Developing a Data-Driven System for Identifying Vulnerable and Resilient Neighborhoods across the United States: Opportunities for Improvement through Small Area Estimation

Katherine Ann (Kate) Willyard, Ph.D.
(Katherine.a.Willyard@census.gov)
Gabriel (Gabe) Amaro, Ph.D.
(Gabriel.amaro@census.gov)
Social, Economic, and Housing Statistics Division
U.S. Census Bureau

2022 Federal Committee on Statistical Methodology Research & Policy Conference
Tuesday, October 25, 3:15 PM, Walter Washington Convention Center, Room 146C
Session C-3 Advances in Small Area Estimation

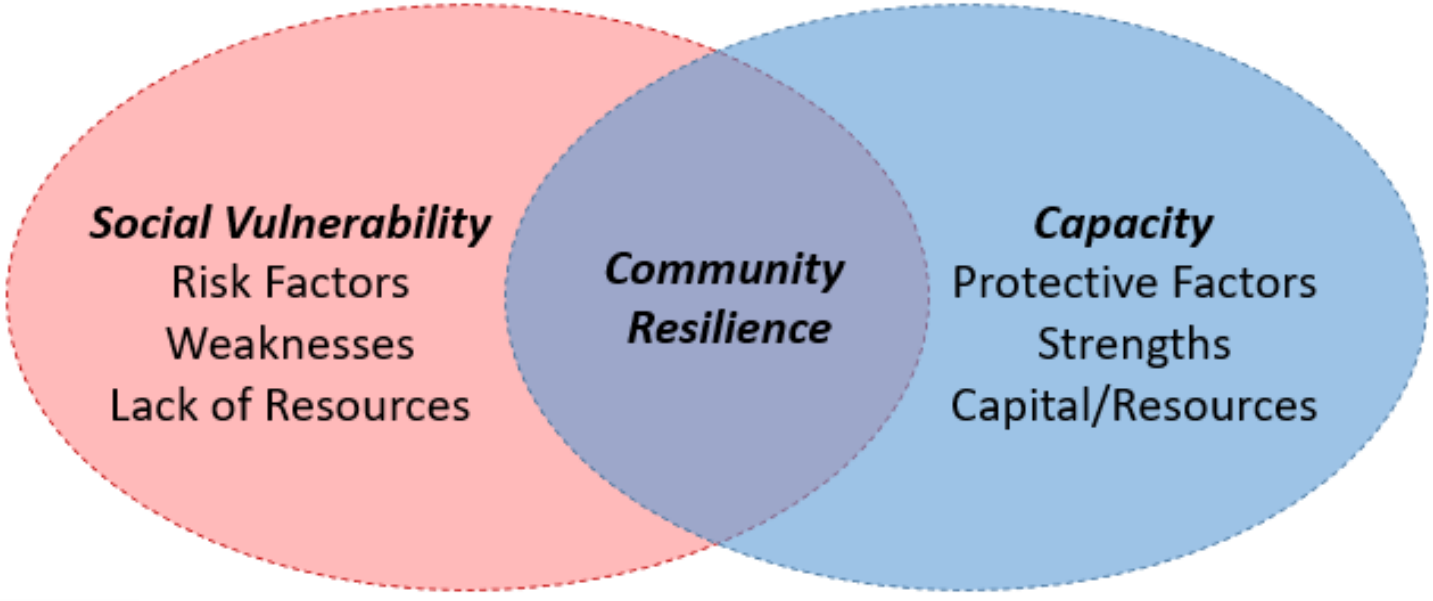
"The Commerce Department is prepared to leverage all of its bureaus, and our dedicated workforce, to ensure this administration and communities across the nation have the data, tools, and resources they need to mitigate the impacts of climate change while building a better, more resilient, America."

-U.S. Secretary of Commerce Gina M. Raimondo on Friday, April 22, 2022, upon the announcement of DAO 216-22 "Addressing the Climate Crisis"

https://www.commerce.gov/news/press-releases/2022/04/secretary-raimondo-establishes-commerce-climate-council-directs

An Introduction to Community Resilience and Social Vulnerability Mapping

- Social vulnerability mapping strengthens community resilience and reduces inequities[1]
  - By helping communities better anticipate, respond, resist, and recover from disasters.
- 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]



1 Van Zandt, Peacock, Henry, Grover, Highfield and Brody 2012; 2 Masterson, Jamie Hicks, Walter Gillis Peacock, Shannon Zandt, Himanshu Grover, Lori Field Schwarz, and John Cooper 2014; 3 Summers, Smith, Harwell, and Buck 2017

The Need for U.S. Census Bureau's Community Resilience Estimates (CRE)

- Prior to 2020, national social vulnerability and community resilience measures were created using methods that do not adequately consider the reliability of the public American Community Survey (ACS) data used as a source

- Other methods do not allow policy makers to use the information to determine if there is a statistically significant difference between two areas or points of time

- It is critical to create measures that allow for statistical comparisons between estimates because they are necessary to determine if policymakers are meeting goals to increase the resilience of underserved communities across the United States

