




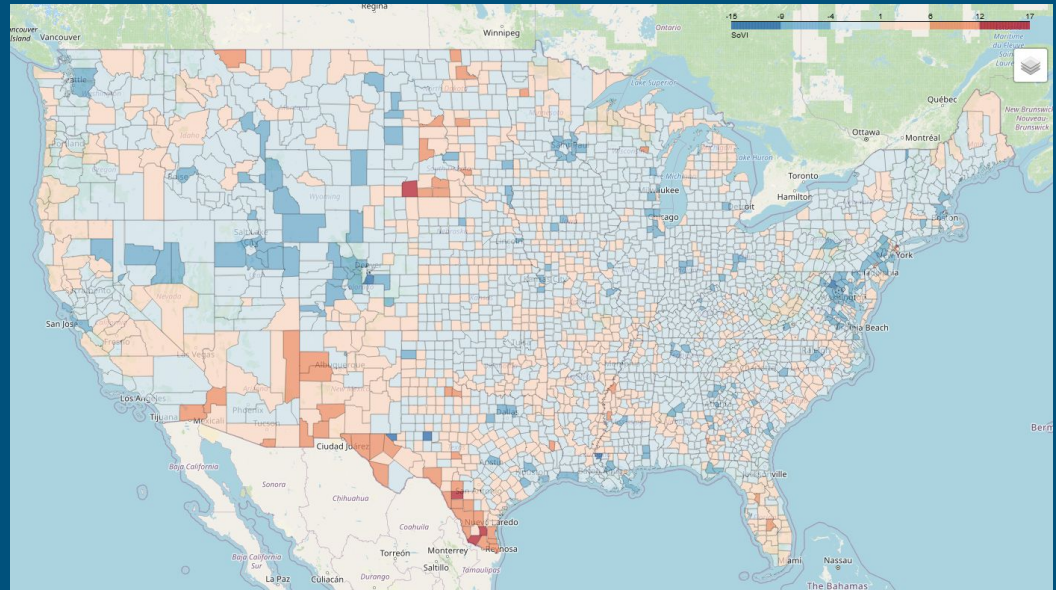
The Social Vulnerability Index: A Replication

Liam Smith and Joseph Holler,
Middlebury College
Peter Kedron, UCSB



Four-Act Story of Open Science

1. Cutter et al (2003) Social Vulnerability Index: SoVI
2. Spielman et al (2020) Replication: 21 different geographic extents
3. Reproduce Spielman et al
4. Replicate over time



1: Cutter et al (2003) SoVI

What is the Social Vulnerability Index (SoVI)?

- Cutter, S. L., Boruff, B. J., & Shirley, W. L. (2003). Social vulnerability to environmental hazards. *Social Science Quarterly*, 84(2), 242–261.
<https://doi.org/10.1111/1540-6237.8402002>
- Geographic measure of social susceptibility to harm
- Widely applied to hazards and climate change planning/policy
- 7,597 Google Scholar citations on 4/12/2024

SoVI Construction

28+ Census Variables by County

Z-score normalization
42 variables in first version

Dimensionality Reduction

Principal Component Analysis
with Varimax rotation

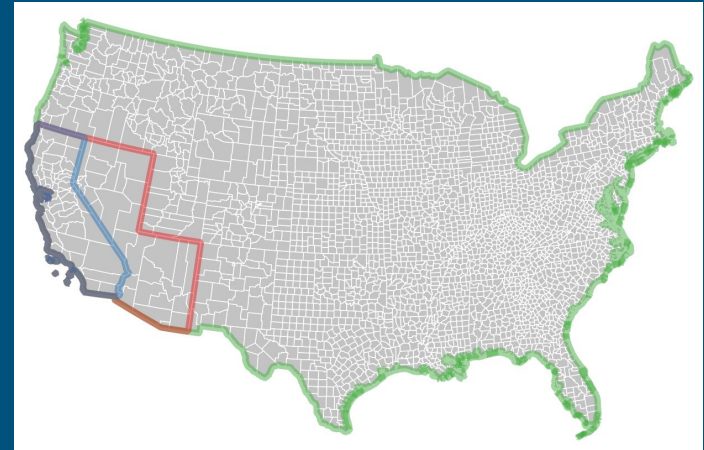
Unweighted Combination

Select components
Invert direction to match
expected vulnerability
Unweighted addition or
average
Z-score normalization

