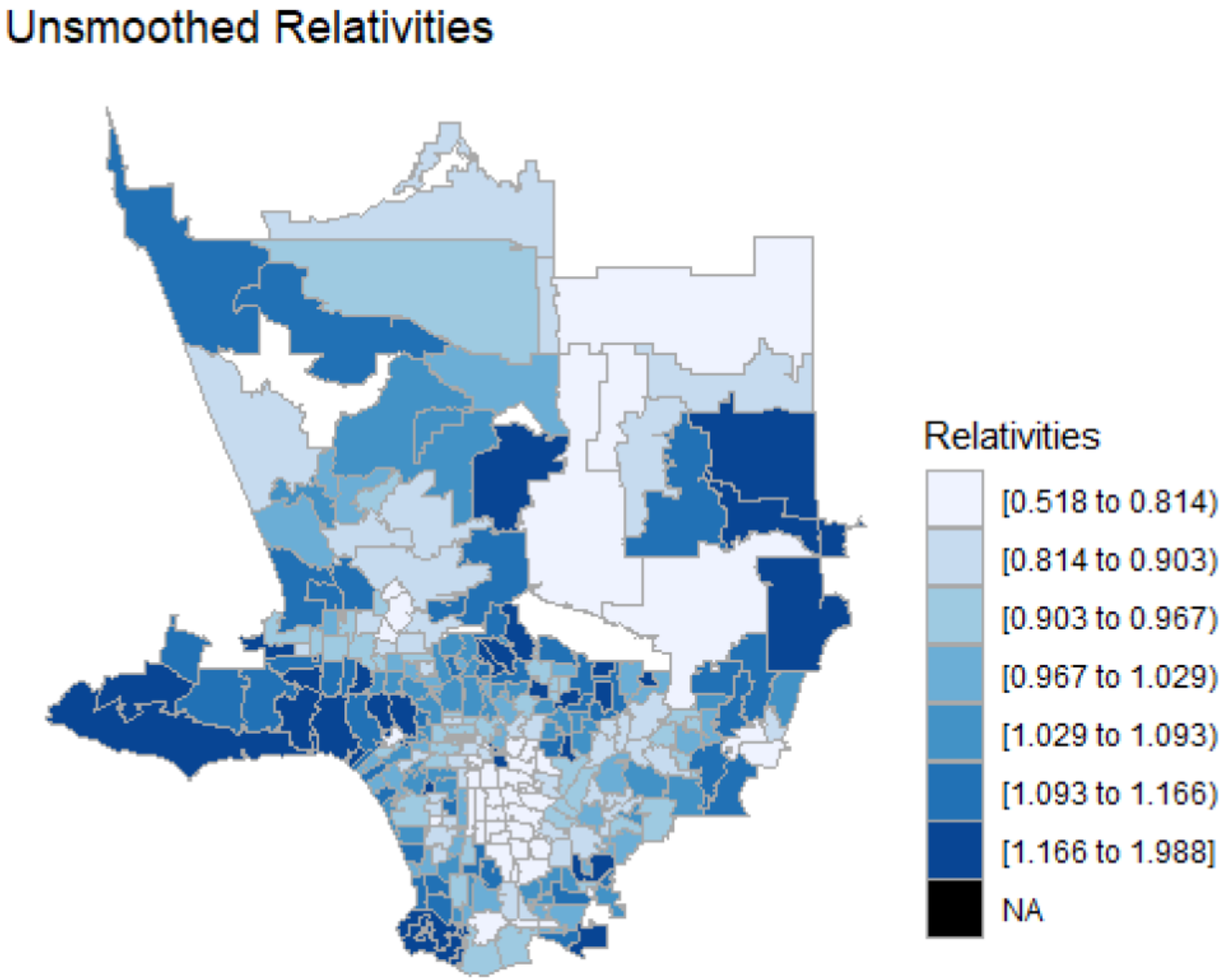
Jump Smoothing Algorithm

Michael Tsiappoutas, PhD

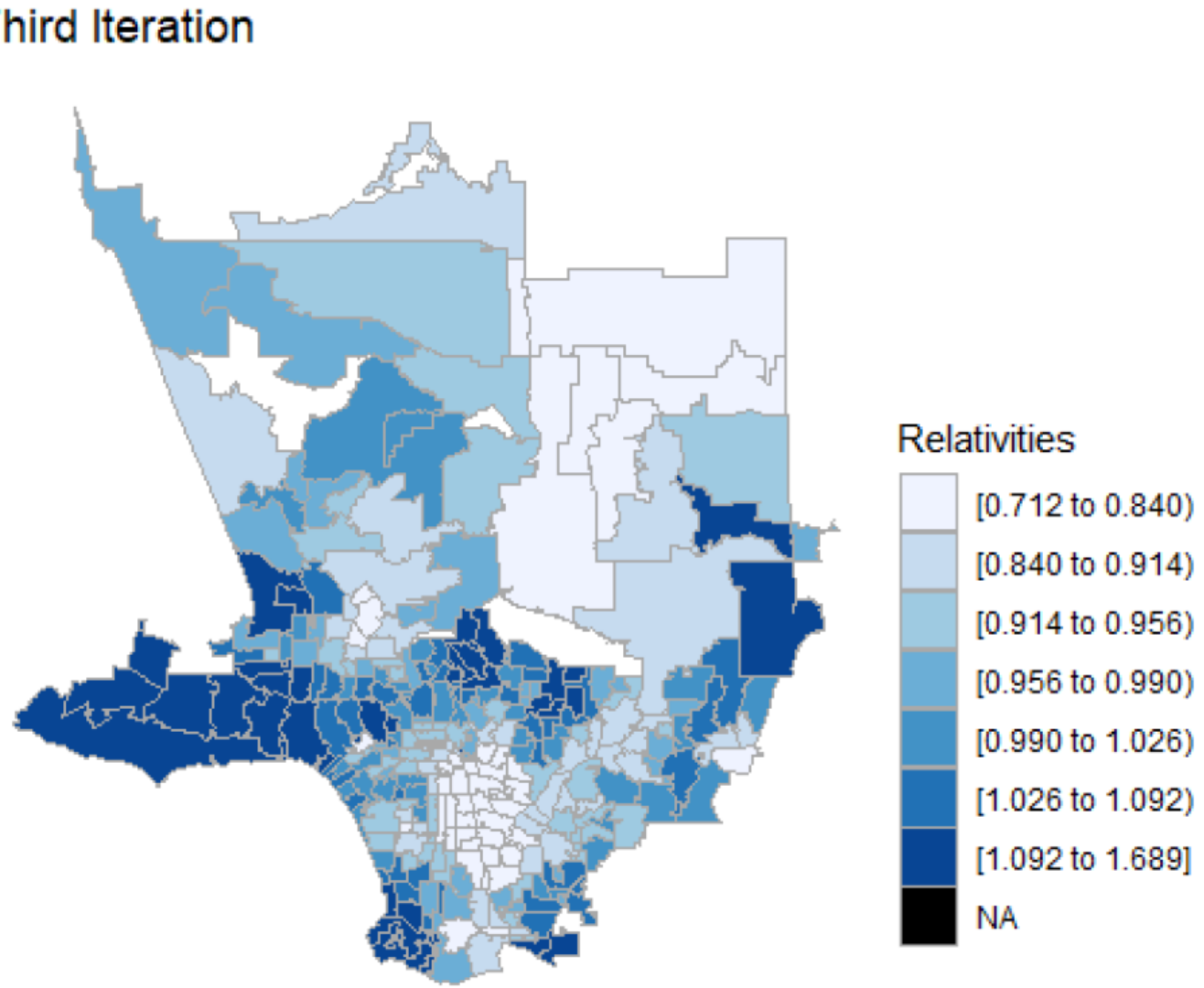
about me

- ✓ PhD Engineering and Applied Physics
- ✓ MS Applied Physics
