2020 Residence Candidate File
 - 2020 Demographic Frame
 - LEHD Origin-Destination Employment Statistics (LODES) jobholder PIKs



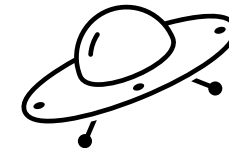
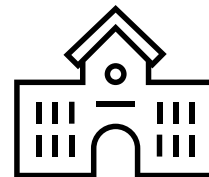
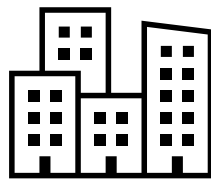
Who are Jobholders?

- Jobholder:
 - Employed on April 1, 2020
 - Anyone who earned at least \$1 in the second quarter
 - Ages 14-99
- Compared jobholders with RCF and Demo Frame



What is the 'best' address?

- RCF and Demo Frame are derived from administrative data
- RCF and Demo Frame utilize a Person-Place Model to select the 'best' address when a person is associated with more than one address
- MAFID identified as 'best' address was selected from each for comparison

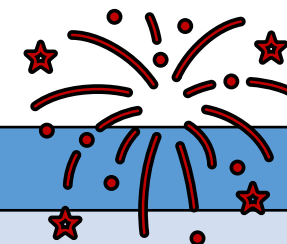


Jobholders in RCF and Demo Frame

	In RCF	% of job PIKs	Not in RCF	% of job PIKs	Total
In Demo Frame	121,500,000	97%	291,000	>1%	121,800,000
Not in Demo Frame	101,000	>1%	3,459,000	3%	3,560,000
Total	121,601,000	97%	3,750,000	3%	125,360,000

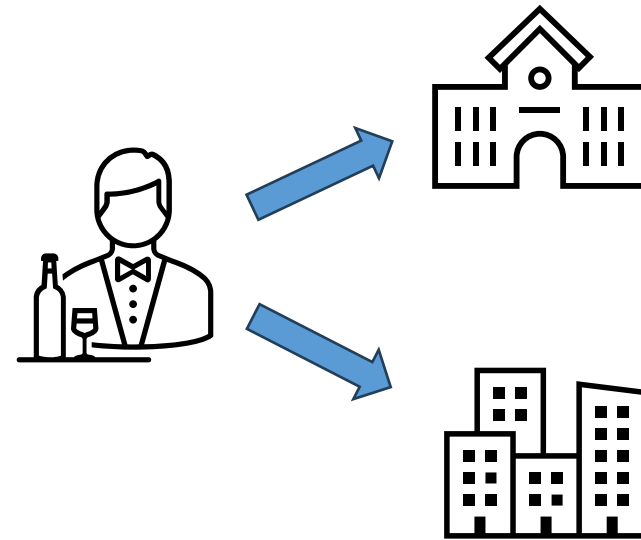


Not in RCF	Not in Demo Frame	Net Gain
291,000	101,000	190,000



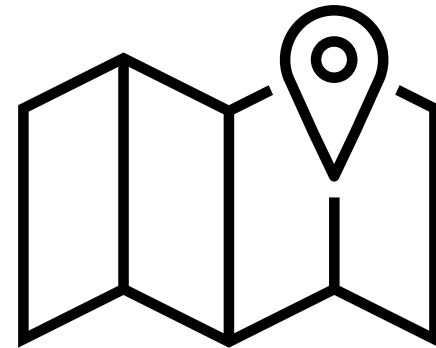
Are jobholders found at the same residence?

