New York City Department of Health and Mental Hygiene

Vulnerable Populations: A Function-Based Vulnerability Measure for the New York City Region

Office of Emergency Preparedness and Response

August, 2013

Executive Summary

Vulnerability generally refers to the likelihood for loss during and after an emergency. Conceptual models of vulnerability distinguish between physical (or bio-physical) vulnerability and social vulnerability. While the former relates to the likelihood of exposure, social vulnerability generally refers to those individual, household and community level characteristics that influence outcomes during and after an emergency. Recently there have been a number of efforts to measure social vulnerability at the local scale through the creation of vulnerability indexes. Critics argue that these efforts overestimate the number of those most at-risk, fail to account for the way that individual level characteristics relate, and do not effectively reflect causality. In addition, these indexes rely on a population or taxonomic approach that is inconsistent with guidance from a range Federal agencies including the Department of Health and Human Services (HHS), the Department of Homeland Security (DHS).

In the following report we address these shortcomings by proposing a function-based conceptual model of social vulnerability and a method for operationalizing this model. The conceptual model argues that social vulnerability is a direct function of an individual's incapacity to perform those functions necessary to maintain health in the event of an emergency. Inability to perform these functional abilities, in turn, is a product of how individual, household, and community level characteristics combine within a particular context. We operationalize this model through a five step process. First we associate individual, household, and community level characteristics with an individual's inability to perform key functions. Second, using data from the American Community Survey Public Use Microdata Sample (ACS PUMS) we calculate individual vulnerability scores for each function. Third, we calculate an overall vulnerability score using methods of Pareto Ranking. Fourth, we identify threshold levels for each functional vulnerability as well as overall vulnerability. Fifth, we identify the number of those who are most vulnerable for each Public Use Microdata Area (PUMA) in the New York City area (n=155).

We find that the results differ depending on whether we consider vulnerability density (the number of vulnerable people divided by area) or vulnerability rate (the number of vulnerable people divided by population count). Vulnerability densities were highest in areas with high population densities including northern Manhattan, central Brooklyn, and Chinatown and the Lower East Side. Vulnerability rates, however, tended to peak in urban areas in northern New Jersey including Clifton, Secaucus, Jersey City, and Newark. The South Bronx had both high vulnerability rates and high vulnerability densities, with vulnerability rates peaking in the southeast Bronx and vulnerability densities peaking in the southwest Bronx. South Brooklyn similarly had both high rates and densities with values for both peaking in Coney Island. Even though they used different underlying variables values for functional rates and densities highly correlated. When we compare our results to four other local scale indexes, we find significant differences.

Introduction

The following memorandum fulfills Milestone xx of the Hazard Vulnearbility Analysis project by presenting 8 all-hazard vulnerability indices for the New York City Combined Statistical Area. The memorandum is divided into a background section, data, methods, and results section. In the background we present 4 other methods for constructing local scale vulnerability indicators. We argue that these methods fail to assess the way that individual characteristics combine, do not differentiate between levels of vulnerability, do not measure the scope of the vulnerable populations, and ultimately provide little guidance for emergency management actions.

In response we propose a method based on the idea of functional vulnerabilities and micro data from the US Bureau of the Census's annual American Community Survey. Using these data we generate vulnerability indices for the rates and density of those with transportation, self-care and communication vulnerabilities, as well as two additional vulnerability indicators, for each of the 156 Public Use Microdata Areas (PUMAS) in the New York City Combined Statistical Area (NYC CSA). In addition to estimating the spatial distribution of vulnerable populations we also estimate a 90% confidence interval for the total population at each degree of vulnerability. After generating these indicators we compare the overall count and density vulnerability indicators with indexes created using four methods presented in the peer reviewed literature.

Background

In general, the term vulnerability refers to the potential for loss (Gall 2007). Despite this seemingly straightforward definition, the term's meaning can vary quite dramatically depending on who is using it and in what context (Alwang, Siegel, and Jorgensen 2001). Given its diverse use, it is not surprising that vulnerability can take on many different meanings. Thywissen (2006), for instance, identifies 35 definitions for vulnerability while Cutter (1996) lists 18. According to The United Nations Development Programme's Bureau for Crisis Prevention and Recovery (UNDP-BCP 2004: 11) vulnerability is "a condition or process resulting from physical, social, economic and environmental factors, which determine the likelihood and scale of damage from the impact of a given hazard." Vulnerability may include damage to "social and economic systems, health status, physical infrastructure and environmental assets." Vulnerability differs from risk in that the latter includes not only the potential for harm but also the probability that an event will occur (Villagran De Leon 2006).

In addition to the multiple definitions of vulnerability, there are many different conceptual models. Although terminology varies, one common thread connecting these conceptual models is a distinction between physical and social dimensions of vulnerability. Physical vulnerability is often synonymous with exposure or process magnitude (Hufschmidt 2011). During coastal storms, for instance, those living in flood plains are more likely to be exposed to inundation. Similarly, those living in urban heat islands are more vulnerable during a heat-wave. Social vulnerability, by contrast, refers to those individual, household, and community characteristics that influence outcomes either during or after an event.

According to Cutter and Emrich (2006) social vulnerability is "the susceptibility of social groups to the impacts of hazards, as well as their resiliency or ability to adequately recover from them …susceptibility is not only a function of demographic characteristics … but also more complex constructs such as health care provision, social capital and access to lifelines." Broadly conceived, social vulnerability determines the difference in outcomes between two individuals who have been equally exposed to the same hazard. If, for example, two neighbors are equally exposed to a coastal storm, the one without insurance, living in a mobile home, without access to transportation, and fewer economic resources will most likely suffer greater losses.

Social vulnerability is a product of processes and factors operating at multiple spatial scales including the body, the household, community, and globe. Of particular importance is the role of social capital. In general social capital refers to "the density of trust, networks, or cooperation within a given community" (Scheffler et al. 2008: 1604). Many empirical studies have found a close relationship between levels of social capital and the ability of a community to respond to an event. In his study of deaths during a Chicago heat wave, for example, Klinenberg (2003) finds that contextual factors like crime prevalence and the presence of vacant lots degraded social capital and ultimately correlated with increased mortality even when controlling for race, income, and other factors. In her study of an E Coli outbreak in Walkerton, Ontario, Murphy (Murphy 2007) similarly concludes that the existence of strong community relations prior to the outbreak created a "good conduit through which people channeled their willingness to help their fellow citizens during the crisis." Murphy further argues that the presence of strong social capital also facilitated rapid recovery. According to Haines, Hurlbert, and Beggs (1996), local social networks, are often more flexible and can mobilize more quickly than a heavily centralized emergency response. During the response to the World Trade Center collapse, for example, co-workers helped each other out of the towers and assisted first responders with evacuation and first aid (Chandra et al. 2010).

The determinants of social vulnerability are often, but not always, hazard-specific. Quality of housing, for instance, may be an important determinant to vulnerability to a flood but has less impact on vulnerability to a draught (Tapsell et al. 2010). There are, however, certain characteristics, such as age, disability or income, which increase social vulnerability regardless of hazard. Brooks, Adger, and Kelly (2005) label the former as hazard-specific and the latter as generic. Generic determinants of vulnerability, in turn, are in line with an all-hazards approach to emergency planning which places an emphasis on those elements of a response that are common across a wide range of hazard types (FEMA 2008).

Vulnerability Indicators

Within the past 10 to 20 years there has become an increased interest in operationalizing conceptual models of vulnerability by creating vulnerability indexes. Many researchers, for example, have become increasingly concerned with identifying nations and geographic areas that are vulnerable to sea level rise, drought and other hazards related to global climate change (Barnett, Lambert, and Fry 2008; Adger et al. 2004; Tapsell et al. 2010). In other instances, researchers are identifying vulnerable areas within national borders (Cutter, Boruff, and Shirley 2003; Cutter and Finch 2008). In addition to these global, continental, and national studies are a number of local scale studies that assess relative vulnerability within a county, city, or metropolitan area (Dwyer 2004; Kleinosky, Yarnal, and Fisher 2006; Flanagan et al. 2011; Cutter, Mitchell, and Scott 2000; Rygel, O'Sullivan, and Yarnal 2006; Clark et al. 1998; Ebert, Kerle, and Stein 2008).

In many cases, these indexes mirror conceptual models by combining indexes of physical vulnerability and social vulnerability. While analysts can use established methods like simulation modeling or proximity to estimate the likelihood of exposure, methods for measuring social vulnerability are still in their infancy. In general, the production of social vulnerability indicators is a two-step process. In the first step, analysts identify variables and data sets that serve as proxies for vulnerable populations. Analysts frequently determine these variables by reviewing the past literature and identifying vulnerable groups. For example, Kleinosky et al (2006) identify poverty, the number of immigrants, age, and disabilities as the primary causes of social vulnerability and then operationalize these categories using 57 census variables. Clark et al (1998) similarly cite age, disabilities, family structure and social networks, housing and the built environment, income and material resources, lifelines (including transportation, communication, utilities, and other services), occupation, and race and ethnicity as the elements that contribute to different abilities to cope.

In the second step, analysts aggregate these variables into one single indicator. Cutter et al (2000) re-scale all 8 of the variables in their analysis on a scale from 0 to 1 and then add them together. Flanagan et al (2011) recommend ranking all of the variables and then summing the rank scores. One common approach is to reduce the number of variables using Principal Component Analysis (Clark et al. 1998; Kleinosky, Yarnal, and Fisher 2006; Cutter, Boruff, and Shirley 2003; Rygel, O'Sullivan, and Yarnal 2006; Dwyer 2004) and then taking either the average of the component scores (Cutter, Boruff, and Shirley 2003), a weighted average of the component scores (Dwyer 2004; Kumpulainen 2006; Brooks, Adger, and Kelly 2005), or optimizing the component scores based on some other criteria (Rygel, O'Sullivan, and Yarnal 2006; Kleinosky, Yarnal, and Fisher 2006; Clark et al. 1998; Ratick and Osleeb 2011). See Appendix D For a more comprehensive discussion of the data and methods used to create local scale indicators.

Adger et al (2004) distinguish between inductive and deductive approaches to creating a social vulnerability index. Inductive studies use data from past events to form a statistical relationship between outcomes and social, economic, and physical characteristics. By contrast, deductive approaches base variable selection on a conceptual model of vulnerability and then promote a largely theoretical method of aggregation. Each approach has its advantages and disadvantages. Inductive methods may be appropriate when examining a relatively frequent and well-documented event at local scales (Hinkel 2011) but are less applicable for rare occurrences where data is not readily available. Conversely deductive approaches may be better suited for predicting the spatial distribution of adverse outcomes for rare events but lack a strong empirical basis particularly in regards to methods of aggregation.

Critiques of Social Vulnerability Indicators

Regardless of the method, local scale vulnerability indicators have been met with substantial criticism. Hinkel (2011) lists at least six purposes for creating vulnerability indicators: identify mitigation targets, identify particularly vulnerable people, regions, or sectors, raise awareness of climate change, allocate adaptation funds, monitor adaptation policies, and conduct scientific research. He finds, however, that vulnerability indicators are only useful for achieving one of these goals – identifying vulnerable populations – and then only for clearly defined subsystems at local scales. Tapsell et al (2010) similarly argue that vulnerability indices have limited utility and that their most effective application is to identify populations for more in-depth and qualitative engagement.

One recurring critique of local scale vulnerability indicators is that they do not identify those who are most at-risk. Whether using deductive or inductive methods almost all social vulnerability indexes identify population groups or population characteristics that have historically suffered greater losses during events. Socio-economic status, age, access to health care, gender, and disability status, for instance, are often understood to be determinants of adverse outcomes. This population-based or taxonomic approach designates everyone who shares a particular characteristic as equally vulnerable. Critics, however, challenge this assumption by claiming that there is substantial variability within each of these population groups. For example, in their planning guidance for at-risk populations and pandemic influenza, the Association of State and Territorial Health Officials (ASTHO 2008: 4) argue that while groups such as the elderly or children have traditionally been designated at-risk "not all elderly individuals or children will necessarily be at greater risk in an influenza pandemic simply due to their age. Many elderly live in home or with families who can provide for them and assist them; children who live with capable adults are not necessarily at risk."

As such, the use of population taxonomies can designate many people as vulnerable when they are not and can lead to a significant overestimation of those most at-risk (Wisner et al. 2004; Tapsell et al. 2010). In a national scale study, Kailes and Enders (2007) capture the impact of these false positives by finding that nearly 50% of the US population falls into at least one traditionally vulnerable group. Identifying such a broad segment of the population makes it difficult to direct limited resources and ultimately undermines emergency managers' ability to respond (Handmer 2003).

A second critique of social vulnerability indicators is that they consider each characteristic in isolation. Rather than the product of a single characteristic, vulnerability usually results from the alignment multiple variables within in a particular context. Wisner et al (2004: 16), for example, compare the vulnerability of young, immigrant, non-english speaking, single mothers living in area's bordering on San Pedro Harbour near Los Angeles with more affluent women in nearby Rancho Palos Verde. The authors argue that "the concatenation of income, age, immigration status, language and single parenthood significantly shifts the meaning of 'gender' as a simple category or box-to-tick in a taxonomy of vulnerability." Although women in San Pedro Harbour have the same gender as their female couterparts in Rancho Palos Verde "in most other respects, they inhabit a separate universe."

Because most taxonomic approaches rely on aggregated data it is impossible to accurately capture the concurrence of attributes. Instead, social vulnerability indexes are restricted to identifying places where there is a high percentage of women, immigrants, single parent households, and those that do not speak English (to use the above example). Yet, just because an area may have a high percentage of these populations, it does not mean that there is a high percentage of people with all of these characteristics combined.

Third, most vulnerability indicators do not distinguish cause. Vulnerability indicators may help identify where there is most likely to be the greatest number or percentage of people at-risk but not why (Wisner et al. 2004). The failure to identify causal processes is particularly problematic because children, the elderly, and those living in poverty (for example) may be at risk for very different reasons (Buckle, Mars, and Smale 2000; Birkmann 2006; Kailes and Enders 2007). For emergency managers, identifying and understanding what makes groups vulnerable is essential for effectively deploying scarce resources. Translators may help reduce the vulnerability of those who do not speak English but be of little utility to the elderly who do not face any language barriers. Further, during an emergency, some sources of vulnerability, like lack of access to transportation or medical care, may be addressed while it may not be possible to address others, like poverty. With knowledge of vulnerability's causes, emergency managers, can target resources where they will have the greatest impact (Buckle, Mars, and Smale 2000).

Function based vulnerabilities

In the following, we propose an alternate approach to measuring social vulnerability based on functional vulnerabilities. The function-based approach to vulnerability comes from recent changes in disability practice and law. Historically, disability was defined in-terms of diagnoses such as blind, deaf or mentally ill. Disability advocates and policy makers, however, have recognized that people with the same diagnosis can have very different needs and capabilities. Accordingly, they have shifted their focus away from an emphasis on diagnosis and towards an individual's ability to perform, or obtain assistance in performing, key day-to-day functions like bathing, eating, or leaving the house. Functional approaches to disability are now common place and have been adapted by the World Health Organization (WHO 2012) and codified in the Americans with Disability Act (*Americans with Disabiliies Act Amendments Act* 2008) .

Recently, those in emergency management have made a similar shift. Rather than focusing on population level characteristics, like age, gender, or income, there is now a greater emphasis on an individual's ability to perform key functions necessary to maintain health and safety before, during, and after an emergency. In the *National Response Framework*, the Federal Emergency Management Agency (FEMA 2008) defines special needs populations as:

"Populations whose members may have additional needs before, during, and after an incident in functional areas, including but not limited to: maintaining independence, communication, transportation, supervision, and medical care. Individuals in need of additional response assistance may include those who have disabilities; who live in institutionalized settings; who are elderly; who are children; who are from diverse cultures; who have limited English proficiency or are non-English speaking; or who are transportation disadvantaged."

Other agencies and organizations including The US Department of Health and Human Services (HHS 2012) and ASTHO (ASTHO 2008) have adopted, FEMA's definition. The relationship between function and vulnerability has also been codified in law in the Pandemic and All-Hazards Preparedness Act of 2006 (*Pandemic and All-Hazards Preparedness Act* 2006).

Aside from compliance with federal directives, adoption of a function based approach has several other advantages. Properly implemented, a function based model should produce significantly fewer false positives than taxonomic approaches, ultimately enhancing emergency managers' ability to identify those most at-risk (Kailes and Enders 2007). A second advantage is that because they make causality explicit (an inability to perform certain tasks), functional approaches provide planners and emergency managers clear guidance for action. Clusters of people who face challenges transporting themselves, for instance, require fundamentally different resources and assistance than those with communication impairments.