COMMUNITY RESILIENCE ESTIMATES

Research Overview

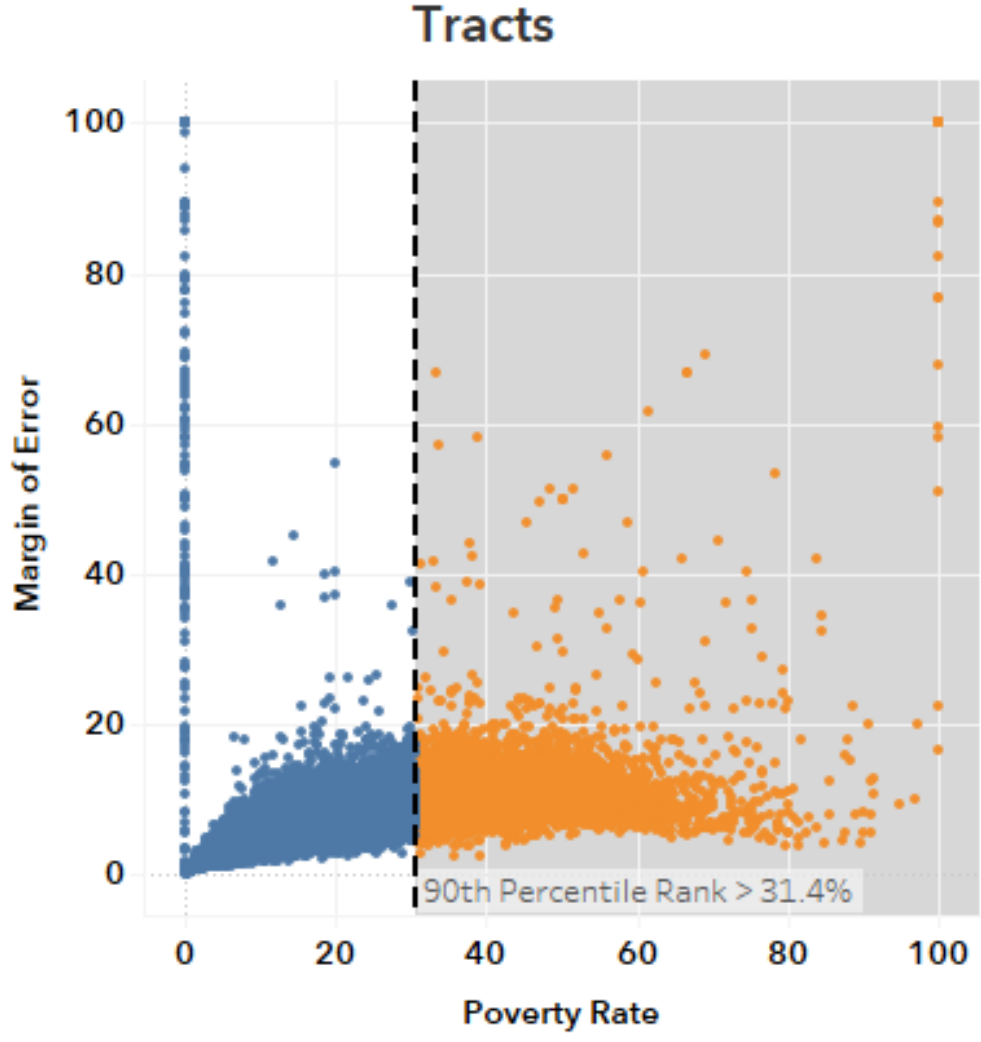
- Objective: To explain issues with how common social vulnerability and community resilience indices use U.S. Census Bureau data and how CRE overcomes these concerns to provide better opportunities for decision makers to evaluate progress on goals of decreasing vulnerability and increasing the resilience of communities across the United States
- Methods: 2019 CRE Case Study
- Key Finding: In comparison to other methods, CRE provides a more precise illustration of how at-risk every neighborhood in the United States is to the impacts of a disaster
- Key Recommendation: Decision makers should use CRE to better measure the resilience and vulnerability of communities across the United States

How Other National Social Vulnerability and Community Resilience Indices Use U.S. Census Bureau Data

- ACS is the premier source of information for building indices
- Other indices treat survey estimates as true parameter values to develop percentile ranks and tag the top percent as “high-risk”
  - For example:
    - Use published ACS 5-year estimates of key economic and social population characteristics to develop vulnerability indicators, like poverty
    - Tag vulnerability indicators ranking in the top 10 percent of all values, i.e., the 90th percentile, for high vulnerability
    - Aggregate vulnerability indicator percentile ranks to create a single score (e.g., socioeconomic theme is the percent in poverty percentile rank plus the unemployment ratio percentile rank plus the per capita income percentile rank plus the percent aged 25+ with no high school diploma percentile rank)

Example of Why Using Percentile Ranking Alone to Build Indices or Describe Survey Data is Extremely Problematic!

