An Evaluation of Social Vulnerability and Community Resilience Indices and Opportunities for Improvement through Community Resilience Estimates

Katherine Ann Willyard, Ph.D.
Gabriel Amaro, Ph.D.
R. Chase Sawyer
Bethany DeSalvo, Ph.D.
Wesley Basel
Social Economic and Housing Statistics Branch
U.S. Census Bureau

SEHSD Working Paper 2022-25
December 15, 2022

Abstract

Purpose:

- This paper describes how the U.S. Census Bureau's Community Resilience Estimates (CRE) program provides an enhanced method of identifying communities most vulnerable and most resilient to a disaster.

Methods:

- Through small area modeling and using auxiliary data sources, the CRE program enhances survey estimates, and reduces margins of error, especially for small geographic areas. CRE are model-based enhancements of American Community Survey (ACS) estimates, created by integrating additional information from the U.S. Census Bureau's Population Estimates Program (PEP). CRE methodology employs statistical modeling techniques to combine supplemental information with survey data to produce estimates that are more reliable. CRE are broadly consistent with ACS direct survey estimates, but with help from other data sources, CRE are more precise than ACS direct survey estimates alone.

Main Points:

- CRE is more precise and timelier than existing measures of social vulnerability and community resilience.
- Because high point estimates are often related to high sampling error, areas described as high-risk using existing measures of social vulnerability and community resilience have higher sampling error than areas not considered high-risk.
- CRE provides a stable measure of social vulnerability and community resilience for planning and to distribute community resources.
- Because it uses microdata, CRE is the only measure to provide both estimates of social vulnerability along with measures of reliability, which are necessary to statistically determine if there is a significant difference between two areas or points of time.

Recommendations:

- Use CRE to make geographic comparisons in community resilience and social vulnerability.
- Define vulnerable communities as tracts or counties with a portion of the population with 3 or more vulnerability indicators higher than the national average.
- Define resilient communities as tracts or counties with a portion of the population with 3 or more vulnerability indicators lower than the national average.

An Evaluation of Social Vulnerability and Community Resilience Indices and Opportunities for Improvement through Community Resilience Estimates

1 | INTRODUCTION

By helping communities better anticipate, respond, resist, and recover from disasters, social vulnerability mapping strengthens community resilience and reduces inequalities.[1] Social vulnerability is the risk of hazards to the physical and socially built environment, while community resilience is the capacity of individuals and households to absorb the stresses from a disaster.[2] To eliminate the need to classify characteristics of an area as contributing to either vulnerability or resilience, resilience and vulnerability are viewed to represent two sides of the same resilience coin.[3] In other words, decreasing the vulnerability of communities to a disaster and making communities more resilient to a disaster is very similar. So, for our purpose of producing simple information that can be easily used by decision makers, resilience and vulnerability are represented as two opposing possibilities.

This report explains how the Census Bureau's new Community Resilience Estimates (CRE)[4] provide an enhanced method of identifying communities most vulnerable to a disaster. CRE are the only source of data for single year estimates of community resilience and social vulnerability for all tracts and counties in the United States except U.S. territories. American Community Survey (ACS) provides detailed 1-year estimates on various characteristics associated with social vulnerability and community resilience for some areas with populations of 65,000 or more.[5] As a data enhancement to ACS, CRE model-based estimates are a vital source of information for *annually* measuring community resilience across *all* counties and tracts in the U.S., except territories. Due to a lack of sufficient quality data, estimates for Puerto Rico and other U.S. territories are not produced.

2 | OVERVIEW

The ACS is the largest U.S. household survey, sampling about 3.5 million addresses per year. Because of the comprehensiveness of ACS, many measures of social vulnerability and community resilience rely on publicly available ACS 5-year direct survey estimates and area-level aggregations of indicators to identify vulnerable or resilient communities. While the ACS is

---

[1] Source: <https://www.tandfonline.com/doi/full/10.1080/10511482.2011.624528>
[2] Source: <https://link.springer.com/book/10.5822/978-1-61091-586-1>
[3] Source: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2016GH000047>
[4]The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product. Data Management System (DMS) number: P-75-17412, Disclosure Review Board (DRB) Approval Number: CBDRB-FY20-305.
[5] In 2019, 825 counties had detailed 1-year estimates of common community resilience or social vulnerability indicators, such as poverty; 1-year estimates are not released for tracts or block groups.

a great source of information for the development of social vulnerability and community resilience indices, there are three aspects where we can improve upon existing measures: using microdata (i.e., person-level data) to build the index, producing more timely estimates, and incorporating sampling error.