2: Spielman et al (2020) Replication

Spielman et. al Replication

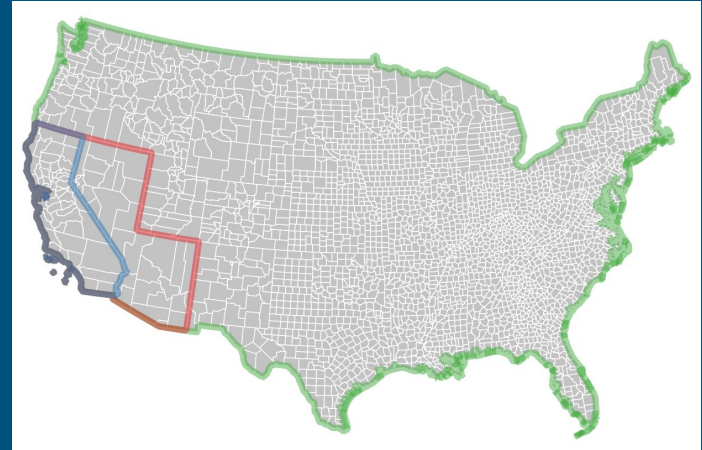
- If the study extent varies...
- Is SoVI *internally* consistent?
 - Vulnerability ranking for **places** is inconsistent
- Is SoVI *theoretically* consistent?
 - Direction and contribution of **variables** is inconsistent
- GitHub repository with data and python code



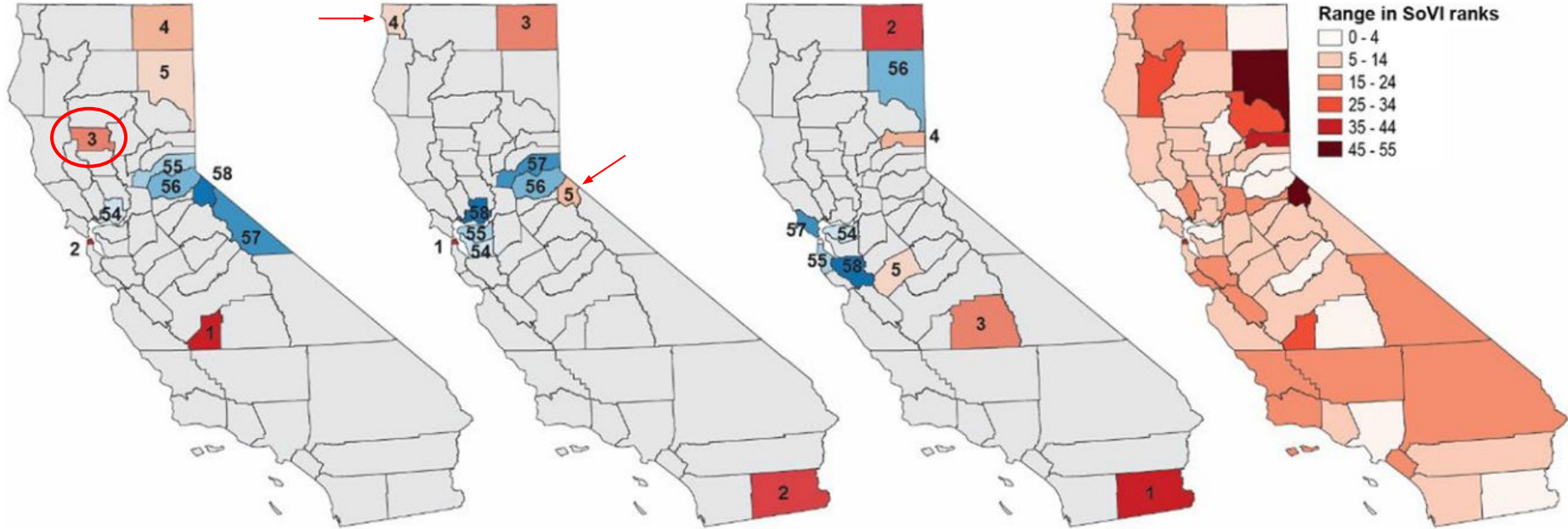
3: Reproduction and Reanalysis of Spielman et. al

Reproduction and Reanalysis Purpose

- Can we *reproduce* Spielman et. al. to confirm their results?
- Reproducible research compendium



Internal Consistency



California

FEMA #9

U.S.A.