Despite these benefits, operationalizing function-based approaches has proven to be a challenge. While data on age, income, or gender are readily available, measures of functional abilities are not. Further, a functional approach explicitly recognizes that people are constantly moving in and out of vulnerable states (Kailes and Enders 2007). While someone may not be functionally impaired prior to an event, they may be after the event has occurred. These temporalities are poorly suited to annual, quinquennial, or decennial surveys. For these, and other, reasons we have been unable to find any examples of studies that operationalize functional approaches. In the following we address this need by combining function and population approaches.

A Conceptual Model of Function-Based Social Vulnerability

Figure 1 shows our conceptual model of social vulnerability. At the heart of the model is the belief that social vulnerability is the direct product of a person's inability to perform those functions necessary to maintain health during and after an emergency. Consistent with Brooks, Adger and Kelly's (2005) distinction between generic and hazard-specific vulnerabilities, necessary functions may be divided between those that are hazard specific and those that are necessary in all situations. In the latter category we include communication, transportation, and maintaining independence. Notably we exclude two functions specified in the FEMA definition: supervision and medical care. We exclude supervision because we consider it to significantly overlap with maintaining independence. Our choice to exclude medical care reflects our belief that medical care is a general need (like food, water, or shelter) and not a function that someone needs to perform.

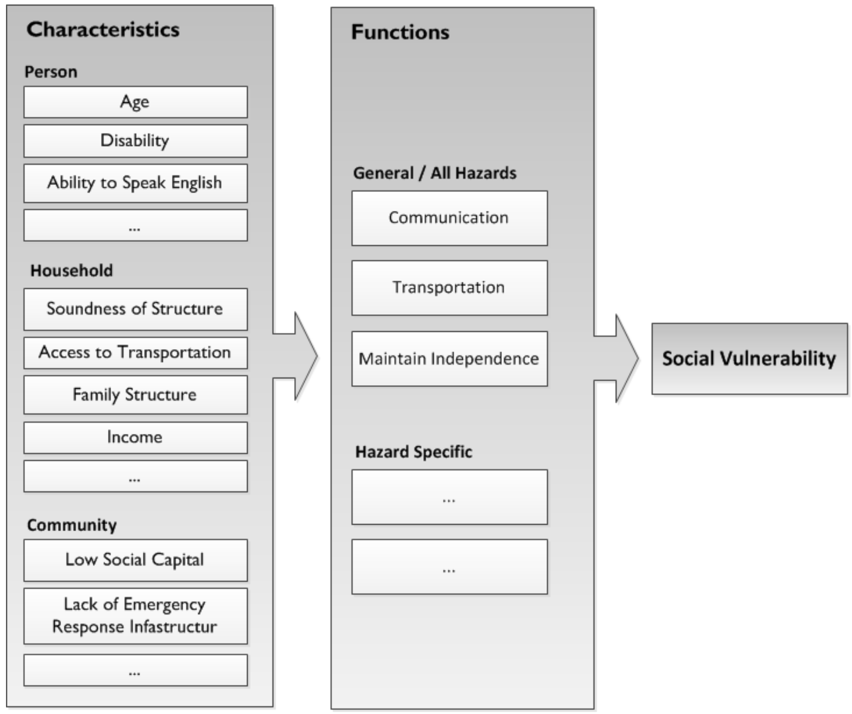


Figure 1. A conceptual model of function-based social vulnerability

Ability to perform each of these functions, in turn, is directly associated with certain characteristics. These characteristics can exist at a range of spatial scales including personal, household, and community. Personal characteristics include age, disability, or an inability to speak English. Household characteristics may include income, the presence of an automobile, or soundness of structure, and community characteristics include such factors as low levels of social capital or lack of emergency response infrastructure. While we specifically mention personal, household, or community level characteristics, processes operating at national, family, or any other of a range of spatial scales may also be relevant. Each characteristic may be, but is not required to be, related to more than one function. Income, for instance, may inhibit someone's ability to transport his or her self or maintain independence but may not impact that person's ability to communicate. Characteristics may also have a multiplicative or compounding effect such that a poor person with a disability and living in an area with low social capital will have significantly more difficulty performing certain functions than a poor person without a disability living in an area with multiple and robust social networks. In the following sections we discuss our efforts to operationalize our conceptual model by creating a function based measures of social vulnerability.

Data

The main data source for the function-based social vulnerability measure is the American Community Survey. Historically the Census Bureau collected data about income, employment status, housing costs, housing conditions, commutes, and disabilities every 10 years through the long-form of the US decennial census. Starting in 2006, the Census Bureau replaced the long-form with the American Community Survey (ACS). Although the ACS is similar to the long-form, the former samples approximately 1 in 6 households every 10 years while the latter samples 1 in 40 addresses annually (US Bureau of the Census 2009).

In addition to publishing ACS summary data for geographic areas (census tracts, counties, states, etc), the Census Bureau also publishes individual and household responses as the ACS Public Use Microdata Sample (ACS PUMS). To protect confidentiality ACS PUMS data excludes names, addresses, and top codes certain clearly identifying values. The census bureau also protects confidentiality by associating each record with a Public Use Microdata Area (PUMA). Each PUMA contains a population of at least 100,000 and as much as possible corresponds with existing city, state, or county boundaries (US Bureau of the Census 2009). PUMAs are far more coarse than census tracts. In New York City Combined Statistical areas, for example, there are 156 PUMAs compared with 2,164 census tracts.

In the 2008 ACS, the Census Bureau redefined existing questions relating to disability. At that time, the Census Bureau included questions specifically related to the following attributes:

- Self-Care Difficulty (difficulty dressing or bathing)
- Hearing Difficulty
- Vision Difficulty
- Independent Living Difficulty (difficulty doing errands alone such as visiting a doctor's office or shopping)
- Ambulatory Difficulty (serious difficulty walking or climbing stairs)
- Cognitive Difficulty (serious difficulty concentrating, remembering, or making decisions)

The Census Bureau does not support comparison with previous disability data, arguing that current disability information is largely incommensurate with prior definitions and questions (US Bureau of the Census 2010).

In order to estimate total population, the Census Bureau assigns a person weight to each individual record, and a housing weight to each household record. Person weights and household weights roughly correspond to the number of people or households that are similar to the specified record. If, for instance, a person record has a person-weight of 34, that record represents 34 individuals. Accordingly, the sum of all weights for records in the CMSA will equal the population of the CMSA. In addition to person and household weights, the ACS PUMS also includes 80 replicate weights. Each of these can be substituted for the person-weight to calculate an alternate population estimate (US Bureau of the Census 2009).

In addition to data from the American Community Survey, the function-based social vulnerability measure also relies on community level data to measure social capital. These data include the County Business Patterns (US Bureau of the Census 2012b). The Bureau of the Census administers the CBP annually to collect subnational data on economic activity. In addition to other information, the CBP includes the number of firms and the number of employees by industry. The other data sets we use to measure social capital are the Mail response rates for the 2010 decennial census (US Bureau of the Census 2012a), and the total number of votes cast by county in the 2008 presidential election (USA Today 2008).

Methods

There is a five-step process for generating function-based social vulnerability measures.

Step 1. Identify Individual Attributes for Each Function

In the first step, we associate individual variables in the American Community Survey Public Microdata Use Sample (ACS PUMS) with each of the three functions mentioned earlier: transportation, self-care, and communication. Variables are divided into categories of individual/personal, household, and community. Individual variables include age (over 80, and under 10), ability to speak English well, whether the person has health insurance, whether the person is living in poverty, and the presence of vision, hearing, independent livig, ambulatory, cognitive, or self care difficulties. Household variables include whether there are incomplete plumbing facilities, whether the household is in a rural or suburban area and has no car, or whether it is a grandparent headed household.

The sole community level variable is whether the community has low social capital. To measure social capital we use methods presented by Rupasingha, Goetz, and Freshwater's (2006) index of social capital. The index is comprised of three parts. First is a measure of associational density which includes the number of civic organizations, bowling centers, golf clubs, fitness centers, sports organizations, religious organizations, political organizations, labor organizations, business organizations, and professional organizations per 100,000 people in each county. Second is the percentage of eligible voters who cast a ballot in the 2008 presidential election. Third is the response rate for the 2010 decennial census, and fourth is the number of tax-exempt non-profit organizations per 100,000 people. We then use methods of principle component analysis (PCA) to extract principal components from these four variables, with the first principal component serving as the index of social capital.

We determined the association between function and characteristic by consulting peer-reviewed literature as well as subject matter experts within the New York City Department of Health and Mental Hygiene. Although, each association is dichotomous (TRUE/FALSE), many of the variables are continuous (such as age) or categorical (like ability to speak English). In these cases we recoded the variable as dichotomous. The variable for age, for example, may be recoded as two separate variables: under the age of 4 and over the age of 80. Similarly, the variable for does not speak English includes responses for does not speak English well and does not speak English at all. We were conservative in making association in the sense that if there was a question whether there was a relationship between function and characteristic, the association was deemed as false. Relationships between population characteristics and functions are shown in table 1.

Characteristic	Transportation	Self-Care	Communication
Individual			
Over 80 yrs old	•	•	
Under 10 yrs old	•	•	
Does not speak English well			•
No health insurance		•	
Living in poverty	•	•	
Vision difficulty	•		•
Hearing difficulty			•
Independent living difficulty	•	•	
Ambulatory difficulty	•	•	
Cognitive difficulty	•	•	•
Self care difficulty		•	
Commute more than 1 hr	•		
Household			
Incomplete plumbing facilities		•	
Rural or Suburban / No Car	•		
Grandparent headed HHD	•		
No telephone service available			•
Community			
Low Social Capital	•	•	

Table 1. Associations between functions and characteristics

Step 2. Calculate individual vulnerability scores

In the second step, we calculate the vulnerability score for each function for each individual record in the ACS PUMS. To calculate a person's vulnerability score we sum that person's number of function related attributes. If, for example, a person is above the age of 80 and has a cognitive disability, the person would receive a transportation score of 2. This scoring process is demonstrated in table 2 which shows the scoring for transportation vulnerability for three hypothetical respondents to the ACS. One way to interpret the table is that Person A has 3, person B has 2, and person C has 4 of the characteristics that will impede their abilities to transport themselves during an emergency.

	Person A	Person B	Person C
Over 80 yrs old	•		
Under 10 yrs old		•	
Living in poverty	•		
Vision difficulty			•
Independent living difficulty	•		
Ambulatory difficulty			
Cognitive difficulty			
Rural or suburban / no car			•
Grandparent headed HHD			•
Low social capital		•	•
Total	3	2	4

Table 2. Hypothetical scoring for transportation vulnerability

Step 3. Generate an overall vulnerability score

In the third step, we generate an overall vulnerability score based on the vulnerability score for each function. While the impulse may be to sum or average the vulnerability scores for transportation, communication, and self-care, there are several problems with this approach. The first issue is that if a person has a high score for one function but a low score for another, the two scores will negate each other. Another concern is that averaging scores requires a weighting scheme of some sort. If no weights are explicitly applied, it is implied that all functions have an equal weight. There may, however, be instances where some functions may be more important than others. Determining weights that reflect these relative values will always be subjective, and, to a certain extent, arbitrary.

In recommending the use of Pareto Ranking alorithms, Rygel, O'Sullivan and Yarnal (2006) provide an alternate approach. Pareto Ranking is a genetic multicriteria optimization algorithm that determines weights based on the concept of domination and non-domination (Fonseca and Fleming 1995). In a Pareto Ranking algorithm, observation A is said to dominate observation B if all values for A are greater than or equal to the same values for B and at least one value of A is greater than the same value for B. Accordingly, if A dominates B, A will be preferable regardless of the weighting scheme.

To generate Pareto ranks, the analyst first identifies all observations that are not dominated by any other observation. These observations are assigned the highest rank and removed from the dataset. It can be argued that regardless of the weighting scheme, at least one of the observations with the highest rank, and none of the observations with a lower rank, could be the most desirable or have the highest score. Next, the analyst again identifies all of the remaining observations that are not dominated by any other observations. These observations are assigned the second highest rank and removed from the dataset. This process continues until all observations have been assigned a rank. Each person's assigned rank is equal to his or her overall vulnerability. For a more complete description of Pareto Ranking methods see Appendix A.

Step 4. Determine vulnerability threshold

In the fourth step, we determine what vulnerability score (or threshold) we consider to be amongst the most vulnerable. To do this we estimate number of people for each vulnerability score for each function and for the overall score. In order to estimate the number of people for each vulnerability score we sum the person weight for all respondents with that vulnerability score or higher and divide by the total population (eq 1).

$$c_s = \frac{\sum_{i=1}^{n} h(s,i) * pwt_i}{\sum_{j=1}^{n} pwt_j} \qquad \textbf{(1)}$$

Where c_s is the percentage of the population with a vulnerability score of at least *s* for a particular function (for example transportation, communication, or self-care), n is the number of respondents to the ACS, h(s,i) is an indicator function which equals 1 if person i has a score of at least s and is otherwise 0, and pwt_i which is a weight of person i. After calculating c_s for all vulnerability scores (s), we can then identify the threshold value (s) that captures a large enough portion of the population to have a meaningful impact but not so many people as to exceed available resources. In making these determinations it may be helpful to calculate a confidence interval for each vulnerability score using methods documented by the Bureau of the Census (US Bureau of the Census 2009).

We illustrate this process using transportation vulnerability scores. Table 3 shows the percentage of people with a transportation vulnerability score greater than equal to each of the 7 possible values. In addition, the table shows the 90% confidence interval for counts at each level.

Vulnerability Score	Low Estimate	Estimate	High Estimate
6	.002 %	.003 %	.005 %
5	.013 %	.017 %	.021 %
4	.561 %	.579 %	.598 %
3	2.139 %	2.185 %	2.232 %
2	9.948 %	10.029 %	10.110 %
1	29.572 %	29.685 %	29.796 %
0	100.000 %	100.000 %	100.000 %

Table 3. Percentage of population with a vulnerability equal to or greater than the number specified. Low estimate and high estimate represent the 90% confidence interval levels

One way to read this table is that we have 90% confidence that between 2.139% and 2.232% of the population have 3 or more of those characteristics that we believe impede their ability to transport themselves during an emergency. Based on these findings we would most likely select a cutoff of 3 because a lower value (like 2) would include too much of the population, while a higher value (like 4) would not include enough.

Step 5. Estimate the number of most vulnerable for each PUMA and map the results

Once the vulnerability scores have been generated, the next step is to estimate the number of people who are most vulnerable for each of the functional categories as well as overall. To do this for each PUMA we sum the person weights of all ACS respondents with a vulnerability score greater than the threshold value (eq 2).

$$MV_p = \sum_{i=1}^{n} h(i, thresh) * pwt_i \quad \textbf{(2)}$$

Where MV_{puma} is equal to the number of most vulnerable in puma *p*, *n* is the number of ACS respondent in puma *p*, *h* is an indicator function which equal 1 if the vulnerability score for person i greater than *thresh* the threshold value for that function, and pwt_i is the person weight for person *i*. Once the number of most vulnerable are calculated the analyst can then map either the density of most vulnerable (by dividing by PUMA area), or the rate (by dividing by the PUMA population).

Results

Threshold Values

Table 4 shows the percentage of the population with the specified vulnerability score or greater for transportation, communication, self care, and overall vulnerability. For example, 3.219% of the population has a transportation vulnerability score of 5 or greater while 2.408% of the population has a communication vulnerability score of 2 or greater. The dark grey cells indicate the threshold value for each type of vulnerability. Appendix B shows the same results with the 90% confidence intervals.

Score	Transportation	Communication	Self Care	Overall
21	NA	NA	NA	0.001
20	NA	NA	NA	0.010
19	NA	NA	NA	0.025
18	NA	NA	NA	0.101
17	NA	NA	NA	0.205
16	NA	NA	NA	0.399
15	NA	NA	NA	0.669
14	NA	NA	NA	1.018
13	NA	NA	NA	1.477
12	NA	NA	NA	2.030
11	NA	NA	NA	2.715
10	NA	NA	NA	3.652
9	0.000	NA	NA	5.167
8	0.047	NA	0.003	7.985
7	0.492	NA	0.104	11.906
6	1.548	NA	0.560	23.372
5	3.219	NA	1.543	34.247
4	8.071	0.100	3.438	64.777
3	25.245	0.629	8.446	72.284
2	66.398	2.408	27.338	99.922
1	99.875	14.127	68.365	100.000
0	100.000	100.000	100.000	NA

Table 4. Cumulative percentage of the population with specified vulnerability score or greater

Spatial Distribution of Vulnerability Results

Figures 2a-b and 3a-b show the spatial distribution of the densities and rates of the overall most vulnerable populations. The overall most vulnerable population refer to the number of people in each PUMA with an overall vulnerability score greater than or equal to the threshold value of 10 (see table 4). Densities are equal to the number of most vulnerable in each PUMA divided by the PUMA's land area, and rates refer to the number of most vulnerable in each PUMA divided by the PUMA's total population. Dark purples denote areas with high values while dark greens indicate areas with lower values.

