Methodological challenges to confirmatory latent variable models of social vulnerability

Zachary T. Goodman¹ · Caitlin A. Stamatis¹ · Justin Stoler^{2,3} · Christopher T. Emrich⁴ · Maria M. Llabre¹

Received: 13 August 2020 / Accepted: 19 January 2021 / Published online: 13 February 2021
© The Author(s), under exclusive licence to Springer Nature B.V. part of Springer Nature 2021

Abstract

Socially vulnerable communities experience disproportionately negative outcomes following natural disasters and underscoring a need for well-validated measures to identify those at risk. However, questions have surfaced regarding the factor structure, internal consistency, and generalizability of social vulnerability measures. A reliance on data-driven techniques, which are susceptible to sample-specific characteristics, has likely exacerbated the difficulty generalizing social vulnerability measures across contexts. This study sought to validate previously published structures of SoVI using confirmatory factor analysis (CFA). We fit CFA models of 28 sociodemographic variables frequently used to calculate a commonly used measure, the social vulnerability index (SoVI), drawn from the American Community Survey across 4162 census tracts in Florida. Confirmatory models generally did not support theory-driven pillars of SoVI that were previously used to study vulnerability in the New York metropolitan area. Modified models and alternative SoVI factor structures also failed to fit the data. Many of the input variables displayed little to no variability, limiting their utility and explanatory power. Taken together, our results highlight the poor generalizability of SoVI across contexts, but raise several important considerations for reliability and validity, as well as issues related to source data and scale. We discuss the implications of these findings for improved theory-driven measurement of social vulnerability.

Keywords Vulnerability · Measurement · Evaluation · Validity · Social indicators · Social vulnerability · SoVI

✉ Zachary T. Goodman
ztg8@miami.edu

¹ Department of Psychology, University of Miami, 5665 Ponce de Leon Blvd, Office 446, Coral Gables, FL 33146-0751, USA

² Department of Geography and Regional Studies, University of Miami, Coral Gables, FL, USA

³ Department of Public Health Sciences, Miller School of Medicine, University of Miami, Miami, FL, USA

⁴ College of Community Innovation and Education, University of Central Florida, Orlando, USA

1 Introduction

Extreme events have inspired decades of research on risk and vulnerability, with the goal of identifying communities with the least capacity to prepare for, respond to, and rebound from disasters. Unfortunately, studies suggest that beyond the impact of physical hazard, socially disadvantaged communities face disproportionate risk when exposed to natural disasters (Laska and Morrow 2006). Similar observations underpinned the development of early social vulnerability indices, which utilize variability in sociodemographic data to quantify population vulnerability to loss in the face of an event such as a hurricane (e.g., Clark et al. 1998; Cutter et al. 2000; Tapsell et al. 2002).

Although the theoretical links between sociodemographic factors and hazard vulnerability are well established, there remains considerable debate around best practices for measuring social vulnerability (Beccari 2016). Several social vulnerability indices have been proposed to quantify this latent construct, but most are ecological in nature (i.e., measuring vulnerability at various levels of census geography or other units), and few studies have attempted to validate them or otherwise assess their relative utility (Fekete 2019). Decision makers have begun to utilize measures of social vulnerability for resource disaster recovery allocation (New York Times 2020), and nearly 40% of states are planning on using vulnerability measures to prioritize COVID-19 vaccine distribution (NPR 2020; Schmidt et al 2020). Such use of this theoretical contract and its empirical manifestations, although seemingly ethical in nature, may have unknown or unintended consequences. Although these measures may have face validity and are being applied, we still don't know how "good" or "useful" these measures are from a variety of perspectives. Here, utility may take several forms, including ease of construction, internal consistency, robustness, transferability, or explanatory power. Arguably, the best model would perform well on each of these criteria, yet most social vulnerability models have not been rigorously tested.

Validation studies aimed at understanding the utility of social vulnerability surrogates, variable aggregates, and models/indices remain difficult to undertake for a variety of reasons, including the lack of available point level data on vulnerability and outcome measures, the ecological fallacy created when aggregating outcomes and social vulnerabilities to various levels of geography and the challenges related to using a multivariate characteristic such as social vulnerability to describe univariate outcomes such as total losses, fatalities, slow recovery and the like. Recently, Rufat et al. (2019) tested the empirical validity of four social vulnerability indices against Federal Emergency Management Agency (FEMA) claims data from Hurricane Sandy (as a proxy for losses) aggregated to census tracts, with mixed results. The four indices they tested were highly variable in their use of data driven, pillar-based, and expert opinion approaches.

One of these indices, the social vulnerability index (SoVI) originally developed by Cutter and colleagues (2003) remains among the most heavily referenced and often utilized for understanding capacity in the face of disaster. SoVI's original formulation—a 42 variable empirical measure of sociodemographic and built environment characteristics of community capacity to prepare for, respond to, and rebound from disaster events—has a long history of utility in both academic settings with more than 4700 citations and applied planning tools (NOAA 2020; Surging Seas 2014). Derived as an empirical method to understand, measure, and monitor difference between places, SoVI is the product of well-known individual/community social inequality measures and place inequality measures or those characteristics contributing to the social vulnerability of places. SoVI, the first large-scale (county level) composite measure capturing the broad suite of social vulnerability

characteristics identified through a literature review, has been largely accepted and applied across academic and practical fields of use. SoVI's more recent incarnations (27–29 variables focused only on sociodemographics) have long filled the need to understand lack of capacity to prepare for, respond to, and rebound from environmental threats. However, an important distinction here is that SoVI's original conceptualization centered on building a model based on a retrospective of disaster impacts and outcomes rather than building the most robust model from a psychometric perspective. Although, the SoVI model has become one of the more notable targets for critical review. Most recently, Spielman and colleagues (2020) observed several problems with two measures used to determine a model's utility, construct validity, and internal consistency. Spielman and colleagues (2020) further argue that SoVI's construction is not consistently robust with respect to traditional psychometric theory. Although the recent criticism is specific to SoVI, it is not clear if any existing social vulnerability measures are generalizable beyond context-specific applications. Rather, these limitations may lie in the approach to measuring and assessing the latent constructs of interest.

