NAS Seminar Sept, 2024

The Effect of Location on Human Decision-Making: Implications for Social and Environmental Sciences

A. Stewart Fotheringham

Regents' Professor of Computational Spatial Science
Director, Spatial Analysis Research Center
School of Geographical Sciences and Urban Planning

Stewart.Fotheringham@asu.edu

Fundamental to our understanding of a great deal of social sciences, health sciences, planning and many environmental issues is the following question…

How do Human Beings make decisions?

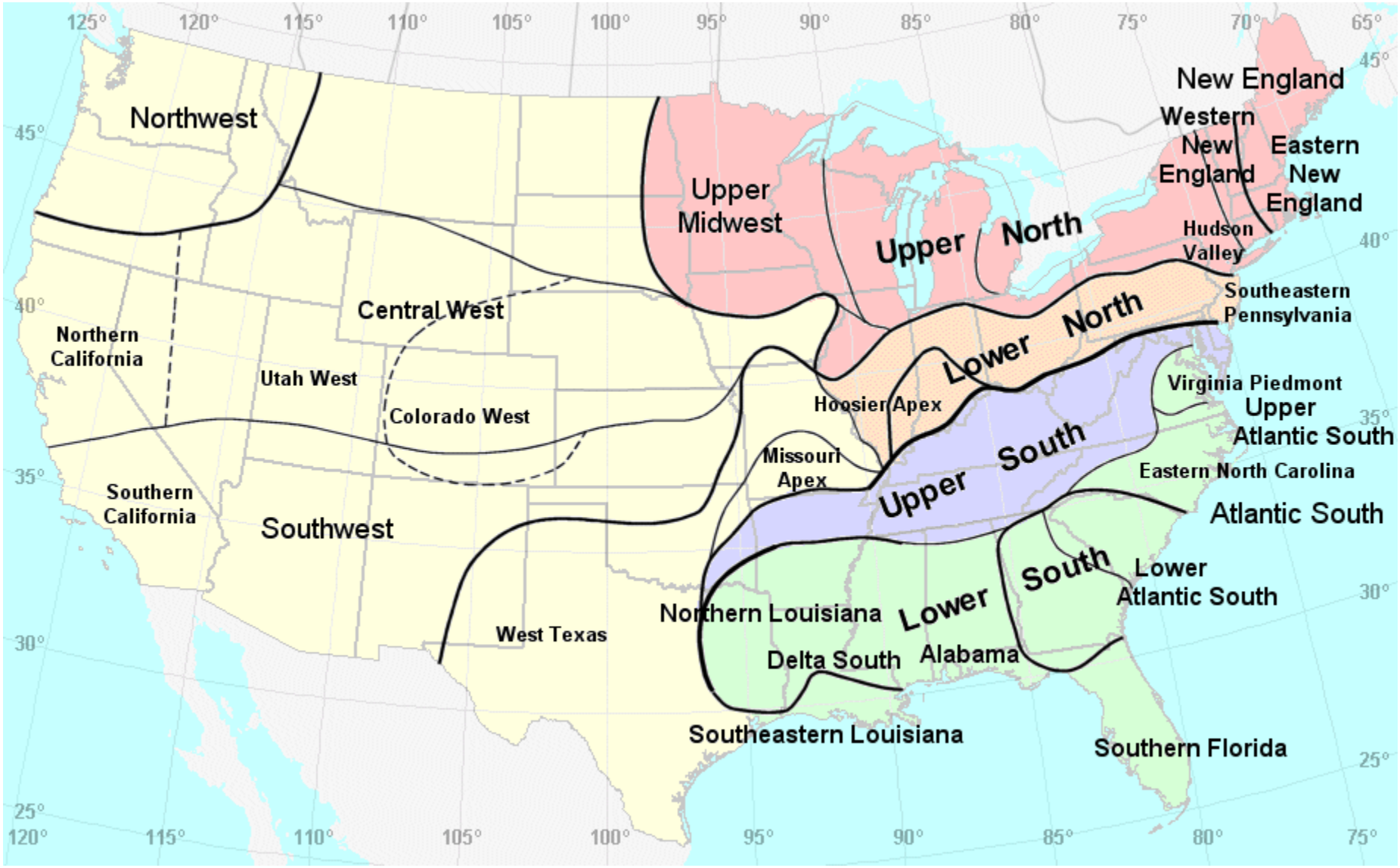
A huge amount of research across many disciplines strongly suggests that the information we receive in order to make decisions and how we process information depends on **who we are**.

- Age
- Ethnicity
- Income
- Education
- Gender
- …

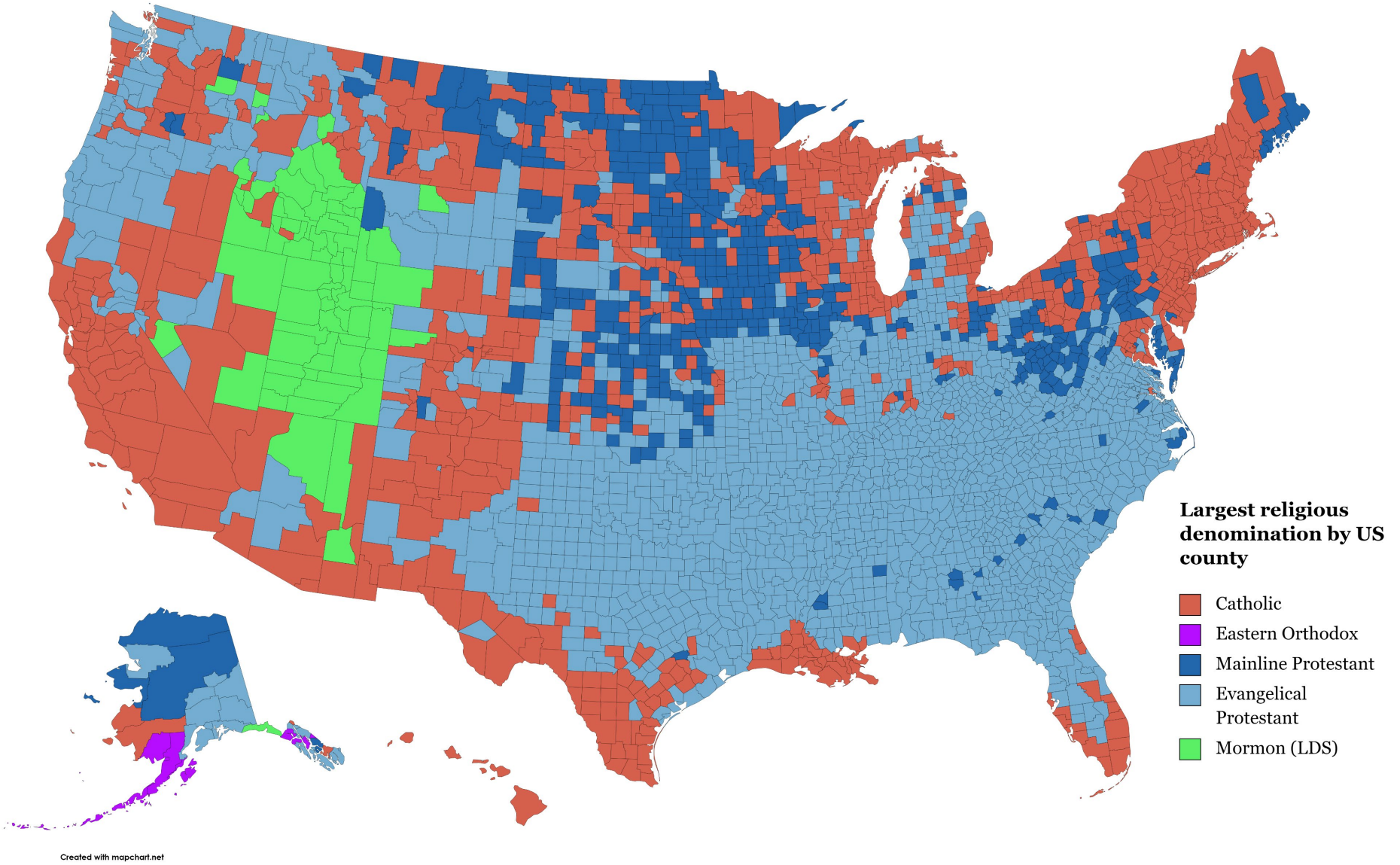
These are the "usual suspects" in any modeling of human behavior…
e.g. SDOH

But, does the information we receive and how we process it also depend on **where we are**?