2.1 | The Importance of Microdata

Individual indicators of social vulnerability (e.g., poverty, unemployment, and disability status) and community resilience (e.g., insured, steady employment, and the establishment of community bonds) are phenomena experienced by people. However, because they are dependent on publicly available data, social vulnerability and resilience indices are not capable of examining data at the person-level and must choose the geographic level, such as counties, tracts, or block groups, to aggregate indicators.

Figure 1 displays the standard hierarchy of geographic entities and the different levels at which social vulnerability and community resilience indicators are aggregated. Without access to restricted person-level data, social vulnerability and community resilience indices built using publicly released ACS data introduce bias through choices in geographic aggregation, leading to areas being misidentified as vulnerable or misidentified as not vulnerable when they should be. For example, when creating an index by aggregating indicators at the area-level, a neighborhood may be flagged as vulnerable because it has a high portion of residents above 65 and a lower median household income. However, this does not necessarily mean the neighborhood contains primarily vulnerable older low-income people. The neighborhood could contain many elderly high-income people with no other vulnerability indicators (who can use financial resources to respond in a disaster), and similarly low-income young people with no other vulnerability indicators (who can use their physical capabilities to respond to a disaster more quickly). An area-level aggregation using publicly available data is limited because it cannot capture how person-level indicators can interact to affect social vulnerability and community resilience.

2.2 | The Importance of Timeliness

Most nationally representative social vulnerability and community resilience indices are constructed using publicly available ACS data. Due to quality standards and confidentiality requirements, ACS does not publicly release 1-year direct survey estimates for all counties, or any tracts or block groups, but it does release 5-year direct survey estimates for these smaller entities. This is because, to reach a sample size large enough to meet Census quality standards and confidentiality requirements, data must be aggregated across multiple years.

However, the use of ACS 5-year data by social vulnerability and community resilience indices creates a limitation in timeliness relative to ACS 1-year data. While ACS 1-year direct survey estimates use data from interviews collected over the course of a single calendar year, ACS 5-year direct survey estimates use data from interviews over the course of five calendar years. So, an interview from January 2016 is included in ACS 5-year data releases for five years, from the 2012-2016 data release thru the 2016-2020 data release.

In comparison to ACS 1-year estimates, ACS 5-year estimates are less timely. The lack of relative timeliness has minimal impact on areas experiencing little social and economic change. However, multiyear estimates can lag in areas experiencing major changes, and the population experiencing social vulnerability indicators can rapidly change. One example of this is after the implementation of the Affordable Care Act, there were rapid changes in the number of people in the United States with health insurance coverage. It took ACS 5-year estimates five years to account for the change. In contrast, programs like the Small Area Health Insurance Estimates (SAHIE) were able to quickly reflect changes for all counties in the nation because of its use of 1-year ACS estimates and small area modeling methodologies.[6]

In addition, Census does not recommend comparing ACS 5-year estimates when those estimates contain overlapping coverage because much of the data underlying the estimate are the same[7]. For example, 2010-2014 ACS 5-year estimates can be compared with 2015-2019 ACS 5-year estimates because they do not share underlying data. However, you cannot compare the 2010-2014 ACS 5-year estimates with 2014-2018 ACS 5-year estimates since the underlying data overlaps with the year 2014. So, when a social vulnerability or community resilience index draws from ACS 5-year data, it takes years before changes in social vulnerability or community resilience can be identified.

2.3 | The Importance of Sampling Error

Published ACS estimates meet Census' high quality statistical standards which include the publication of sampling error. Sampling error is the uncertainty that comes from the fact that a survey is based on a sample, rather than all housing units or individuals. The amount of error is directly related to the size of the sample, as well as the variability. Of course, the larger the sample, the less error but we also need to ensure that the sample has accurately captured the population characteristics. These include characteristics such as age, sex, race, and ethnicity to name a few. Are all these groups accurately represented in the sample? If not, that variation will impact the size of error in the sample.