How Other Indices Tag High-Risk Areas Example- Poverty Indicator Tag

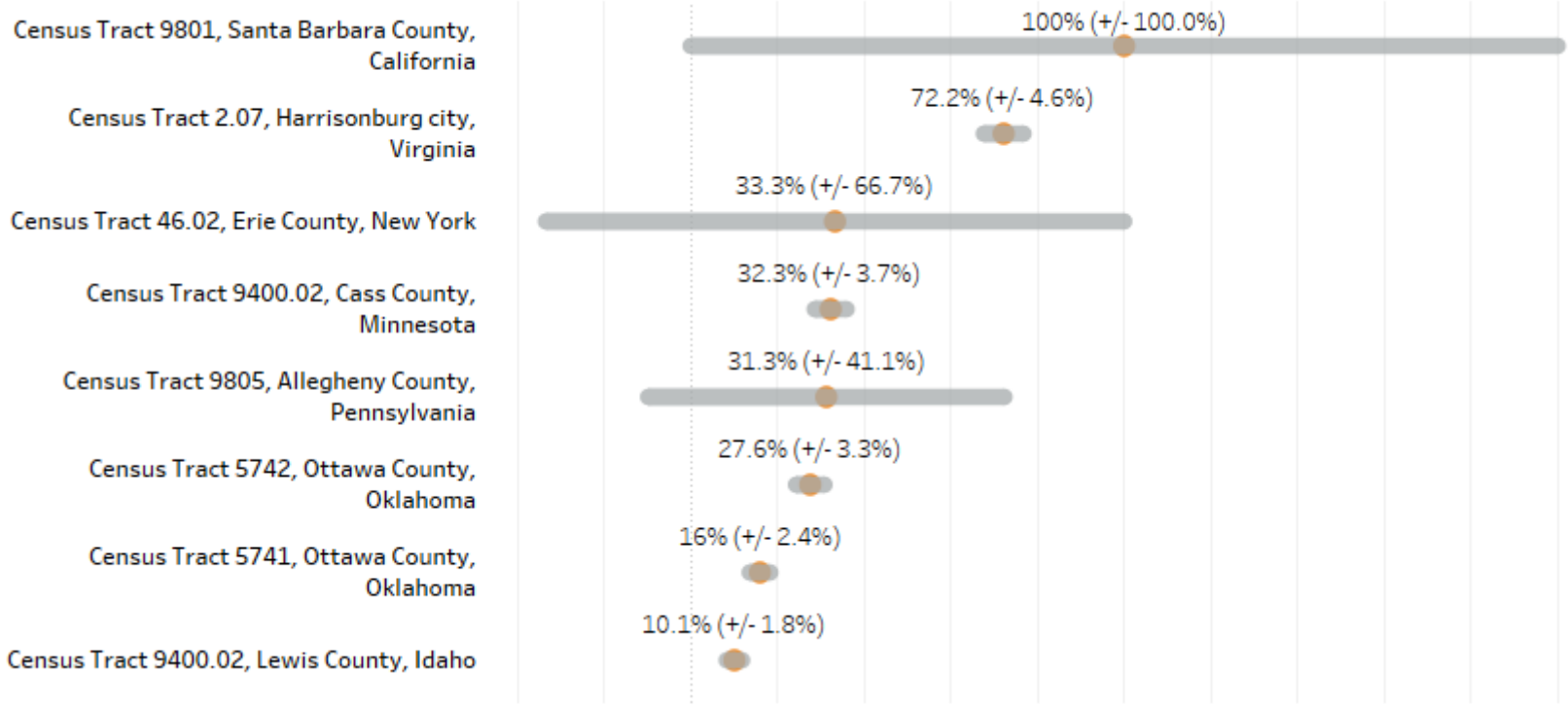


United States Census Bureau

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

2015-2019 ACS Tract Poverty Rate Estimates and Margins of Error
Source: 2015-2019 American Community Survey 5 Year Estimates

Correct Use of Survey Data-
Includes Estimates and Margins of Error

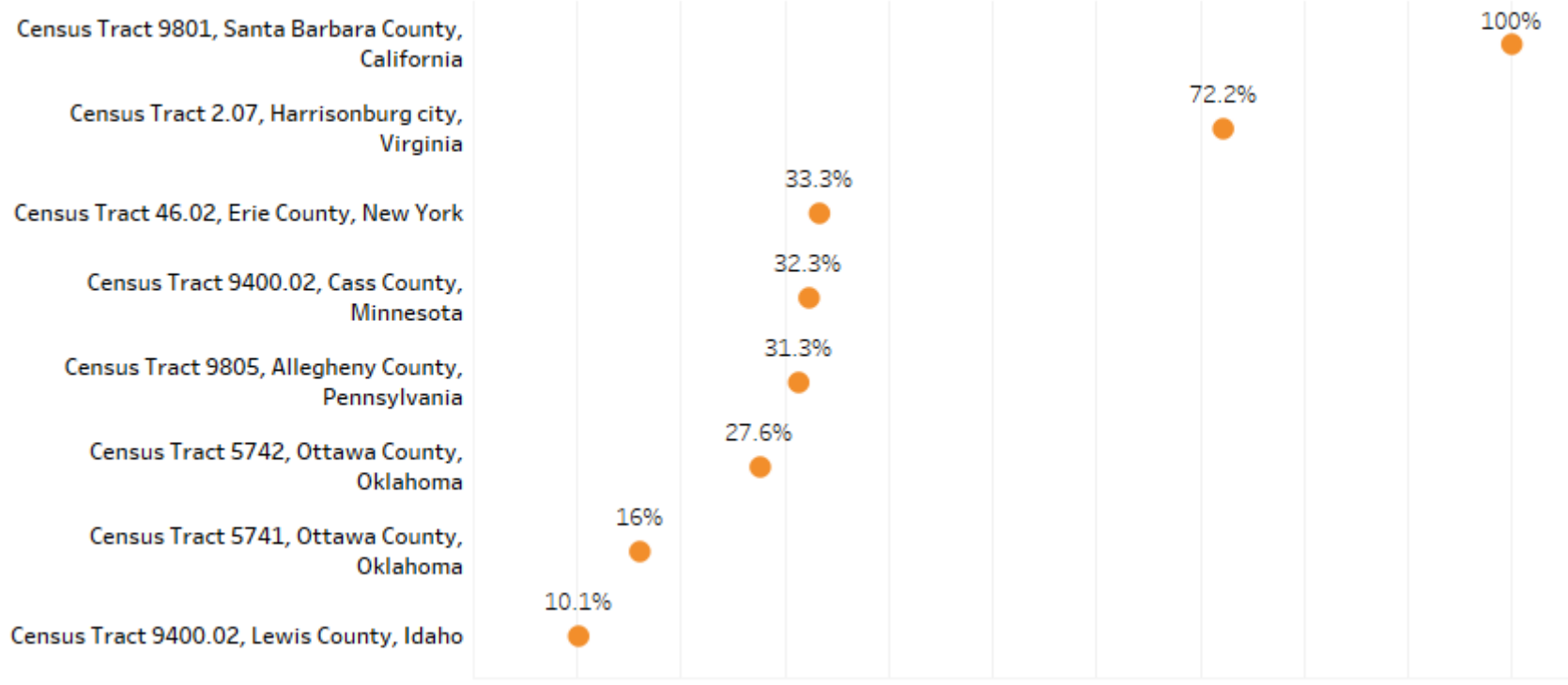


United States Census Bureau

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

2015-2019 ACS Tract Poverty Rate Estimates and Margins of Error
Source: 2015-2019 American Community Survey 5 Year Estimates