Although the advent of social vulnerability measurement has yielded some important insights in the disaster risk literature, several limitations remain in the ways which researchers have attempted to quantify social vulnerability. On the one hand, the concept is theoretically robust; it has long been appreciated that social factors influence one's experience of hazards and stressors (Turner et al. 2003). Simultaneously, data-driven and pillar-based (i.e., construct-based) approaches, such as those tested by Rufat and colleagues (2019), appear to be limited in their generalizability. This hinders progress toward the ultimate goals of developing portable metrics for understanding who is affected by disasters and mitigating related losses.

The development of more generalizable social vulnerability indices may benefit from statistical approaches that combine theoretical or conceptual knowledge of vulnerability with empirical confirmation. Data-driven approaches, such as principal component analysis (PCA) and exploratory factor analysis (EFA), are known for their susceptibility to sample-specific characteristics of the data (Widaman 2018). Models informed by theory, such as the pillar-based approach by Rufat and colleagues (2019), can help reduce sample specificity and overfitting by imposing constraints consistent with theory. At the same time, purely pillar-based or expert-based models account for theoretical knowledge but minimize nuances in the data pertinent to a given disaster or a given community. Moving forward, this area of research can rely on current confirmatory latent variable approaches that incorporate the best of both worlds.

Latent variable models can enhance the generalizability of social vulnerability indices. In particular, confirmatory factor analysis (CFA) is a framework that uses theory-guided a priori knowledge to specify a latent construct or constructs hypothesized to underlie the covariation among a set of indicators (Llabre and Fitzpatrick 2012). Applying CFA to social vulnerability indexing would involve specifying indicators, or observed variables, that are presumed to reflect the construct or constructs of interest. CFA can test single factor or multifactor models, and for each factor or latent variable, the indicators reflecting the construct are specified. For example, if theory suggests that a socioeconomic factor is important for social vulnerability indices, we can specify census measures—such as per capita income, percent living below the poverty line, and median house value—as indicators of the latent socioeconomic construct. CFA models partition the variability in an indicator into that which is derived from the latent variable and the residual which contains specific variance and error. The model yields estimate of the loadings associated with each indicator. These loadings represent the correlation between the latent variable and the

indicator and provide a measure of the adequacy of the indicator in reflecting the construct (Kline 2016; Labre and Fitzpatrick 2012). Further, squared standardized factor loadings provide the proportion of explained variance of the observed variable. Tests of fit between the model specified and the data provide important information about the adequacy of the theory in a given situation, as well as the adequacy of the indicators. Applying a confirmatory measurement modeling approach to social vulnerability indices may provide a deeper understanding of the latent social vulnerability construct. Additionally, this approach may identify additional social characteristics that can best inform policies to decrease disaster vulnerability.

In the present study, our overarching objective was to apply confirmatory factor analysis (CFA) to evaluate SoVI model construct validity and generalizability following separate natural disasters (Rufat et al. 2019). This study represents a departure from the purely data-driven approaches typically used to quantify social vulnerability by applying a theoretically-based pillars approach to SoVI's raw underlying data. Our overall intent is to leverage the best aspects of prior social vulnerability indices and incrementally create a revised index with stronger empirical validity for our study area, Florida. Our first aim was to fit confirmatory measurement models of each of the five pillars identified by Rufat and colleagues (2019), including socioeconomic status, population structure, race and ethnicity, access and needs, and housing structure (Fig. 1). Assuming we could attain acceptable fit for each measurement model, our second aim was to integrate the individual measurement models into a five-factor model representing the overall SoVI construct and ultimately test its association with FEMA claims data from Hurricane Irma, the fifth costliest hurricane to hit the mainland United States (National Hurricane Center 2018). But, we did not get that far our analysis encountered several obstacles highlighting some of the underlying challenges that have afflicted past attempts to validate social vulnerability indices. This paper focuses on those methodological issues and discusses the implications for generalizable social vulnerability measures with recommendations for future attempts to quantify and validate this construct.

2 Method

2.1 Data and data screening

This study analyzed the 28 sociodemographic census variables commonly used as candidates in the computation of SoVI (Table 1), gathered from the American Community Survey, five-year product (2014–2017) and aggregated to 4,246 Florida census tracts. We could have used alternate spatial units, such as counties or metropolitan statistical areas, but census tracts are the smallest geographic unit for which demographic variables are considered to be statistically reliable (Spielman et al. 2014). Of the initial sample, 84 tracts were excluded due to zero residents, resulting in an effective sample of 4162 tracts.

In preparation for testing confirmatory factor models, we examined the variability, skewness, kurtosis, and scaling of each SoVI variable before proceeding with further analyses. Low variability indicates a variable that is relatively stable across census tracts. In the extreme case, when there is no variability for a given measure (i.e., the variable was effectively constant), it was impossible to assess the variable's quality as an indicator of a latent construct or relationships to other variables.

Fig. 1 Theoretical correlated factor structure of SoVI

In addition to variability, we also addressed variable scaling, as estimation difficulties can arise when variables have drastically different ranges. We chose to rescale percentage variables by multiplying them by 100, and we divided absolute measures such as per capita income, median house value, and median rent by 1000. Measures of skewness and kurtosis provide information about the extent to which variables meet the normality assumptions required by maximum likelihood estimation. The recommended cut points are ± 3 for skewness and 8 for kurtosis (Kline 2016).