ACS goes through substantial efforts to identify, reduce, and measure error. While the creators of many social vulnerability and community resilience indices publish the margins of error developed by the Census Bureau along with their respective ACS estimates, margins of error are often not included in the development of the social vulnerability or community resilience index itself.[8][9][10] Without the production of margins of error along with estimates, a statistically significant difference between places or across time cannot be found. Since other methods of quantifying social vulnerability and community resilience do not produce margins of

[6] Source: <https://www.census.gov/content/dam/Census/library/working-papers/2016/demo/SEHSD-WP2016-16.pdf>
[7] Source: < https://www.census.gov/newsroom/blogs/random-samplings/2022/03/period-estimates-american-community-survey.html#:~:text=How%20should%20users%20compare%205,any%20overlapping%20years%20of%20data>
[8] Source: <https://svi.cdc.gov/Documents/Publications/CDC_ATSDR_SVI_Materials/SVI_30April2013.pdf>
[9] Source: < https://www.epa.gov/sites/default/files/2021-04/documents/ejscreen_technical_document.pdf>
[10] Source: <http://artsandsciences.sc.edu/geog/hvri/sovi%C2%AE-error-discussion>

error, they cannot be used to determine if there is a statistically significant difference between two areas or points of time.

Common methods of quantifying social vulnerability and community resilience without consideration of margins of error, combined with the distribution of ACS sampling error, can lead to problems in identifying high-risk communities.

Figure 2 shows the relationship between 2015-2019 ACS 5-year estimates of the percentage of the population whose income in the past 12 months is below the poverty level and margins of error for tracts.

When combined with index creation methods that use only the estimates in a percentile rank, the visible clustering of error at the extremes can be problematic. For example, with SVI methods of tagging the top ten percent of estimates as vulnerable, a tract with a poverty estimate of 100 percent and a 100 percent margin of error will be tagged as vulnerable. Although this tract, with its high margin of error, could have 0 percent of the population living below the poverty level, it could knock out a tract very near the top ten percent threshold that has much a lower margin of error and thus less uncertainty.

Figure 3 shows the distribution of sampling error among the top ten percent of 2015-2019 ACS 5-year estimates of poverty for tracts (i.e., SVI vulnerable tracts) as well as the distribution of the estimates for the remaining tracts. T-test results comparing the sampling error of vulnerable and not vulnerable poverty indicator tracts with a 90 percent confidence interval shows vulnerable tracts have a higher sampling error than tracts that are not vulnerable (See **Appendix A**). Because poverty rate estimates have a positive, high degree of correlation with survey error (See **Appendix B**), it could be sampling error, not actual vulnerability, is reflected in the point estimate.

To correctly interpret the estimate, data users need to incorporate the margin of error which would show that the estimate has a large range for which the actual population statistic exists. Since indices that rely on publicly available data don't utilize the margins of error, they don't reflect these intricacies in their analysis.

3 | IMPROVING COMMUNITY RESILIENCE ESTIMATES USING SMALL AREA MODELING TECHNIQUES

Due to increased demand for more timely and precise information about small populations, Census has established small area methods for estimating key social, economic, and housing statistics. For example, the Small Area Income and Poverty Estimates (SAIPE) program produces annual child poverty estimates for school districts across the United States, which is used for Title I allocations. Additionally, the Centers for Disease Control and Prevention's (CDC) National Breast and Cervical Cancer Early Detection Program uses Census' SAHIE Program to allocate funds.

Building upon established Census small area estimation methods, the 2019 CRE was released in July 2021. CRE are modeled estimates of social vulnerability in the population, based on the number of vulnerability indicators individuals within the population have. Along with high-risk population estimates, CRE produces margins of error for the population estimates at the 90 percent confidence interval. Building upon established Census small area estimation methods, the CRE incorporates the importance of microdata, the importance of timeliness and the importance of sampling error.

CRE overcomes issues with aggregating aggregates by working with the microdata. By using restricted microdata, CRE can account for the interaction of risks experienced by individuals and households, as well as the associated sampling error.