Incorrect Use of Survey Data-
Treats Survey Point Estimates as True Parameter Values Without Considering Margins of Error



United States Census Bureau
U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

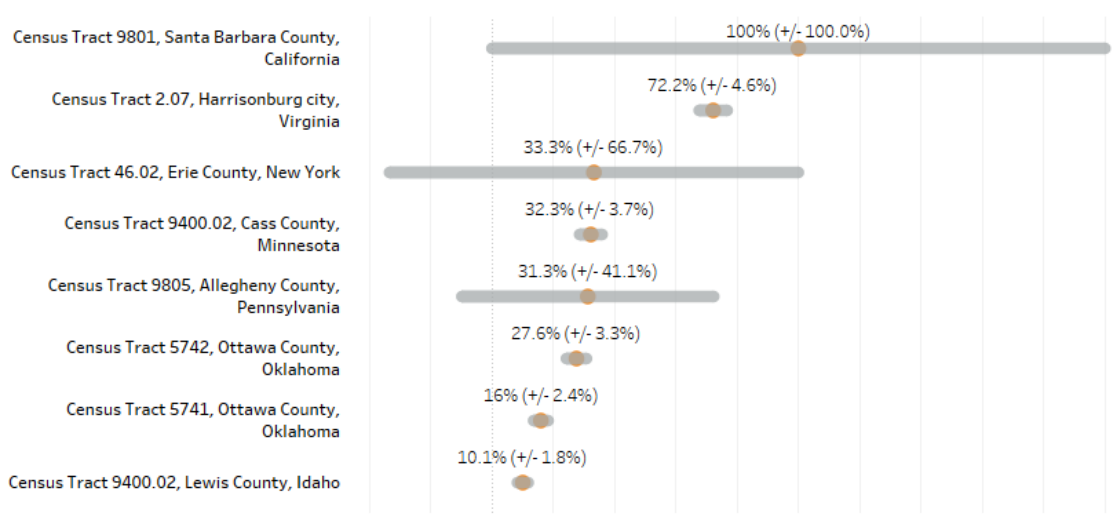
2015-2019 ACS Tract Poverty Rate Estimates
Source: 2015-2019 American Community Survey 5 Year Estimates

Incorrect Use of Survey Data-
Transforms Survey Estimates into Percentile Ranks Without Considering Margins of Error

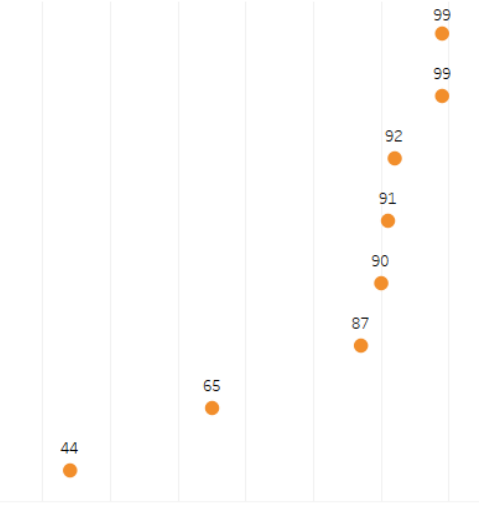
United States Census Bureau
U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

2015-2019 ACS Tract Poverty Rate Percentile Rank
Source: 2015-2019 American Community Survey 5 Year Estimates

Summary of Why Using Percentile Ranking Alone to Build Indices or Describe Survey Data is Extremely Problematic



≠



United States Census Bureau

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

2015-2019 ACS Tract Poverty Rate Estimates with Margins of Error and Percentile Rank

Source: 2015-2019 American Community Survey 5 Year Estimates

Percentile Ranking Methods Using ACS Data is Problematic Because the Margin of Error is Related to High Estimates

The margin of error is higher in areas tagged as high-risk using common percentile ranking methods.

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

The margin of error is highly correlated with percentile rank.