- ✓ MS Quantitative Psychology
- ✓ BS in Psychology and Physics
- ✓ 14 years in insurance

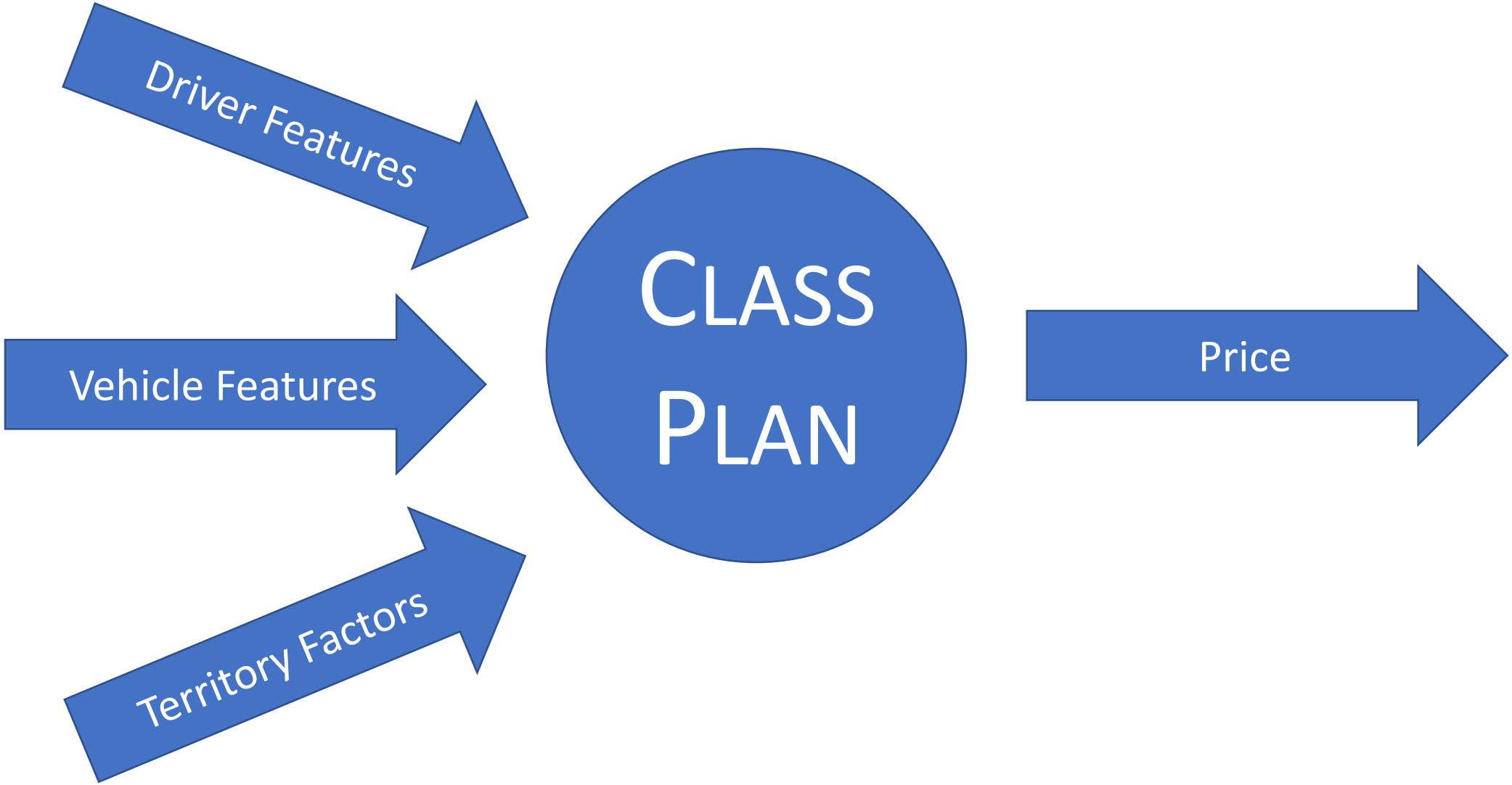
what *is* smoothing?



what *is* smoothing?



who uses smoothing?



why smoothing?