Table 1 Descriptive statistics for each of the SoVI census variables at the census tract level

Factor	Census variable	Variable label	<i>M</i>	<i>SD</i>	Skew	Kurtosis
Socioeconomic status	Per capita income (\$)	PERCAP	29,065.44	16,955.50	2.59	10.87
	Living under poverty line (%)	POVTY	0.16	0.11	1.56	4.26
	Unemployed (%)	UNEMP	0.28	0.09	1.25	3.30
	Median house value (\$)	MHSEVAL	205,014.73	165,567.79	4.05	29.31
	Annual income over \$200,000 (%)	RICH200K	0.05	0.07	2.76	10.04
	Working in service industry (%)	SERV	0.20	0.08	0.73	2.44
	Working extractive industry (%)	EXTRCT	0.01	0.03	6.56	60.41
Population structure	Less than 12 years of education (%)	ED12LES	0.13	0.09	1.32	2.36
	Median rent (\$)	MDGRENT	1161.64	413.43	1.67	5.31
	Median age of population (#)	MEDAGE	43.81	10.62	0.75	0.36
	Population with age-based dependents (%)	AGEDEP	0.27	0.13	1.77	3.83
	People per housing unit (#)	PUNIT	2.64	0.56	0.53	0.37
	Female head of households (%)	FHH	0.13	0.08	1.06	1.58
	Female population that is employed (%)	FEMLEBR	0.48	0.06	0.06	1.78
Race and ethnicity	Population that is female (%)	FEMALE	0.51	0.05	-2.08	14.57
	Families that are married (%)	MARFAM	0.64	0.21	-0.57	-0.06
	Population that is Asian American (%)	ASIAN	0.02	0.03	2.44	9.22
	Population that is African American/Black (%)	BLACK	0.15	0.20	2.07	3.89
	Population that is Hispanic/Latino (%)	HISP	0.22	0.23	1.68	2.19
	Population that is native American (%)	NATAM	<0.01	0.01	9.95	170.57
	Residing in nursing homes (%)	NURSE	<0.01	0.01	5.32	44.84
Access and needs	Receiving social security benefits (%)	SSBEN	0.37	0.16	0.72	0.60
	English as a second language (%)	ESL	0.11	0.13	2.04	4.35
	Without automobiles (%)	NOAUTO	0.07	0.07	2.24	6.99
	Uninsured (%)	UNINSURED	0.15	0.08	0.80	1.08

Table 1 (continued)

Factor	Census variable	Variable label	<i>M</i>	<i>SD</i>	Skew	Kurtosis
Housing structure	Renters (%)	RENTER	0.30	0.18	0.84	0.21
	Unoccupied housing units (%)	UNOCCHU	0.18	0.13	1.80	4.43
	Population residing in mobile homes (%)	MOHO	0.09	0.16	2.13	4.19

2.2 Statistical analyses

2.2.1 Confirmatory factor analysis

We conducted a series of confirmatory factor analyses (CFA) with full-information maximum likelihood (FIML) estimation to test the structure of theoretically driven models of social vulnerability. All statistical analyses were conducted in *R* version 3.6.3, and CFA was conducted via the *lavaan* package (Rosseel 2012). For each model, we explicitly specified which indicators (i.e., variables) served as observed measures of a given latent construct. Measures of model fit allowed us to evaluate how well the theoretically specified model was supported by the data. Each CFA was evaluated for overall model fit and reliability, with model fit evaluated according to the following standard criteria: nonsignificant chi-square test (χ^2), root mean square error of approximation (RMSEA) less than .08, standardized root mean square residual (SRMR) less than .06, and comparative fit index (CFI) greater than .95 (Kline 2016). A standardized factor loading (λ) less than .30 suggests the indicator is a weak reflection of the latent construct. We also examined Cronbach's alpha (α), a commonly used metric of internal consistency, for the five measurement models (i.e., for each pillar); α greater than .70 generally indicates acceptable reliability. Indicators which loaded negatively onto a latent variable were reversed (i.e., multiplied by -1) before calculating reliability. Fit indices for CFA models are reported in Table 2. Standardized and unstandardized factor loadings as well as associated significance tests are reported in Table 3.

2.2.2 Initial single-factor measurement model specification

To test our first aim, we specified a set of confirmatory single-factor measurement models based on the pillars suggested by Rufat and colleagues (2019) (Table 1). Socioeconomic status was indicated by *per capita income*, *percent living under the poverty line*, *median house value*, *percent earning over \$200,000 per year*, *percent working in the service sector*, *percent working in the extractive sector*, *percent with less than 12 years of education*, and *median rent*. Population structure was indicated by *median age*, *percent living with a child younger than 5 years or adult older than 65 years*, *people per housing unit*, *percent of female-headed households*, *percent of females employed*, *percent of population which is female*, and *percent of married families*. Race and ethnicity was indicated by four variables: *percent of population which is African American*, *Hispanic/Latino*, *Native American*, and *Asian American*. Access and needs was indicated by *percent living in a nursing home*, *percent receiving social security benefits*, *percentage of population in which English is a second language*, *percent of population who do not have an automobile*, and *percent uninsured*. Housing structure was indicated by *percent of units occupied by renters*, *percent of vacant houses*, and *percent living in a mobile home*.