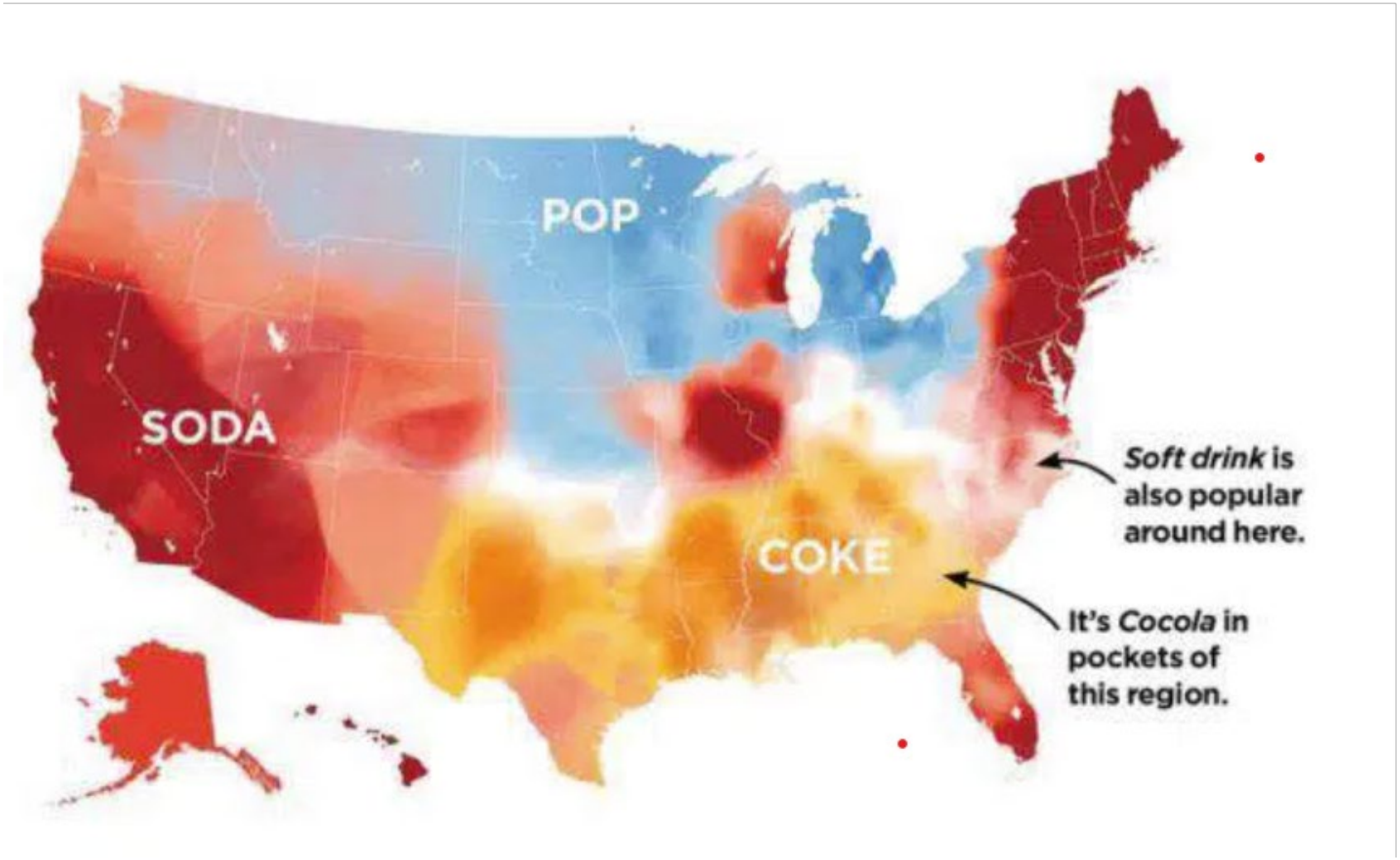
For instance, it is clear that some personal attributes depend more on where we are than who we are…e.g. Accents



Religious Affiliation



What do you call a “soft drink”



Three fundamental (*transformative*) research questions

1. Does location affect decision-making, independently of who we are?

2. If so, how important a factor is it?

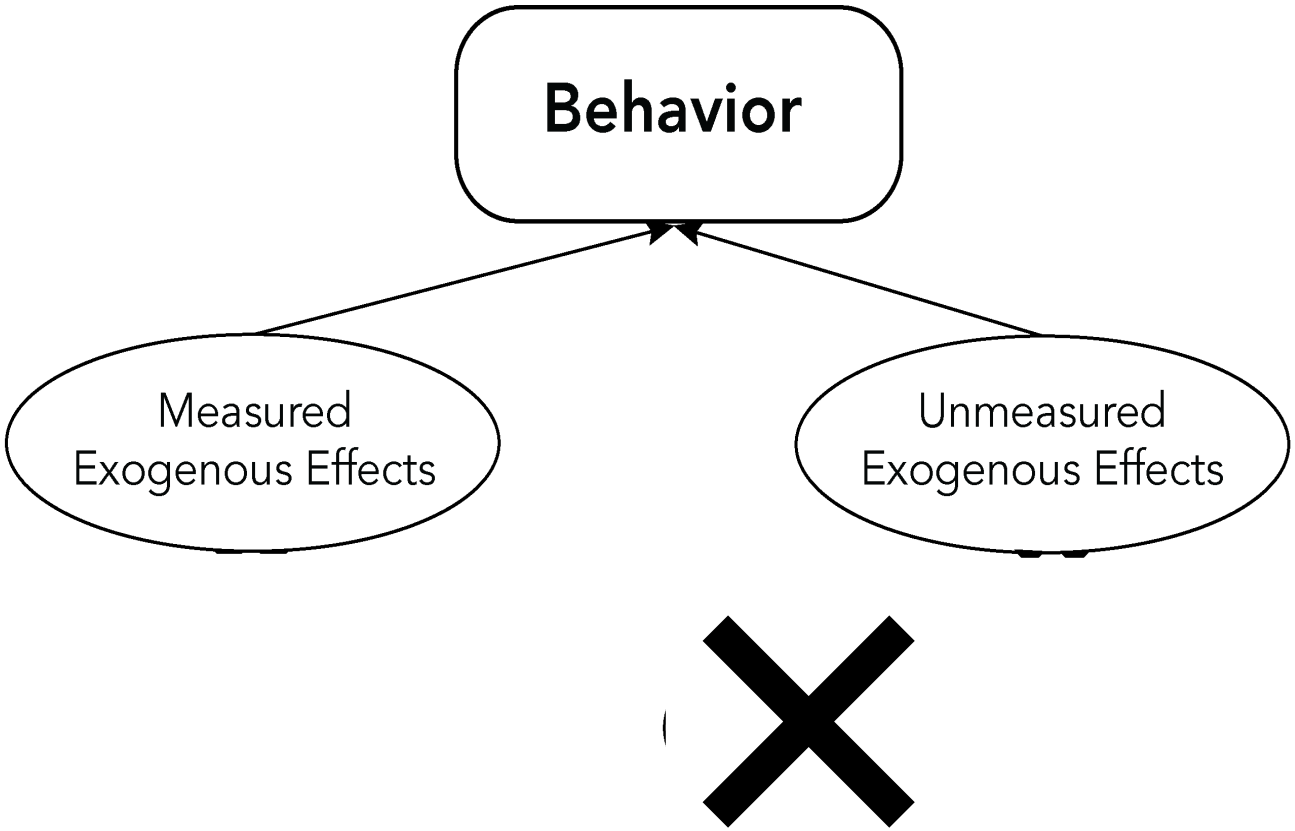
3. If it is important, how can we incorporate it into models of human behavior?

The answers to these questions relate to a VERY wide range of applications/problems.

"Context"

The *raison d'être* of place-based analysis of human behavior is that location has an effect on decision-making which is ***independent of the identifiable factors that describe both a location and its inhabitants.*** We call this "Geographical Context" or just **"Context"**

In fact there are two types of contextual influence that we need to be aware of when we model human behavior…



e.g. older voters - preference for Democrats in FL; preference for Republican party in TX, *ceteris paribus*

e.g. stronger preference for the Democratic party in California than in Alabama, *ceteris paribus*

Models Incorporating Contextual Effects

1. Multiscale Geographically Weighted Regression (MGWR)

$$y_i = \beta_{0i} + \beta_{1i}x_{1i} + \beta_{2i}x_{2i} + \ldots + \beta_{ki}x_{ki} + \varepsilon_i$$

Intrinsic contextual effects

Behavioral contextual effects

www.routledge.com/9781032564227

20% Discount with Discount Code.