Internal Consistency

Table 2 Spearman's rank correlation coefficient for SoVI values calculated using nested subsets of a file describing all counties in the US

	FEMA region									
	I	II	III	IV	V	VI	VII	VIII	IX	X
All US counties input file versus all counties in a state input file	0.75	0.79	0.68	0.50	0.50	0.62	0.90	0.61	0.53	0.66
All counties in a FEMA region versus counties in a state within the FEMA region input file	0.94	0.61	0.90	0.80	0.34	0.65	0.82	0.87	0.69	0.88
State used for comparison	Composite of ME, NH, MA	NY	VA	GA	IL	TX	MO	SD	CA	ID

$p < 0.01$ for all values

Theoretical Consistency

- Identical results
- Only four variables load consistently positive
- On average, the 28 variables:
 - Change direction 5.4 times
 - Range of ranks is 23.7

	National Model	Reversals	Min	Average	Max	Range
QAGEDEP_ACS	+	0	1.0	10.90	24.0	23.0
QFEMALE_ACS	+	9	1.0	8.33	28.0	27.0
QSERV_ALT	+	3	1.0	12.48	26.0	25.0
QHISP_ACS	+	3	2.0	11.00	28.0	26.0
QFEMLBR	+	4	1.0	11.62	25.0	24.0
QNATAM_ACS	+	9	1.0	9.62	28.0	27.0
QESL_ALT	+	4	1.0	11.29	27.0	26.0
QSSBEN	+	0	3.0	11.76	21.0	18.0
QNOAUTO_ALT	+	0	1.0	14.05	28.0	27.0
QMOHO	-	12	2.0	15.29	27.0	25.0
QPOVTY	+	0	3.0	11.71	27.0	24.0
QNRRES_ACS	+	4	1.0	10.19	22.0	21.0
- QFAM	+	2	1.0	13.67	28.0	27.0
QUNOCCHU_ACS	+	6	1.0	10.71	20.0	19.0
- PERCAP_ALT	+	3	9.0	18.48	28.0	19.0
BLACK_ACS	-	12	3.0	17.29	28.0	25.0
- MDGRENT_ALT	+	1	8.0	18.29	27.0	19.0
- QRICH200K	+	6	3.0	19.19	27.0	24.0
MEDAGE_ACS	+	1	8.0	18.19	28.0	20.0
QFHH_ACS	+	6	2.0	17.38	28.0	26.0
PRENTER_ACS	+	7	3.0	18.57	27.0	24.0
POPDENS	+	3	4.0	15.86	26.0	22.0
QCVLUN	-	15	1.0	13.38	23.0	22.0
- MHSEVAL_ALT	+	4	5.0	19.29	28.0	23.0
QED12LES_ALT	+	2	4.0	15.00	28.0	24.0
QEXTRCT_ALT	+	6	1.0	17.71	28.0	27.0
QASIAN_ACS	-	14	4.0	18.71	28.0	24.0
QPUNIT_ACS	+	14	2.0	16.05	28.0	26.0