CRE overcomes issues with timeliness by drawing only from ACS 1-year microdata. CRE methodology combines the 1-year ACS estimates with other data sources to provide more timely, precise, and stable estimates than any index or estimate that uses publicly available data. For this reason, communities across the U.S. that experience the most change are captured in the CRE and not hidden among an aggregation of other years of data. In addition, if CRE replicates methods to produce estimates for two years (e.g., 2019 and 2020), it is possible to compare them and determine if significant changes occur.

While other indices do not incorporate ACS sampling error in their methods of quantifying social vulnerability and community resilience, CRE overcomes this issue by working with microdata and incorporating administrative records required to produce estimates of high statistical quality. To allow policy makers to make decisions that increase the resilience of underserved communities across the United States, reliable measures of social vulnerability and community resilience must be developed. Because it uses microdata, CRE is the only measure to provide both estimates of social vulnerability along with measures of reliability. Measures of reliability accompanying estimates are necessary to statistically determine if there is a significant difference between two areas or points of time. Without an accompanied measure of reliability, statistical comparisons cannot be made.

CRE creates more precise estimates than direct survey estimates alone. T-test results comparing the relative error of 2019 ACS direct estimates of tract high-risk populations to the relative error of 2019 CRE high-risk populations with a 90 percent confidence interval shows that small area methods significantly reduce the relative error of high-risk populations (See **Appendix C**). **Figure 4** describes the amount that the relative error of 2019 ACS direct estimates of the high-risk population are reduced through the small area modeling techniques employed to create the 2019 CRE. In comparison to 2019 ACS direct estimates, on average, small area modeling reduces the coefficient of variation of high-risk population estimates by 26 percent. In all cases, the error is reduced through the small area modeling techniques employed to create the 2019 CRE. In 279 cases, the relative error is cut at least in half. In 40 cases, the relative error is reduced by over 70 percent.

Figure 5 shows the relationship between CRE high-risk population estimates and margins of error for Census tracts in the United States. In addition to reducing relative error, **Figure 5** does not show the margin of errors clustered at the extremes, as we saw in **Figure 2**.

In short, in comparison to other methods, CRE's method to identify socially vulnerable populations is often more reliable, precise, and timely. Because CRE is also more stable, it provides decision makers with an effective tool to better plan and respond to disasters.

3.1 | CRE Methods

CRE is created by first cumulatively tagging vulnerability indicators[11] to individuals within ACS microdata. Individuals within the ACS microdata are then categorized as low-risk (0 vulnerability indicators), medium-risk (1-2 vulnerability indicators), or high-risk (3 or more vulnerability indicators). Next, using traditional direct survey methods, tabulations for states, counties, and tracts for the number of people at low-, medium-, and high-risk are estimated. These traditional direct survey estimates are then used to inform the small area model.

Then, to create the small area estimates, CRE fits an empirically optimal shrinkage model, which is made up of a combination of regression estimation techniques and shrinkage techniques. Traditional direct survey estimates are used as the dependent variable of the regression model which inform estimates. Using Census' Population Estimates Program (PEP) postcensal population estimates as independent variables, a regression "prediction" is obtained. These regression-based predictions are then combined with direct sample estimates, with each of the two parts receiving a weight and each of the two weights adding up to one.

The weight of a model prediction component is the ratio of the sampling variance of the direct estimate to the total variance of the direct estimate. So, when direct survey methods are more precise, the direct survey estimate receives a greater weight; when direct survey methods are less precise, the modeled estimate receives a greater weight. Using this strategy, CRE produces nationally representative estimates of social vulnerability and community resilience to hazards with smaller standard errors than direct survey estimates alone.