1ST EDITION

Multiscale Geographically Weighted Regression

Theory and Practice

By **A. Stewart Fotheringham**, **Taylor M. Oshan**, **Ziqi Li**

Multiscale Geographically Weighted Regression (MGWR) is an important method that is used across many disciplines for exploring spatial heterogeneity and modeling local spatial processes. This book serves as definitive guide to local regression modeling and the analysis of spatially varying processes, a very cutting-edge, hands-on, and innovative resource. The authors start with the basic ideas and

2. Spatial Error Models (SEM)

$$\boldsymbol{y} = \beta_0 + \boldsymbol{X\beta_{\neq 0}} + \boldsymbol{u}$$

$$\boldsymbol{u} = \lambda \boldsymbol{W}_i \boldsymbol{u} + \boldsymbol{\varepsilon}$$

where $\boldsymbol{u}$ is the unfiltered error term, λ is the parameter for the spatial dependency in the error term, $\boldsymbol{W}$ is an *n* by *n* spatial weights matrix *(defined a priori),* and $\boldsymbol{\varepsilon}$ is the remaining random error.

Combining the above two equations, gives:

$$\boldsymbol{y} = (\lambda \boldsymbol{W}_i \boldsymbol{u} + \beta_0) + \boldsymbol{X\beta_{\neq 0}} + \boldsymbol{\varepsilon}$$

Intrinsic contextual effects - similar to local intercept in MGWR

3. Multi-Level Models (MLM)

Example of a two-level hierarchical model

$$y_{ip} = \beta_{0p} + \beta_{1p} x_{1ip} + \beta_{2p} x_{2ip} + \varepsilon_{ip}$$

where y_{ip} is the dependent variable for observation *i* that belongs to region *p* (*defined a priori*), β_{0p} is the intercept term for region *p*, x_{1ip} and x_{2ip} are the covariate values for observation *i* in region *p*, β_{1p} and β_{2p} are the slopes for region *p*, and ε_{ip} is the random error.

The intercept and slope parameters can vary between regions:

$$\beta_{0p} = \beta_0 + \mu_{0p}$$ ← Intrinsic contextual effects

$$\beta_{1p} = \beta_1 + \mu_{1p}$$ ← Behavioral contextual effects

$$\beta_{2p} = \beta_2 + \mu_{2p}$$

Testing the Models: Voting in the 2020 US Presidential Election

In 2020 there was a *de facto* survey of 159,633,396 people (66.7% of the voting age pop) across the US who revealed their preference for either a Democrat (51.3%) or a Republican (46.8%) to be President (3rd party votes were around 1.8%)

These data were aggregated to counties with the dependent variable being the **%D**

Along with these data, we have data for each county on 14 covariates thought to influence voting preferences

Basic Question

Democrat
or
Republican?

Intrinsic contextual effects

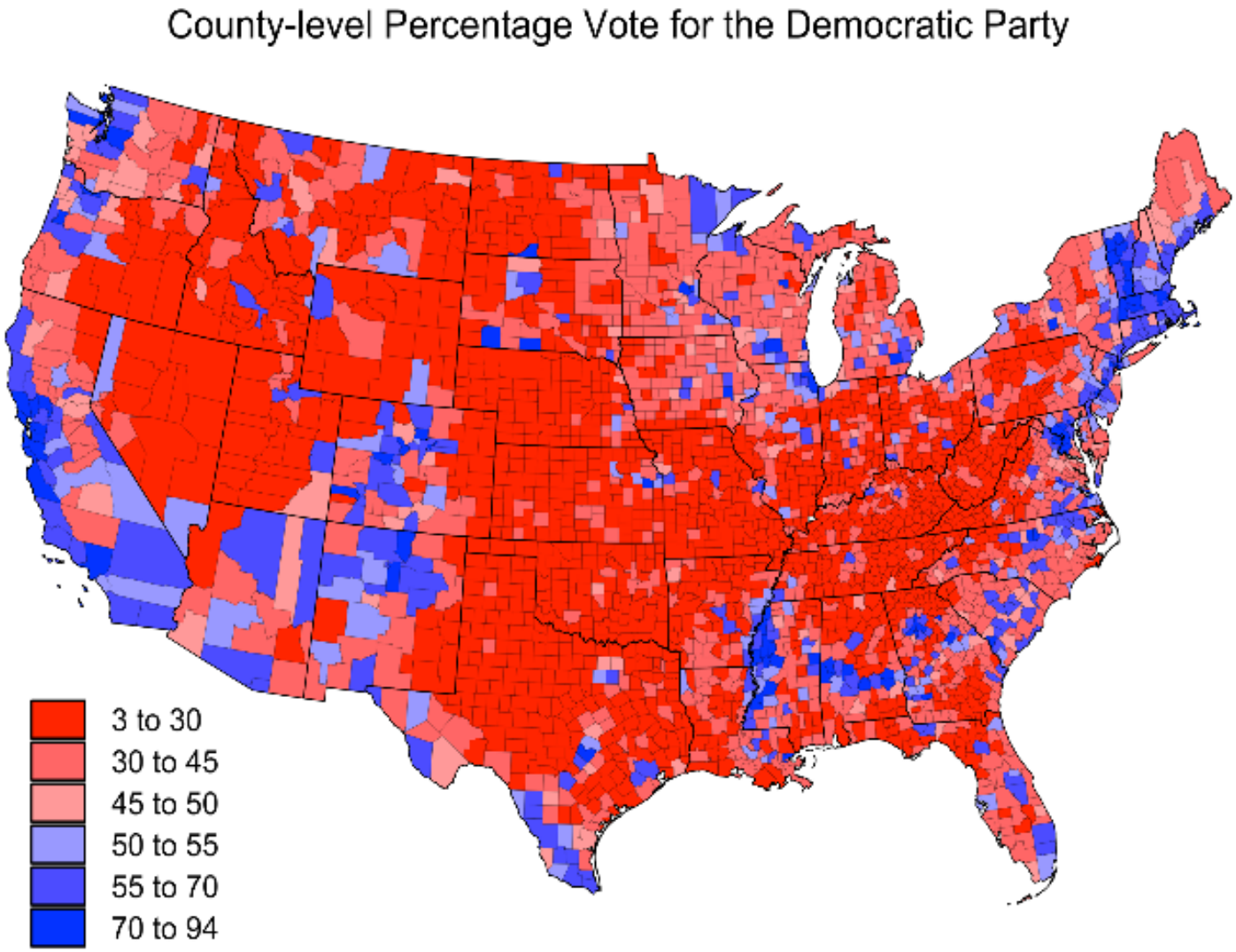
Behavioral contextual effects

Who we are

Where we are

Does where you live influence how you vote?
If so, how much of your voting behavior is determined by location?

2020 US Presidential Election: Biden vs Trump



From the literature and news media, 14 socio-economic variables appear to be influential in determining voting behavior

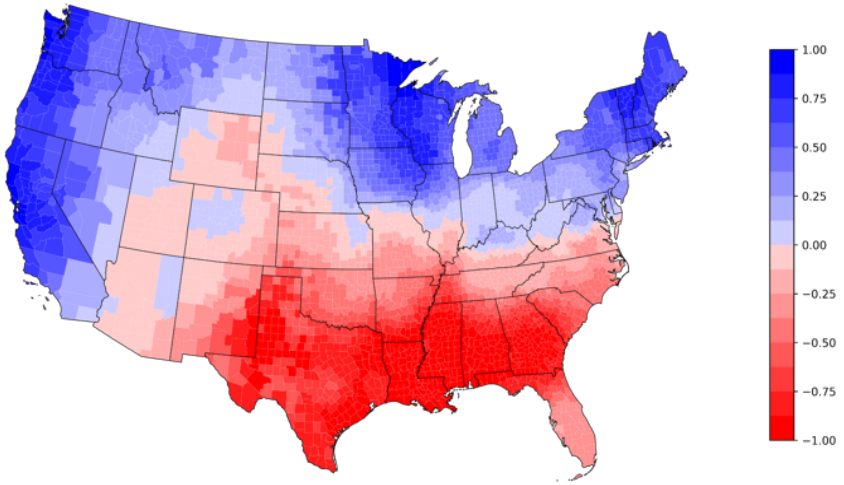
- **DEMOGRAPHIC**: Age (young and old); gender; ethnicity (Hispanic, Black); foreign born; education; pop density
- **ECONOMIC:** Income; Income disparities; health insurance; manufacturing employment
- **VOTING**: Turnout; 3rd party voting

For a general comparison of goodness-of-fit, the R^2 values are:

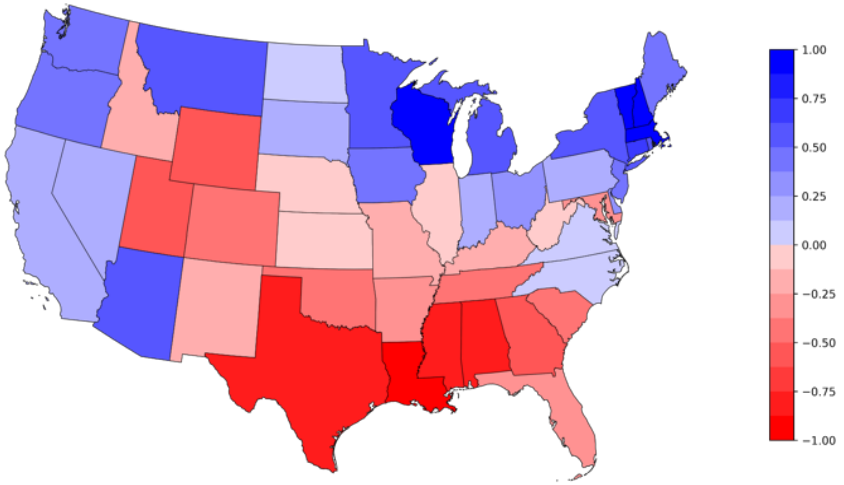
MGWR	0.95
MLM	0.93
SEM-KNN	0.88

Intrinsic Contextual Effects

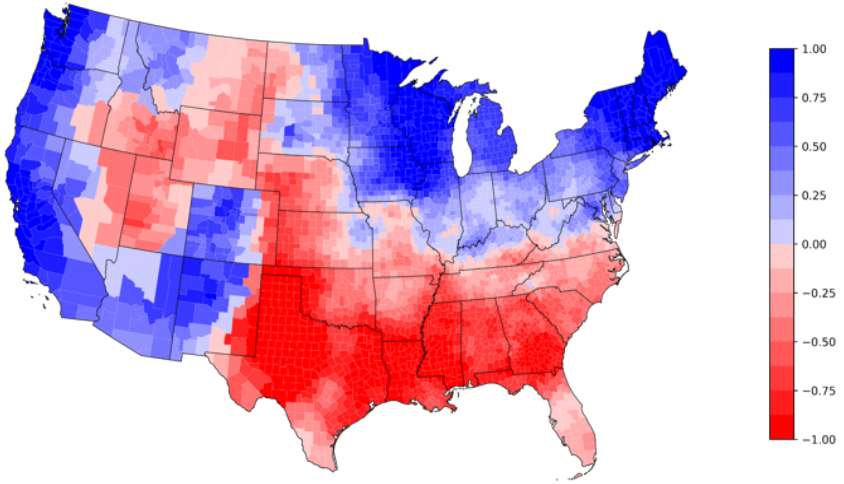
(a) MGWR local intercept



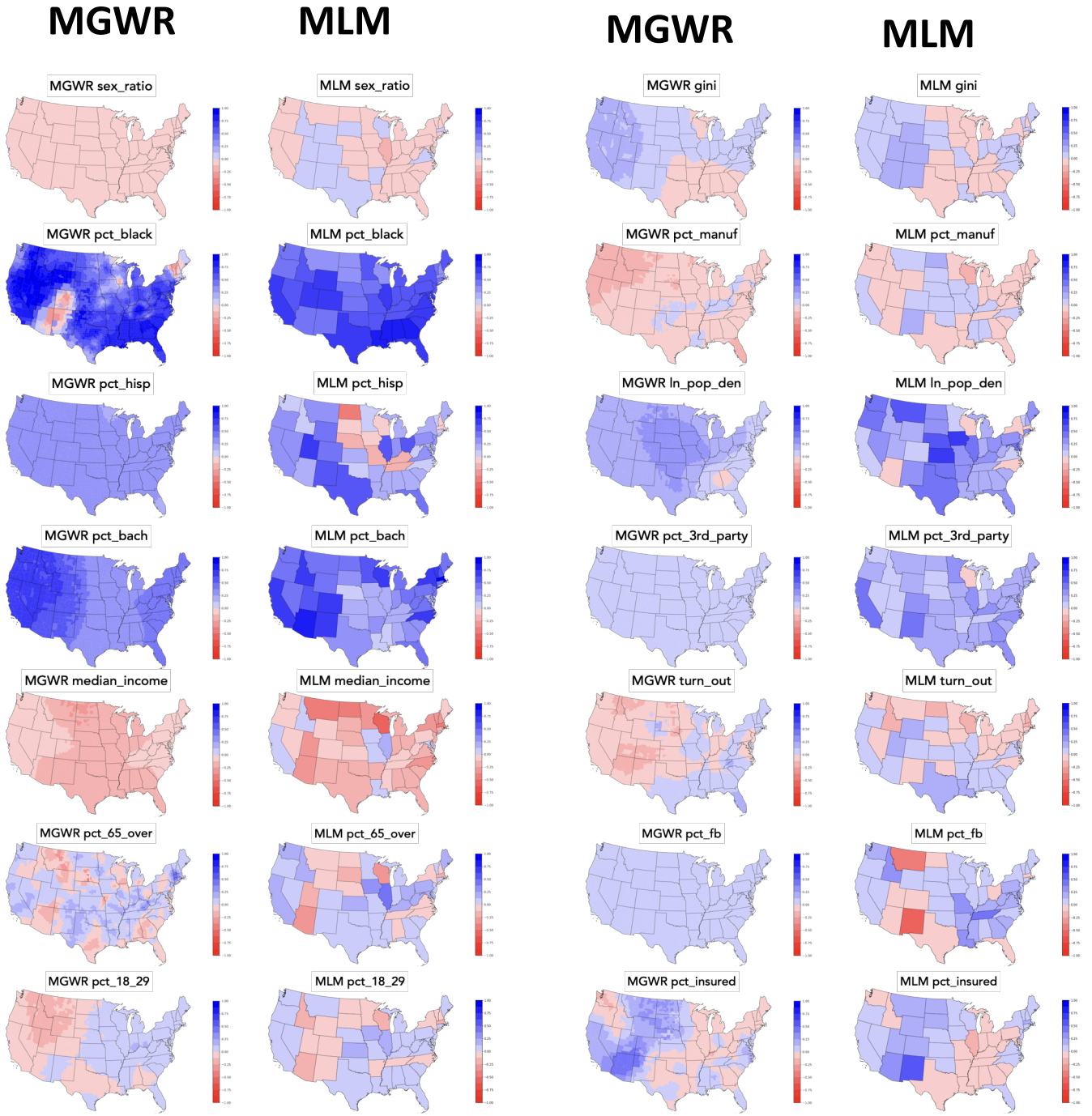
(b) MLM state-level varying intercept



(d) SEM (KNN) spatial autoregressive error



Behavioral Contextual Effects



So, it appears we **can** identify the role of place

BUT

How important is it?????

For this, one of the three models provides the answer…

MGWR not only identifies both intrinsic and behavioral contextual effects but also to quantifies the separate effects of place vs people. To see this, consider the local model written in terms of standardized variables (0,1): $y_i^* = (y_i - \bar{y})/\sigma_y$

$$y_i^* = \alpha_i + \sum_k \beta_{ik} x_{ik}^*$$

Rewriting this in terms of the original, unstandardized variables,

$$(y_i - \bar{y})/\sigma_y = \alpha_i + \sum_k \beta_{ik} \ (x_{ik} - \bar{x}_k) \ /\sigma_{x_k}$$