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

Source: 2015-2019 American Community Survey 5 Year Estimates

CRE Methods

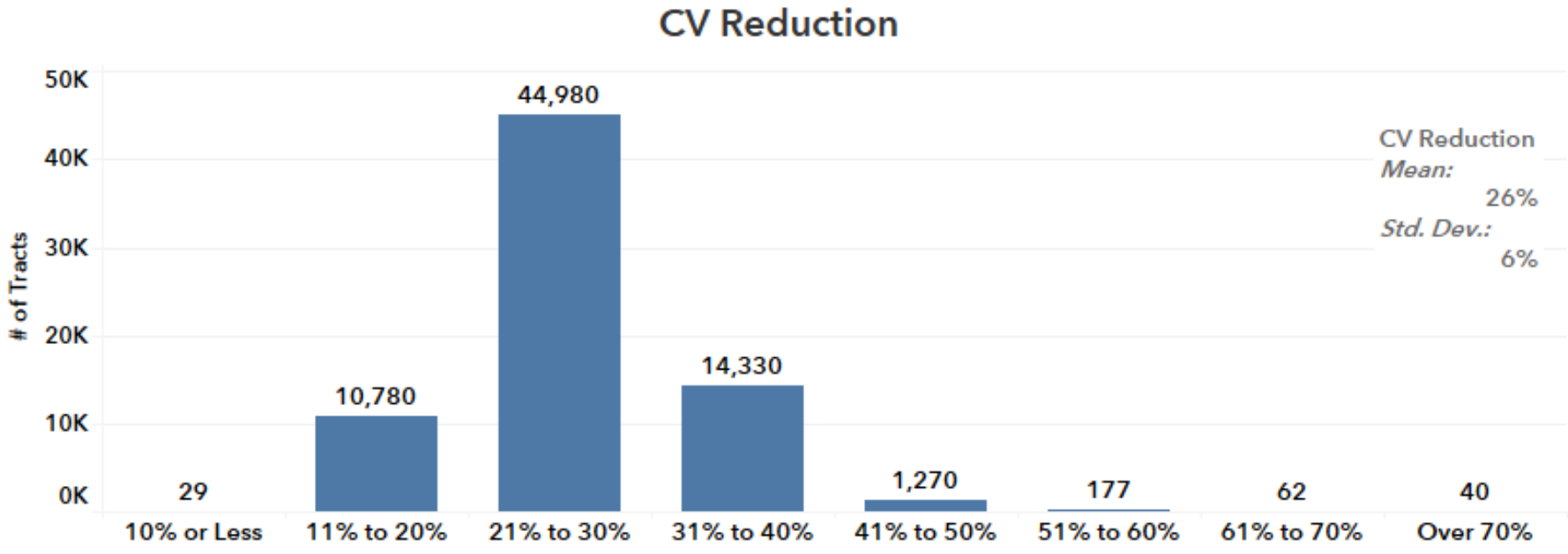
- 2019 CRE used restricted U.S. Census Bureau data
  - 2019 American Community Survey
  - 2019 Population Estimates Program
  - 2010 Decennial Census
- CRE uses established U.S. Census Bureau small area estimation methods
  - Small Area Income and Poverty Estimates (SAIPE)
  - Small Area Health Insurance Estimates (SAHIE)
- Vulnerability indicators are aggregated at the person-level within the microdata and survey weights are used in combination with auxiliary data to create small area estimates of individuals that are low-risk (0 vulnerability indicators), medium-risk (1-2 vulnerability indicators), and high-risk (3 or more vulnerability indicators)

CRE Vulnerability Indicators

1. Households with an income-to-poverty ratio < 130 percent
2. Only one or no individuals living in the household are aged 18-64
3. Household crowding defined as > 0.75 persons per room
4. Household with a communication barrier defined as either limited English-speaking households or households where no one over the age of 16 has a high school diploma
5. No one in the household is employed full-time, year-round. The flag is not applied if all residents of the household are aged 65 or older
6. Individual with a disability posing constraint to significant life activity
7. Individual with no health insurance
8. Individuals aged 65 or older
9. Households without a vehicle
10. Household without broadband internet access

CRE Reduces ACS Error

Description of Percent Reduction in the Relative Error of High-Risk Population for Populated Tracts



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	25.95%	25.91%	25.99%
Standard Deviation	6.20%	6.17%	6.23%
	D.F.	T Value	Pr > |t|
	7,242	65,020	2.15