3.2 | CRE is Adaptable to a Wide Variety of Indicators

While there are limitations to the construction of social vulnerability and community resilience indices relying on publicly available ACS data, CRE provides a drastic advancement. CRE is adaptable and can be easily modified for a broad range of natural disasters, such as hurricanes, tornadoes, and floods. For example, while the 2018 experimental CRE released in June 2020 focused on the COVID-19 Pandemic, the 2019 CRE was adapted to be applicable to a wider range of disasters. Based on user feedback, five key changes were made between the 2018 experimental CRE and the 2019 CRE: (1) The health condition indicators in the 2018 experimental CRE were dropped from the 2019 CRE. (2) While 2018 experimental CRE included household-level and tract-level crowding, the 2019 CRE crowding indicator only

[11] Vulnerability indicators from the 2019 ACS include: Income to Poverty Ratio; Single or Zero Caregiver Household; Crowding; Communication Barrier; Households without Full-time, Year-round Employment; Disability; No Health Insurance; Age 65+; No Vehicle Access; No Broadband Internet Access.

involved household-level crowding. (3) While the 2018 experimental CRE included an unemployment indicator (i.e., in households with at least one person under the age of 65, all individuals did not have current employment during the time of the survey), the 2019 CRE included an indicator for no stable employment (i.e., in households with at least one person under the age of 65, no individual was employed full-time, year-round). (4) An indicator for households with no vehicle access was added to the 2019 CRE. (5) An indicator for households with no broadband internet access was added to the 2019 CRE.

4 | IMPLICATIONS AND FUTURE RESEARCH STRATEGIES

CRE provides an adaptable method to quantify vulnerable and resilient populations. By working with microdata and incorporating administrative records, CRE produces estimates of high statistical quality. In comparison to CRE, other existing measures of social vulnerability and community resilience are less timely and less precise. Because high point estimates are often related to high sampling error, areas described as vulnerable using existing measures of social vulnerability and community resilience have higher sampling error than areas not considered vulnerable. In comparison to CRE, if existing measures of social vulnerability and community resilience were to be used to distribute resources, communities would have more difficulty planning because the estimates are less reliable.

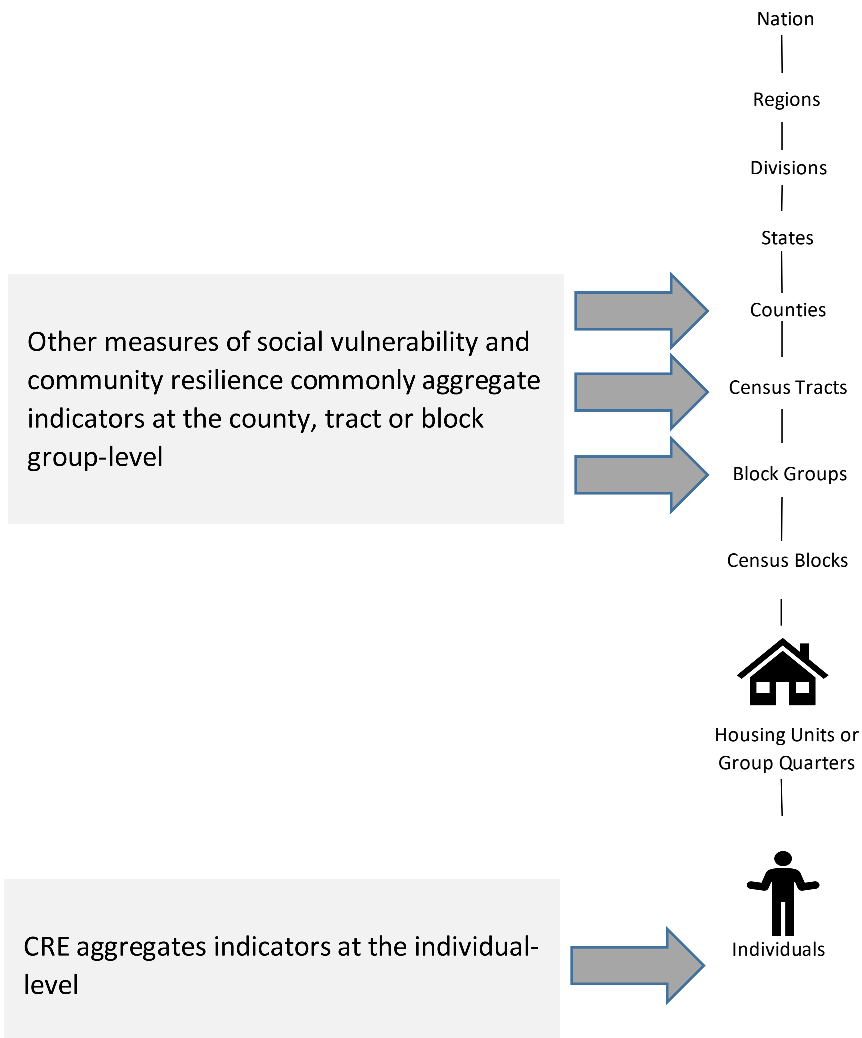
Because it uses microdata, CRE is the only measure to provide both estimates of social vulnerability and resilience, along with measures of reliability, which are necessary to statistically determine if there is a significant difference between two areas or points of time. This allows researchers to quantify the portion or number or high-risk residents in different places and make statistical comparisons. Ongoing research at Census is comparing the portion of high-risk residents in different places, like between different Census regions and divisions, between historically disenfranchised and not historically disenfranchised communities, and between toxic communities and not toxic communities. Without CRE, which provides a population estimate accompanied with a measure of reliability, statistical comparisons cannot be made. To determine if decision makers are meeting goals of increasing the resilience of underserved communities across the United States, measures, like the CRE, must allow for statistical comparisons between estimates.