And then rearranging,

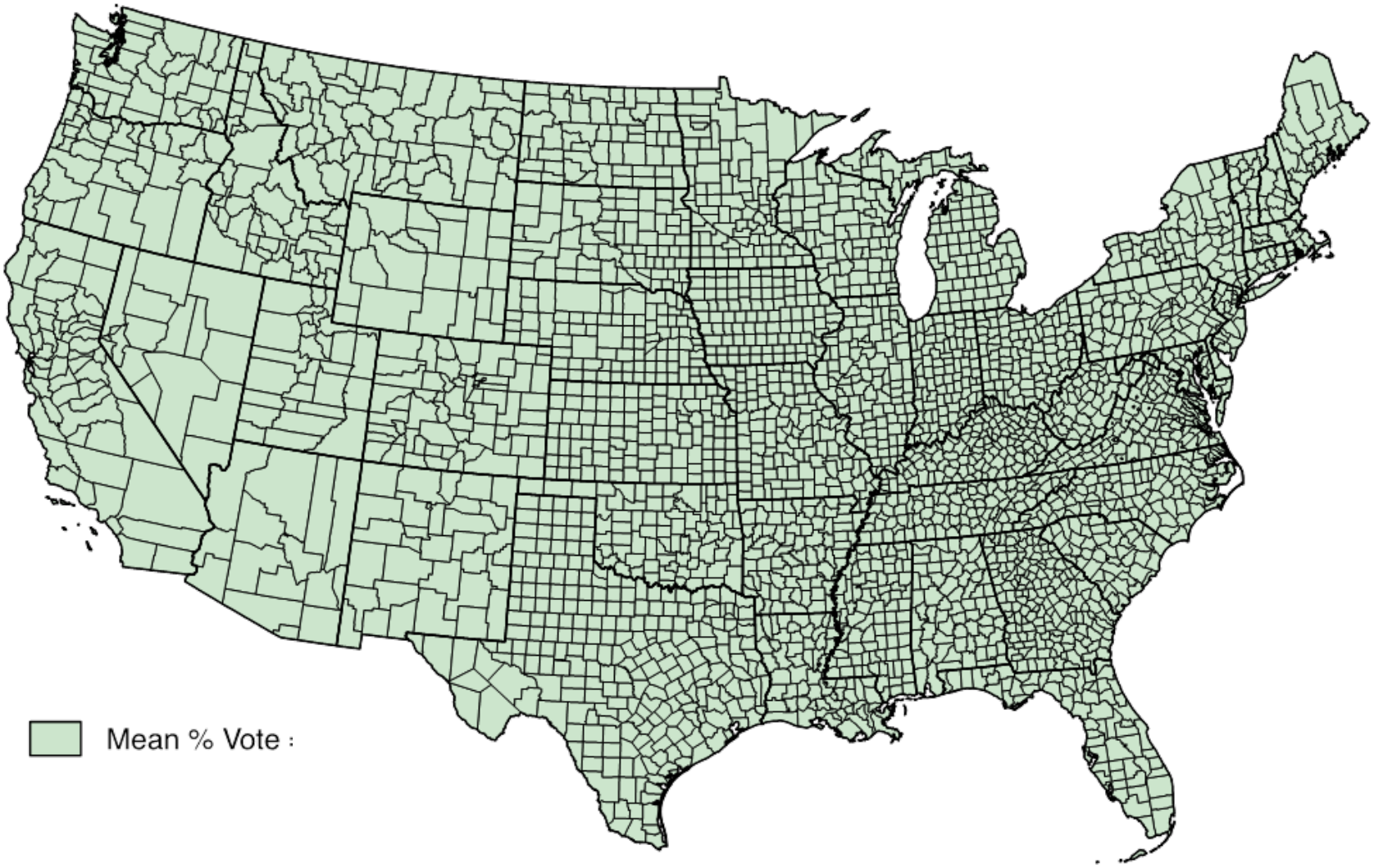
$$y_i = \bar{y} + \alpha_i \sigma_y + \sigma_y \sum_k \beta_{ik} \ (x_{ik} - \bar{x}_k) \ /\sigma_{x_k}$$

Mean level

Contribution due to place

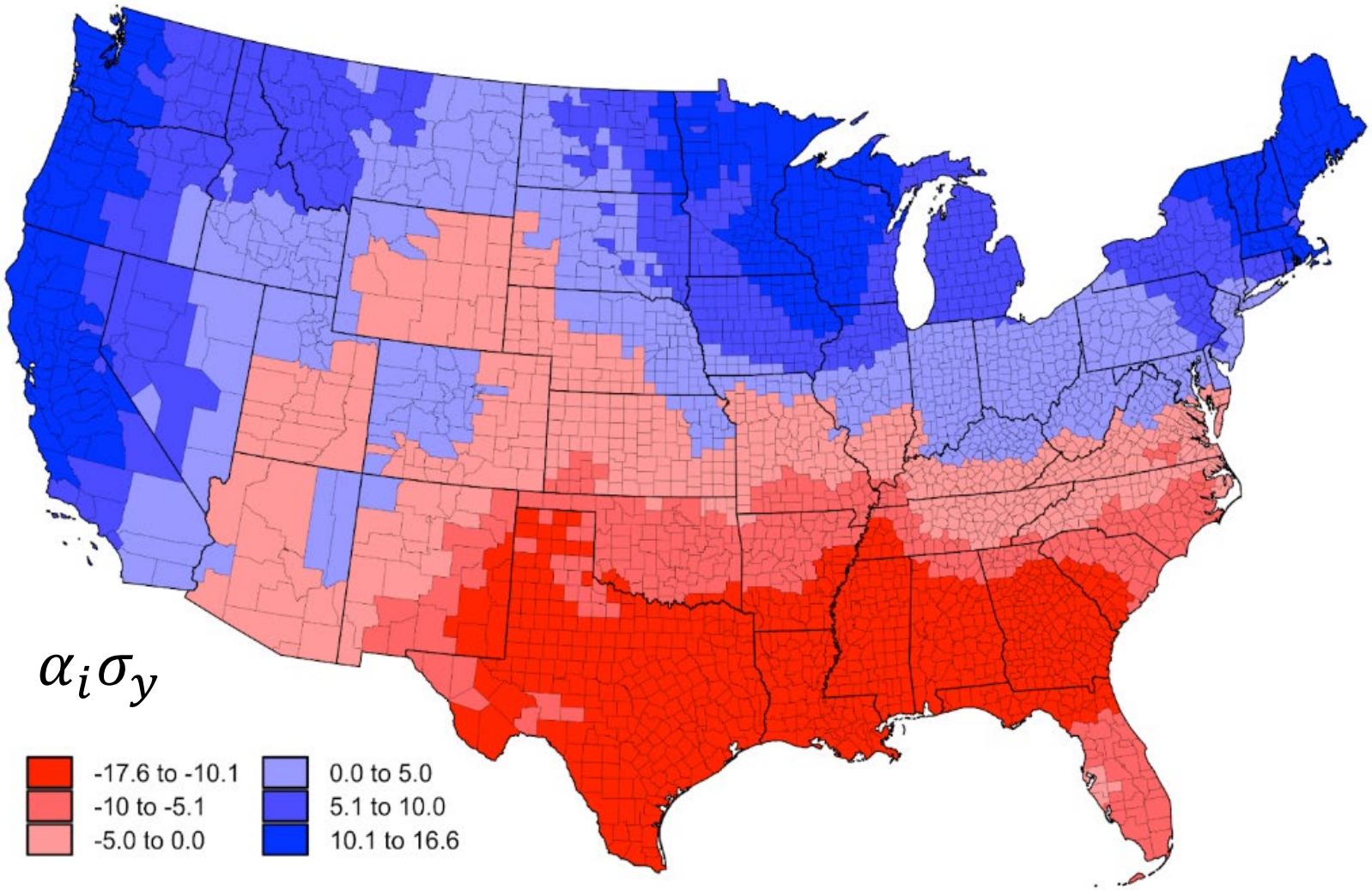
Contribution due to pop. comp.