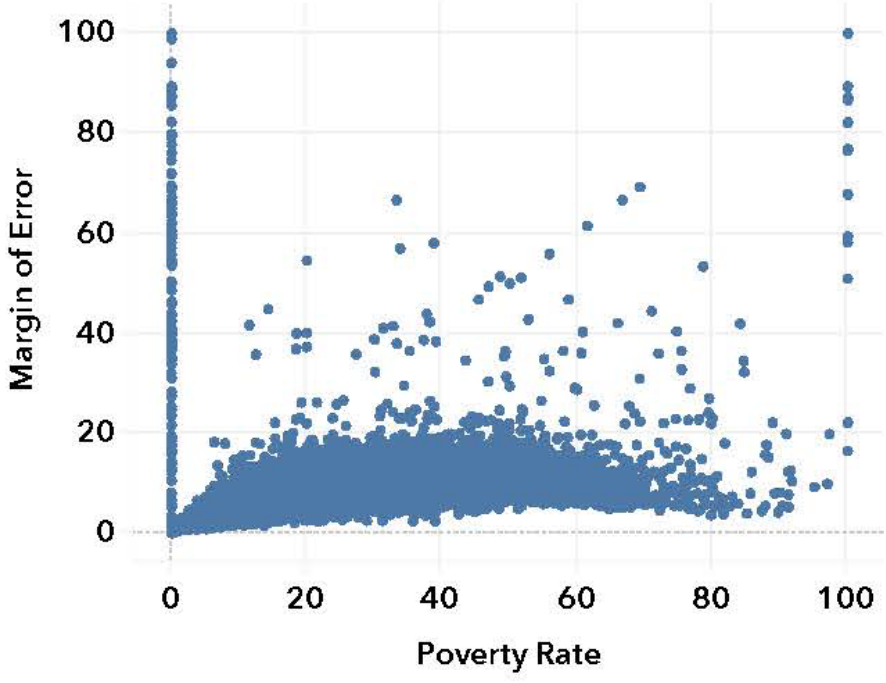
United States Census Bureau

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

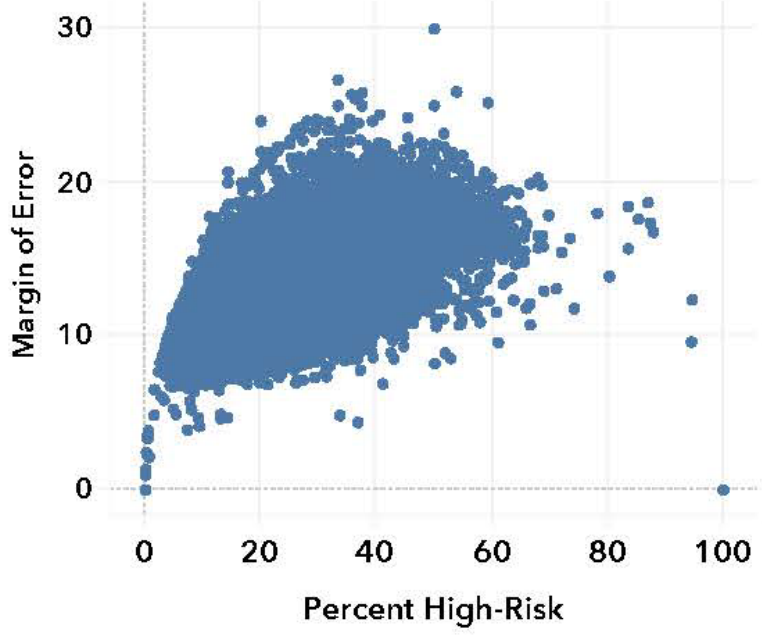
Source: 2019 CRE Connected to Restricted Complete 2019 1-Year ACS Estimates

CRE does not have the same clustering at extremes

2015-2019 ACS Poverty Rate Estimates and Margins of Error



2019 CRE High-Risk Population Rate Estimates and Margins of Error



Source: 2019 CRE and 2015-2019 5-Year American Community Survey

CRE is the ONLY Index that Allows for Statistical Comparisons Between Places

- CRE can answer statistical questions that other indices cannot:
  - What portion of the population has a high-level of socioeconomic vulnerability to disasters (i.e., high-risk population rate) in the United States? What is the rate by region or division?
  - How many people live in vulnerable communities? How about resilient ones?
  - In comparison to small rural, large rural and urban communities, do isolated communities have a higher portion of their population that has a high-level of socioeconomic vulnerability?
  - To what extent is the high-risk population rate of communities in persistent poverty greater than communities not in persistent poverty? How about historically disenfranchised communities?
  - To what extent is the high-risk population rate of toxic communities greater than non-toxic communities?

National, Regional and Divisional High-Risk Population Rates for All, Vulnerable*, and Resilient** Tracts in the United States

	All Tracts	Vulnerable* Tracts		Resilient** Tracts	
	High-Risk Population Rate	Total Population	High-Risk Population Rate	Total Population	High-Risk Population Rate
US	21.57% (+/- 0.24%)	135,039,620	31.10% (+/- 0.28%)	181,584,808	14.49% (+/- 0.46%)
Midwest	20.28% (+/- 0.51%)	24,129,496	30.09% (+/- 0.60%)	41,751,403	14.58% (+/- 0.87%)
East North Central	20.72% (+/- 0.61%)	17,347,131	30.43% (+/- 0.71%)	27,945,385	14.68% (+/- 1.07%)
West North Central	19.32% (+/- 0.91%)	6,782,365	29.23% (+/- 1.10%)	13,806,018	14.37% (+/- 1.49%)
Northeast	22.10% (+/- 0.57%)	22,682,531	32.60% (+/- 0.67%)	31,255,462	14.49% (+/- 1.03%)
Middle Atlantic	23.24% (+/- 0.64%)	18,342,456	33.17% (+/- 0.73%)	21,324,514	14.76% (+/- 1.26%)
New England	18.91% (+/- 1.21%)	4,340,075	30.18% (+/- 1.61%)	9,930,948	13.92% (+/- 1.82%)
South	23.23% (+/- 0.39%)	61,951,852	31.37% (+/- 0.42%)	59,004,105	14.77% (+/- 0.87%)
East South Central	25.05% (+/- 0.86%)	11,213,647	31.08% (+/- 0.91%)	7,130,015	15.82% (+/- 2.20%)
South Atlantic	21.96% (+/- 0.56%)	29,009,873	30.60% (+/- 0.62%)	34,316,830	14.67% (+/- 1.12%)
West South Central	24.45% (+/- 0.69%)	21,728,332	32.54% (+/- 0.71%)	17,557,260	14.55% (+/- 1.74%)
West	19.68% (+/- 0.55%)	26,275,741	30.11% (+/- 0.64%)	49,573,838	14.09% (+/- 0.93%)
Mountain	19.76% (+/- 0.94%)	8,614,103	30.95% (+/- 1.08%)	15,463,055	13.48% (+/- 1.68%)
Pacific	19.64% (+/- 0.67%)	17,661,638	29.70% (+/- 0.80%)	34,110,783	14.36% (+/- 1.11%)

*Vulnerable tracts are those with a high-risk population rate above the national rate.

** Resilient tracts are those with a high-risk population rate below the national rate.

Source: U.S. Census Bureau 2019 Community Resilience Estimates

Comparison of High-Risk Population Rates in Urban, Large Rural, Small Rural, and Isolated Tracts

	Total Population	High-Risk Population Rate	T-Test Comparion Z-Score			
			Urban	Large Rural	Small Rural	Isolated
Urban	276,217,366	20.91% (+/- 0.28%)				
Large Rural	25,617,761	24.58% (+/- 0.70%)	8.04*			
Small Rural	12,052,184	26.59% (+/- 0.91%)	9.84*	2.89*		
Isolated	9,201,005	26.68% (+/-0.87%)	10.34*	3.09*	0.11	

Source: U.S. Census Bureau 2019 Community Resilience Estimates and a Four Category Rural Urban Classification Chart using Rural Urban Commuting Area Codes

* Significantly different based on a t-test with a 90% confidence interval

Comparison of High-Risk Population Rates in Historically Disenfranchised and Persistent Poverty Tracts+

	Total Population	High-Risk Population Rate	Z-Score
Historically Disenfranchised Tracts			
Yes	1,974,058	28.26% (+/- 2.93%)	
No	4,680,685	21.13% (+/- 2.29%)	
			3.15*
Persistent Poverty Tracts			
Yes	1,851,614	32.26% (+/- 2.71%)	
No	4,800,459	19.26% (+/- 2.40%)	
			5.67*

Source: U.S. Census Bureau 2019 Community Resilience Estimates Connected to 2022 United States Department of Transportation Historically Disenfranchised Areas

+Areas were defined using maps produced by the Department of Transportation in accordance to 2021 Consolidated Appropriations Act requirements https://usdot.maps.arcgis.com/apps/dashboards/d6f90dfcc8b44525b04c7ce748a3674a.