Residential location in RCF and Demo Frame	Count	Percentage
Residential location matches	108,600,000	89%
Residential location does not match	12,950,000	11%
Total of PIK that overlap	121,550,000	



Non-matching MAFIDs

- Where residential location does not match, at what level of geography does the Demographic Frame location match the RCF location?
- Different use cases may not need specific MAFID match
- RCF MAFID is the reference location



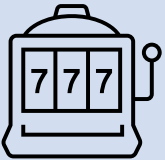
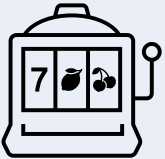
Geographic Matches



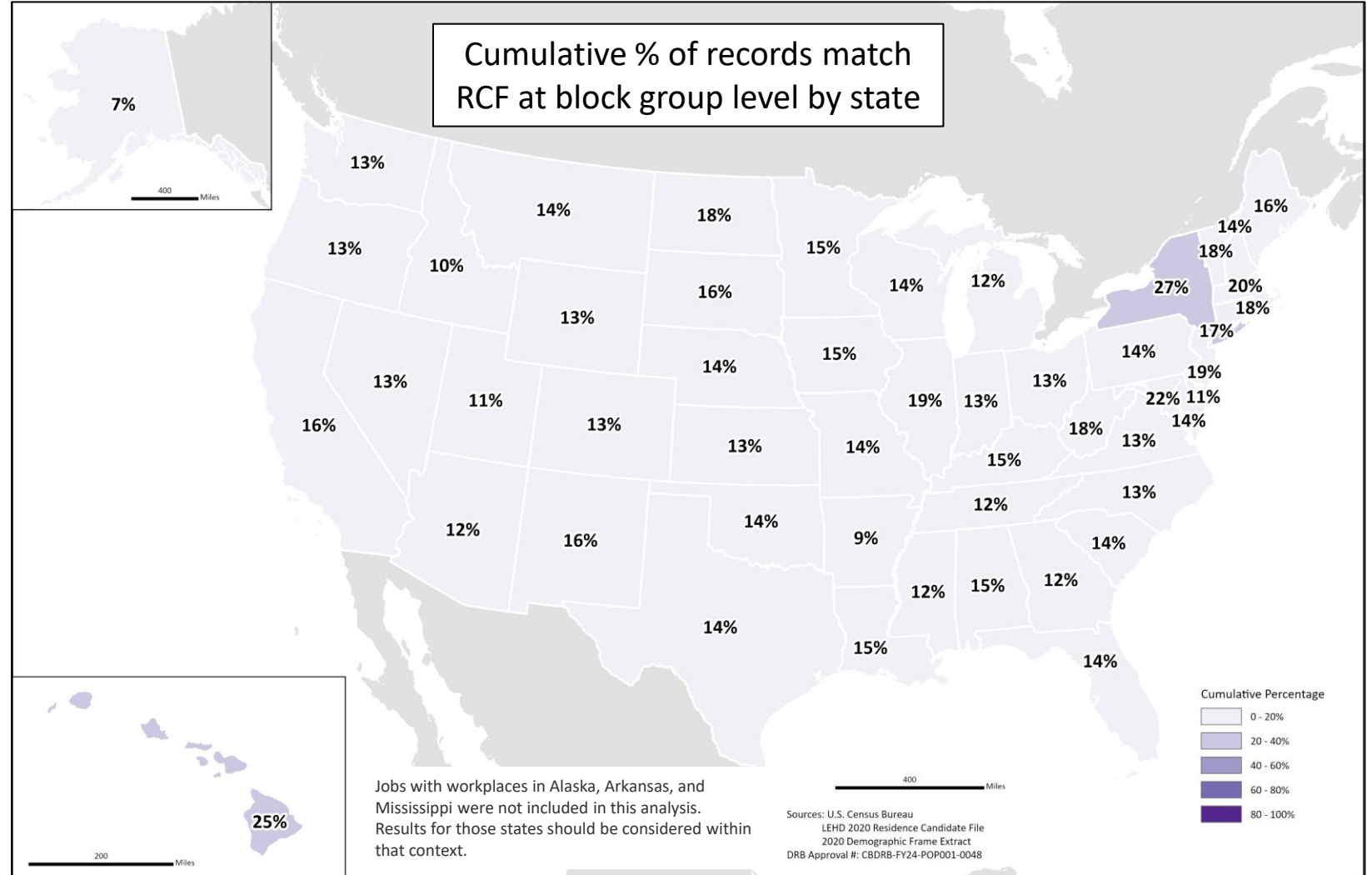
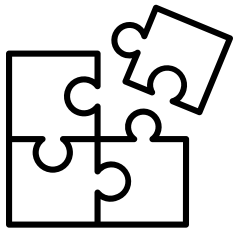
Geography	Count	Cumulative Count	Percent	Cumulative Percent
Block	1,416,000	1,416,000	11%	11%
Block Group	523,000	1,939,000	4%	15%
Tract	421,000	2,360,000	3%	18%
County	5,742,000	8,102,000	44%	63%
State	2,811,000	10,910,000	22%	84%
States Differ	1,868,000	12,780,000	14%	99%
State Missing	167,000	12,950,000	1%	100%