2.2.3 Single-factor model respecification

If specified single-factor models did not fit the data well, we explored modifications to improve fit. While model respecification is an empirically guided step, we sought to make modifications consistent with theory, in line with our overarching aim of increasing the generalizability of the SoVI model. To make modifications, we examined

Table 2 Fit indices for single-factor CFA models with subsequent modified models including residual covariances between indicators

Factor	χ^2	<i>df</i>	<i>p</i>	CFI	RMSEA	SRMR
<i>Theoretical factor structures</i>						
Socioeconomic status						
Initial model	1872.85	20	< .001	.91	.15	.05
+ PERCAP & MDGRENT	1323.39	19	< .001	.94	.13	.04
+ MHSEVAL & RICH200K	1035.75	18	< .001	.95	.12	.04
+ MHSEVAL & MDGRENT	840.01	17	< .001	.96	.11	.04
+ PERCAP & RICK200K	648.75	16	< .001	.97	.10	.04
+ PERCAP & MHSEVAL	525.11	15	< .001	.98	.09	.03
+ UNEMP & SERV	372.43	14	< .001	.98	.08	.03
Population structure						
Initial model	4092.57	14	< .001	.68	.27	.15
+ FEMALE & FEMLBR	2780.00	13	< .001	.78	.23	.11
+ FHH & MARFAM	2101.53	12	< .001	.84	.21	.09
+ FHH & PUNIT	1175.68	11	< .001	.91	.16	.09
Race and ethnicity						
Initial model	3.72	2	.156	.99	.01	.01
Access and needs						
Initial model	271.77	2	< .001	.89	.18	.06
+ UNINSURED & NOAUTO	65.22	1	< .001	.98	.12	.03
<i>Alternative factor structures</i>						
Correlated factors						
Model ^a	31,683.32	233	< .001	.60	.18	.17
Bifactor model						
Model ^b	34,372.41	251	< .001	.58	.18	.17
Single-factor model						
Initial model ^c	44,939.25	252	< .001	.43	.21	.17
Final model ^d	9245.93	111	< .001	.85	.14	.11

CFI, comparative fit index; RMSEA, root mean square error approximation; SRMR, standardized root mean residual

^aModel includes each theorized factor with correlations between factors

^bModel includes each theorized factor model and a general factor indicated by all SoVI variables

^cAll SoVI variables loading onto a broad, single factor

^dSingle-factor model with 23 residual covariances included

problematic indicators (e.g., low factor loadings) and modification indices suggested by *lavaan*. Typically, model respecification was done by removing indicators or correlating residual variances between indicators, with the latter modification suggesting that two variables share some relationship above and beyond the covariance explained by the latent variable. In the case of correlating residual variances, it is generally the responsibility of the analyst to interpret the covariance and theorize what external factor might contribute to these associations.

Table 3 Unstandardized and standardized factor loadings, with associated test statistics, for the final single-factor models

Indicator	λ	b	SE	z	p
<i>Socioeconomic status</i>					
PERCAP	.92	0.43	0.01	74.04	< .001
POVTY	−.80	−0.52	0.01	−61.04	< .001
CVLUN	−.40	−0.13	0.01	−27.31	< .001
MHSEVAL	.66	0.46	0.01	45.19	< .001
RICH200K	.79	0.74	0.01	58.10	< .001
SERV	−.68	−0.30	0.01	−49.13	< .001
ED12LES	−.84	−0.60	0.01	−65.72	< .001
MEDRENT	.65	0.11	< 0.01	45.10	< .001
<i>Population structure</i>					
MEDAGE	.99	0.22	< 0.01	78.82	< .001
AGEDEP	.85	0.34	0.01	63.88	< .001
PUNIT	−.58	−0.12	< 0.01	−39.88	< .001
FHH	−.58	−0.37	0.01	−39.71	< .001
FEMPLBR	−.05	−0.01	< 0.01	−3.36	.001
FEMALE	.13	0.01	< 0.01	8.34	< .001
MARFAM	.19	0.09	0.01	11.98	< .001
<i>Race and ethnicity</i>					
ASIAN	.24	0.17	0.04	4.56	< .001
BLACK	.18	0.22	0.05	4.39	< .001
HISP	.41	0.40	0.09	4.69	< .001
<i>Access and needs</i>					
UNINSURED	.63	0.37	0.01	29.93	< .001
SSBEN	−.43	−0.20	0.01	−23.36	< .001
ESL	.79	0.79	0.02	33.42	< .001
<i>Housing structure</i>					
RENTER	.43	0.30	0.02	15.16	< .001
UNOCCHU	−.49	−0.33	0.02	−15.69	< .001
MOHO	−.40	−.60	0.04	−14.80	< .001

λ , standardized factor loadings; b , unstandardized factor loadings; SE , standard error

2.2.4 Correlated factor model

Although we observed that the single-factor models were not a good fit, we tested our second aim by fitting a correlated five-factor model that tested the overall fit of the theoretical SoVI structure. All five of the single-factor models described above were included in a model simultaneously, and latent variables were allowed to correlate. We tested the correlated factor model first using the original single-factor models (i.e., with no modifications). We then tested the model again including the respecifications from the single-factor models to see if this would improve fit.

2.2.5 Alternative models

To examine alternative factor structures of SoVI variables, we tested several additional models. These included a single-factor model with a broad social vulnerability latent variable indicated by all previously mentioned indicators and a bifactor model which tested both a general and domain-specific latent variable. A single-factor model implies all census variables represent a common, social vulnerability construct rather than representing distinct but related subfactors, as discussed above—this solution would reject the theory-based pillars of past literature. A bifactor model is a combination of the single-factor model and the specific factor models discussed above and implies all census variables represent two constructs: a general social vulnerability latent variable and a domain-specific latent variable (e.g., socioeconomic status).