Reproduction Findings & Outcomes

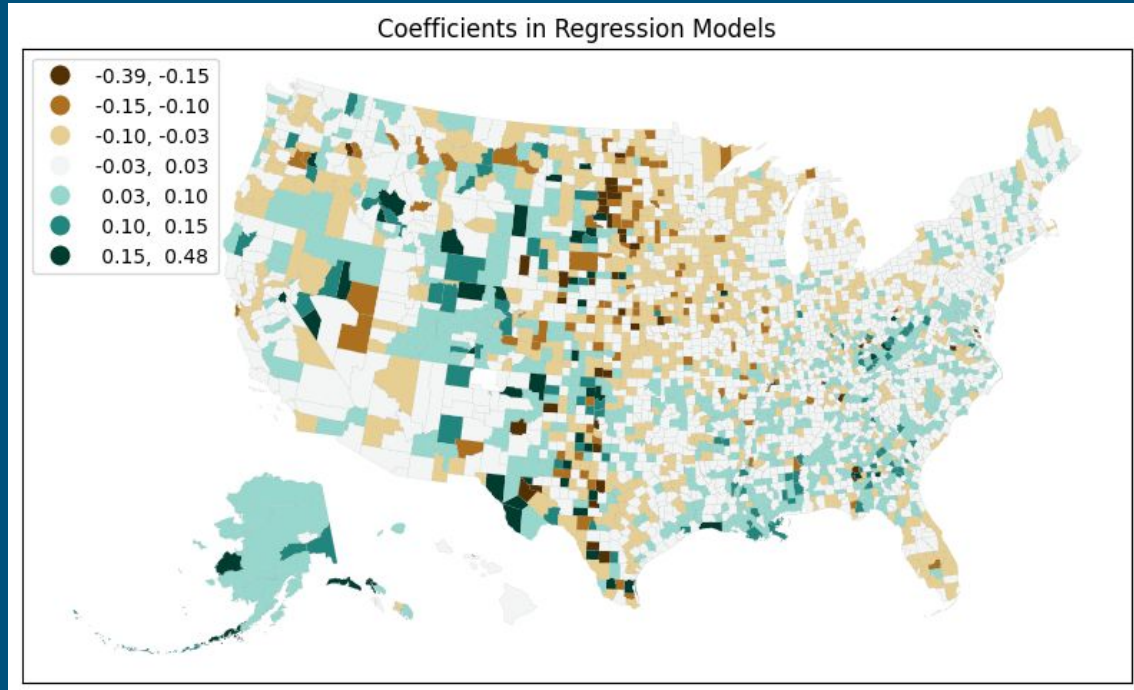
- Reproduction is not trivial
 - Outdated packages
 - Extraneous data and code
 - File organization
 - Code for figures
- Reproducible research compendium
 - <https://osf.io/dzpe9/>
- Identically reproduced output data
- Subtle differences in results
- Confirm findings
 - internal inconsistency
 - theoretical inconsistency