Geographic Matches

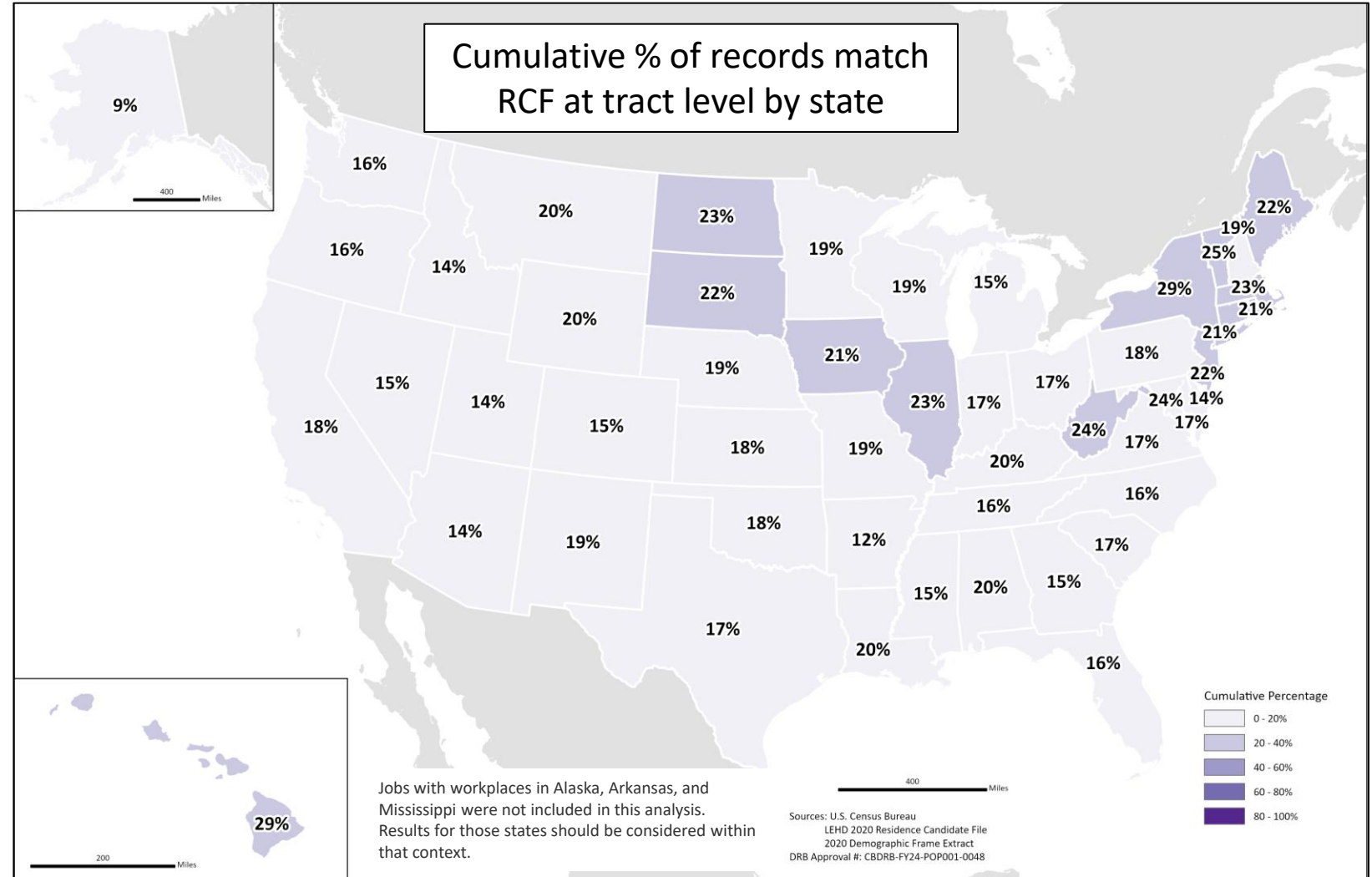
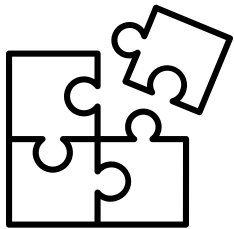
When comparing match rates at geographic level, we gain an addition of 9% with state and lower-level matches.