Suggested Citation

K. Willyard, Amaro G., Sawyer R. C., DeSalvo B., Basel W., “An Evaluation of Social Vulnerability and Community Resilience Indices and Opportunities for Improvement through Community Resilience Estimates,” *SEHSD Working Paper Series*, 2022-25, U.S. Census Bureau, Washington, D.C., 2022.

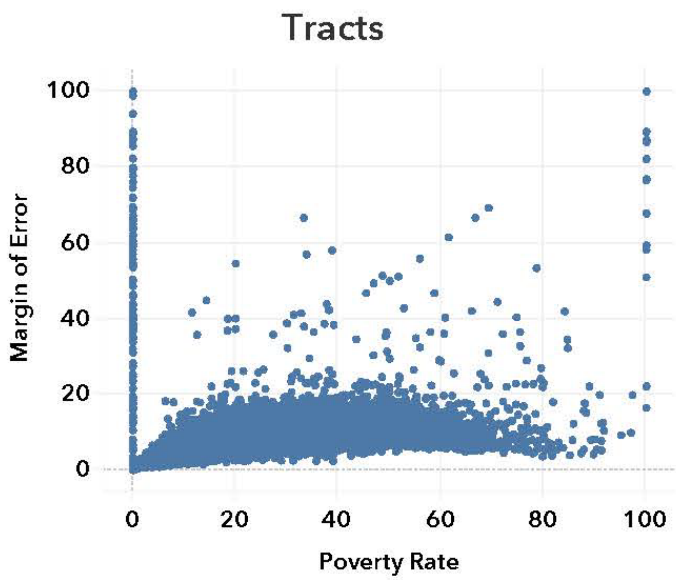
FIGURES

Figure One: Social Vulnerability and Community Resilience Indicator Aggregation Levels