Component 1 Mean Vote (ẏ) 35%

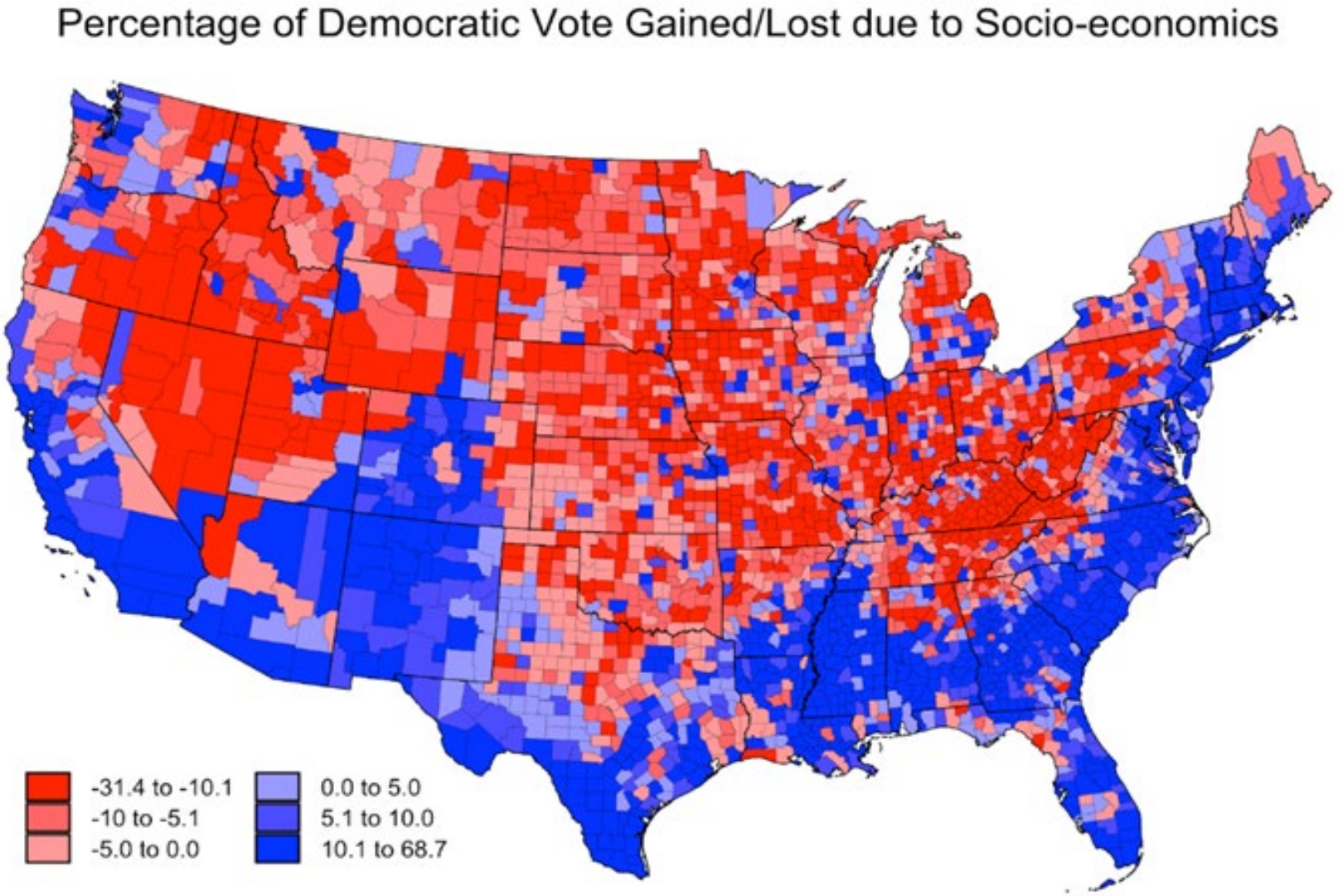


Component 2

Percentage of Democratic Vote Gained/Lost due to Geography

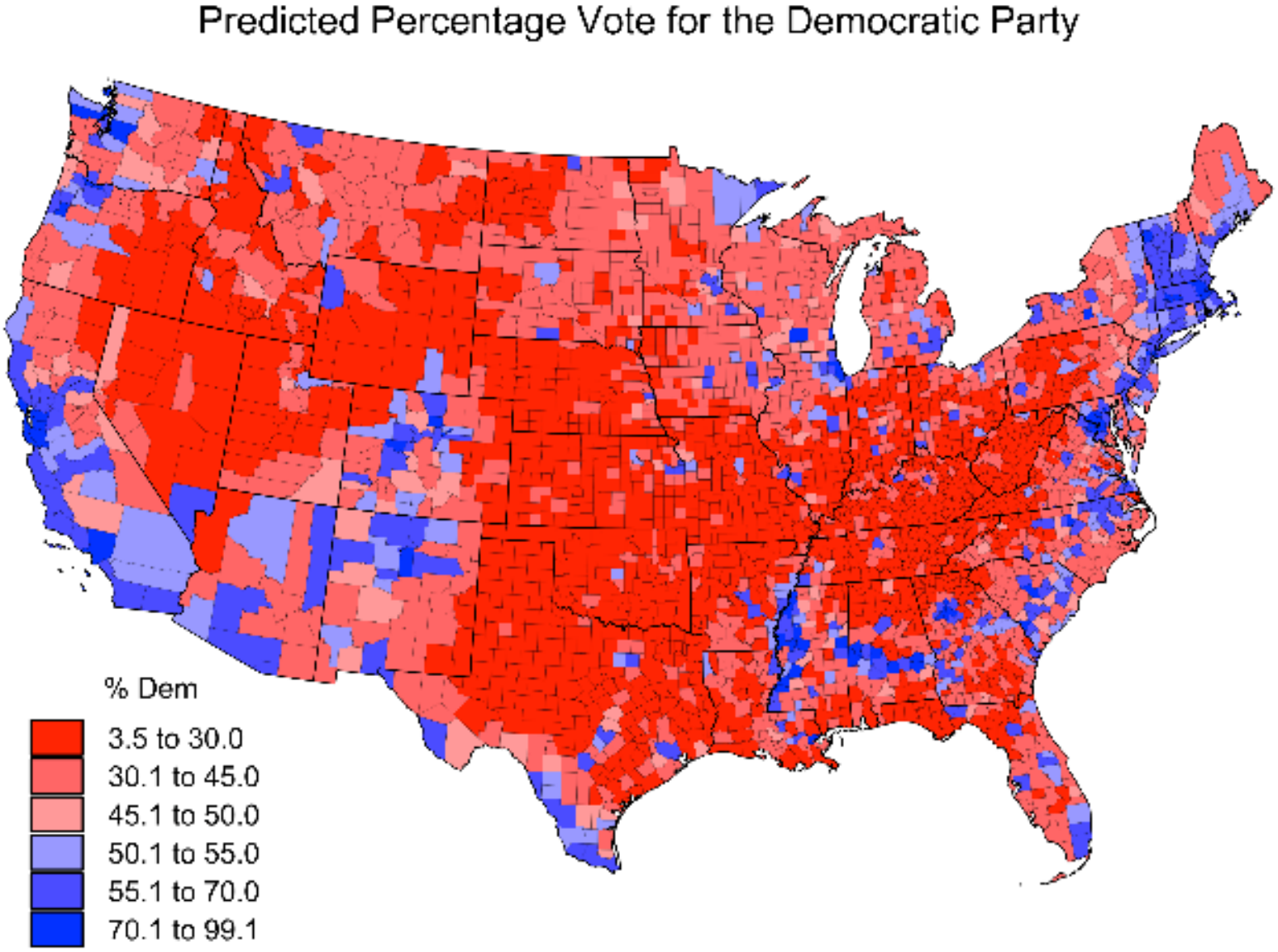


Component 3

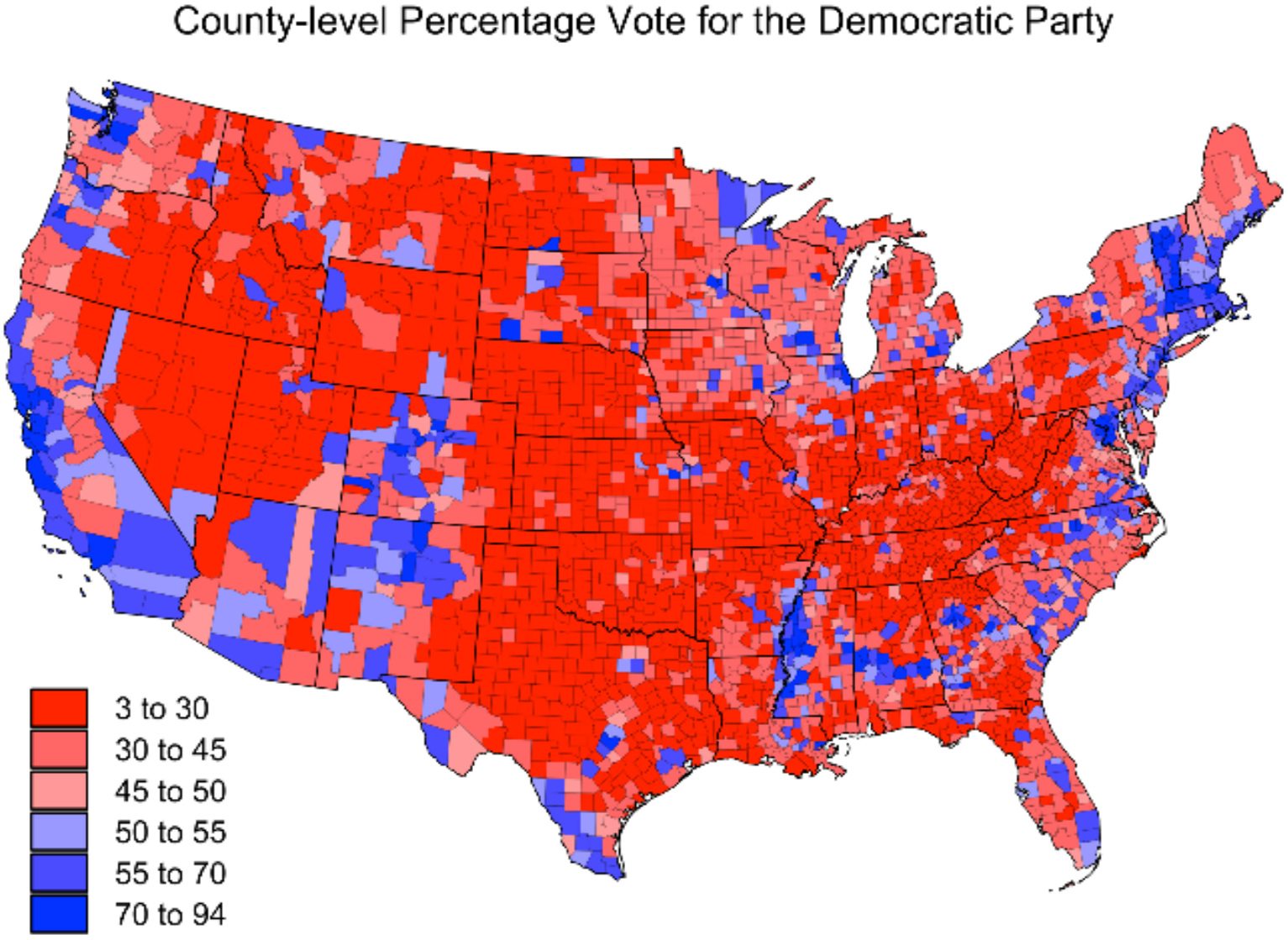


$$\sigma_y \sum_k \beta_{ik} \ (x_{ik} - \bar{x}_k) \ / \sigma_{x_k}$$

Component 1 + Component 2 + Component 3



2020 US Presidential Election: Biden vs Trump



Modeling voting behaviour with MGWR allows us to consider two counterfactuals:

Scenario 1: **What would happen if every county had the same mix of population?**

In this case, the influence of socio-economic factors would be zero (every county would have the average of every covariate)

So,

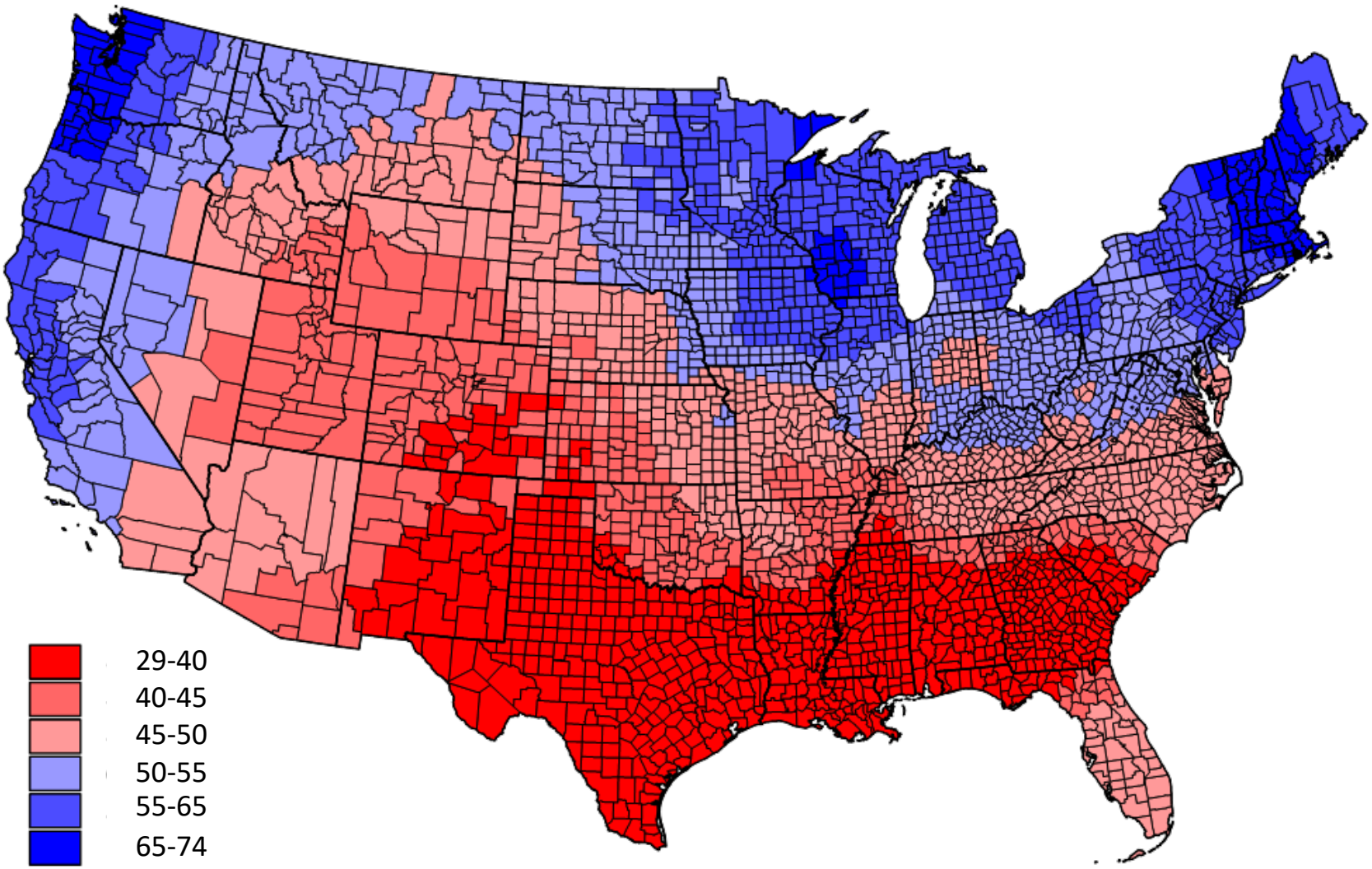
$$y_i = \dot{y} + \alpha_i \sigma_y + \sigma_y \sum_j \beta_{ij} (x_{ij} - \dot{x}_j)/\sigma_{xj}$$

Mean vote or baseline

Increment due to geographical context

Increment due to population composition

This is how we would vote if population composition were uniform across the US



Scenario 2: **What would happen if geographical context did not influence voting?**

In this case, the influence of context would be zero

So,

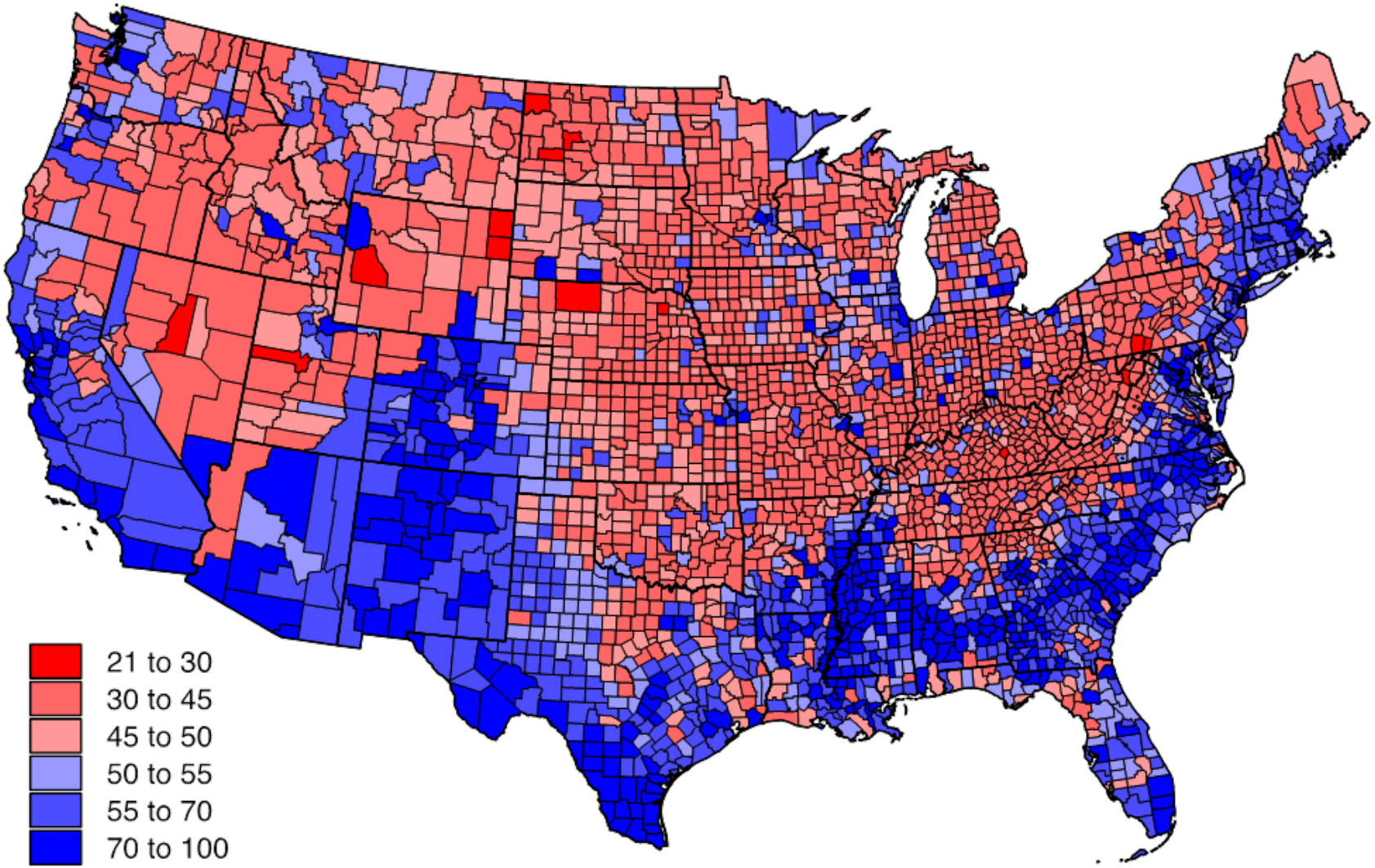
$$y_i = \dot{y} + \alpha_i \sigma_y + \sigma_y \sum_j \beta_{ij} (x_{ij} - \dot{x}_j)/\sigma_{xj}$$