	Count	% of total Jobholders
Residential location matches exactly	108,600,000	 89%
Additional matches at state level	10,910,000	 9%

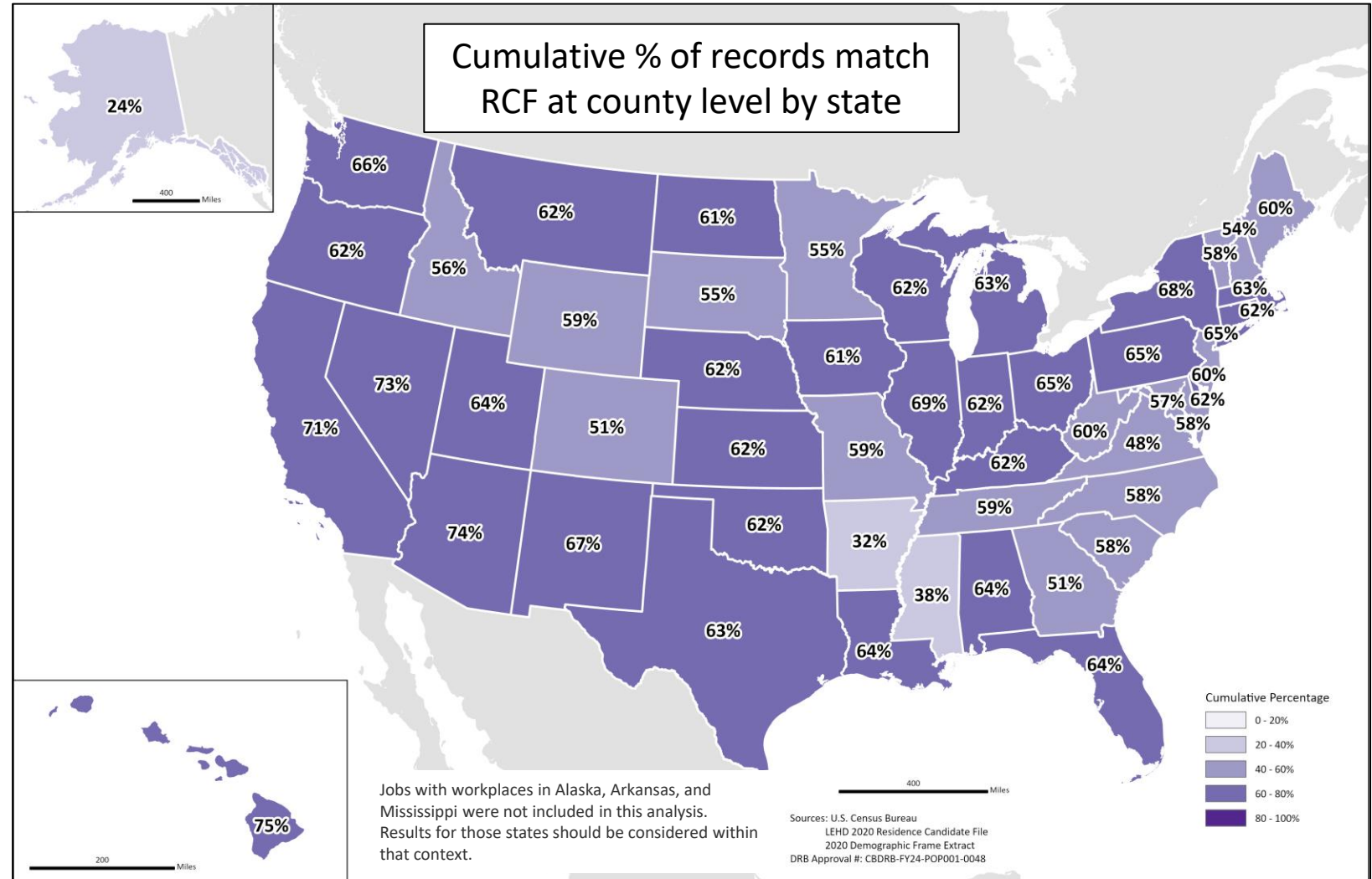
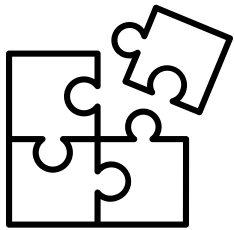
Within each state, how many of the Demo Frame MAFIDs match the RCF MAFIDs at the **block group** and lower level?