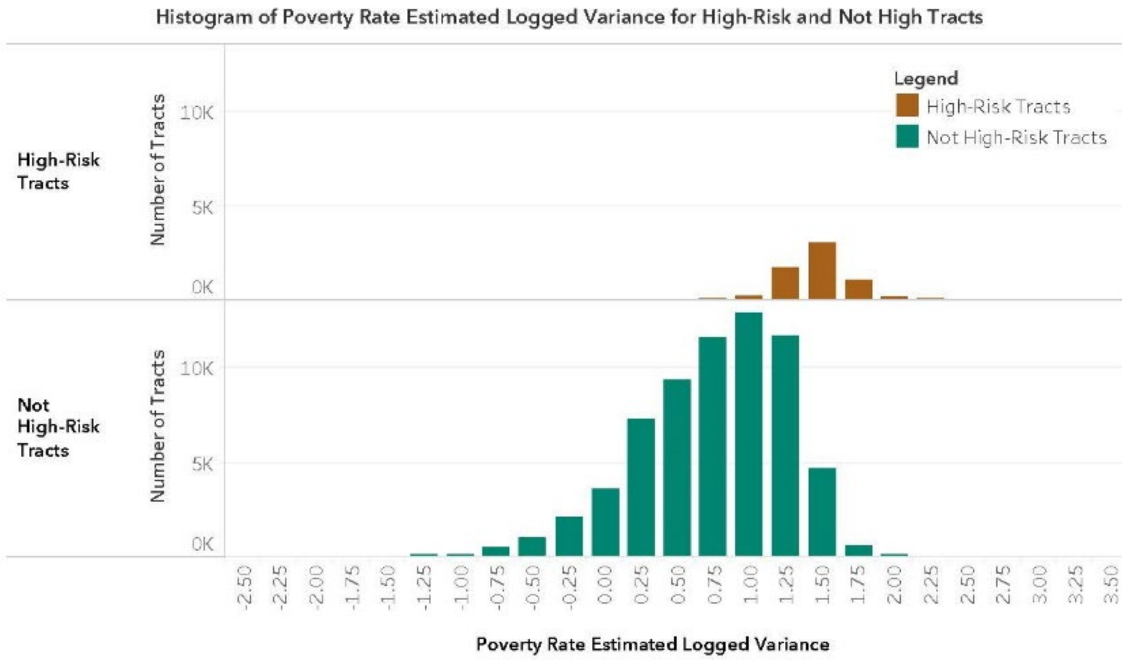
Source: U.S. Census Bureau, Geography Division

Figure Two: 2015-2019 ACS 5-Year Poverty ate Estimates and Margins of Error



Source: 2015-2019 American Community Survey 5-Year Estimates, Subject Table S0601

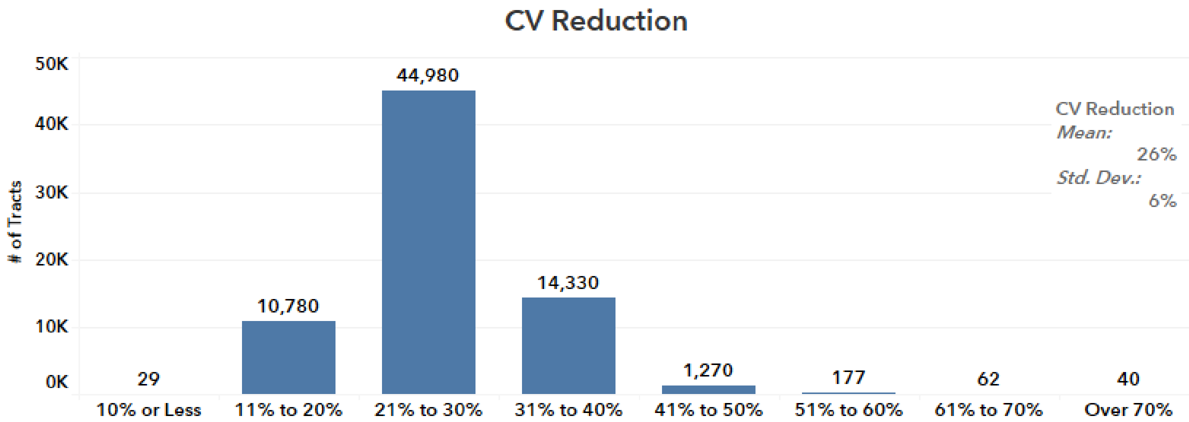
Figure Three: Distribution of 2015-2019 ACS 5-Year Poverty Sampling Error for Vulnerable and Not Vulnerable* Tracts



* Vulnerable poverty indicator communities are defined using SVI's percentile ranking method, which flags the top ten percent of estimates