4: Replication of Spielman et al over time

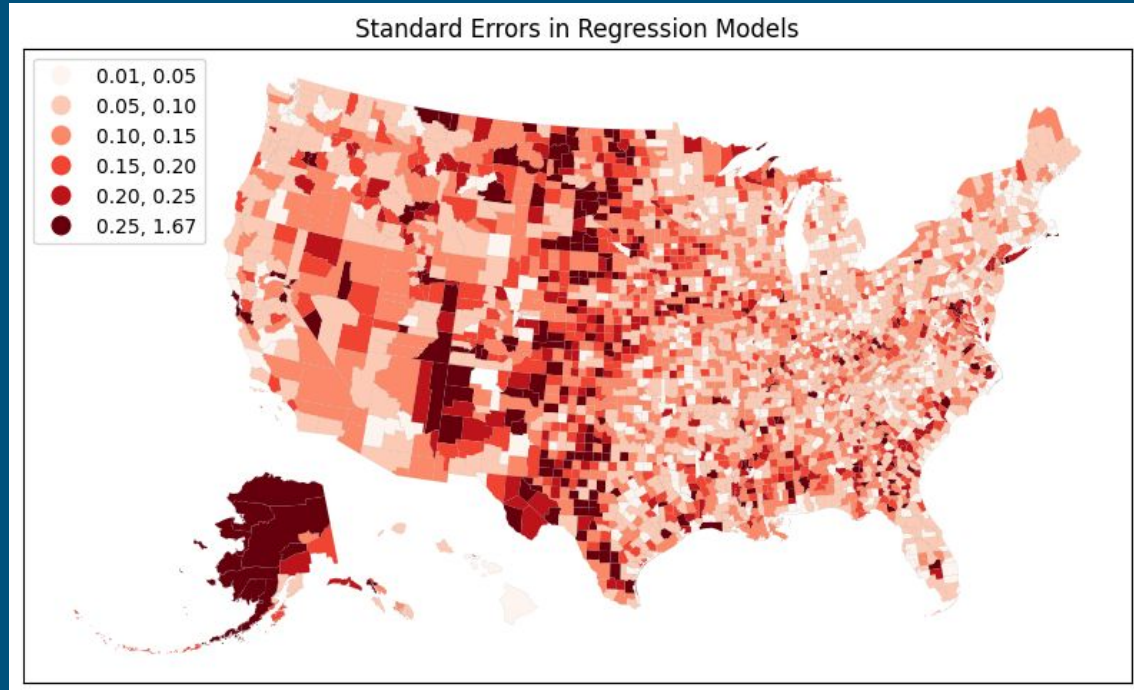
Replication Data & Methods

- Temporal support
 - 5-year ACS Estimates
 - 10 years: 2012 - 2021
- Geographic support
 - USA counties
- Internal consistency: do places exhibit linear trends in SoVI over time?
- Theoretical consistency: do variables contribute similarly to SoVI over time?

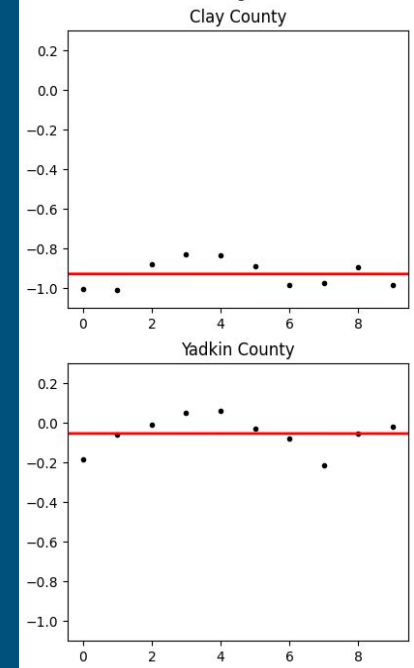
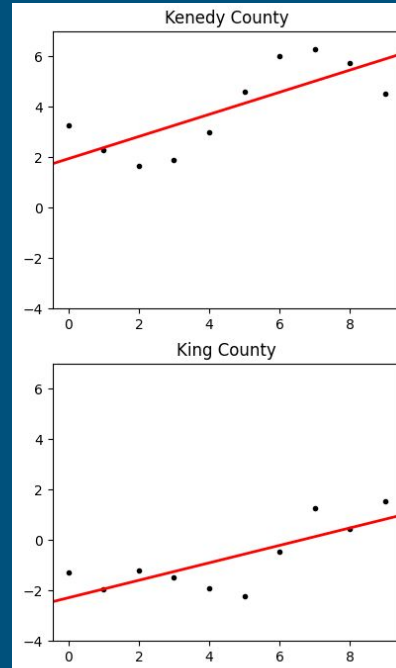
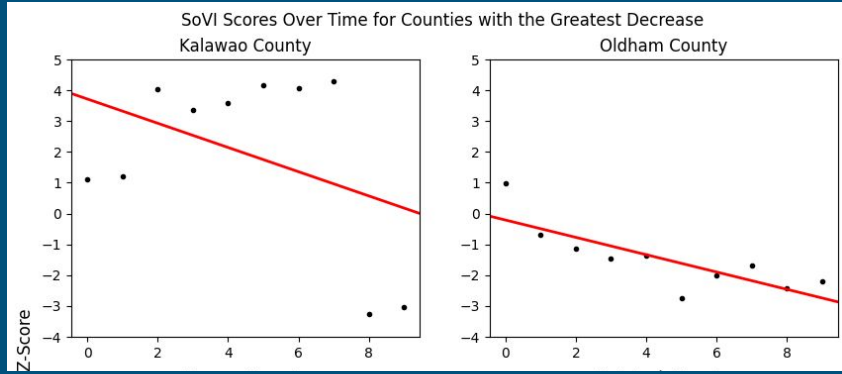
County SoVI Change over Time



County SoVI Change over Time



Changes in counties with high error



Conclusions

- Reproduction & replication studies assess the consistency of prior research
- Open science reduces barriers to validating and extending prior research
- Confirmed Spielman et al's finding of internal and theoretical inconsistency in SoVI over geographic extent
- Extended methods and findings to temporal extent
- Question application of SoVI to monitor change in social vulnerability over space and time

Acknowledgements

NSF award BCS-2049837:

Transforming theory-building and STEM education through reproductions

Full study repository and reports: <https://osf.io/dzpe9/>

Reproducible research template: <https://osf.io/w29mq/>

HEGSRR.github.io