* Significantly different based on a t-test with a 90% confidence interval

Comparison of High-Risk Population Rates in Environmentally Toxic+ and Not Environmentally Toxic Tracts

	Not Environmentally Toxic Tracts		Environmentally Toxic Tracts		
	Total Population	High-Risk Population Rate	Total Population	High-Risk Population Rate	Z-Score
Environmental Toxin Type					
2017 Air Toxics Cancer Risk	259,094,926	19.65% (+/- 0.28%)	64,112,323	29.34% (+/- 0.49%)	28.05*
2017 Air Toxics Respiratory HI	258,799,646	19.65% (+/- 0.28%)	64,407,603	29.30% (+/- 0.49%)	27.99*
2017 Diesel Particulate Matter	256,634,031	19.67% (+/-0.28%)	66,573,218	28.91% (+/-0.49%)	27.04*
Traffic Proximity	259,322,102	19.83% (+/- 0.28%)	63,885,147	28.65% (+/- 0.49%)	25.58*
Wastewater Discharge	276,145,933	20.55% (+/- 0.27%)	47,061,316	27.62% (+/- 0.57%)	18.43*
Superfund Proximity	259,200,388	19.92% (+/- 0.28%)	64,006,861	28.28% (+/- 0.50%)	24.04*
RMP Facility Proximity	258,652,070	19.74% (+/- 0.28%)	64,555,179	28.94% (+/- 0.48%)	27.01*
Hazardous Waste Proximity	257,600,115	19.89% (+/- 0.28%)	65,607,134	28.17% (+/- 0.48%)	24.39*
Ozone	257,541,399	19.61% (+/- 0.28%)	65,665,850	29.26% (+/- 0.49%)	28.17*
Particulate Matter 2.5	257,494,471	19.65% (+/- 0.28%)	65,712,778	29.11% (+/- 0.49%)	27.51*
Underground Storage Tanks	256,160,115	19.47% (+/- 0.29%)	67,047,134	29.61% (+/- 0.46%)	30.57*

Source: U.S. Census Bureau 2019 Community Resilience Estimates connected to 2021 EJSCREEN

+ Environmentally toxic communities are those with a percentile ranking of 80 or above https://www.epa.gov/ejscreen/frequent-questions-about-ejscreen#q5

* Significantly different based on a t-test with a 90% confidence interval

Conclusion

Main Points

- Other existing measures of social vulnerability and resilience are less timely and precise and cannot be used to make statistical comparisons between places
- CRE is more reliable to distribute resources and funding
- CRE is the only measure that uses microdata and can thus provide estimates of social vulnerability and community resilience, along with measures of reliability

Key Recommendations

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

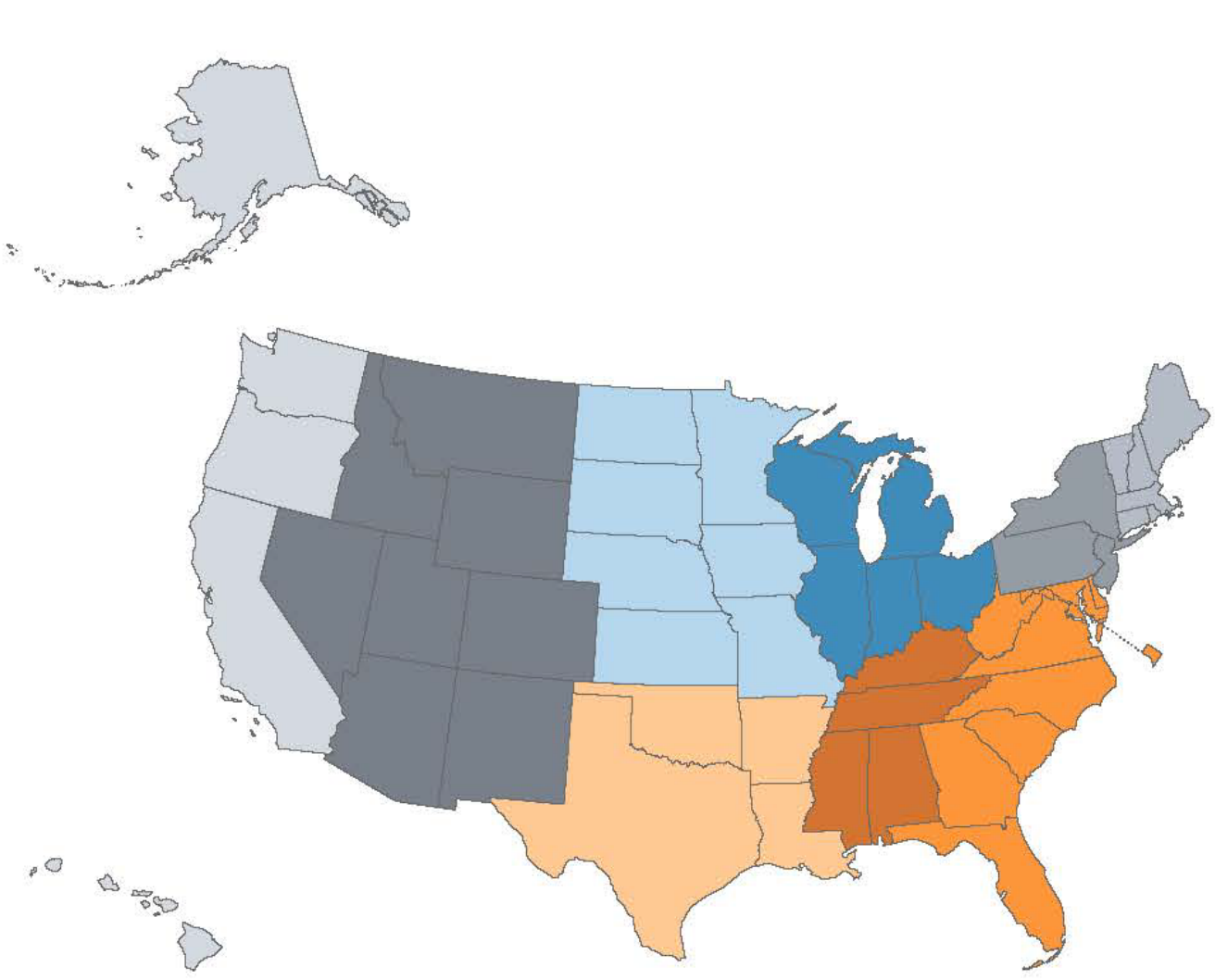
Thank you!