In general, density values follow population densities and are highest in the city with peaks in the Bronx, Manhattan, and eastern Queens and eastern Brooklyn. Within these areas the highest values are in the south Bronx, with a peak in southwest Bronx, and Chinatown and the Lower East Side in southern Manhattan. There are also elevated values across northern Manhattan, with a peak in West Harlem, Central Brooklyn, South Brooklyn, and Coney Island, and in northwestern Queens, including Astoria. Vulnerability rates are also highest in the south Bronx although peak values shift eastward. The elevated values in Brooklyn are still pronounced, although less so compared with densities, with peak values in Coney Island and the Rockaways. Areas of Manhattan that showed high densities of vulnerable populations, including northern Manhattan and Chinatown and the Lower East Side, did not show elevated rates. Several small and midsized urban areas in New Jersey including Secaucus, Clifton, and Newark also show elevated rates.

Figures C1 – C6 (Appendix C) show the spatial distribution of the densities and rates for each of the functional vulnerabilities. As was true with the overall density, the PUMAs with the highest functional densities cluster in and near Manhattan. In the Southwest Bronx, the density values for all three functional vulnerabilities are more than 2.5 standard deviations greater than the mean. There were also consistently high values in Chinatown and the Lower East Side, upper Manhattan, Central and South Brooklyn, and Coney Island. Rate values for all three functional vulnerabilities were more dispersed with consistent peaks in the Southeast Bronx and Coney Island and elevated values in the Rockaways and Central Brooklyn. There was also a consistent peak near Clifton, NJ with sporadically elevated values in other urban areas in New Jersey including Newark, Jersey City, Hoboken, and Secaucus.

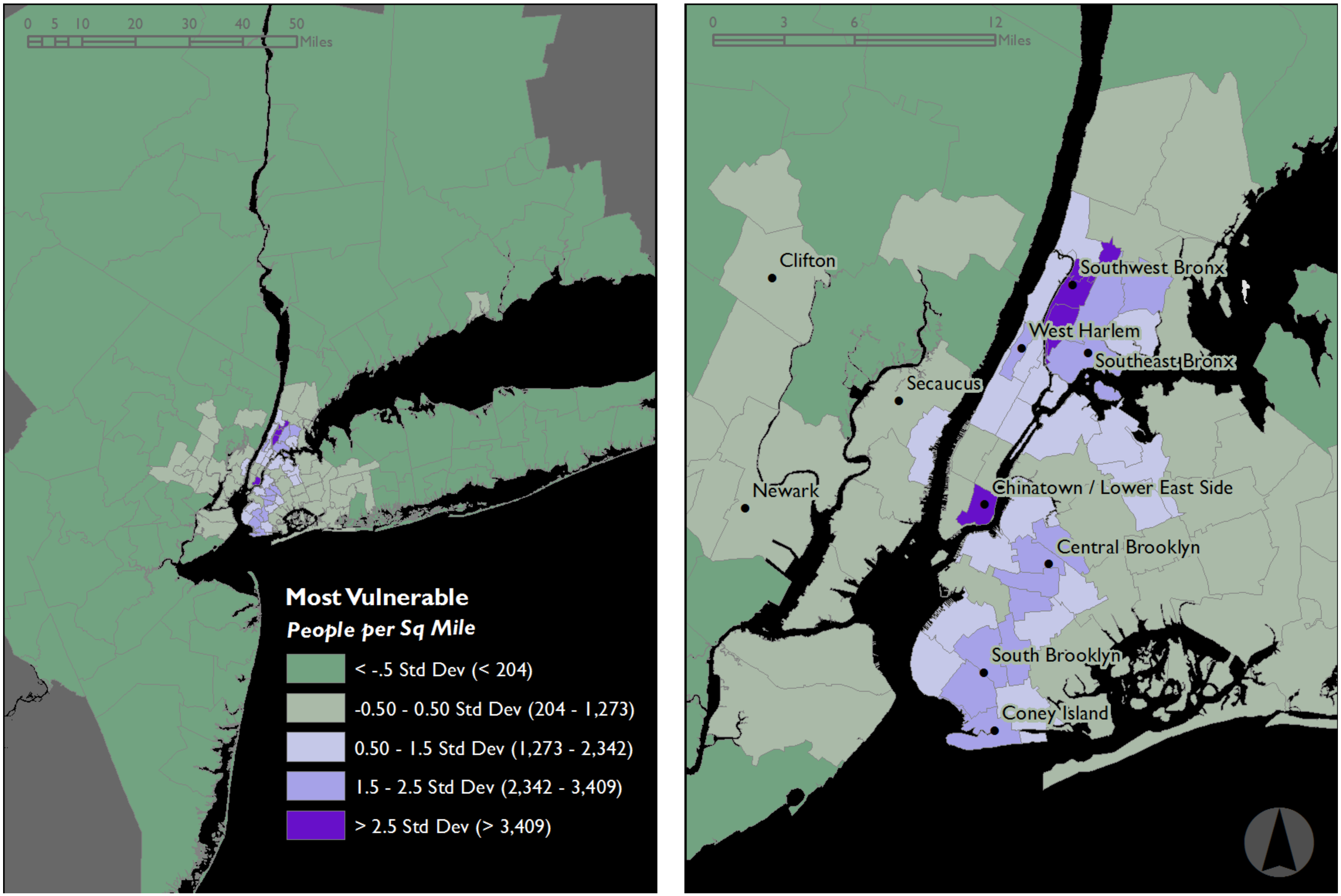


Figure 2a-b. Most Vulnerable Population densities

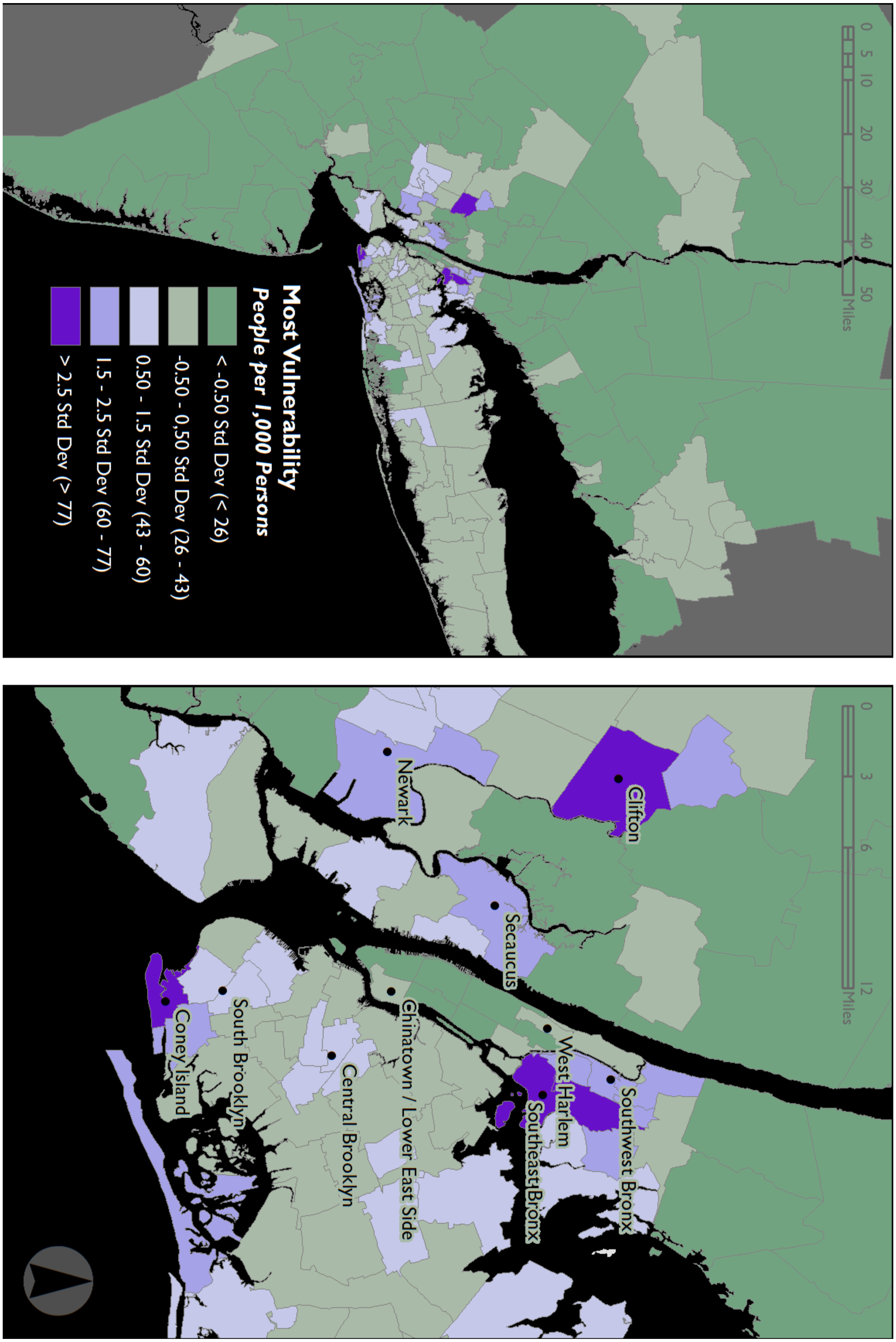


Figure 3a-b. Most Vulnerable Population densities

Relationships between functional measures of vulnerability

Table 5 shows the correlation of measures for transportation, communication, self-care, and overall functional vulnerabilities for both rates and densities. Overall, table 5 shows that these measure are all highly related. The upper left quadrant shows the relationship of density measures to other density measures. These values are highly correlated with values ranging between .858 for the relationship between communication density and transportation density and .994 for the relationship between self care densities and overall density. The overall density, in particular, relates strongly to other measures with correlation values of .984, .911, and .994 for transportation, communication, and self care respectively.

		Density				Rate			
		Transportation	Communication	Self Care	Overall	Transportation	Communication	Self Care	Overall
Density	Transportation	1.000	0.858	0.979	0.984	0.556	0.564	0.672	0.687
	Communication	0.858	1.000	0.898	0.911	0.357	0.702	0.543	0.552
	Self Care	0.979	0.898	1.000	0.994	0.447	0.558	0.641	0.627
	Overall	0.984	0.911	0.994	1.000	0.461	0.570	0.632	0.637
Rate	Transportation	0.556	0.357	0.447	0.461	1.000	0.579	0.875	0.928
	Communication	0.564	0.702	0.558	0.570	0.579	1.000	0.682	0.725
	Self Care	0.672	0.543	0.641	0.632	0.875	0.682	1.000	0.971
	Overall	0.687	0.552	0.627	0.637	0.928	0.725	0.971	1.000

Table 5. Correlation of Functional Measures

The lower right quadrant of table 5 shows the relationships between the rate measures. Although these measures clearly correlate the relationships are not as strong as those for the density measures with values ranging from .579 for the relationship between transportation rates and communication rates and .971 for the relationship between self care rates and overall rates. As was true with density measures the overall value correlated highly with other measures with correlations of .928, .725, and .971 for the correlation between the overall rate and transportation, communication, and self care rates respectively. These relationships are not terribly suprising given that many of the measures rely on the same variables (table 5) and that the overall measures are a composite of the other 3.

The lower left quadrant and the upper right quadrant show the relationships between the densities and rate measures. In general there is a moderate relationship between the two. Most measures moderately relate to their counterpart. The correlation between transportation rates and transportation density, for example, is .556 while those for communication rates and communication density is .702. The overall measures show a similar relationship with the correlation between overall density and overall rate equal to .637. What this indicates is that results may differ quite substantially whether population or area is in the measure denominator, and that while some PUMAs may rank highly for both rate and density, others may rank highly for one and not the other.

The spatial distribution of these relationships is shown in figure 4. Darker green colors indicate areas where overall vulnerability densities are more than .5 standard deviations below the mean, and darker purple colors indicate areas where overall vulnerability densities are more than .5 standard deviations above the mean. Small white circles indicate areas where the overall vulnerability rates are more than .5 standard deviations below the mean and black circles indicate areas where overall vulnerability rates are more than .5 standard deviations above the mean. As can be seen in figure 4a, the majority of areas outside of New York City have both low values for the overall rate-based measure and low values for the overall density-based measure. Within New York City both high values for the overall rate-based measure and high values for the overall density-based measure can be found in south Brooklyn, central Brooklyn, and the south Bronx. With the exception of the Lower East Side, areas in southern Manhattan tend to have higher densities and lower rates. Many areas in northern New Jersey have the reverse with higher rate and lower densities.

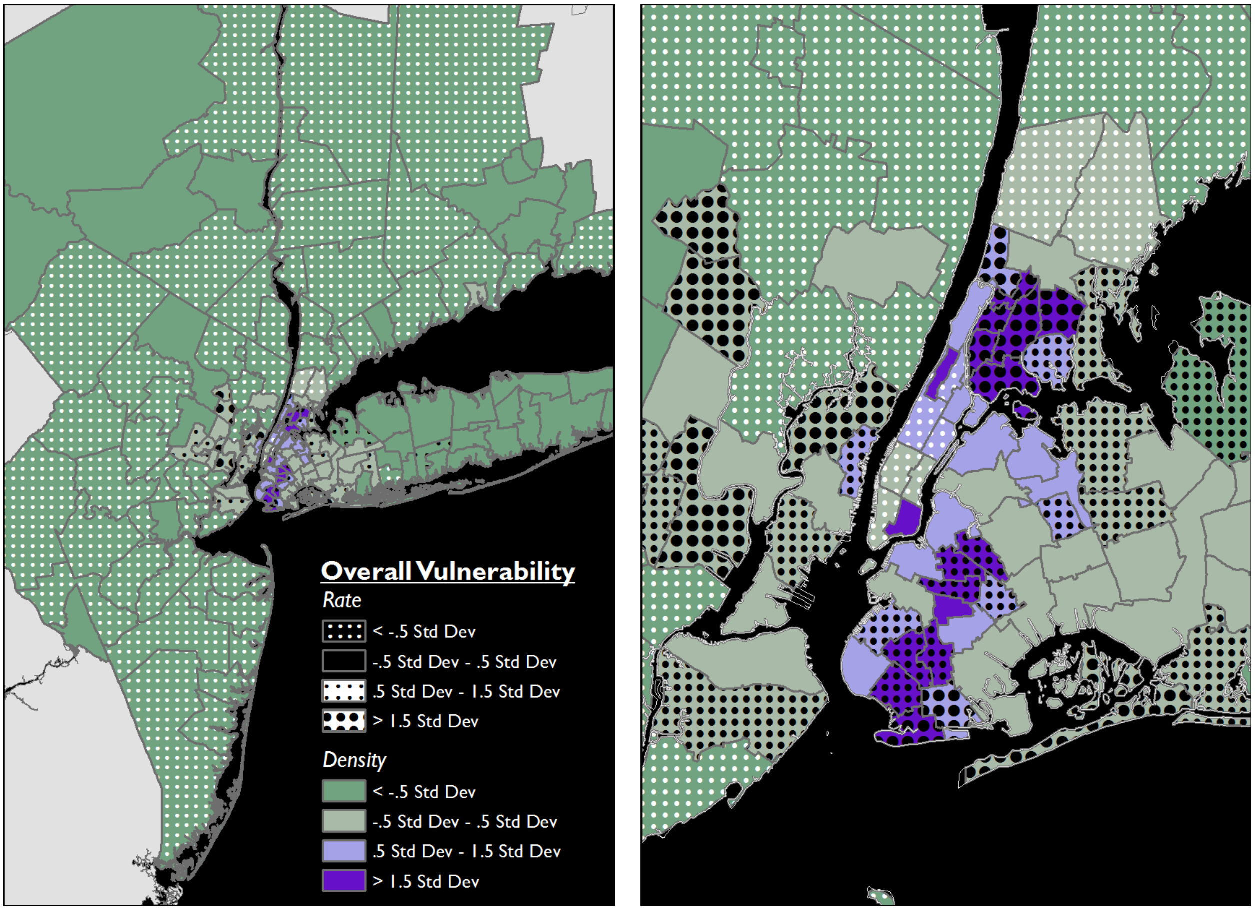


Figure 4a-b. Comparison of overall vulnerability rates and densities

Rate and Density Measures Compared with Vulnerability Indices

Table 6 shows the correlation between the rate and density measures and five local scale vulnerability indexes (see Appendix D for indicator description and derivation). Overall the vulnerability indexes do not correlate very well with each other. Aside from the meta index, the only two indicators that relate are Rygel and Flanagan (r=.545). Given that it is an average of the other indexes, it is not surprising that the meta index correlates with all of the other four indexes.