- ✓ Encouraged by regulators
- ✓ ‘Fair Discrimination’

‘traditional’ smoothing

Model-based, with two flavors:

1) Smooth existing geo factors directly

2) Smooth geo factor features and hope the predicted geo factors will be smooth.

credibility

Do you have enough data
to trust results?

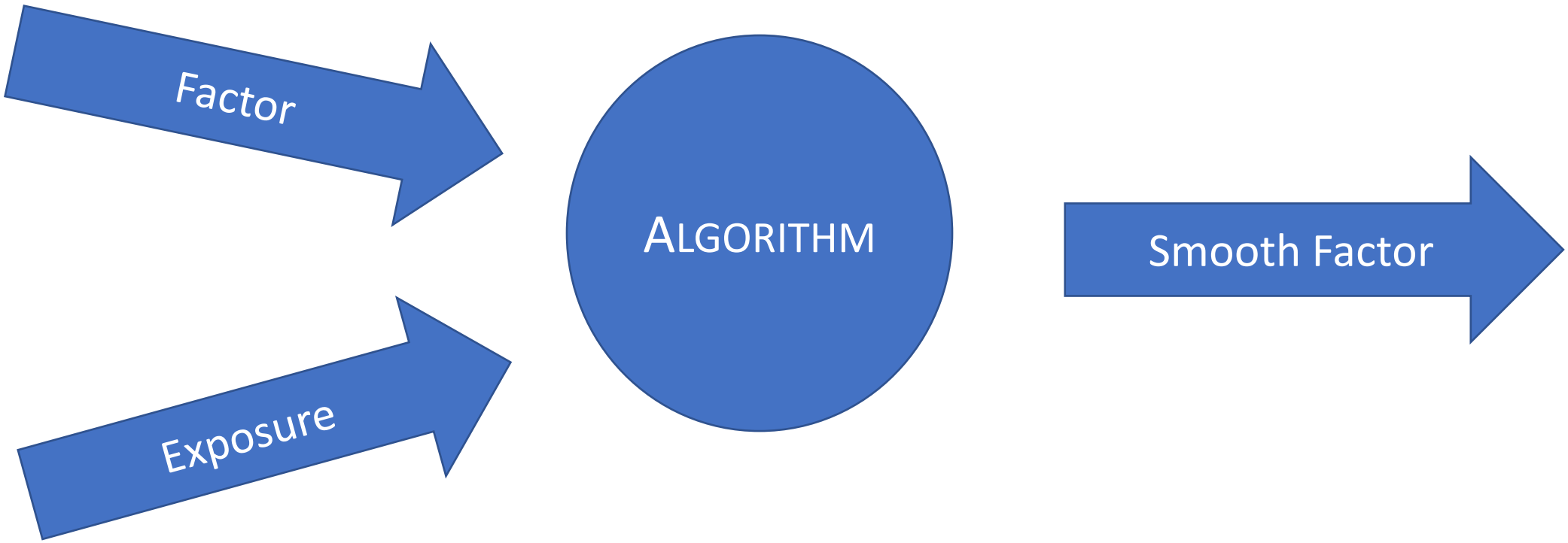
problem

If your modeling data is *not* credible,
do you trust its smoothing?

algorithmic approach

To my knowledge, this is the first smoothing *algorithm.*

algorithm inputs/output



(GOOD RESOURCE: PRINCIPLES OF RATEMAKING, CHARLES L. MCCLENAHAN)

big idea

- ✓ find all neighbors of a zip
- ✓ take factor difference
- ✓ if difference is smaller than a ‘jump’ threshold, ignore it
- ✓ take avg of zip factors
- ✓ avg original factor with above

data example

First zip code with all its neighbors (pairs1 dataset)

zip	nbr	zip.exp	zip.rel	nbr.exp	nbr.rel	jump	nbr.rel.jump
90001	90002	0.5411	0.7146	0.4853	0.6850	0	0.6850
90001	90003	0.5411	0.7146	0.6278	0.7065	1	0.7146
90001	90011	0.5411	0.7146	0.9843	0.7038	0	0.7038
90001	90058	0.5411	0.7146	0.0305	0.6984	0	0.6984
90001	90255	0.5411	0.7146	0.7112	0.7817	0	0.7817

First zip code after calculating smooth relativity (wtd.avgs1 data set)

zip	zip.rel	zip.exp	wtd.nbr.rel	avg.nbr.exp	smooth.rel1
90001	0.7146	0.5411	0.7224	0.5678	0.7186