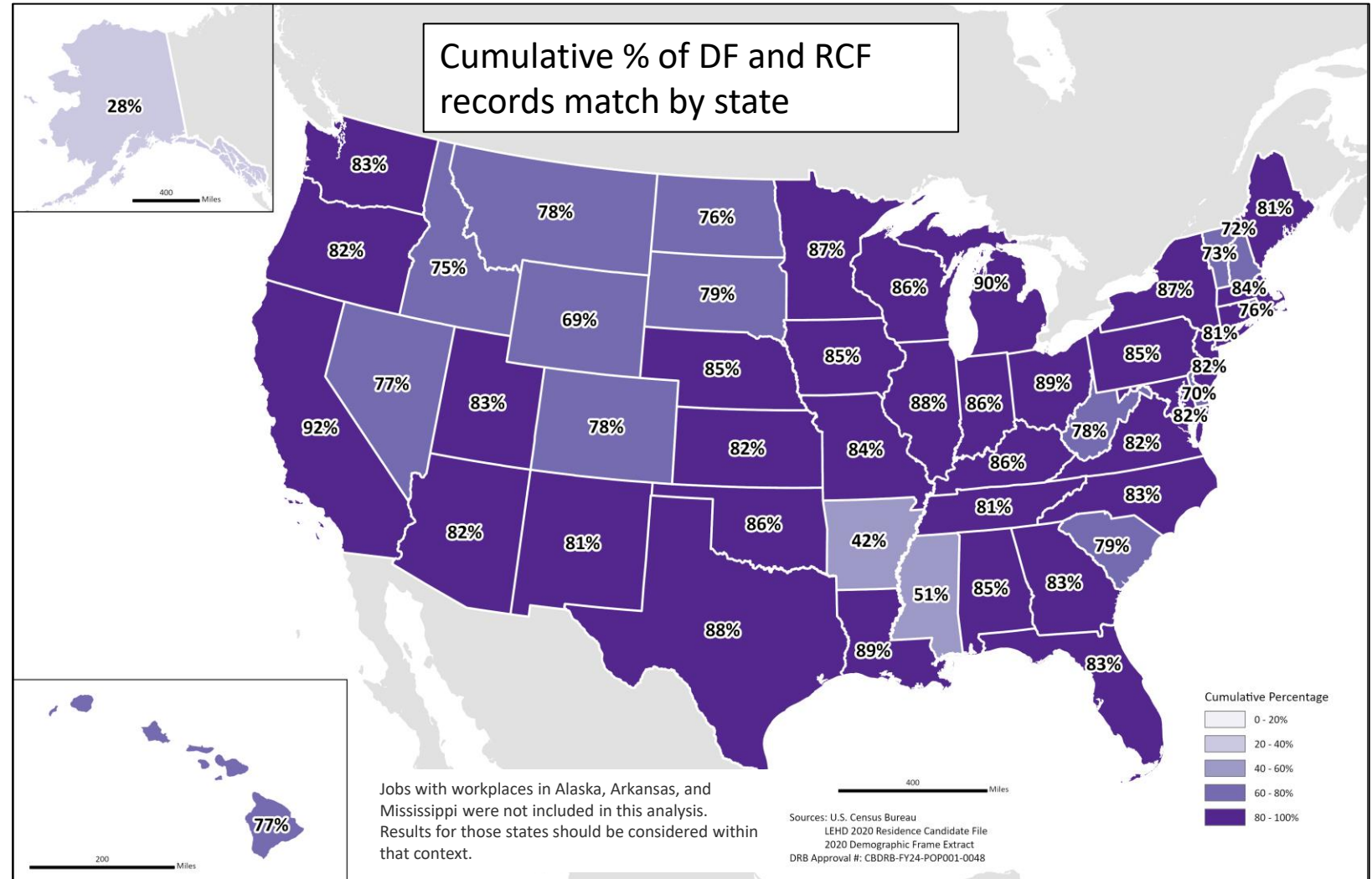
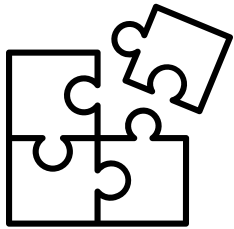
Within each state, how many of the Demo Frame MAFIDs match the RCF MAFIDs at the **tract** and lower level?