		Cutter 2000	Cutter 2003	Rygel	Flanagan	Meta
Indexes	**Cutter 2000**	1.000	0.103	0.232	0.193	0.573
	Cutter 2003	0.103	1.000	0.173	0.310	0.595
	Rygel	0.232	0.173	1.000	0.545	0.731
	Flanagan	0.193	0.310	0.545	1.000	0.768
	Meta	0.573	0.595	0.731	0.768	1.000
Density Measures	**Transportation**	0.064	0.434	0.515	0.633	0.617
	Communication	0.096	0.463	0.566	0.556	0.631
	Self Care	0.094	0.472	0.548	0.601	0.643
	Overall	0.094	0.468	0.543	0.621	0.647
Rate Measures	**Transportation**	-0.105	0.314	0.256	0.492	0.359
	Communication	0.042	0.309	0.437	0.480	0.476
	Self Care	-0.056	0.496	0.406	0.474	0.495
	Overall	-0.045	0.450	0.395	0.557	0.509

Table 6. Comparison of local scale vulnerability indexes with function-based rates and densities

With the exception of Cutter 2000, the density measures correlate reasonably well, but not perfectly, with the vulnerability indices. The overall density measure best correlates with the Meta (r=.647), Flanagan (r=.621), and Rygel (r=.543) indexes, less well with Cutter 2003 index (r=.468), and not at all with Cutter 2000 index (r=.094). The vulnerability indices do not relate to the rate measures as well as they do the density measures. Flanagan is the only vulnerability index that shows a strong relationship with the overall rate measures (r=.557) while the relationships between the Cutter 2003 (r=.450) and the Rygel (r=.395) indexes are much weaker and the relationship with the Cutter 2000 index (r=-.045) is non-existent.

To help explain the reasons behind the differences, we collected all PUMA level variables (n=156) used in the vulnerability indexes and then compressed them into four factors using factor analysis with Varimax rotation. The four factors explained 68% of the variance in the dataset and can be described as follows:

Factor 1: High density, urban, low percentage of car ownership, high percentage of renters, and high percentage of public transit users

Factor 2: young, lower percentage of people over the age of 65, high percentage of people living in poverty, lower median household incomes, lower percentage of high school graduates, higher percentage of single mothers

Factor 3: higher percentage of all disability types, lower percentage of adults in the workforce

Factor 4: higher percentage of Asians, new immigrants, and people who do not speak English

Table 7 shows the correlation between the density measures, rate measures, indices and the above factors. Although the Rygel (r=.673), Flanagan (r=.540), and Meta (r=.631) indexes highly correlate with factor 1 (high density), they do not correlate as strongly as the transportation density (r=.779), communication density (r=.829), self care density (r=.819), and overall density (r=.827) measures. Factor 2 (high percentage young, poorer) reveals another key difference between the vulnerability indices and the density measures. Neither the local scale indexes nor the function-based measures highly correlate with poverty. The only vulnerability indexes that relates to the factor is Flanagan (r=.490) and, to a lesser extent, Cutter 2003 (r=.332). The relationships between factor 2 and transportation density measure (r=.432) and self care density measure (r=.366) is not much stronger. Factor 2 shows a similar relationship with the rate measures, particularly the transportation rate (r=.454) and the self care rate (r=.372). Factor 3 (high percentage of people with disability) reveals another key difference between the rate measures and the vulnerability indices. For the most part, the vulnerability indices show no relationship or a very weak relationship with factor 3. By contrast the communication rate measure (r=.500), the self care measure (r=.472), and the overall rate measure (r=.489) show a moderate relationship with this factor.

		Factor 1	Factor 2	Factor 3	Factor 4
Indexes	Cutter 2000	0.145	0.038	0.082	0.086
	Cutter 2003	0.326	0.332	0.284	0.059
	Rygel	0.673	-0.102	0.295	0.350
	Flanagan	0.540	0.490	0.325	0.020
	Meta	0.631	0.285	0.370	0.193
Density Measures	Transportation	0.779	0.432	0.186	0.001
	Communication	0.829	0.175	0.299	0.110
	Self Care	0.819	0.366	0.214	0.018
	Overall	0.827	0.380	0.214	0.016
Rate Measures	Transportation	0.178	0.454	0.331	0.058
	Communication	0.392	0.169	0.500	0.182
	Self Care	0.334	0.372	0.472	0.091
	Overall	0.324	0.442	0.489	0.090

Table 8. Correlation of vulnerability indexes and function based rates and densities with population factors

To better understand the spatial relationships between the vulnerability indexes and the rate and density measures, we first standardized the meta index, overall density measure, and overall rate measure by subtracting the mean from each value and dividing the resulting difference by the standard deviation. We then subtracted the standardized meta index from the standardized overall density measure and subtracted the standardized meta index from the overall rate measure. The results are shown in figure 5a-d. We label all PUMAs with a value greater than one as places where the overall density or rate measure is higher than the meta index and all PUMAs with values less than -1 as places where the meta index is higher than the overall rate or density measures.

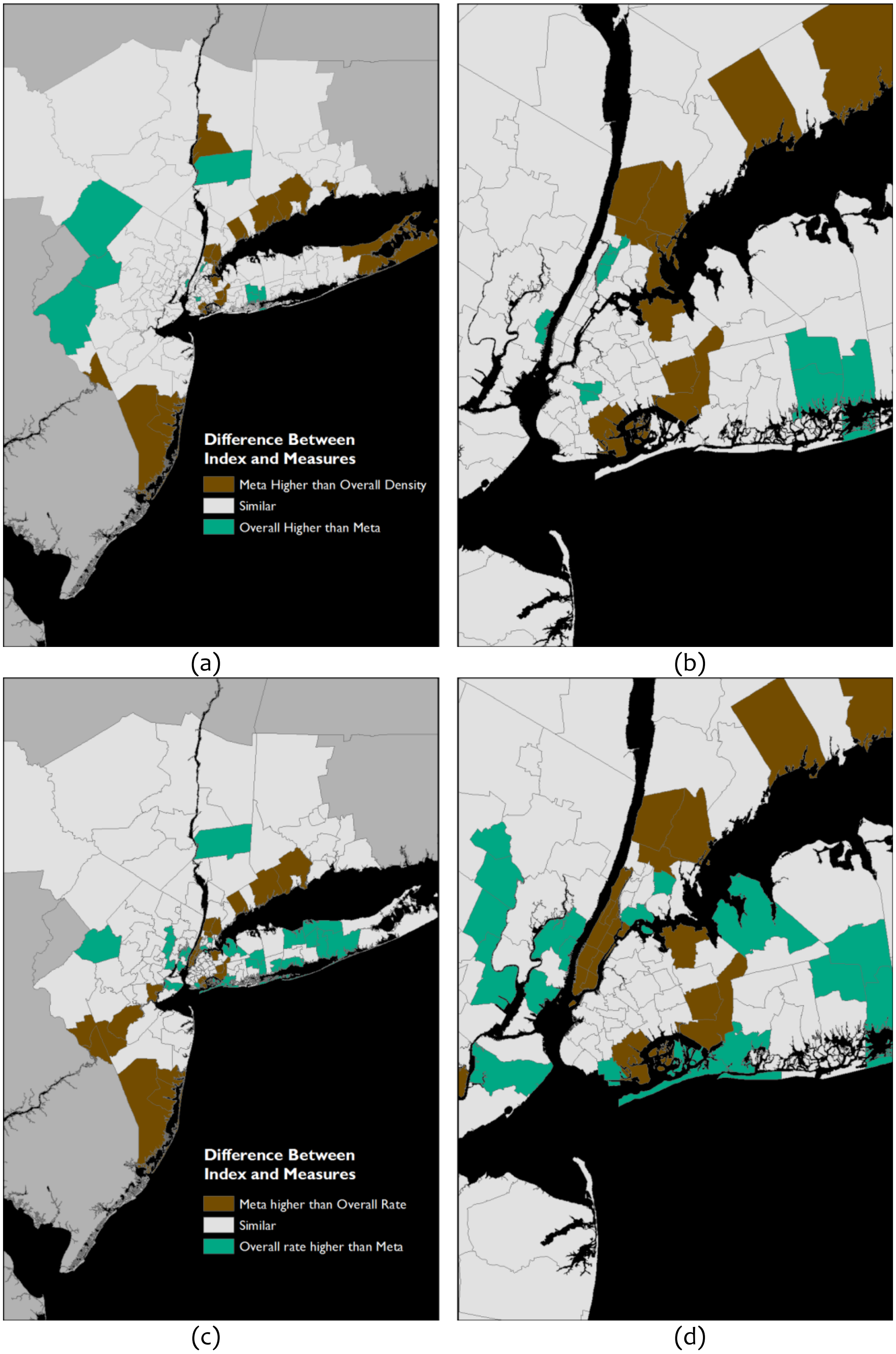


Figure 5a-d. Differences between vulnerability indexes and function-based densities and rates

In general, the meta index is higher than the overall density measures in central New Jersey, eastern Suffolk County in Long Island and intermittently along the Atlantic coast from southern West Chester through southern Connecticut. The overall density is higher than the meta index in fewer places including western New Jersey, central Brooklyn, Jersey City, eastern Nassau county and the south eastern edge of the Bronx. The differences between the overall rate measure and the meta index largely mirror the differences between the overall density measure and the meta index with clusters in central New Jersey and along the Atlantic Coast north of New York City. Notably, the entirety of Manhattan and eastern Nassau County show similar differences. This most likely reflects the Meta index's greater correlation with population densities. There is a cluster in northern New Jersey from Jersey City and Bayonne up to Passaic where overall rates are higher than Meta Index values.

Discussion

The most vulnerable areas are in New York City and urban areas in Northern New Jersey

As seen in figures 2, c1, c2, and c3, peak values for all functional based vulnerabilities cluster in and around Manhattan. Rates and densities for communication, transportation, self-care, and overall vulnerabilities were consistently highest in the south Bronx with higher density values in the southwest Bronx and higher rate values in the southeast Bronx. South Brooklyn is another area with both high rate and density measures, with peak values in and around Coney Island. Central Brooklyn had consistently high values, but this peak was less pronounced when examining rates. There were consistently high density values in upper Manhattan, with peaks in West Harlem, but those areas were close to the area-wide mean for vulnerability rates. Chinatown and the Lower East Side similarly had high density values but lower rate values. Conversely, many of the urban areas in northern New Jersey including Clifton, Newark, Jersey City, Hoboken, and Secaucus had relatively low density values but higher rate values.

Communication, Transportation, and Self Care Vulnerabilities Are Closely Related

Although we stress the differences between the different types of functional vulnerabilities the three type of vulnerability we discuss in this paper: communication, transportation, and self care closely correlate. This is not surprising given the significant overlap between the characteristics associated with each. Of the characteristics identified in table 2, roughly half (8 out of 17) are associated with two or more functions. Aside from the overlap of characteristics, the reasons that functional measures relate differ for density and rate measures.

For density measures, one reason that functional vulnerabilities are closely related is the influence of population density. As shown in table 7, all functional measures strongly correlate with high population densities. This factor is also associated with a high density of almost all at risk individuals including the very old, the very young, those with disabilities, and those living in poverty. This indicates that the density of vulnerable populations roughly follows the density of the general population, and that one is most likely to find vulnerable groups where there are the most people. This is also evident in figures 2a, c1a, c2a, and c3a where peak density values are highest in New York City. As strong as this relationship may be, however, it is far from perfect and there are some areas, like Clifton, NJ, where there may be a high density of vulnerable populations even though there is not a high population density. Conversely, there are some areas, like lower Manhattan, with very high population densities but low densities of those with functional vulnerabilities.

For rate measures, the strong relationship between transportation, communication, and self care vulnerabilities is likely driven by the fact that many people have multiple disabilities. Roughly half of the 2.3 million people with disabilities in the New York City area have one or more disabilities. For example, slightly more than 125,000 people (roughly 5% of those with disabilities) simultaneously have self care, independent living, ambulatory and cognitive difficulties. So even if two functional vulnerabilities are associated with different disabilities it is likely that the same people will score highly for both.

The above findings are, of course, dependent on this report's operationalization of different functional vulnerabilities. If functional vulnerabilities were associated with different individual, household, or community characteristics, it is entirely possible that the relationships between them would also differ. It is also important to note that even though functional vulnerabilities correlate, they do not do so perfectly. There are therefore potential and important differences between the measures, and analysts should exercise caution in substituting one for the other.

Analysts should use both rates and densities

In the New York City area vulnerability rate measures and vulnerability density measures overlap but differ in important ways. As shown in table 5, there is a clear relationship between the overall vulnerability density and overall vulnerability rate (r=.637). There are, however, key differences between the two. Throughout much of northern New Jersey, for example, there are a number of areas where vulnerability rates are quite high but overall vulnerability densities are not. Conversely, in places like the Upper East Side or Upper West Side in Manhattan there are a number of places where overall vulnerability density is high but overall vulnerability rates are not.

In these instances, it may be difficult for emergency managers to determine whether to rely on rates or densities. There is a compelling argument for either. If an emergency manager relies on rates, interventions are more likely to reach intended audiences. If she instead depends on densities to determine the allocation of resources, those interventions are more likely to impact a greater number of people. In general, the selection of the appropriate measure is dependent upon the task at hand. Public messaging, for example, may best be targeted where densities are highest, while outreach may be better in areas with higher rates.

Because the two measures provide different insights, they can also be used in combination. Consider, for example, an area like the Upper East Side that has relatively high density of vulnerable people but relatively low rates. The combination of the two measures reveal that the high density is most likely the result of high population densities and where there are high densities there are most likely to be many people who are vulnerable. Conversely, in a place like northern New Jersey vulnerability densities are low but vulnerability rates are high. The higher vulnerability rates in that area indicates that low vulnerability densities are due to lower population densities in general and not a lower percentage of those that are most vulnerable. A third example of how rates and densities can be used in combination are in places like the southern Bronx and Coney Island where rates and counts are both high. These are most likely the areas that demand the most attention.

There is no consensus understanding of vulnerability

Exsiting vulnerability indicators tell very different stories. As shown in table xx many of the indexes do not correlate. The one exception to this is Rygel and Flanagan (r=.545) and that may, in large part, be due to population density (table 8). For the most part, however, there is little agreement between existing vulnerability indicators and the areas they identify as most vulnerable. The differences between vulnerability indicators in part reflect their intended purposes. While both Flanagan and Cutter 03 created national scale indices, Rygel's index was created for Hampton's Roads area of Virginia and Cutter 00's indexe was created for Georgetown County, South Carolina.

In each instance, the creators of the index incorporated factors that may be relevant to their study areas but that may not exert as great an influence elsewhere. All indexes, for example, include those without a car as particularly vulnerable. In many areas of New York City, however, there is a robust public transportation infrastructure and car ownership may be of little value in an emergency.   The vulnerability of agricultural lands and areas with low population densities may also vary significantly between places.   Differences may also arise depending on whether the index was created to measure vulnerability to a particular hazard or is general. For example, it is logical that both Rygel and Cutter include the percentage of mobile homes because each was designed to capture vulnerability to coastal storms. For other hazards, like a pandemic outbreak, the prevalence of mobile homes may exert less influence on event outcomes.

Another factor that may cause a difference between the indices is the relative importance of rates, counts, and densities. Some of the indexes, like Flanagan, include only percentage of those considered vulnerable, while others, like Cutter 00, include only the counts of historically vulnerable groups, while still others, like Flanagan and Rygel, combine both counts and rates. As mentioned earlier, however, results can vary substantially depending on whether one is examining rates or count. While indexes that use counts or densities tend to identify areas that are urban and have higher population densities as the most vulnerable, indexes that use rates are more likely to emphasize the vulnerability of rural and agricultural areas.

It is important that vulnerability measures and indexes directly relate to emergency preparedness and response activities

In almost all cases, existing indexes identify those areas deemed most vulnerable. These indexes , however, are of little value to emergency managers either before, during, or after a response for at least three reasons. First, they do not clearly identify either the level of vulnerability they are measuring or the size of the at-risk population. This oversight often leads to an overestimation of those most at-risk and ultimately may impede emergency managers' ability to effectively deploy scarce resources. Second, these indexes are difficult to interpret. In many cases, index authors do not clearly state whether they are measuring rates or counts. Each measure, however, serves a different purpose. Further, an index alone is difficult to gauge. If one area has an index score of 50 and another has a score of 25, does it mean that the first is twice as vulnerable as the second? Third, indexes capture areas where there is either a high number or a high percentage of people who are members of traditionally vulnerable groups, but fail to measure the coincidence of factors that make people vulnerable. This oversight is problematic because vulnerability's cause is often the alignment of multiple characteristics within a particular context. An elderly person who is blind, for example, is more vulnerable than a similar elderly person without visual impairments. Fourth, indexes conflate vulnerabilities' causes. The failure to identify cause makes it difficult for emergency managers to deploy the resources necessary to mitigate potential loss.