3 Results

3.1 Data screening

Many SoVI candidate variables yielded minimal variability across census tracts in Florida; several percentage variables had standard deviations as low as 0.01 (Table 1). Normality was a concern for many of the variables included in SoVI (Table 1 and Fig. 2). Skewness ranged from -2.08 (percent of females employed) to 9.95 (*percent of population that is Native American*), with a median skewness of 1.68 ($SD=2.27$). Similarly, kurtosis ranged from -0.06 (*percent of married families*) to 170.57 (*percent Native American*), with a median kurtosis of 4.04 ($SD=33.63$). Consequently, all variables were transformed logarithmically to resolve complications in model estimation that can arise due to non-normality. Due to a lack of variability and concerning high skew, three census variables were

Fig. 2 Frequency distributions for all 28 census tract-level SoVI indicators before conducting logarithmic transformations

removed from subsequent analyses: *percent working in the extractive sector*, *percent of the population which is Native American*, and *percent living in a nursing home*.

3.2 Single-factor confirmatory factor models

3.2.1 Socioeconomic status

The initial CFA demonstrated poor overall fit, although the SRMR was acceptable (Table 2). Several respecifications based on modification indices were conducted to reach acceptable model fit, and improvements between each respecified model were marginal. Residual covariances were added one at a time between the following pairs of census variables, with model fit reassessed after each addition: *per capita income* and *median rent*; *median house value* and *percent earning over \$200,000*; *median house value* and *median rent*; *per capita income* and *percent earning over \$200,000*; *per capita income* and *median house value*; and *percent unemployed* and *percent working in the service sector*. The final respecification step improved model fit to acceptable cutoffs and produced a reliable latent variable ($\alpha = .88$). Note that all of these respecifications are data driven and require additional validation.

3.2.2 Population structure

The initial single-factor model demonstrated very poor fit to the data (Table 2). A residual covariance between percent of females employed and percent of the population which is female improved fit substantially, although fit was still poor according to standard criteria (Kline 2016). While the addition of a residual covariance between *percent of female-headed households* and *percent of married families* improved fit, as did the inclusion of a covariance between *percent of female-headed households* and *percent of people per housing unit*, attempts to include additional parameters produced negative variances. Consequently, the final model did not achieve acceptable fit indices as demonstrated by marginal reliability ($\alpha = .67$).

3.2.3 Race and ethnicity

The indicator representing *percent of the population that is Native American* was removed due to minimal variability, such that the race and ethnicity factor had only 3 indicators. While factor loadings for a single-factor model with three indicators can be estimated, fit indices cannot be produced because a minimum of four indicators is required in order to test model fit. Consequently, we focused on the factor loadings, which revealed several concerns. The standardized factor loadings of *the percent of the population identifying as Asian American* ($\lambda = .24$) and *African American* ($\lambda = .18$) were low, suggesting that the indicators do not share significant common variance. Additionally, these indicators demonstrated highly skewed and leptokurtic distributions, which persisted after logarithmic transformations. Unsurprisingly, the reliability of the race and ethnicity factor was extremely low ($\alpha = .17$), which may mirror the transition from historic settlement patterns shaped by racial capitalism to newer, evolving patterns of residential segregation (Ellis et al. 2018).

3.2.4 Access and needs

The initial model demonstrated poor fit to the data (Table 2). Of note, the standardized factor loading of *percent receiving social security benefits* ($\lambda = -.34$) was negative. The inclusion of a covariance between *percent uninsured* and *percent without an automobile* improved fit; however, this addition resulted in a low standardized factor loading for percent without an automobile ($\lambda = .25$). Removal of *percent without an automobile* resulted in a model with three indicators, so fit could not be examined. The loading of *percent receiving social security benefits* remained negative ($\lambda = -.43$), and the final model demonstrated low reliability ($\alpha = .60$).

3.2.5 Housing structure

As with the race and ethnicity model, the housing structure latent variable only included three indicators, such that model fit could not be tested. While the standardized factor loadings of the housing structure latent variable were moderate ranging from .40 to .49, the reliability of this latent variable was extremely poor ($\alpha = .35$).

3.3 Combined models

Given that all but one of the single-factor models demonstrated poor fit to the data and/or low reliability, it was unlikely that a correlated factor model would be tenable. However, in keeping with our second aim, we tested a correlated five-factor model, wherein each of the hypothesized latent variables was estimated simultaneously. As expected, the model demonstrated very poor fit to the data (Table 2), indicated errors in estimation and produced correlations between factors exceeding 1. Attempting to estimate the correlated factor model using a combination of the modified single-factor models produced similar results. Given these empirical limitations, a correlated factor structure of SoVI was not supported.

Next, we tested a general single-factor model, wherein every indicator loaded onto one social vulnerability factor. Again, this model demonstrated poor fit (Table 2), although there were no difficulties in model estimation. Despite removing several indicators with low loadings and the addition of 23 residual covariances, model fit remained poor.

Finally, we explored a bifactor model, wherein each indicator belonged not only to its previously identified specific factor, but also to a general social vulnerability factor. Factors in bifactor models are required to remain uncorrelated. The bifactor model demonstrated poor fit to the data (Table 2) and resulted in an estimated factor loading greater than 1, suggesting an improper solution.

4 Discussion

This study applied confirmatory factor analysis to test the generalizability and construct validity of SoVI, a commonly used measure of social vulnerability and its component input variables. We demonstrated that the previously theorized pillars of social vulnerability, established with indicators drawn from the US Census data, were not well

supported in a latent variable framework. Only one of the five pillar models (Socio-economic Status) achieved acceptable fit, while two others had marginal fit even after significant model respecification that may not generalize to new samples. Our results underscore challenges with the construct validity of SoVI (Spielman et al. 2020), which remain an obstacle to achieving the goal of specifying a widely applicable vulnerability index. At the same time, we also face the reality that social vulnerability to hazards and disasters may be more context-specific than many other public health constructs, and that a generalizable indicator may be unnecessary for locally identifying vulnerable communities and facilitating effective disaster preparedness, response, and recovery efforts. Simply put, space matters, and social vulnerability manifest itself differently across study/impact areas. As many scholars have previously found, a one size fits all approach to social vulnerability may not be achievable (Cutter and Emrich 2006; Cutter and Finch 2008; Tapsell et al. 2010; Emrich and Cutter 2011; Tate 2012).