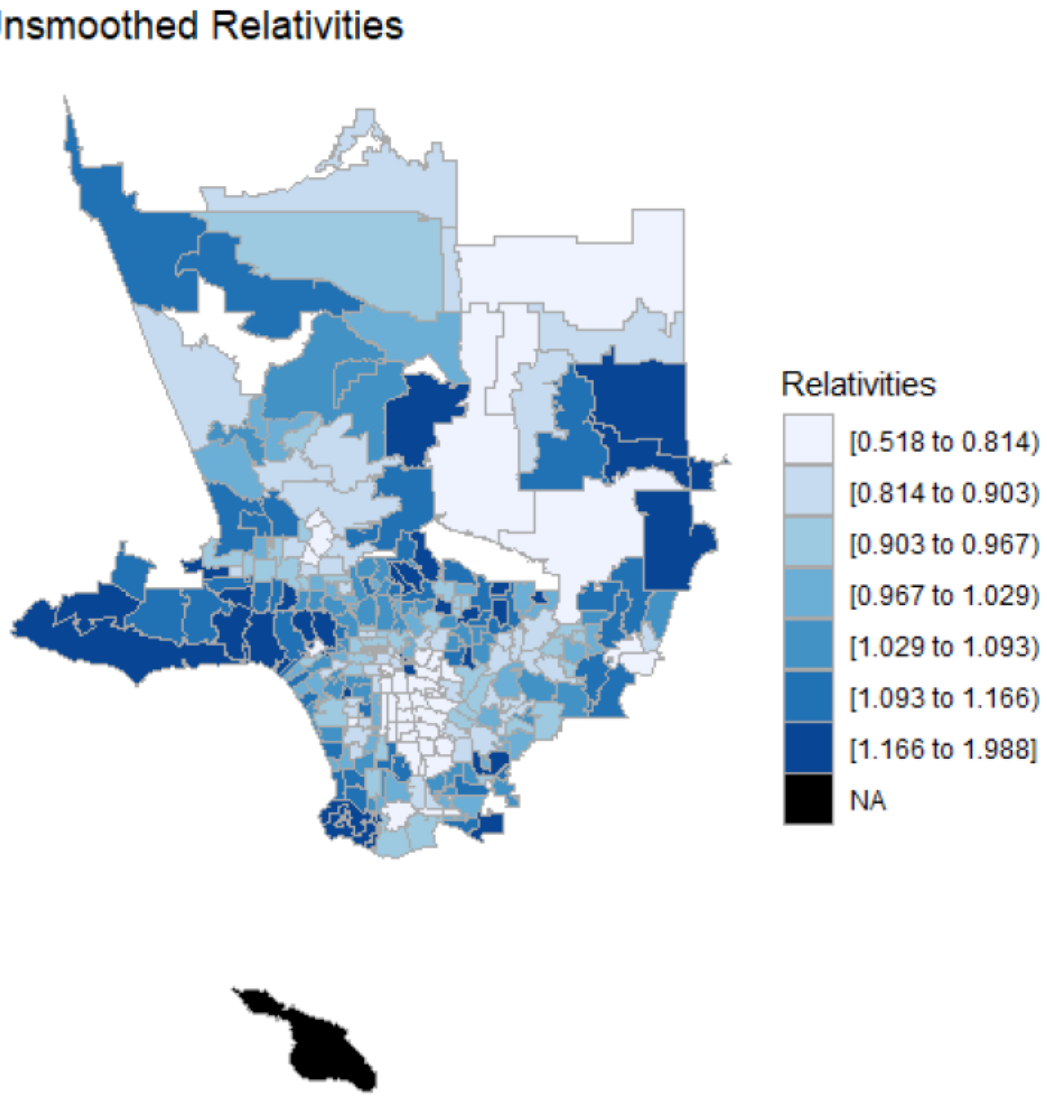
Algorithm

1. Import exposure and relativity by zip code dataset.
2. Import 'zip pairs', a dataset of zip codes and all their neighboring zip codes.
3. Merge (1) and (2).
4. Calculate 'jump', the difference between each zip pair.
5. If jump is small (less than a threshold), replace the neighbor relativity with the main zip relativity. That is to say, ignore any neighbors whose difference with the main zip code is less than a threshold ('jump').
6. Calculate the exposure-weighted average relativity of all zip codes touching the main zip code.
7. Calculate smooth relativity using Equation 1 below.