Mean vote or baseline

Increment due to geographical context

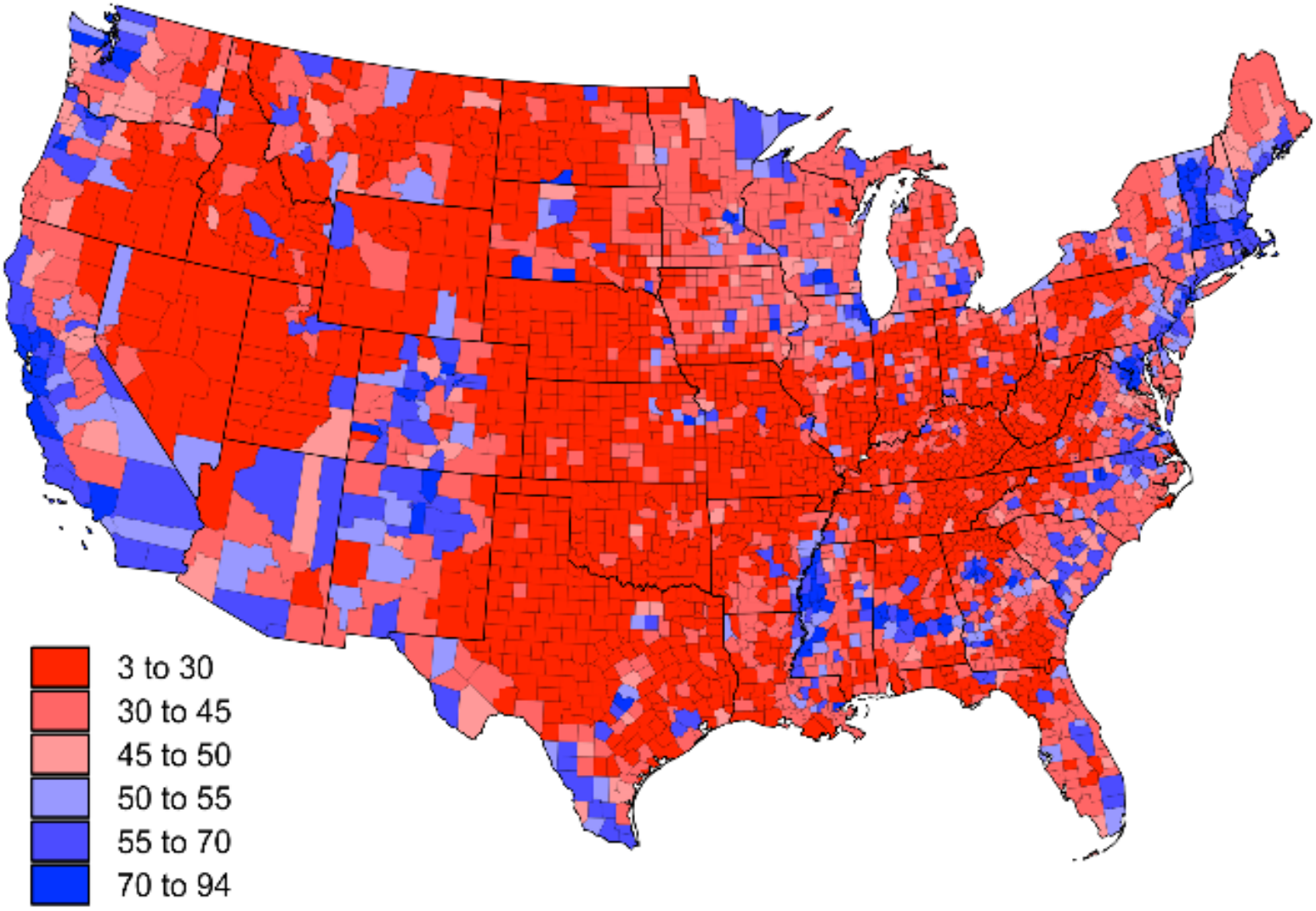
Increment due to population composition

This is how we would have voted **without** geographical context…



2020 US Presidential Election: Biden vs Trump

County-level Percentage Vote for the Democratic Party



5. Summary and Implications

1. Place or Geographical context can have a significant impact on human behavior. This has important implications for the replicability of model calibration results.

2. Contextual effects can be intrinsic and behavioral. We try to minimize the former but we can't do anything about the latter.

3. Context or Place needs to be included, rather than ignored, in models of behavior otherwise inferences drawn about other determinants of behavior may be misleading

.

The End

Basic Question

Intrinsic contextual effects

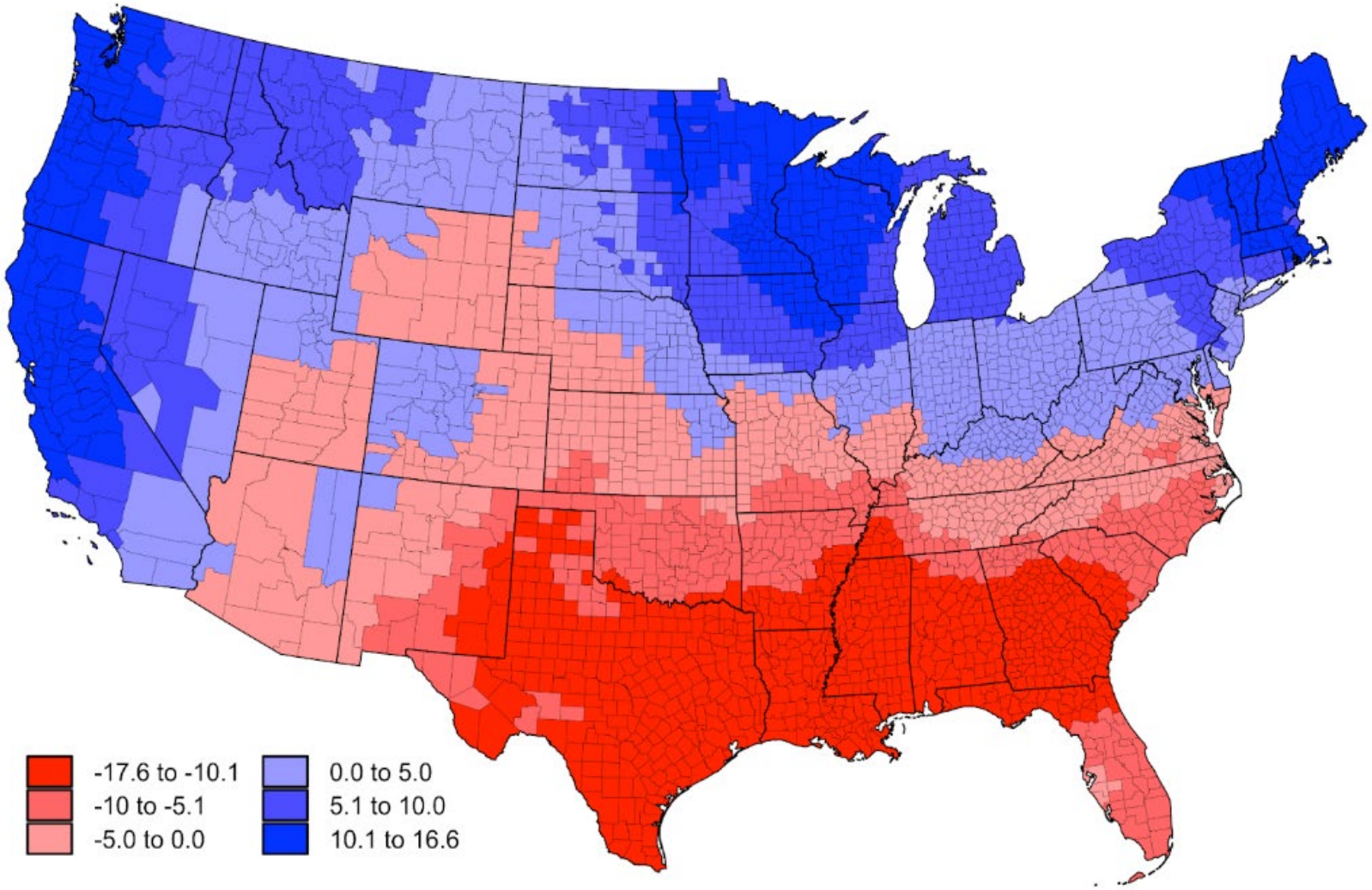
Behavioral contextual effects

Who you are

Where you are

Does where you live influence how healthy you are?
If so, how important is location compared to SDOH?
Does your location affect how SDOH influence your health?

Percentage of Democratic Vote Gained/Lost due to Geography



Division of States during the Civil War