At both the pillar level and the overall SoVI construct level, confirmatory models did not fit the data and showed low reliability, indicating that the SoVI model applied to the New York region following Hurricane Sandy was not portable to a statewide vulnerability analysis in Florida. The establishment of well-fitting confirmatory factor models is the minimum requirement to establishing unidimensional, reliable indices of social vulnerability which can then be examined for generalizability. Results from our study indicated that this minimum requirement is not met for the pillar approach to SoVI put forth by Rufat and colleagues (2019), at least not with the proposed set of indicators. The consistently poor fit of our attempted models suggests that the theorized SoVI pillars may not adequately capture latent commonalities between groups of census variables that are thought to measure common vulnerabilities. While the socioeconomic status latent variable did achieve acceptable fit and reliability after considerable modification, this is likely a consequence of overfitting the model. When a model is excessively modified, it will fit the observed data very well; however, the likelihood that such a data-driven model will generalize to new samples is substantially lower than if the theorized model had fit the data well without modification (Babak 2004). This general finding aligns with criticism of the internal consistency and construct validity of SoVI (Spielman et al. 2020), further suggesting that all indicators of social vulnerability may not be applicable to every community. On the other hand, achieving an acceptable model does provide some degree of optimism that at least the socioeconomic dimension of SoVI may be generalizable across contexts.

While a number of factors likely explain why these models failed to fit the data, one reason stems from properties of the indicator variables, pointing to important considerations for future attempts to fit SoVI models. A critical step in these attempts is an examination of the distribution of each indicator. In our data, several variables were constants or near constants for Florida (e.g., percent working in the extractive sector; percent Native American). Variables with no or low variability do not contribute to covariation and detract from the fit of the model. Moreover, a variable with low to no variance cannot serve as an informative predictor of some outcome, such as hardships during disaster recovery. Finding differences in the variation of indicators across datasets, as is the case when our data are compared to the data on Hurricane Sandy presented by Rufat and colleagues (2019), suggests that indicators may be region and disaster specific, limiting the generalizability of social vulnerability models in general and SoVI in this case.

Given that generalizability is essential for prediction, it seems important to identify a minimum set of indicators that are likely useful across situations and therefore generalizable. To the extent, it is possible to identify a minimum number of indicators that are generalizable, these indicators could serve as anchor items (i.e., items used across all regions and

contexts in the assessment of vulnerability to disasters). For each use case, the generalizable anchor items could be combined with specific indicators that are more variable across certain disasters and regions. Multiple group latent variable models represent a promising methodology for pursuing such an approach, as they allow a combination of invariant and region/disaster specific indicators. In order to test the notion of a SoVI that includes a set of fixed generalizable indicators and a subset of specific ones that may vary across disasters or regions, researchers would need data across multiple disasters and regions, all contributing a comparable set of measures.

Many indicators depend on context for their valence in social vulnerability and may not be theorized to consistently increase or decrease social vulnerability. For example, a greater *percent Hispanic* may result in greater vulnerability in a less diverse region such as northern Florida, but may buffer households from vulnerability in more diverse places like South Florida, particularly in Florida's most populous county, Miami-Dade, where the majority of the population is Hispanic. Consequently, while Rufat and colleagues' (2019) SoVI pillars made theoretical sense for the New York metropolitan area, Florida's sociodemographic context may present a challenge to constructing them using this set of 27 indicators.

Other researchers have also raised the possibility that the indicators that are typically combined during social vulnerability analyses might be better used as independent measures, rather than combined in an index (Fekete 2019). Such a formative modeling approach would test common indicators as a proxy for social vulnerability, rather than attempt to directly approximate this latent construct. Either way, this approach requires a "best" set of indicators, which remains elusive throughout the vulnerability case study literature, and certainly no set of indicators has been validated across a region as large and diverse as the US. Given the limitations described earlier, it is possible that the social vulnerability of large, diverse regions cannot be reduced to a small set of generalizable indicators with any meaningful degree of precision that would shape public policy in the planning and response to disasters. It is more likely that a big data approach leveraging a very large set of indicators would improve precision but at greater cost. For local and regional planning purposes, perhaps the local scale is what matters most.

Variable selection may also be limited by the classic modifiable areal unit problem, i.e., the ecological fallacy that can inadvertently arise when grouping data into regions at different geographic scales (Openshaw 1984). When conceptualizing social vulnerability, the intent is to identify individuals who are at the greatest risk of hardship following a natural disaster. To operationalize this, a predictive model would require individual-level data, which is generally unavailable due to important ethical restrictions on the use of personal identifiable information. In homogenous regions, census tract-level data may approximate individual-level data. But in heterogeneous areas such as Florida, particularly urban areas with severe residential segregation over short distances, the census tract level of analysis may not accurately reflect the individual social vulnerability. Moreover, such levels of analyses may put decision makers at risk of ecological fallacies, in which conclusions based on group-level data are erroneously applied to individuals within that group. Consequently, indices relying on census tract data may provide valuable risk prediction in regions with homogenous and stable populations but may not be specific enough in more dynamic areas.