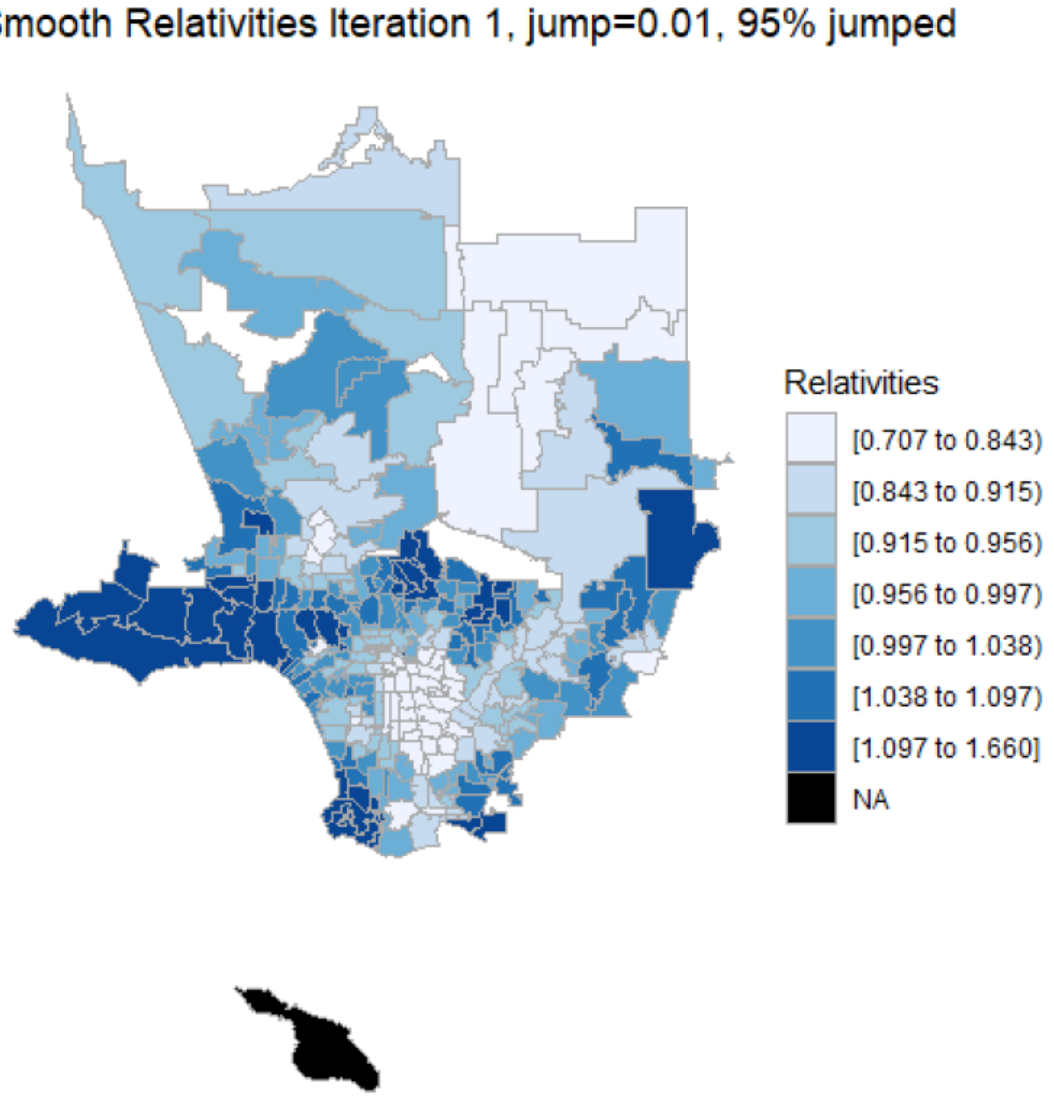
control convergence speed

- ✓ jump threshold
- ✓ iterations

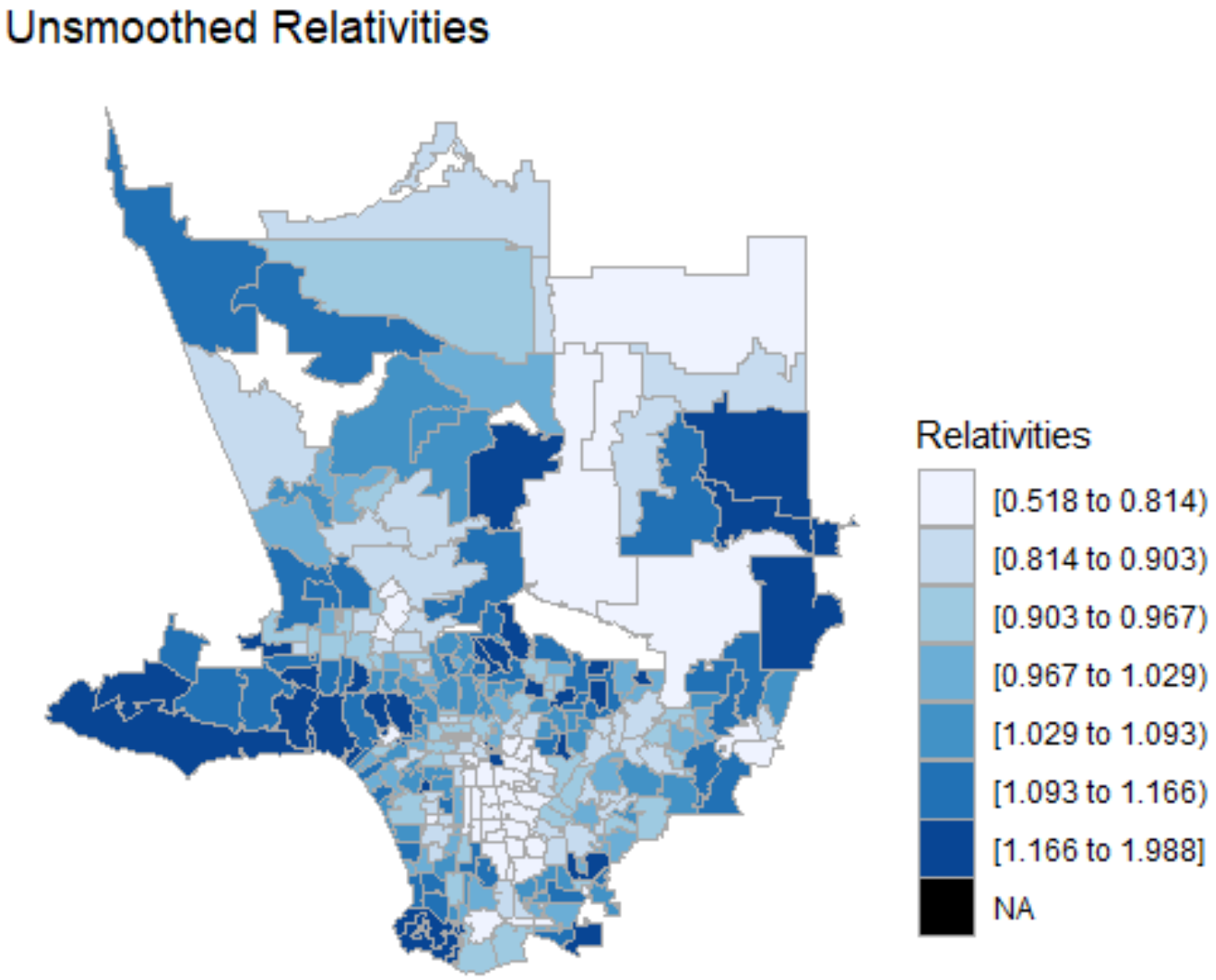
jump coefficient



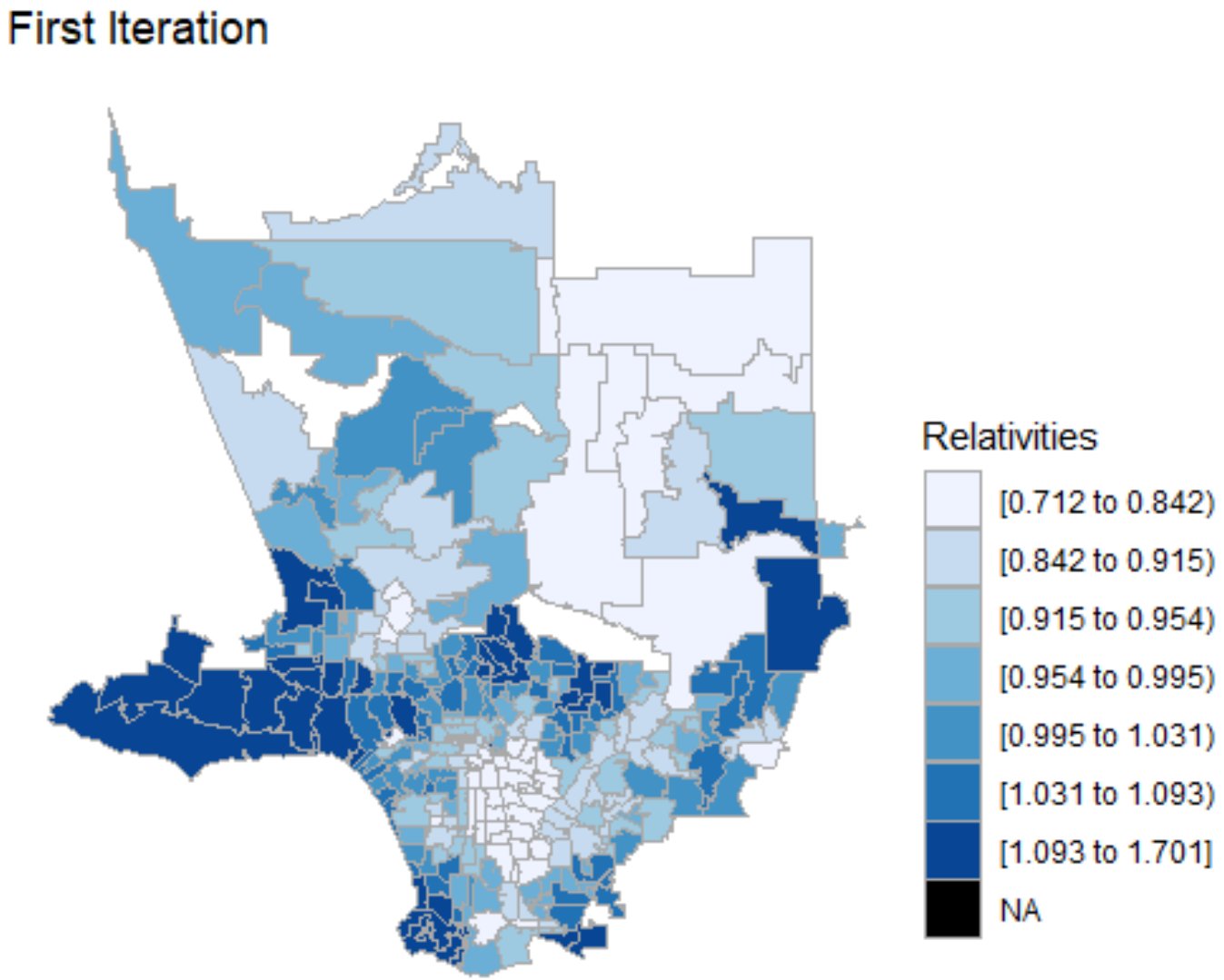