Contact Information

Katherine Ann Willyard (Katherine.a.Willyard@census.gov)

Gabriel Amaro (Gabriel.amaro@census.gov)

Census Regions and Divisions



United States Census Bureau

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

Source: U.S. Census Bureau, Geography Division

Rural Urban Commuting Area (RUCA) Codes and Four Category Rural-Urban Classification Chart

Category	RUCA Code and Description
Urban	1.0: Metropolitan area core: Primary Flow with an Urbanized Area (UA), No additional code
	1.1: Metropolitan area core: Primary Flow with an UA, Secondary flow 30% to 50% to a larger UA
	2.0: Metropolitan area high commuting: Primary Flow 30% or more to a UA, No additional code
	2.1: Metropolitan area high commuting: Primary Flow 30% or more to a UA, Secondary flow 30% to 50% to a larger UA
	3.0: Metropolitan area low commuting: Primary Flow 10% to 30% to a UA, No additional code
	4.1: Micropolitan area core: Primary flow within an urban cluster of 10,000 to 49,999 (large UC), Secondary flow 30% to 50% to a UA
	5.1: Micropolitan high commuting: primary flow 30% or more to a large UC, Secondary flow 30% to 50% to a UA
	7.1: Small town core: primary flow within an urban cluster of 2,500 to 9,999 (small UC), Secondary flow 30% to 50% to a UA
	8.1: Small town high commuting: primary flow 10% to 30% to a small UC, Secondary flow 30% to 50% to a UA
	10.1: Rural areas: primary flow to a tract outside a UA or UC, Secondary flow 30% to 50% to a UA
Large Rural	4.0: Micropolitan area core: Primary flow within a large UC, No additional code
	4.2: Micropolitan area core: Primary flow within a large UC, Secondary flow 30% to 50% to a large UC
	5.0: Micropolitan high commuting: primary flow 30% or more to a large UC, No additional code
	5.2: Micropolitan high commuting: primary flow 30% or more to a large UC, Secondary flow 30% to 50% to a large UC
	6.0: Micropolitan low commuting: primary flow 10% to 30% to a large UC, No additional code
	6.1: Micropolitan low commuting: primary flow 10% to 30% to a large UC, Secondary flow 30% to 50% to a UA
Small Rural	7.0: Small town core: primary flow within a small UC, No additional code
	7.2: Small town core: primary flow within a small UC, Secondary flow 30% to 50% to a larger UC
	7.3: Small town core: primary flow within a small UC, Secondary flow 30% to 50% to a smaller UC
	7.4: Small town core: primary flow within a small UC, Secondary flow 30% to 50% to a smaller UC
	8.0: Small town high commuting: primary flow 10% to 30% to a small UC, No additional code
	8.2: Small town high commuting: primary flow 10% to 30% to a small UC, Secondary flow 30% to 50% to a larger UC
	8.3: Small town high commuting: primary flow 10% to 30% to a small UC, Secondary flow 30% to 50% to a smaller UC
	8.4: Small town high commuting: primary flow 10% to 30% to a small UC, Secondary flow 10% to 30% to a large UC
	9.0: Small town low commuting: Primary flow 10% to 30% to a small UC, No additional code
	9.1: Small town low commuting: Primary flow 10% to 30% to a small UC, Secondary flow 30% to 50% to a UA
	9.2: Small town low commuting: Primary flow 10% to 30% to a small UC, Secondary flow 30% to 50% to a large UC
Isolated	10.0: Rural areas: primary flow to a tract outside a UA or UC, No additional code
	10.2: Rural areas: primary flow to a tract outside a UA or UC, Secondary flow 30% to 50% to a UA
	10.3: Rural areas: primary flow to a tract outside a UA or UC, Secondary flow 30% to 50% to a large UC
	10.4: Rural areas: primary flow to a tract outside a UA or UC, Secondary flow 30% to 50% to a small UC
	10.5: Rural areas: primary flow to a tract outside a UA or UC, Secondary flow 10% to 30% to a large UC
	10.6: Rural areas: primary flow to a tract outside a UA or UC, Secondary flow 10% to 30% to a small UC

Source: WWAMI Rural Health Research Center