In order to maximize utility, it is recommended that indexes authors work closely with emergency managers to create tools that provide the information that emergency managers need. One way to approach this problem is to adopt a function-based conceptual model of vulnerability. Not only are function-based conceptual models consistent with current federal guidance, when properly implemented, they can also directly tie into existing mitigation and response activities. Knowing where people are likely to have difficulty communicating, transporting, or caring for themselves, for instance, provides emergency managers important and actionable information. These functions can be further broken down to relate to specific tasks like self-decontamination or navigating an emergency shelter.

Future Research

Apply to specific hazards

The above indexes are generic in the sense that they are applicable to a range of hazards. Future research, however, should apply this conceptual model to specific hazard-scenarios. Doing so would entail identifying those specific functions that are required for the event. Application of the above method would also entail a more comprehensive conceptual model of vulnerability one that considers not only social vulnerability but also other factors that may influence outcomes including exposure and susceptibility.

Validate results

As mentioned above the results of the function-based approach differ substantially from four local scale vulnerability indexes and these indexes, in turn, differ from each other. These differences make it difficult for emergency managers to know which index or measure to use in the case of an emergency. One way to assess which measure or index best captures vulnerability is to compare the indexes against health outcomes from past events. One of the challenges to such a study, however, is the need to identify which health outcomes are associated with the event and which are part of day-to-day operations. A second potential issue is that vulnerability may manifest in several different outcomes ranging from fatalities to a visit to a hospital or a physician down to outcomes that do not require medical attention. A third concern is that there may be insufficient data from past events, or a small number of events, to accurately assess the spatial distribution of outcomes.

Create a more comprehensive method for associating characteristics with functions

Another possible future research endeavor is to create a more robust method for associating each characteristic with a function. One possibility is to consult the peer review literature for associations. These studies, however, may exclude important population groups such as those with an ambulatory disability because those data were not readily available when the study was conducted. An alternative is to interview or survey experts in the field. While we took this approach in the above study, these assessments would benefit from both a more formal process as well as broader breath of respondents including, but not limited to, health care professionals, public health officials, and advocates for traditionally vulnerable groups including those with disabilities, immigrants, and the homeless.

Conclusions

In the above report we present the following:

- A function-based conceptual model of social vulnerability
- A method for implementing this model
- Results of this process for the New York City metropolitan area including rate and density measures for transportation, communication, and self-care vulnerabilities as well as an overall measure of vulnerability

By creating the ability to target those most vulnerable, identifying the cause of the vulnerability, and measuring the concurrence of multiple attributes, the proposed process addresses several existing limitations of existing social vulnerability indexes. The proposed measures also have the advantage of being in compliance with several Federal directives including those from the Department of Health and Human Services, the Department of Homeland Security and Association of State and Tribal Health Organizations.

We have found that the above measures closely correlate with population densities. This finding is hardly surprising given that one is most likely to find those who are vulnerable where there are the most people. For instance, upper Manhattan, including the Upper West Side, the Upper East Side, and East and West Harlem consistently have the highest densities of vulnerable populations in the area, as does the Chinatown and the Lower East Side despite the fact that the proportion of those who are vulnerable is not that high. Conversely, urban areas in northern New Jersey including Hoboken, Secaucus, Jersey City, and Clifton have very high rates but lower densities. There are, however, a few areas which have both high rates and high densities. The South Bronx in particular is an area of particular concern, as is south Brooklyn particularly in Coney Island, and, to a lesser extent, Central Brooklyn. We believe these are the areas where the need for vulnerability related resources will be most pronounced.

There are, of course, several limitations to these findings. The measures themselves are not hazard specific and may therefore exclude important groups or functions. Further, the methods for associating hazards to population group still require additional refinement. As is always the case when relying on national scale data, we are limited in selecting population characteristics and therefore must exclude important groups, like those who are immuno compromised, from our analysis. Finally, we do not know how well the above results relate to outcomes from past events.

Appendix A – Pareto Ranking Methods

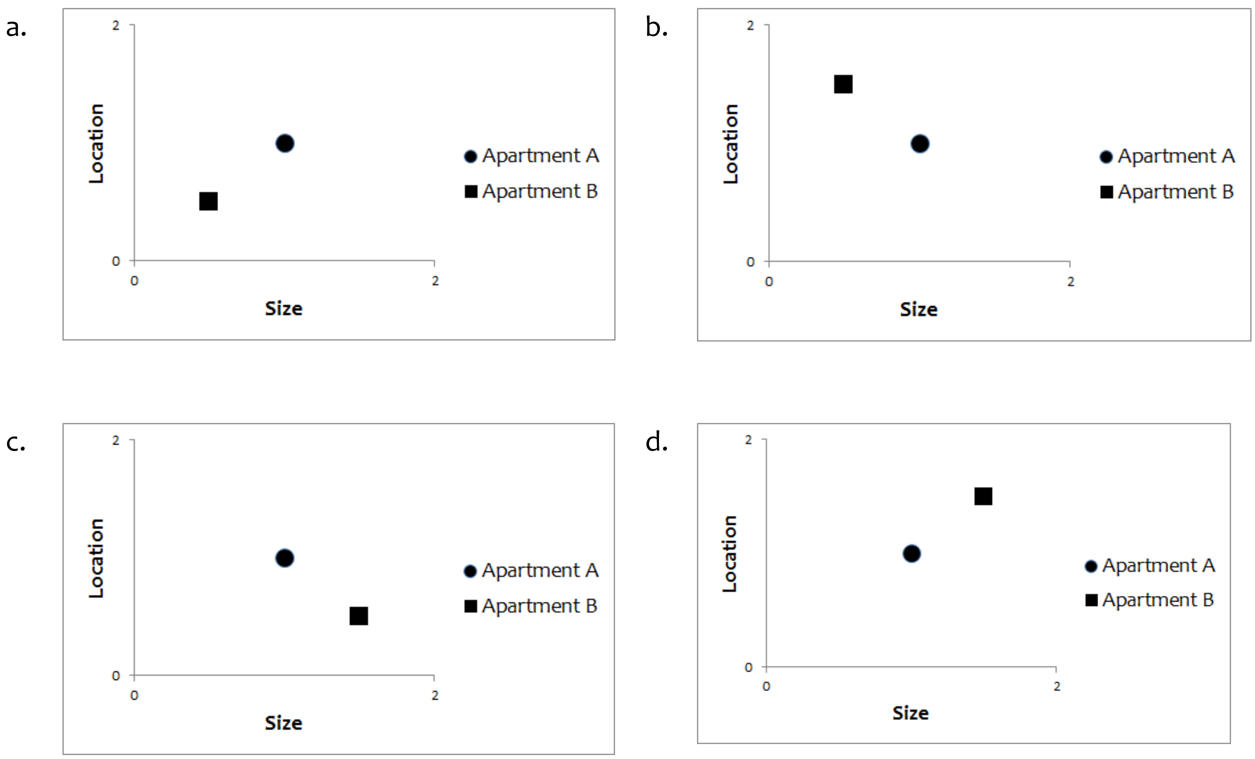
Pareto ranking methods are widely used method for simultaneously optimizing two or more conflicting objectives also known as multi-criteria optimization. In a multi-criteria optimization problem, several different observations are each given a fitness score for one or more criteria. To illustrate this problem consider a person searching for an apartment. There are several different apartments (observations) and the apartment seeker evaluates each alternative based on several different criteria including size, cost, location, amenities, and condition. The apartment seeker can then assign a score for each apartment for each criteria. Collectively, the scores for any one criteria are called a fitness vector.

Pareto ranking is built upon the concept of Pareto Optimization. To illustrate the concept of Pareto Optimization with the above example, consider comparing two apartments (apartment A and apartment B) using just two criteria: cost and location. Each apartment is assigned a score for cost and location with higher values indicating a greater preference. In this scenario, the apartment hunter will select apartment A if it is both larger and in a better location. In this case, apartment A dominates apartment B. Conversely, the apartment hunter will select apartment B if it is both larger and in a better location. In this case, apartment B dominates apartment A.

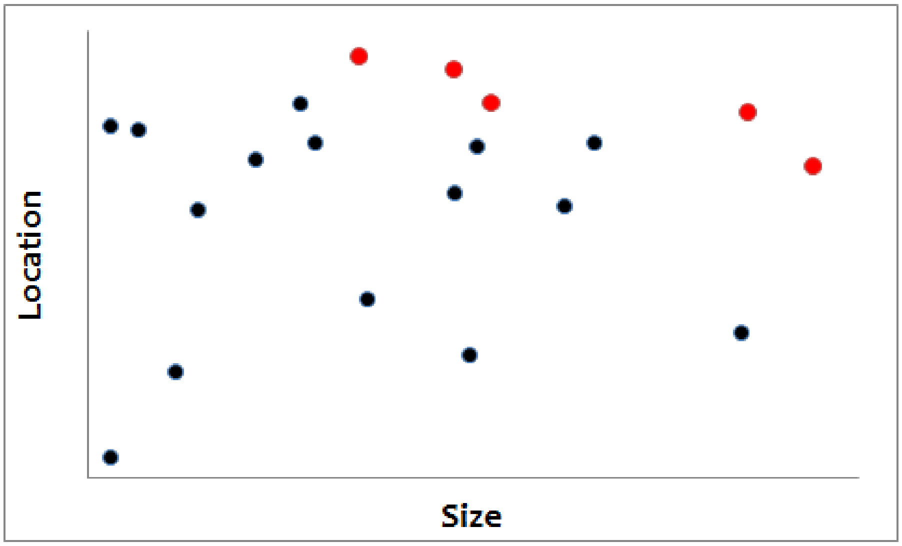
If, however, apartment A is larger but in a worse location,, the choice depends on how the apartment hunter weights these criteria. If location is more important than size, the apartment seeker will select apartment B. If size is more important than location, the apartment seeker will select A. Similarly if

	$LOCATION_A > LOCATION_B$	$LOCATION_A < LOCATION_B$
$SIZE_A > SIZE_B$	Select apartment A	*Uncertain*
$SIZE_A < SIZE_B$	*Uncertain*	Select apartment B

Figure A1a-d expresses these relationships graphically. In each figure apartments A and B are graphed according to their fitness to the two criteria. In figure a, the apartment seeker will select apartment A and in figure d the apartment seeker will select apartment B. In both figures the dominant apartment is above and to the right of the dominated apartment. In figures a1c and a1b, there are no points above and to the right of each of the points, and they are therefore said to be non-dominated.



In the above example, we only consider two apartments. Figure A2 shows the same logic extended to 20 apartments, each plotted based on location and size. Those marked in red are non-dominated because there are no apartments to the above or to the left. Regardless of the weighting scheme used one of these apartments will be most preferable, and they form what is called the "Pareto Frontier." Conversely, each of the points in black is dominated by one or more points along the frontier.



In a Pareto Ranking algorithm each of the points along the Pareto frontier is assigned the highest or most preferable value. These points are then removed from the data set and a new Pareto frontier is identified. Each of the points along this second frontier is assigned the second highest fitness value. This process continues until all points have been assigned a value. This logic can be extended to incorporate many more fitness vectors. Using the above example, the apartment seeker may also want to include criteria for condition, price, and amenities. No matter how many criteria, the logic remains the same.

sa

Appendix B – Confidence Intervals for Population Sizes for Function-Based Scores

Score	Low Estimate	Estimate	High Estimate
21	0.00	0.00	0.00
20	0.01	0.01	0.02
19	0.02	0.02	0.03
18	0.08	0.10	0.12
17	0.17	0.21	0.24
16	0.35	0.40	0.45
15	0.60	0.67	0.74
14	0.93	1.02	1.11
13	1.36	1.48	1.59
12	1.89	2.03	2.17
11	2.54	2.71	2.88
10	3.45	3.65	3.85
9	4.93	5.17	5.40
8	7.72	7.99	8.24
7	11.59	11.91	12.22
6	23.04	23.37	23.69
5	33.92	34.25	34.56
4	64.64	64.78	64.91
3	72.14	72.28	72.43
2	99.93	99.92	99.91
1	100.00	100.00	100.00

Table B1. 90% confidence intervals for percent of population with an overall vulnerability score equal to or greater than the specified value

Score	Low Estimate	Estimate	High Estimate
5	0.00	0.00	0.00
4	0.09	0.10	0.11
3	0.59	0.63	0.67
2	2.31	2.41	2.51
1	13.93	14.13	14.32
0	100.00	100.00	100.00

Table B2. 90% confidence intervals for percent of population with a communication vulnerability score equal to or greater than the specified value

Score	Low Estimate	Estimate	High Estimate
8	0.00	0.00	0.01
7	0.09	0.10	0.12
6	0.52	0.56	0.60
5	1.47	1.54	1.62
4	3.31	3.44	3.56
3	8.26	8.45	8.63
2	27.13	27.34	27.54
1	68.30	68.37	68.43
0	100.00	100.00	100.00

Table B3. 90% confidence intervals for percent of population with a transportation vulnerability score equal to or greater than the specified value

Score	Low Estimate	Estimate	High Estimate
9	0.00	0.00	0.00
8	0.04	0.05	0.06
7	0.46	0.49	0.53
6	1.48	1.55	1.62
5	3.10	3.22	3.33
4	7.89	8.07	8.24
3	25.04	25.24	25.45
2	66.32	66.40	66.47
1	99.89	99.87	99.86
0	100.00	100.00	100.00

Table B4. 90% confidence intervals for percent of population with a self-care vulnerability score equal to or greater than the specified value