jump coefficient



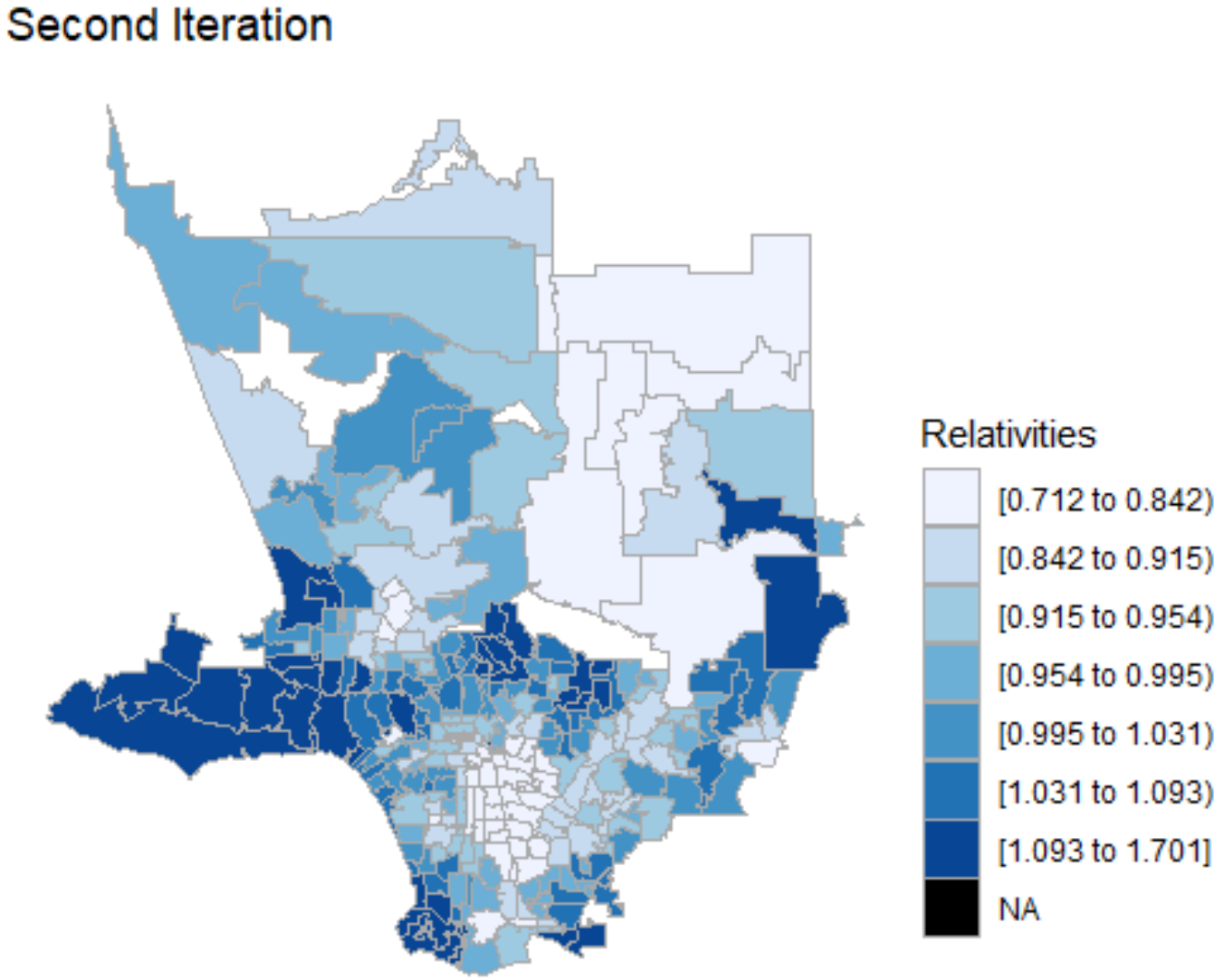
iterations



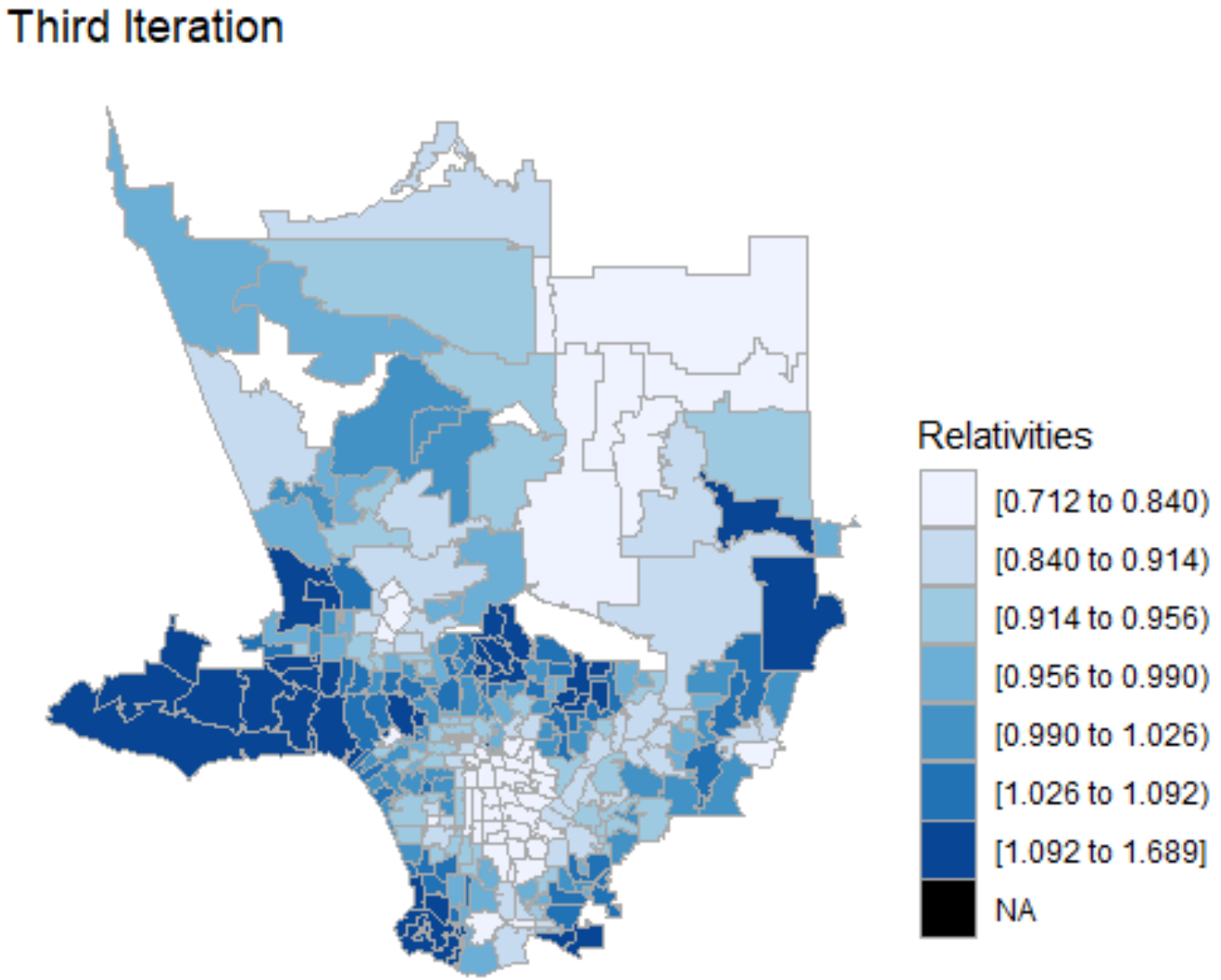
iterations



iterations



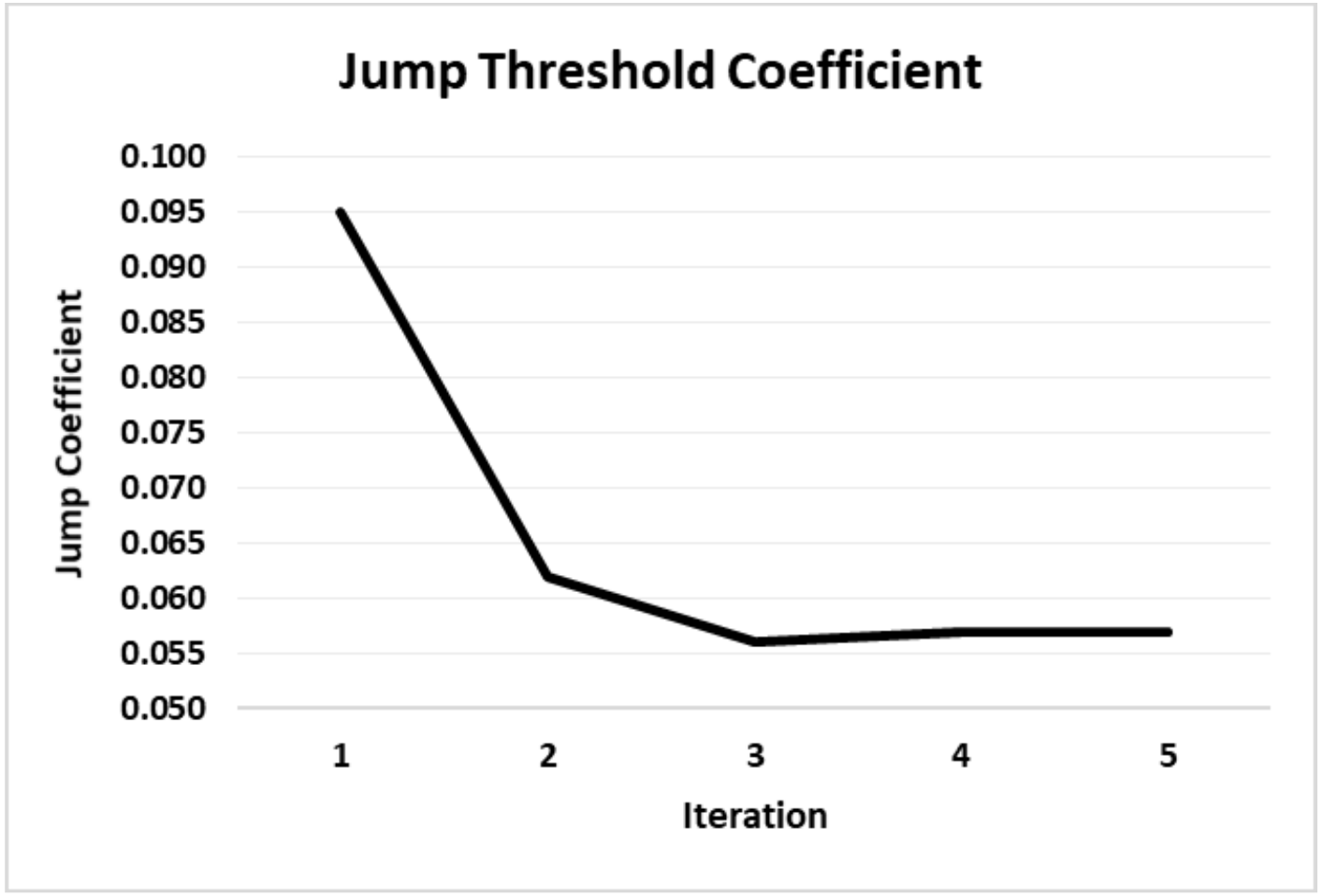
iterations



when do you stop?