Within each state, how many of the Demo Frame MAFIDs match the RCF MAFIDs at the **county** and lower level?



Within each state, how many of the Demo Frame MAFIDs match the RCF MAFIDs at the **state** and lower level?

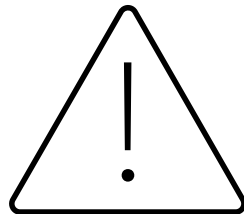
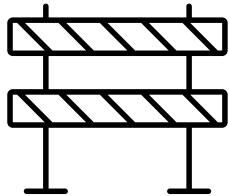


Summary

- High rate of overlap between RCF and Demo Frame PIKs
- Net gain of PIKs from Demo Frame
- 89% of MAFIDs match, but we can add 9% state level matches
- Highest geographic level matches at county level, may be due to missing lower levels of geography

Limitations

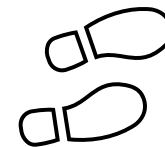
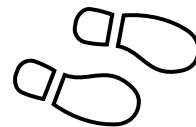
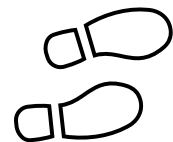
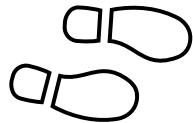
- Missing data for lower levels of geography
- Modeling differences may be part of result in non-matches



Next Steps



- Where MAFIDs do not match, how often is the RCF MAFID in the Demo Frame, but not the 'best' MAFID?
- Investigate 'apartment confusion'
- Adjust PPM process to incorporate earlier years in administrative records



Acknowledgements

- Matt Graham (Center for Economic Studies) – RCF guru



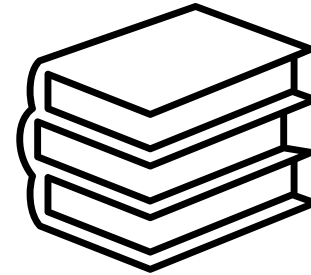
- Josh Wixom (Geography) - maps



Resources

Demographic Frame:

[Frames Program](#)



LEHD/RCF:

[Developing a Residence Candidate File for Use With Employer-Employee Matched Data](#)

[LEHD/Job Frame: Draft Requirements for Residence Data](#)

LODES:

[Longitudinal Employer-Household Dynamics](#)

Contact



DELIVERANCE.BOUGIE@CENSUS.GOV

How far is the Demo Frame location from the RCF location?

Percentile	Distance (miles)	Distance when “states differ” is removed
5 th	<1	<1
25 th	2	1
50 th	6	5
75 th	25	12
95 th	904	79



How far is the Demo Frame location from the RCF location?