A related issue that may account for differences in indicator variability across contexts is that different relative scales are appropriate for different regions with different costs of living. A household income that suggests vulnerability in one area, such as New York City or Miami, may present little risk in nearby rural or suburban areas such as Central Florida. One approach is to normalize indicators that are in absolute scales in

order to standardize their meaning across regions. While past studies have used Z-scores as an attempt to address this issue (Emrich et al. 2020), Z-scores computed relative to the overall population still reflect regional heterogeneity. Future studies may consider the use of multilevel models, which would account for effects relative to the group mean (e.g., household income relative to others in Central Florida rather than the country overall). Hierarchical models can improve the accuracy of comparisons by controlling for regional homogeneity and could be further refined to adjust for spatial interaction effects between regions (e.g., Tate 2012; Dong et al. 2020). Hierarchical models also enable the simultaneous modeling of indicators and their interactions across multiple scales, an approach likely to bring new insights to social vulnerability analysis.

These issues of scale and level of analysis have led others to call for improvements to social vulnerability analysis that differentiate urban and rural places and that conceptually separating social vulnerability at national, community, and individual scales (Fekete 2019). We must acknowledge that Florida is, both empirically and experientially, an unusual place that does not uniformly conform to the theoretical manifestation of social vulnerability. For reasons already noted, Florida's most populated county, Miami-Dade, is so different from the rest of the state that it may bias our attempts to reconstruct the SoVI pillars used by Rufat and colleagues (2019). That said, the New York metropolitan area used by Rufat and colleagues is itself a diverse urban area, and if the pillars and indicators were robust enough, we should have experienced fewer modeling challenges.

Natural hazards require a coordinated response from federal to local scales, and scale matching in disaster response has long been a challenge for environmental policy (Baker and Refsgaard 2007). Ultimately the implementation of disaster response occurs locally, often implemented by federal agencies, but under direction from state and local officials who know their constituencies best. By using a context-driven set of indicators, vulnerability models could be tailored to pinpoint local risk and protective factors. Social vulnerability measures can tell us about differential recovery. SoVI and its input variables are highly utilized in emergency management (Tierney and Oliver-Smith 2012) and have proven useful in linking outcomes to underlying sociodemographic in several disaster specific case studies (Cutter and Finch 2008; Domingue and Emrich 2019; Emrich et al. 2020), demonstrating how baseline conditions are influential across the emergency management cycle. From this perspective, we can identify interventions and pathways toward hazard mitigation by setting aside the theoretically wobbly ideal of a single vulnerability model—one that can be broadly applied to different types of disasters at different geographic scales—and instead focusing on the links between variables and outcomes.

This study faced several important limitations. We tested the portability of a prior analysis of the New York metropolitan area to the state of Florida using a defined set of indicators and pillar design. When our modeling efforts failed, we did not explore other indicators or pillars in the absence of new theory. It is possible that a data-driven reconfiguration of the pillars might have improved our modeling approach, but such an approach would abandon longstanding theory about the drivers of social vulnerability. Further, Florida's peninsular shape, which creates problems for geospatial analysis, may also lead to challenges in underlying data representations, estimations, and variation. Finally, because SoVI and disasters are both place-specific in nature, perhaps a tract-level index computed at the state level should make little empirical sense given the local scales at which processes such as residential segregation operate. Perhaps SoVI's utility as a measure of outcomes would be much more appreciable at smaller scales of analysis such as disaster impact areas or informing city level programs and regional intervention.

Future research on the development and validation of vulnerability should consider the kind of evidence needed to validate an index across multiple contexts and how a balance of universal and context-specific indicators may be integrated. A social vulnerability index might be reconceptualized using different mixes of theoretical and data-driven approaches to deriving its underlying pillars, and with potentially different indicators and pillars for different scales of analysis, and for different types of disasters. Future social vulnerability measures will likely evolve along with changes in data availability and more robust conceptual and practical ties between vulnerability and outcomes. Ultimately new approaches must strike the balance between theoretical rigor and the reality that—despite the challenges outlined in this paper and others—applications of social vulnerability measures, including SoVI, continue to provide decision makers with empirical, pragmatic measures of a complex concept.

5 Conclusion

This study sought to confirm the construct validity of theoretically driven organizations of social vulnerability indices, with the goal of empirically testing the generalizability of SoVI in diverse contexts. The lack of alignment between previous formulations of SoVI and census tract-level data for Florida, as demonstrated in this study, does not support the generalizability of SoVI to other contexts—at least in SoVI's current form. We believe several limitations in the data contributed to the poor psychometric evidence produced from this study. The poor support for the theorized pillars of SoVI and alternative model structures suggest the need to reconsider the input variables and scale of analysis, if the goal is to measure a generalizable construct (or subconstructs) of social vulnerability. This study highlights the challenges inherent in measuring social vulnerability across different geographic settings and types of natural disasters and the need for more robust assessment of confirmatory models.

Author contributions All authors made substantial contributions.

Funding Zachary T. Goodman is supported by the National Institutes of Health (T32-HL007426).

Availability of data and material All data used in this study are publicly available, and analytical code is available upon request.

Compliance with ethical standards

Conflicts of interest The authors have no conflicts of interest to report.