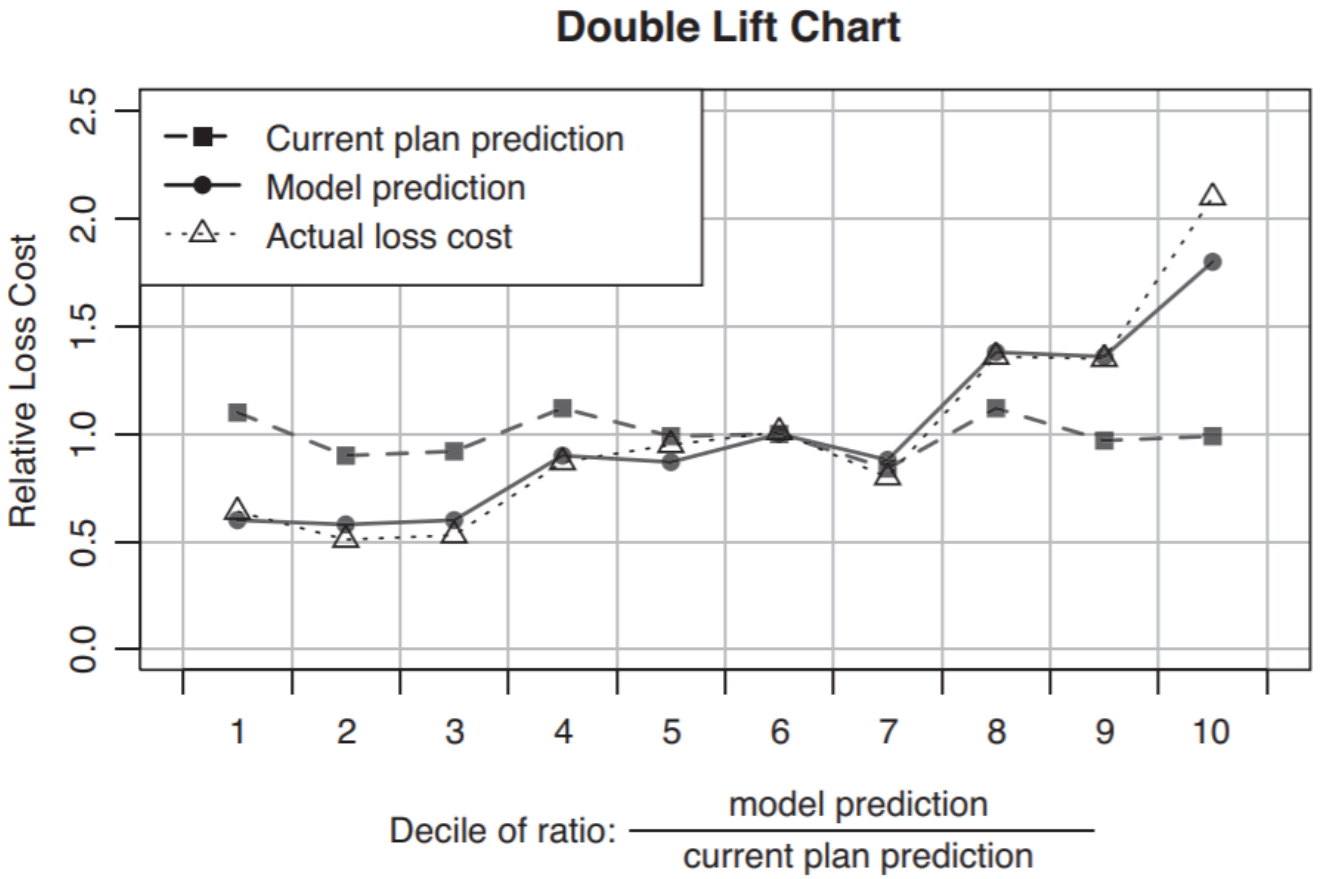
With each successive iteration, differences between main and smooth factors required to achieve the jump rate, should shrink.

when do you stop?



does it work?

Figure 23. A Sample Double Lift Chart



(SOURCE: GENERALIZED LINEAR MODELS FOR INSURANCE RATINGS, CAS MONOGRAPH SERIES NO. 5.

GitHub

- ❑ code
- ❑ data
- ❑ whitepaper
- ❑ presentation

www.GitHub.com/Tsiappoutas/Smoothing

Q&A