Appendix C – Spatial Distribution of Function-Based Measures

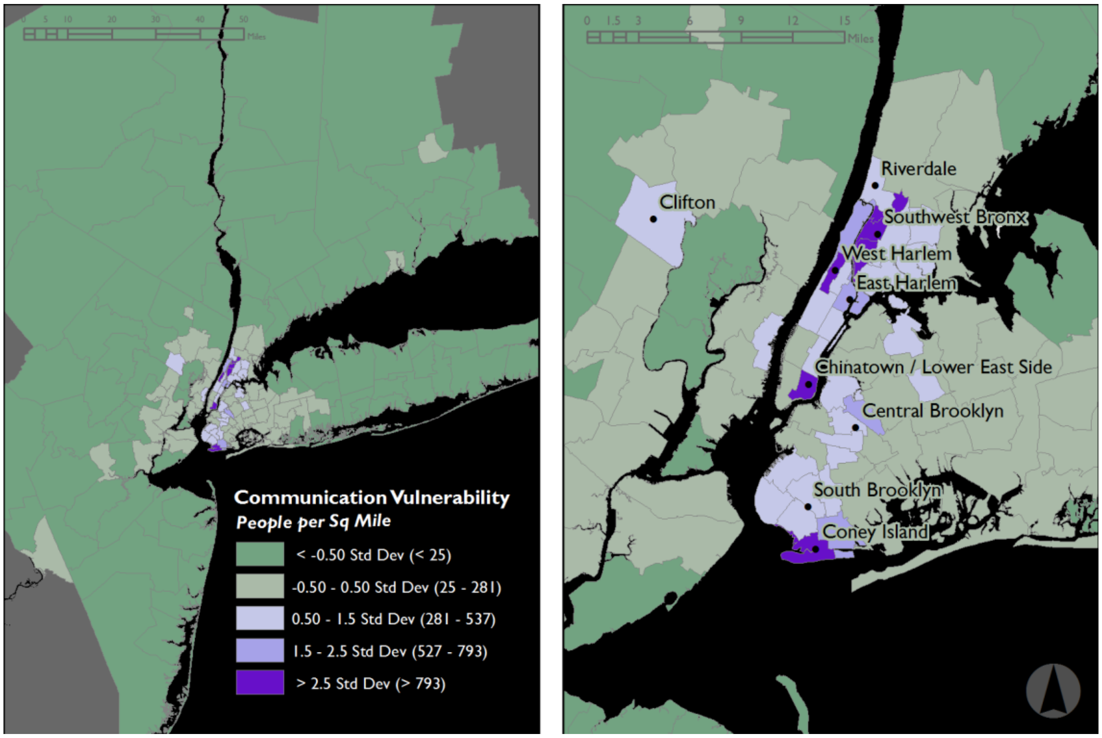


Figure C1a-b. Communication vulnerability density

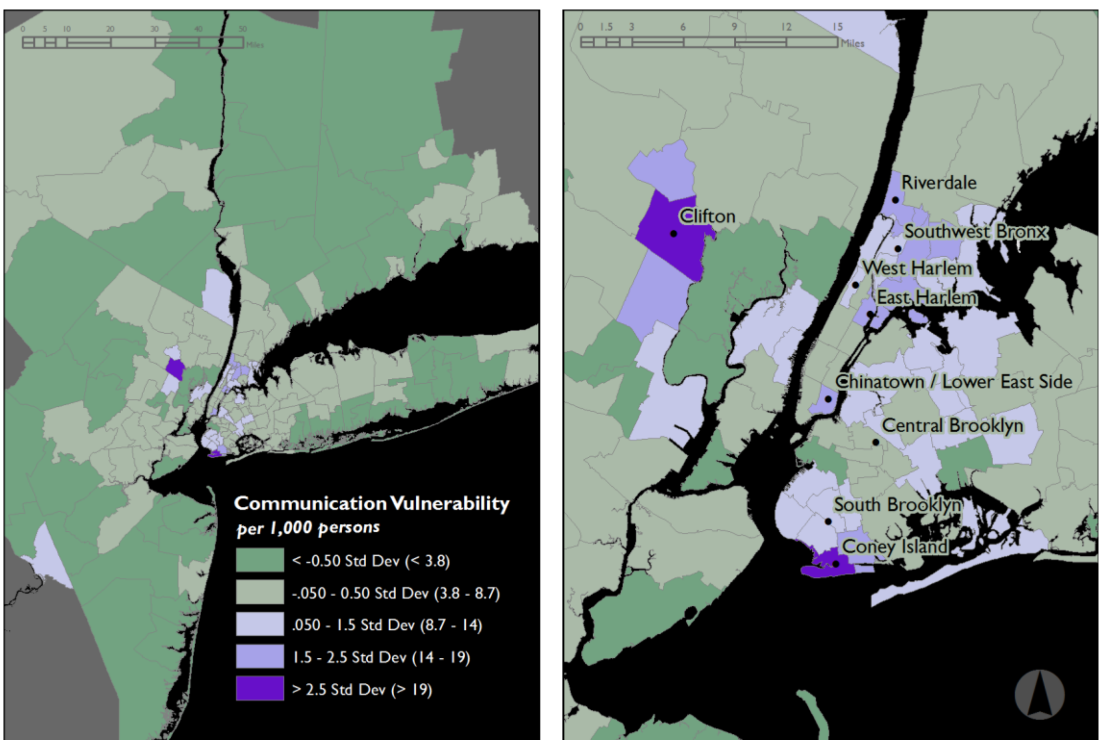


Figure C2a-b. Communication vulnerability rate

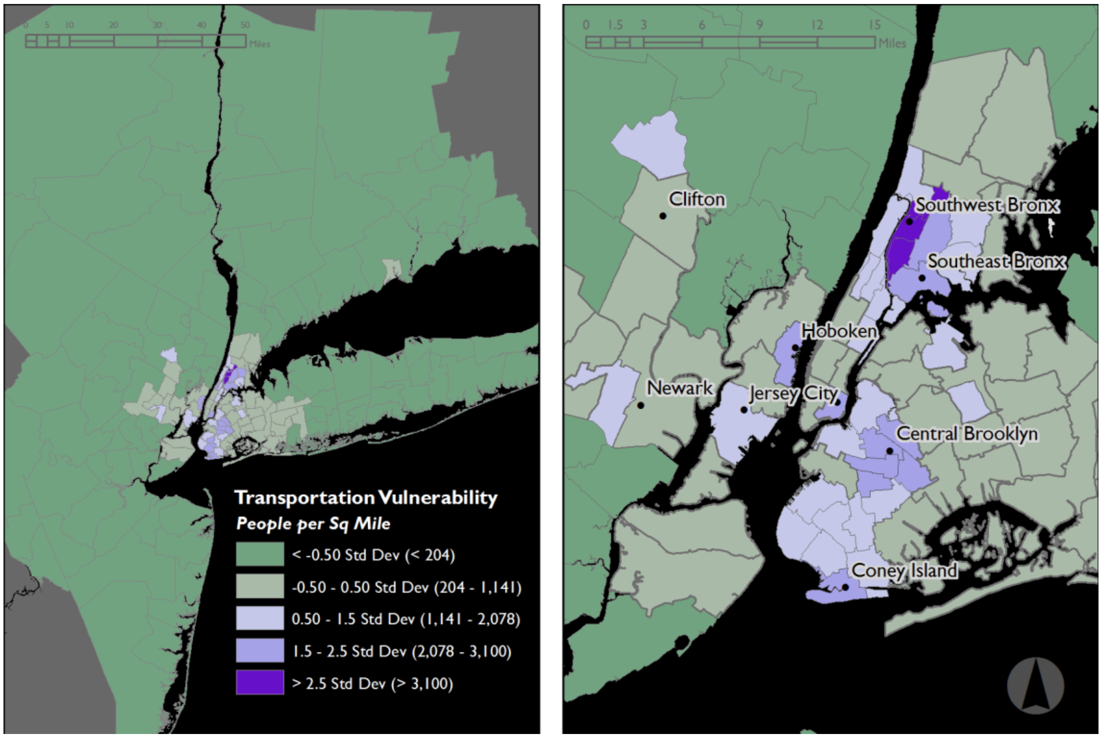


Figure C3a-b. Transportation vulnerability density

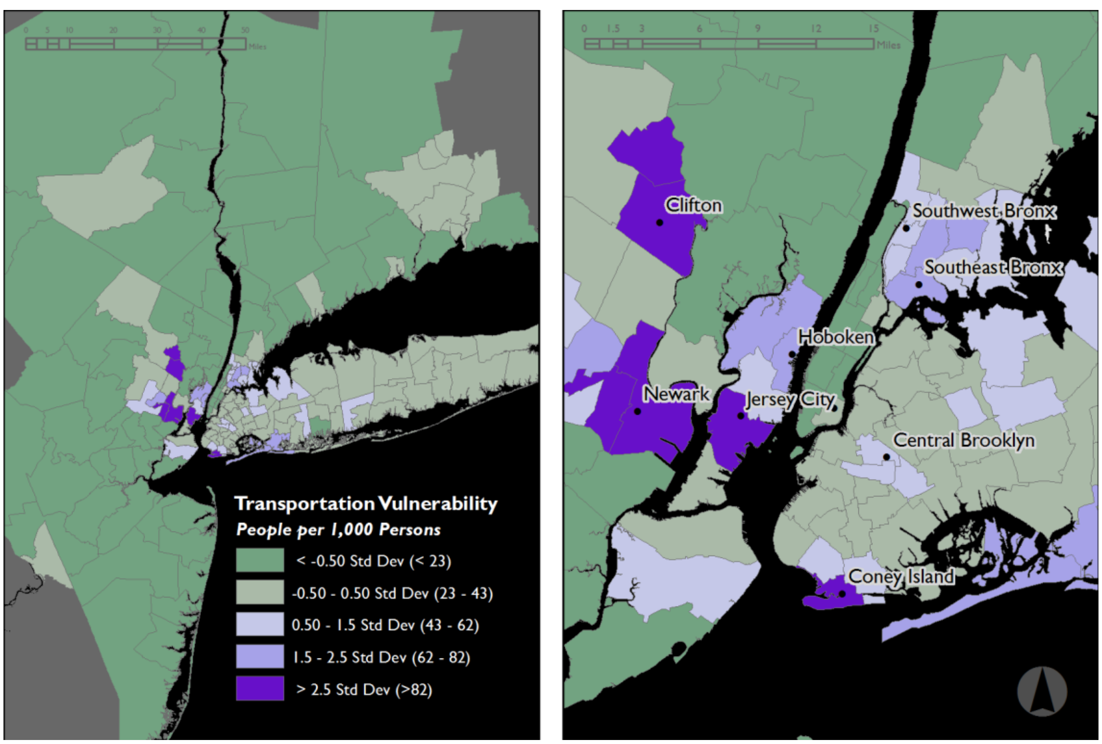


Figure C4a-b. Transportation vulnerability rate

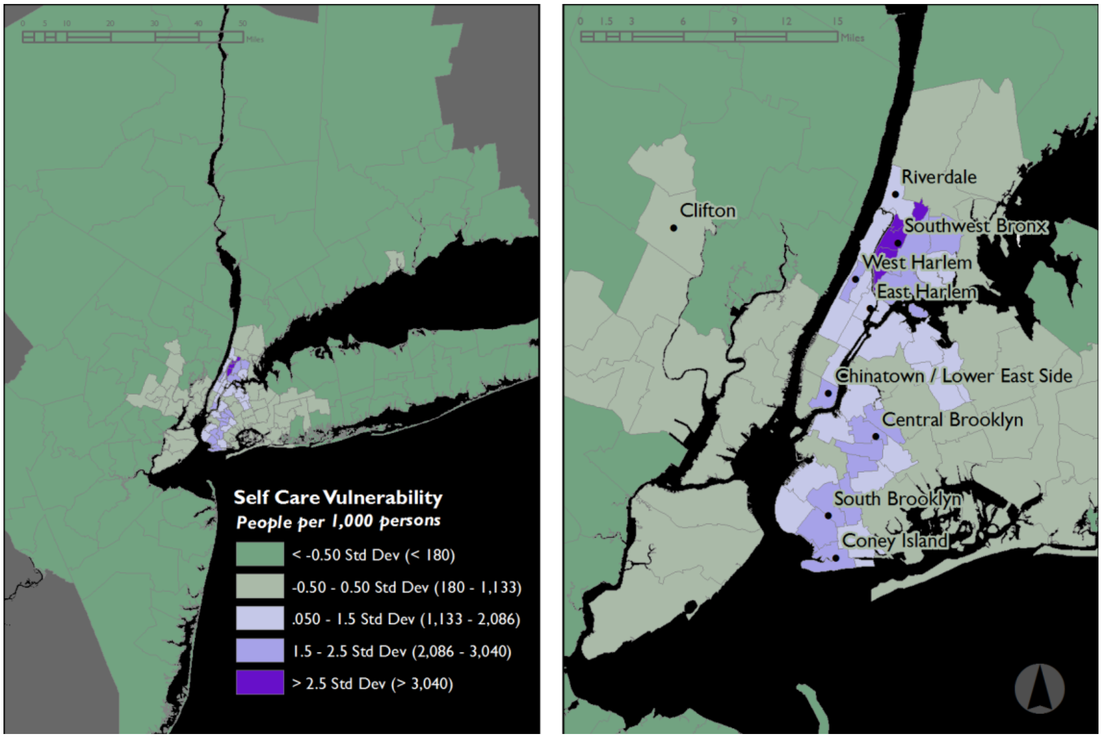


Figure C5a-b. Self care vulnerability density

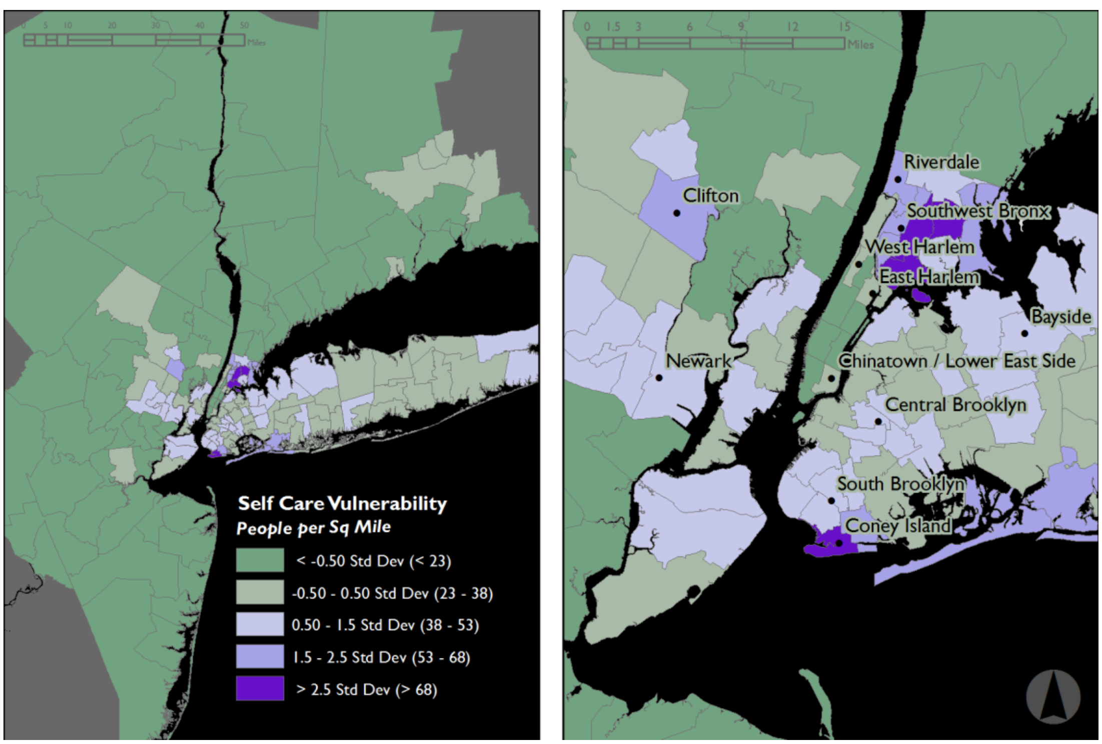


Figure C6a-b. Self care vulnerability rate

Appendix D – Local Scale Vulnerability Indices

D1. Cutter 03

Data

Table A1.1 the variables and data sources for the variables used in this index 39 of the 42 variables listed in Cutter and Bornuff (2003). Data for three variables – net international migration, percent rural farm population, and percent urban population – were only available for the 2000 decennial census and were thus excluded from analysis.

Name	Description	Source
MED_AGE	Median Age	ACS 08-10
PERCAP	Per Capita Income	ACS 08-10
MVALOO	Median Value owner-occupied housing units	ACS 08-10
MEDRENT	Median Rent	ACS 08-10
PHYSICN	Number of physicians for 100,000 population	ACS 08-10
PCTVOTE	Vote cast for president, percent voting for leading party	http://www.usatoday.com/news/politics/election2008/president.htm
BRATE	Birth Rate (per 1,000 population)	ACS 08-10
MIGRA	Net International Migration	
PCTFARMS	Land in Farms as Percent of total land	Agricultural Census 2007
PCTBLACK	Percent African American	ACS 08-10
PCTINDIAN	Percent Native American	ACS 08-10
PCTASIAN	Percent Asian	ACS 08-10
PCTHISPANIC	Percent Hispanic	ACS 08-10
PCTKIDS	Percent Population Under 5	ACS 08-10
PCTOLD	Percent Population Over 65	ACS 08-10
PCTUVLUN	Percent Civilian Labor Force Unemployed	ACS 08-10
AVGPERHH	Average Number of People in a Household	ACS 08-10
PCTHH75	Percent of houshelod earning more than $75,000	ACS 08-10
PCTPOV	Percent living in poverty	ACS 08-10
PCTRENTER	Percent Reneter occupied housing units	ACS 08-10
PCTRFRM	Percent Rural Farm Population	
DEBREV	General local government debt to revenue	State and County Survey of Local Government Finance 2007
PCTIMOBL	Percent households that are mobile homes	ACS 08-10
PCTNONHS	Percent population 25 or older with no high school diploma	ACS 08-10

HODENUT	Number of housing units per square mile	ACS 08-10
HUPTDEN	Number of housing permits per new residential construction per square mile	2010 Housing Permit Survey
MAESDEN	Number of manufacturing establishments per square mile	Economic Census 2007
EARNDEN	Eranings (in $1,000) all industries per square mile	Economic Census 2007
COMDEVDN	Number of commercial establishments per square mile	Economic Census 2007
RPROPDEN	Value of all property and farm products sold per square mile	Agricultural Census 2007
CVBRPC	Percent of the population participating in the labor force	ACS 08-10
FEMLBR	Percent females participating in the civilian labor force	ACS 08-10
AGRIPC	Percent employed in primary extractive industries (farming, fishing, mining, and forestry)	ACS 08-10
TRANPC	Percent employed in transportation, communications, and other public utiliites	ACS 08-10
SERVPC	Percent employed in service occupations	ACS 08-10
NRRESPC	Per capita residents in nursing homes	2010 Census, Summarized at County Level
HOSPTPC	Per capita number of community hospitals	CT DEP, PA Dept of Health and Hospitals, NJ Office of Information Technology, HANYS
PCCHGPOP	Percent population change	ACS 05, ACS 10
PCTURB	Percent urban population	
PCTFEM	Percent females	ACS 08-10
PCFTD_HH	Percent female headed households	ACS 08-10
SSBENPC	Per capita social security recipients	ACS 08-10

Table A1.1 Variables for Cutter 03 Index

In many cases data were only available at the county scale. As all of these variables were a rate (either by area or by population), those PUMAs entirely contained by a county were assigned the value of the county that contain them. In some instances PUMAs extend across county lines. In these cases, we estimated PUMA values by taking a weighted average of the counties the PUMA intersects. Depending on the variable, weights were equal to either the percent of the PUMA's population (A1.1) or area (A1.2) in each of the intersecting counties.

$$\bar{V}_i = \sum_{c=1}^{n} \frac{V_c(POP_i \cap POP_c)}{POP_i} \quad \textbf{(A1.1)}$$

$$\bar{V}_i = \sum_{c=1}^{n} \frac{V_c(AREA_i \cap AREA_c)}{AREA_i} \quad \textbf{(A1.2)}$$

Where V_i is the estimated value for variable *V* in Puma *i,* n is the number of counties, V_C is the value for variable V in county c, $POP_i \cap POP_c$ is the population of Puma i in county c, $AREA_i \cap AREA_c$ is the land area of county c in Puma i, POP_i is the population of PUMA i and AREA i is the land area of PUMA i.