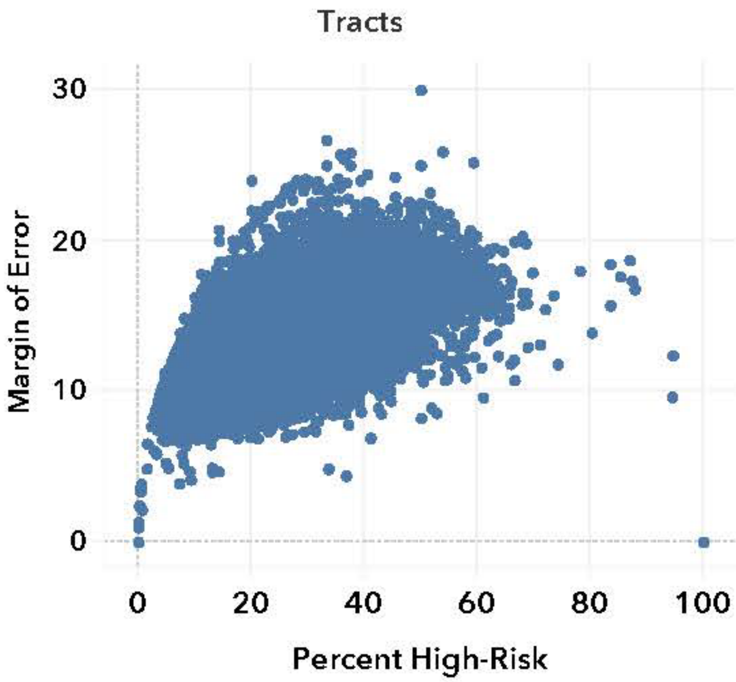
Source: 2015-2019 American Community Survey 5-Year Estimates, Subject Table S0601

Figure Four: Description of Percent Reduction in the Relative Error of High-Risk* Population Estimates for Populated Census Tracts



* Individuals with 3 or more vulnerability indicators are high-risk.
Source: 2019 American Community Survey 1-Year Estimates, 2019 Community Resilience Estimates

Figure Five: CRE High-Risk* Population Estimates and Margins of Error



* Individuals with 3 or more vulnerability indicators are high-risk.
Source: 2019 Community Resilience Estimates

Appendix A

Two-Sample T-Test Comparing Poverty Rate Margin of Error for Vulnerable and Not Vulnerable* Poverty Indicator Tracts

TTEST Procedure – Variable: Poverty Rate Margin of Error

Vulnerable	**Method**	**Mean**	**90% Confidence Level Mean**		**Std. Dev.**	**90% Confidence Level Std. Dev.**	
0		5.3154	5.2913	5.3395	3.7309	3.714	3.748
1		10.9226	10.8169	11.0283	5.4698	5.396	5.5457
Diff (1-2)	Pooled	-5.6072	-5.6875	-5.5269	3.9399	3.923	3.9571
Diff (1-2)	Satterthwaite	-5.6072	-5.7156	-5.4988			

	Method	**Variances**	**D.F.**	**T Value**	**Pr > |t|**
	Pooled	Equal	72,261	-114.89	<.0001
	Satterthwaite	Unequal	8,009.70	-85.07	<.0001

Equality of Variances

	Method	**Num D.F.**	**Den. D.F.**	**F Value**	**Pr > F**
	Folded F	7,242	65,019	2.15	<.0001

* Vulnerable poverty indicator communities are defined using SVI's percentile raking method, which flags the top ten percent of percent in poverty estimates

Source: 2015-2019 American Community Survey 5-Year Estimates, Subject Table S0601

Appendix B

Correlation Between Tract Poverty Rate Estimates and Margins of Error

CORR Procedure – Simple statistics for Poverty Rate Estimates and Margin of Error

Variable	**N**	**Mean**	**Std. Dev.**	**Sum**	**Min.**	**Max.**
Poverty Rate Margin of Error	72,263	5.8774	4.2847	424,719	0.1000	100.0000
Poverty Rate Estimate	72,263	14.6479	11.5919	1,058,502	0.0000	100.0000

Pearson Correlation Coefficients, N = 72,263

	Poverty Rate Estimate
Poverty Rate Margin of Error	0.6038
Prob > |r| under H0: Rho=0	<.0001

Appendix C

One-Sample T-Test of the Reduction in Relative Error for High-Risk* Population Estimates through Small Area Modeling

TTEST Procedure – Variable: Percent Reduction in Relative Error
(N~71,670 Tracts)

		90% Confidence Level	
Mean	25.92%	25.91%	25.99%
Standard Deviation	6.20%	6.17%	6.23%

D.F.	**T Value**	**Pr > |t|**
7,242	65,020	2.15

* Individuals with 3 or more vulnerability indicators are high-risk.
Source: 2019 American Community Survey 1-Year Estimates, 2019 Community Resilience Estimates