References

- Babyak MA (2004) What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. *Psychosom Med* 66:411–421
- Baker D, Refsgaard K (2007) Institutional development and scale matching in disaster response management. *Ecol Econ* 63:331–343
- Beccari B (2016) A comparative analysis of disaster risk, vulnerability and resilience composite indicators. *PLoS Curr*. <https://doi.org/10.1371/currents.dis.453df025e34b682e9737f95070f9b970>

- Clark GE, Moser SC, Ratlick SJ et al (1998) Assessing the vulnerability of coastal communities to extreme storms: the case of revere, MA., USA. *Mitig Adapt Strat Glob Change* 3:59–82. <https://doi.org/10.1023/A:1009609710795>
- Climate Central. (2014). Surging seas.
- Cutter SL, Boruff BJ, Shirley WL (2003) Social vulnerability to environmental hazards*. *Soc Sci Q* 84:242–261. <https://doi.org/10.1111/1540-6237.8402002>
- Cutter SL, Emrich CT (2006) Moral hazard, social catastrophe: the changing face of vulnerability along the hurricane coasts. *Ann Am Acad Political Soc Sci* 604:102–112
- Cutter SL, Finch C (2008) Temporal and spatial changes in social vulnerability to natural hazards. *Proc Natl Acad Sci* 105:2301–2306
- Cutter SL, Mitchell JT, Scott MS (2000) Revealing the vulnerability of people and places: a case study of Georgetown County, South Carolina. *Ann Assoc Am Geogr* 90:713–737. <https://doi.org/10.1111/0004-5608.00219>
- Domingue SJ, Emrich CT (2019) Social vulnerability and procedural equity: exploring the distribution of disaster aid across counties in the United States. *Am Rev Public Adm* 49:897–913
- Dong G, Ma J, Lee D et al (2020) Developing a locally adaptive spatial multilevel logistic model to analyze ecological effects on health using individual census records. *Ann Am As Geogr* 110:739–757
- Ellis M, Wright R, Holloway S, Fiorio L (2018) Remaking white residential segregation: metropolitan diversity and neighborhood change in the United States. *Urban Geogr* 39:519–545
- Emrich CT, Cutter SL (2011) Social vulnerability to climate-sensitive hazards in the Southern United States. *Wea Climate Soc* 3:193–208. <https://doi.org/10.1175/2011WCAS1092.1>
- Emrich CT, Tate E, Larson SE, Zhou Y (2020) Measuring social equity in flood recovery funding. *Environ Hazards* 19:228–250
- Fekete A (2019) Social vulnerability (re-)assessment in context to natural hazards: review of the usefulness of the spatial indicator approach and investigations of validation demands. *Int J Disaster Risk Sci* 10:220–232. <https://doi.org/10.1007/s13753-019-0213-1>
- Kline RB (2016) Principles and practices of structural equation modeling, 4th edn. American Psychological Association, Washington, DC
- Laska S, Morrow BH (2006) Social vulnerabilities and Hurricane Katrina: an unnatural disaster in New Orleans. *Marine Technology Society Journal* 40:16–26. <https://doi.org/10.4031/002533206787353123>
- Llabre MM, Fitzpatrick SL (2012) Revisiting measurement models in psychosomatic medicine research: a latent variable approach. *Psychosom Med* 74:169–177. <https://doi.org/10.1097/PSY.0b013e3182433a30>
- National oceanic and atmospheric administration, office for coastal management. “Digital Coast.”. Charleston, SC: NOAA Office for Coastal Management. Accessed June 2020 at <https://coast.noaa.gov/digitalcoast/>
- National Hurricane Center (2018) 2018 Atlantic Hurricane Season. Accessed June 2020 at <https://www.nhc.noaa.gov/data/>
- New York Times (2020) Will the hardest-hit communities get the coronavirus vaccine? <https://www.nytimes.com/2020/10/30/health/covid-vaccine-racism.html>. Accessed 10 Jan 2021
- NPR (2020) California may consider ‘historical injustice’ when allocating COVID-19 vaccine. Accessed January 2021 at <https://www.npr.org/sections/health-shots/2020/12/16/942452184/california-may-consider-historical-injustice-when-allocating-coronavirus-vaccine>
- Openshaw S (1984) Ecological fallacies and the analysis of areal census data. *Environ Plan A* 16:17–31. <https://doi.org/10.1068/a160017>
- Rosseel Y (2012) lavaan: an R package for structural equation modeling. *J Stat Softw* 48:1–36
- Rufat S, Tate E, Emrich CT, Antolini F (2019) How valid are social vulnerability models? *Ann Am As Geograph* 109:1131–1153
- Schmidt H, Weintraub R, Williams MA et al (2020) Equitable allocation of Covid-19 vaccines: an analysis of the initial allocation plans of CDC’s jurisdictions with implications for disparate impact monitoring. *SSRN Elect J*. <https://doi.org/10.2139/ssrn.3740041>
- Spielman SE, Folch D, Nagle N (2014) Patterns and causes of uncertainty in the American community survey. *Appl Geogr* 46:147–157. <https://doi.org/10.1016/j.apgeog.2013.11.002>
- Spielman SE, Tuccillo J, Folch DC et al (2020) Evaluating social vulnerability indicators: criteria and their application to the social vulnerability index. *Nat Hazards* 100:417–436
- Tapscell S, McCarthy S, Faulkner H, Alexander M (2010) Social vulnerability to natural hazards. State of the art report from CapHaz-Net’s WP4 London
- Tapscell SM, Penning-Rowsell EC, Tunstall SM, Wilson TL (2002) Vulnerability to flooding: health and social dimensions. *Philos Transact Math Phys Eng Sci* 360:1511–1525

- Tate E (2012) Social vulnerability indices: a comparative assessment using uncertainty and sensitivity analysis. *Nat Hazards* 63:325–347. <https://doi.org/10.1007/s11069-012-0152-2>
- Tierney K, Oliver-Smith A (2012) Social dimensions of disaster recover. *Int J Mass Emerg Disasters* 30:123–146
- Turner BL, Kasperson RE, Matson PA et al (2003) A framework for vulnerability analysis in sustainability science. *Proc Natl Acad Sci U S A* 100:8074–8079. <https://doi.org/10.1073/pnas.1231335100>
- Widaman KF (2018) On common factor and principal component representations of data: implications for theory and for confirmatory replications. *Struct Equ Model Multidiscip J* 25:829–847. <https://doi.org/10.1080/10705511.2018.1478730>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.