Method

We used Principle Component Analysis (PCA) with varimax rotation to compress the above variables into a fixed number of factors. We then calculated a vulnerability score for each PUMA by adding the factor scores of those factors that clearly correlate with increased vulnerability, subtracting the scores of those factors that inversely correlate with increased vulnerability, and adding the absolute value of the factor scores that have an ambiguous relationship with vulnerability.

Factor Analysis Results

Based on the factor analysis we identified 11 factors that explain 86% of the variability. It was difficult to determin the appropriate number of factors to retain. Ultimately, the decision was made based on the final groupings and scores. Cutter and Boruff's nationwide, county-based analysis also included 11 factors. Narrative descriptions of the factors are included below.

1. **Dense Businesses**: Positive correlation with a high density of commercial and manufacturing firms firm earnings, and the number of housing units per square mile. Also correlates with percentage of votes for Obama, and has a negative correlation with percent of social security recipients. Action: add absolute value.

2. **Wealthy**: High correlation with per capita income, percentage of physicians, high rents and high house value, percentage of households earning more than $75,000/ yr, and inverse correlation with percent unemployed. Action: subtract.

3. **Youth**: High correlation with birthrate, percent under 5, percent in poverty. Negative correlations with median age, and elderly. Action: add absolute value.

4. **Employed:** Correlates with percent of the population participating in the labor force and percent of females in the labor force. Action: subtract.

5. **Farmland**: Correlates with value of all farm products sold, and percent of land in farms.

   Action: add.
6. **Women:** Correlates with the percentage of women and the percentage black residents. Action: add.

7. **Dense Housing:** Correlates with the average number of people per household and the percent of female headed households. Action: add.

8. **Not Asian:** Negative correlation with percent Asian and permits for new construction. Action: subtract.

9. **Nursing Homes:** Positive correlation with the percent of people in nursing facilities. Action: Add

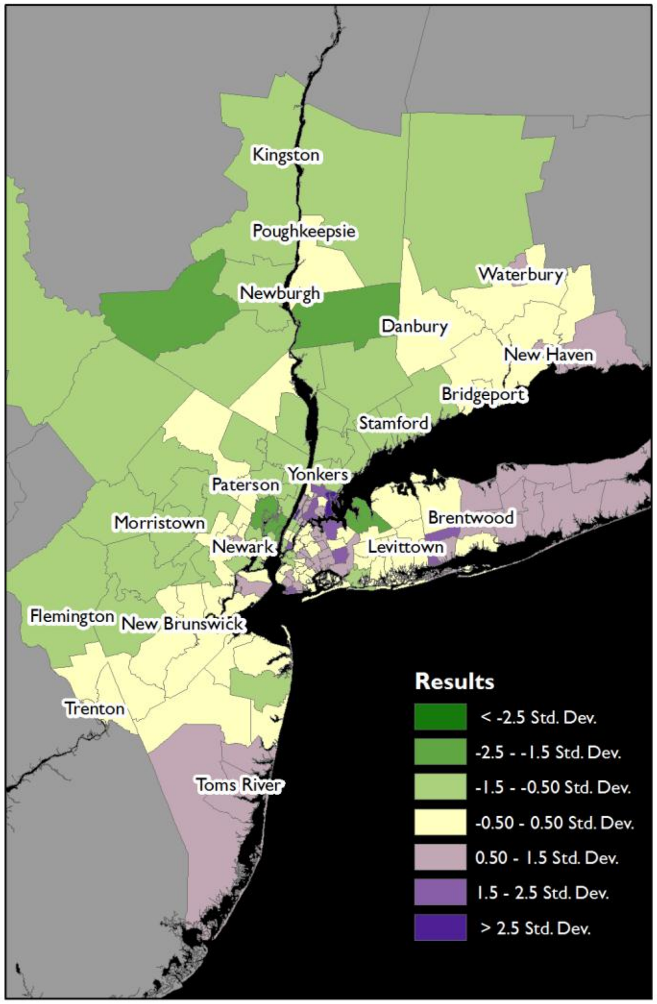
10. **Native Americans:** Positive correlation with the percentage of native americans. Action: add.

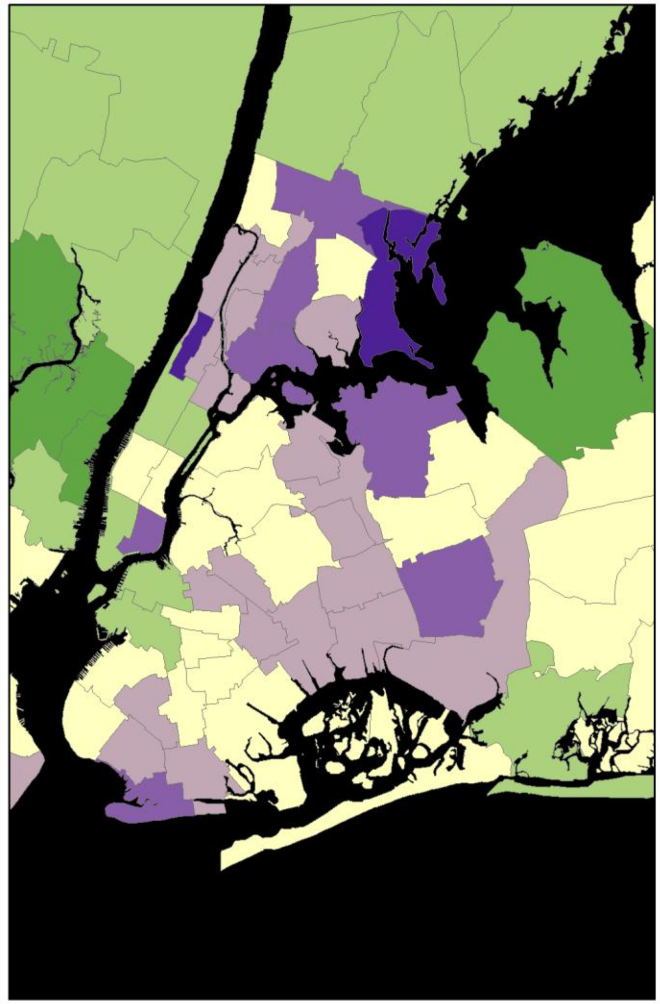
11. **Shrinking:** Negative correlation with the percent population increase. Action: add absolute value.

Map Results

Figures A1.1a and A1.1b show the index results for the entire region and the city. Regionally there are four clusters of high vulnerability: central New Jersey, near New Haven in Connecticut, New York City, and Suffolk County. Elevated vulnerability in central New Jersey can be attributed to a high value for factors 3 and 4 (youth and employed), New Haven can similarly be attributed to a high values for factors 3 and 7 (Youth and Dense Housing), and Suffolk County to factor 5 (farmland).

Many areas in New York City were outliers particularly across the Bronx, western Queens, south Brooklyn, and Northern Manhattan. Almost every outlier had high values for factors 2,3, and 4 (wealth, youth, and employed), with high values in other factors scattered throughout including factor 11 (shrinking) in Coney Island and western Queens, factor (6) women in western queens, factor 10 (native americans) in bayside, queens, and Pelham Bay, Bronx, Nursing Homes in Pelham Bay, and business density, in northern Manhattan.


Kingston
Poughkeepsie
Newburgh
Waterbury
Danbury
New Haven
Bridgeport
Stamford
Yonkers
Paterson
Morristown
Brentwood
Newark
Levittown
Flemington
New Brunswick
Trenton
Toms River
Results
< -2.5 Std. Dev.
-2.5 - -1.5 Std. Dev.
-1.5 - -0.50 Std. Dev.
-0.50 - 0.50 Std. Dev.
0.50 - 1.5 Std. Dev.
1.5 - 2.5 Std. Dev.
> 2.5 Std. Dev.



A2. Rygel

Data

Table A2.1 shows the variables used for the Rygel index. Data for all variables came from the 2010 American Community Survey 3 year public micordata use sample, with the exception of nursing home populations which came from the 2010 census. The Rygel index uses disability indicators from the 2000 decennial census. In 2008, the census bureau changed these questions. As best as possible we matched 2010 ACS disability variables with their 2000 counterparts. It should be noted, however, that the census bureau does not support comparison between pre and post 2008 values. In one case – work disability – there was no post 2008 counterpart. This variable was therefore left out of analysis.

Variable Name	Description
PC_YOUNG	Percentage of total population that is 17 years of age or younger
YOUNG_DENS	Number of people, per square kilometer, who are 17 years of age or younger
PC_OLD	Percentage of total population that is 65 years of age or older
OLD_DENS	Number of people, per square kilometer, who are 65 years of age or older
PC_FEMALE	Percent of total population that is female
FEMALE_DENS	Number of females, per square kilometer
PC_BLACK	Percent of total population that is black or African-American
BLACK_DENS	Number of people, per square kilometer, who are black or African-American
PC_AMERINDIAN	Percent of total population that is American Indian or Alaska native
AMERINDIAN_DENS	Number of people, per square kilometer, who are American Indian or Alaska native
PC_ASIAN	Percent of total population that is Asian
ASIAN_DENS	Number of people, per square kilometer, who are Asian
PC_HAWAII	Percent of total population that is Native Hawaiian or other Pacific Islander
HAWAII_DENS	Number of people, per square kilometer, who are Native Hawaiian or other Pacific Islander
PC_OTHER	Percent of total population that is some other (non-white) race
OTHER_DENS	Number of people, per square kilometer, who some other (non-white) race
PC_GTE2RACE	Percent of total population that belongs to two or more races

GTE2RACE_DENS	Number of people, per square kilometer, who belong to two or more races
PC_NEWIMMIG	Percent of total population that is foreign-born and has entered the United States between 1995 and 2000
NEWIMMIG_DENS	Number of people, per square kilometer, who are foreign-born and have entered the United States between 1995 and 2000
PC_NOENGLISH	Percent of population (age 5 and over) that speaks English "not well" or "not at all"
NOENGLISH_DENS	Number of people (age 5 and over), per square kilometer, who speak English "not well" or "not at all"
PC_NURSEHOME	Percent of total population that resides in nursing homes
NURSEHOME_DENS	Number of people, per square kilometer, who reside in nursing homes
PC_SENSDISABIL	Percent of population (age 5 and over) with a sensory disability
SENSDISABIL_DENS	Number of people (age 5 and over), per square kilometer, with a sensory disability
PC_PHYSDISABIL	Percent of population (age 5 and over) with a physical disability
PHYSDISABIL_DENS	Number of people (age 5 and over), per square kilometer, with a physical disability
PC_MENTDISABIL	Percent of population (age 5 and over) with a mental disability
MENTDISABIL_DENS	Number of people (age 5 and over), per square kilometer, with a mental disability
PC_SLFCRDISABIL	Percent of population (age 5 and over) with a self-care disability
SLFCRDISABIL_DENS	Number of people (age 5 and over), per square kilometer, with a self-care disability
PC_HOMEDISABIL	Percent of population (age 16 and over) with a go-outside-home disability
HOMEDISABIL_DENS	Number of people (age 16 and over), per square kilometer, with a go-outside-home disability
PC_EMPLDISABIL	Percent of population (ages 16-64) with an employment disability
EMPLDISABIL_DENS	Number of people (ages 16-64), per square kilometer, with an employment disability
PC_NODIP	Percent of adults (age 25 and over) with no high school diploma
NODIP_DENS	Number of adults (age 25 and over), per square kilometer, with no high school diploma
PC_POVERTY	Percent of population (for whom poverty status has been determined) living below the poverty line
POVERTY_DENS	Number of people (for whom poverty status has been determined), per square kilometer, living below the poverty line
PC_PUBTRANS	Percent of workers (age 16 and over) that rely on public transportation to get to work
PUBTRANS_DENS	Number of workers (age 16 and over), per square kilometer, who rely on public transportation to get to work

PC_SINGLMOMS	Percent of households composed of a female head-of-household and children under age 18
SINGLMOMS_DENS	Number of households, per square kilometer, composed of a female head-of-household and children under age 18
PC_RENT	Percent of occupied housing units that are renter-occupied
RENT_DENS	Number of occupied housing units, per square kilometer, that are renter-occupied
PC_MOBLHOME	Percent of occupied housing units that are mobile homes
MOBLHOME_DENS	Number of occupied housing units, per square kilometer, that are mobile homes
PC_NOPHONE	Percent of occupied housing units with no telephone service
NOPHONE_DENS	Number of housing units, per square kilometer, with no telephone service
PC_NOCAR	Percent of occupied housing units with no vehicle available
NOCAR_DENS	Number of occupied housing units, per square kilometer, with no vehicle available
POP_DENSITY	Total number of people, per square kilometer
PERCAPITA	Per capita income (in dollars), 1999
MED_EARN	Median earnings (in dollars) (for the population with earnings, age 16 and over), 1999
MEDHS_INCOME	Median household income (in dollars), 1999
MEDHS_VALUE	Median value (in dollars) of specified owner-occupied housing units

Table A2.1 Variables for Rygel index

Method

Rygel proposes a two step method for index creation. In the first step, the analyst uses factor analysis to compress the above 56 variables into a smaller number of factors. In the second step, the analyst uses a pareto ranking method to rank the factor scores for each PUMA from step 1. Briefly, pareto ranking is a genetic multi-criteria optimization (GMCO) algorithm. Under a Pareto Ranking, that no item with can be the most preferable relative to all those items with a higher rank. If, for example, an item has a rank of 7 it can be asserted that at least one item with a rank of 8 will be more preferable regardless of the weighting scheme. It may, however, be possible to devise a weighting scheme whereby an item 7 could be more preferable to an item with rank 8. Pareto ranking is therefore advanced as an alternative to generating often subjective factor weightings.

Results

Factor Analysis Results

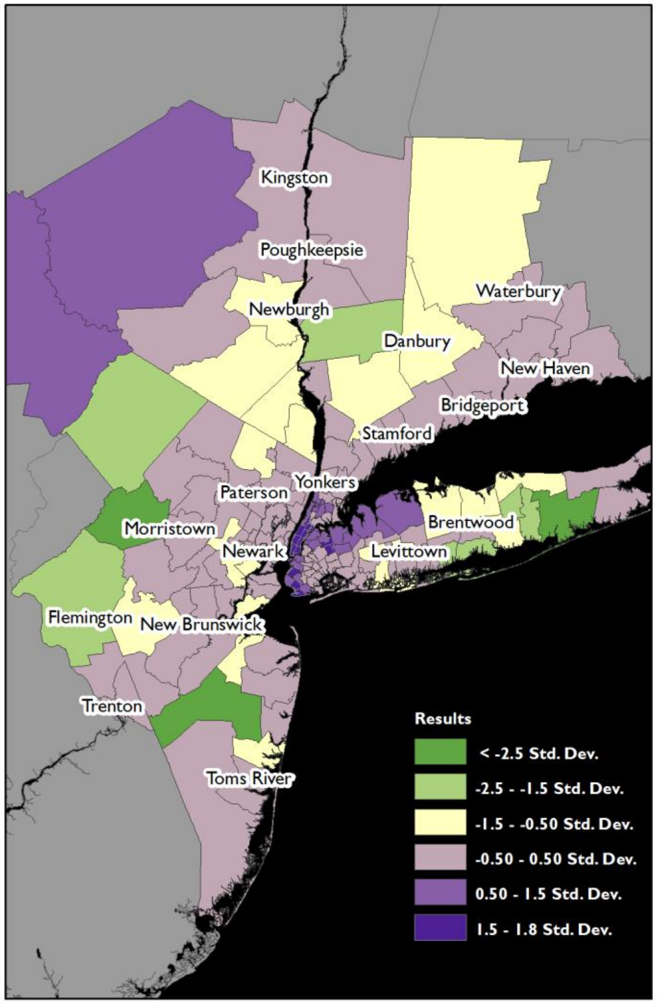
The above variables were compressed into four factors that cumulatively explained 75% of the variance in the underlying data set. The four factors could be described as follows:

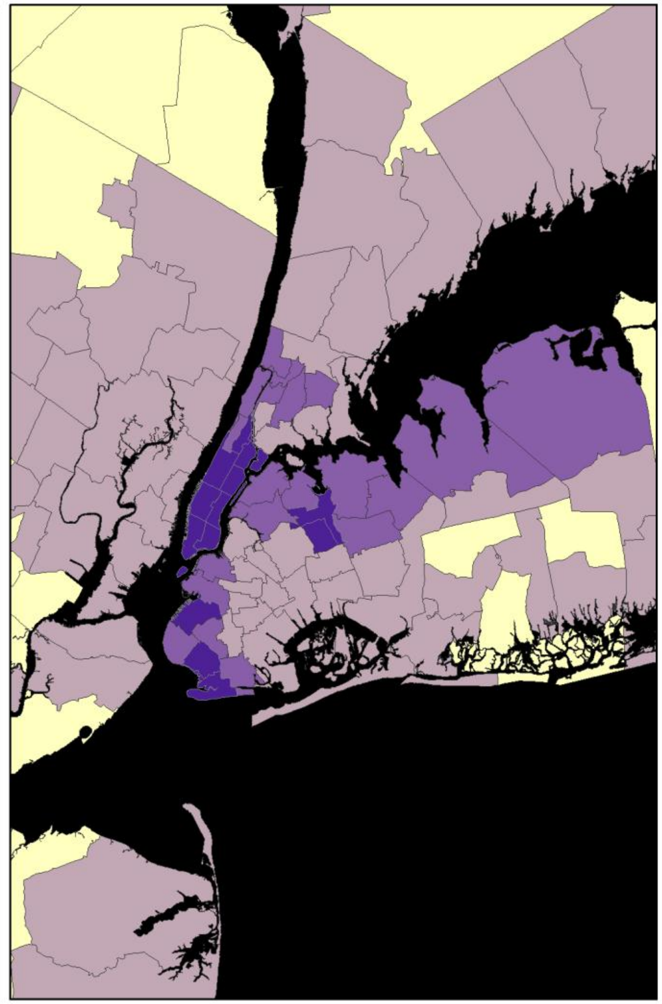
1. **High Density** – correlates with almost all density variables as well as the percentage of renters, the percentage who take public transit to work, and inversely correlates with car ownership.
2. **Wealth** – Correlates with poverty and youth and inversely correlates with per capita income, median earnings, and median household income.
3. **Disability** – Correlates with the percentage of all disability variables.
4. **Asian** – Correlates with percent Asian, percent new immigrants, percent who do not speak English well, and density of Asians.

Map Results

Figure A2.1 shows the results of the index mapped for the region and for the New York City area. There are three clusters: northwestern Pennsylvania, New York City, and northern Queens/Suffolk county. The northwester Pennsylvania cluster is clearly driven by high factor 3 (disability) scores. The Queens, Suffolk county, and Brooklyn clusters are driven by high factor 4 (Asian) scores, with the exception of Coney Island which has a very high factor 3 (disability) score. All Manhattan has very high factor 1 (density) scores. There are also a number of areas with elevated values throughout w Jersey and along the Connecticut coast. These areas should be taken with a grain of salt, however, as each index has a score of 1 to 7 with frequencies shown in table A2.2 below.

Rank	**Count**
1	3
2	7
3	22
4	40
5	49
6	22
7	14


Kingston
Poughkeepsie
Newburgh
Waterbury
Danbury
New Haven
Bridgeport
Stamford
Yonkers
Paterson
Brentwood
Morristown
Newark
Levittown
Flemington
New Brunswick
Trenton
Toms River
Results
< -2.5 Std. Dev.
-2.5 - -1.5 Std. Dev.
-1.5 - -0.50 Std. Dev.
-0.50 - 0.50 Std. Dev.
0.50 - 1.5 Std. Dev.
1.5 - 1.8 Std. Dev.



A3. Flanagan 2011

Data

The Flanagan index is based on 15 variables divided into four groups or domains - socioeconomic status, household composition/ disability, minority status/ language, housing/transportation. All data come from the 2010 American Community Survey 3 year public use microdata sample. Variables and domains are shown in table A3.1.

Domain 1. Socioeconomic Status	
POVERTY	Percent Individuals below poverty
UNEMPL	Percent civilian unemployed
INCOME	Per capita income
NOHS	Percent persons with no HS diploma
Domain 2. Household Composition/Disability	
OLD	Percent persons 65 years of age or older
YOUNG	Percent persons 17 years of age or younger
DISABL	Percent persons more than 5 yrs old with a disability
NOSPOUSE	Percent male or female householder with no spouse present with children under 18
Domain 3. Minority Status/Language	
MINORITY	Total of the following:“black or African American alone” +“American Indian and Alaska Nativealone” + “Asian alone” + “NativeHawaiian and other Pacific Islanderalone” + “some other race alone” + “twoor more races” + “Hispanic or Latino – white alone.”
LANGUAGE	Percent persons 5 years of age or older who speak English less than "well"
Domain 4. Housing/Transportation	
MUNIT	Percent housing units with 10 or more units in structure.
MOBILE	Percent housing units that are mobile homes.
CROWD	At household level, more people than rooms. Percent total occupied housing units (i.e., households) with more than one person per room.
NOCAR	Percent households with no vehicle available.
GQ	Percent of persons who are in institutionalized group quarters (e.g., correctional institutions, nursing homes) and non-institutionalized group quarters (e.g., college dormitories, military quarters).

Table A3.1 Variables and domains for Flanagan index

Method

There are four steps to calculate the Flanagan index. In the first step, each of the 15 variables are ranked from lowest to highest (with the exception of per capita income which is ranked from highest to lowest). In the second step a percentile rank was calculated for each variable. The percentile rank is equal to the ratio of the variable's rank less one to the total number of observations less one (eq A3.1). The observation with the highest value for a particular variable will have a percentage rank score of 1, and the observation with the lowest value will have a percentage rank score of 0.

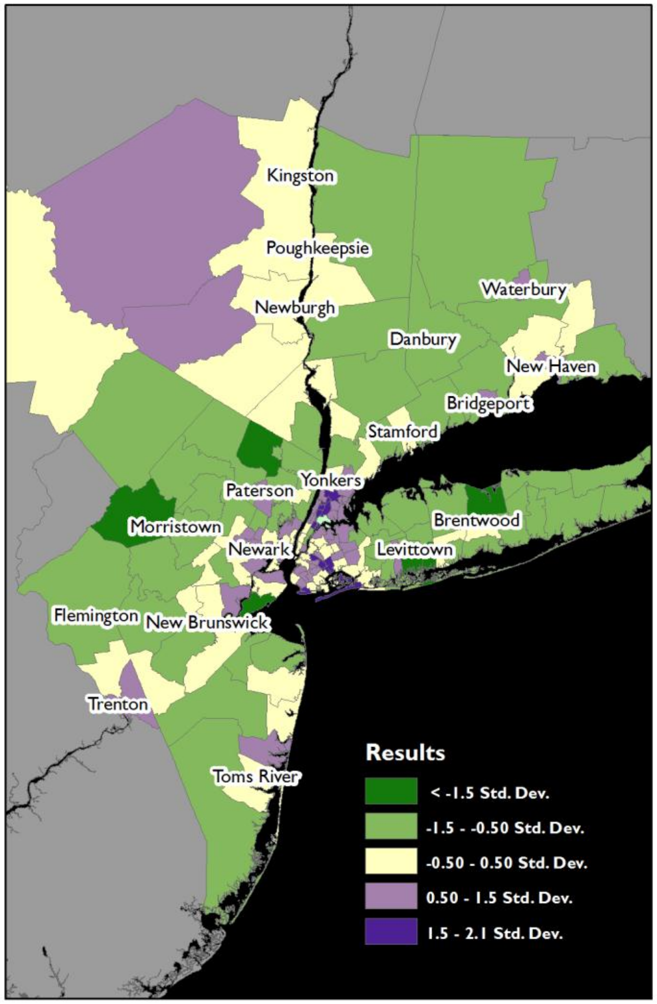
$$Percentile\ Rank_{i,v} = \frac{\left(Rank_{i,v} - 1\right)}{(N - 1)} \qquad \textbf{(A3.1)}$$

Where Percentile $Rank_{i,v}$ is equal to the percentile rank for puma i for variable v, $Rank_{i,v}$ is equal to the rank of puma i for variable v, and N_v is the number of pumas.

After calculating the percentile rank for each variable, the analyst then sums the percentile ranks for each of the variables in each of the four domains. The analyst then calculates percentile ranks for each domain based on the results of step 3. In the fifth and final step, the analyst adds the percentile ranks for each domain (that were calculated in step 4). This value is equal to the overall vulnerability.

Results

Unlike Cutter03 and Rydel, the Flanagan index more closely follows the expected spatial distribution of vulnerability with peaks in major clusters of urban poverty including Newark, Trenton, New Haven, and Patterson. There were also peaks in northwestern New Jersey and near Tom's River. The northwestern New Jersey cluster is attributed to high percentile ranks for domains 2 and 4, while the Tom's River cluster is attributed to high percentile ranks in domains 1 and 2. Within New York City there are elevated clusters in South Bronx, Central Brooklyn, Coney Island and Rockaways.


Kingston
Poughkeepsie
Newburgh
Waterbury
Danbury
New Haven
Bridgeport
Stamford
Yonkers
Paterson
Brentwood
Morristown
Newark
Levittown
Flemington
New Brunswick
Trenton
Toms River
Results
< -1.5 Std. Dev.
-1.5 - -0.50 Std. Dev.
-0.50 - 0.50 Std. Dev.
0.50 - 1.5 Std. Dev.
1.5 - 2.1 Std. Dev.

A4. Cutter 00

Data

Table A4.1 shows the 8 variables that make up the Cutter 00 vulnerability index. All data for these variables came from the 2010 3 year American Community Public Microdata Use Sample.

Population and Structure	
POP	Total Population
HU	Total Housing Units
Differential access to resources / Susceptibility due to physical weakness	
FEM	Number of females
NONWHITE	Number of non-white residents
OLD	Number of people over the age of 65
YOUNG	Number of people under the age of 18
Wealth or Poverty	
HV	Mean house value
Level of Physical or Structural Vulnerability	
MOBILE	Number of Mobile Homes

Table A4.1 Variables for Cutter 00 Index

Method

The Cutter 00 index is comprised of the sum of each of the above variables re-scaled from 0 to 1. To re-scale each variables, with the exception of mean house value, analysts divide the observed value for a variable by the maximum observed value for that variable. The analyst then essentially repeats the above process by dividing the result of the first step for each variable by the maximum from step 1. The resulting value will be between 0 and 1. This process is documented in table A4.2 below.

PUMA	# of Mobile Homes in PUMA	# of Mobile Homes in Study Area	Ratio of PUMA to Study Area (X)	Mobile Home Vulnerability Index (X / maximum X)
A	125	3,500	0.036	1.00
B	76	3,500	0.022	.61
C	4	3,500	0.001	.03
D	21	3,500	0.006	.17

Table A4.2 Example mobile home vulnerability score calculations

For house values, the analyst first subtracts the observed mean house value for each puma from the study area mean.   The result is X. The analyst then calculates Y by adding each value x to the maximum of the absolute value of x. In the final step, the analyst calculates the Mean House Value Vulnerability score by dividing each value y by the maximum value of Y. This process is shown in table A4.3.

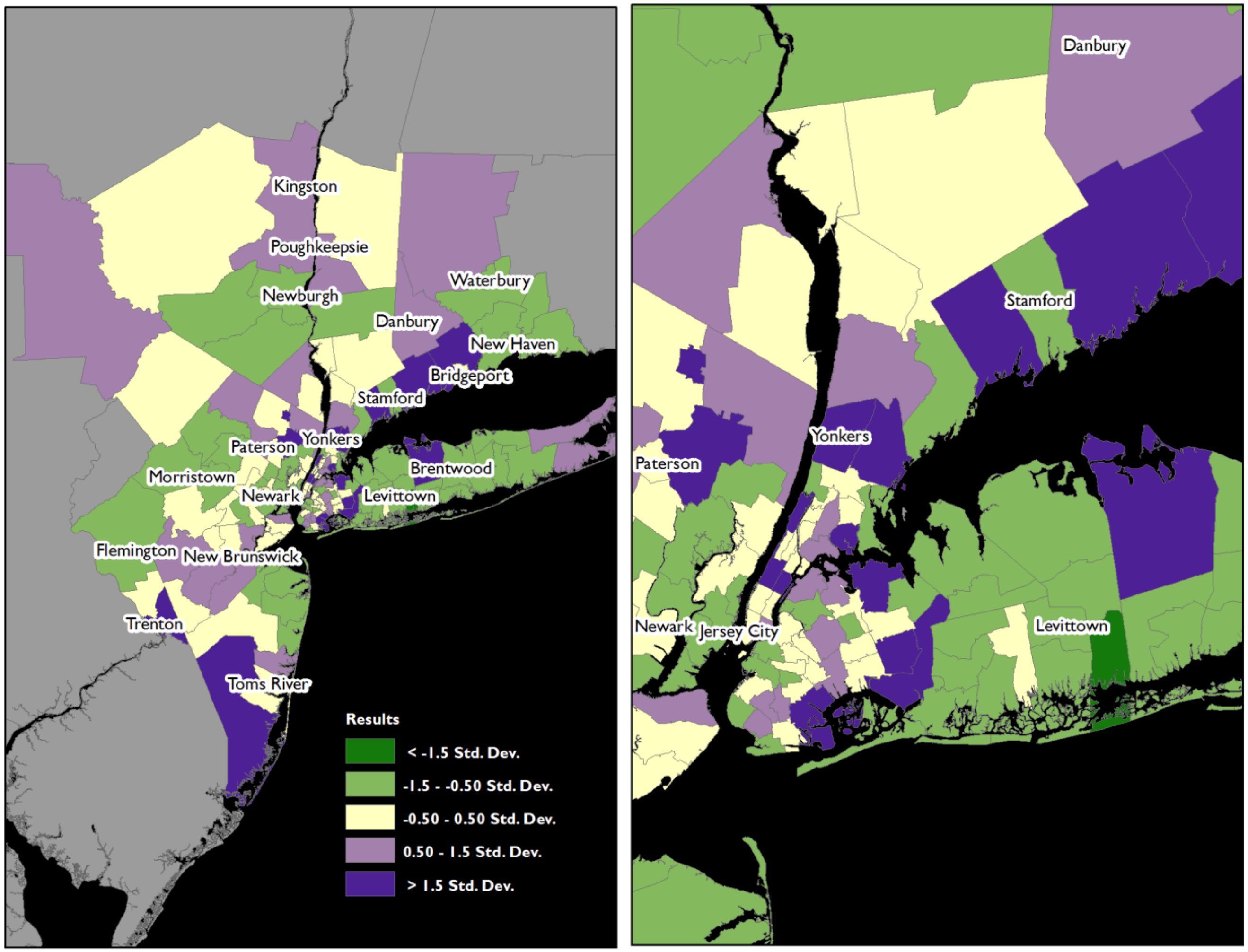
PUMA	Mean House Value in PUMA	Mean House Value in Study Area	Value Difference ($) of County and Block (X)	X + Maximum of Absolute Value of X (Y)	Mean House Value Vulnerability Score (Y / maximum Y)
A	41,286	75,000	33,714	69,364	1.00
B	110,650	75,000	-35,650	0	.00
C	76,776	75,000	-1,776	33,874	.49
D	64,900	75,000	5,100	40,750	.58

Table A4.3 Example mean house value vulnerability score calculation

As a final step, the analyst calculates an overall vulnerability score by summing the eight variables' vulnerability scores.

Results

The results of the Cutter 00 index are shown in figure A4.1. There are four clusters with outlier values: eastern Connecticut from New Haven to Stamford, Southeastern New Jersey, Trenton, and eastern Queens and Brooklyn. The city map also reveals clusters in Yonkers, and Patterson. Within the city there are high values on the upper east side and upper west side, Pelham Bay in the Bronx, Willets point and eastern Brooklyn and eastern Queens.


Kingston
Poughkeepsie
Newburgh
Waterbury
Danbury
New Haven
Bridgeport
Stamford
Yonkers
Paterson
Morristown
Newark
Brentwood
Levittown
Flemington
New Brunswick
Trenton
Toms River
Results
< -1.5 Std. Dev.
-1.5 - -0.50 Std. Dev.
-0.50 - 0.50 Std. Dev.
0.50 - 1.5 Std. Dev.
> 1.5 Std. Dev.
Danbury
Stamford
Yonkers
Paterson
Newark
Jersey City
Levittown

A5. Meta-Index

Data

Data for the meta index come from the results for the four indexes listed above.

Method

Two meta-indexes were created. For both indexes all of the above indexed values were standardized by subtracting the index mean and then dividing by the index standard deviation. For the, average value index all of the standardized values were averaged. The high value count index is equal to the count of the number of indexes with a value greater than .5 standard deviations above the mean.

Results

The meta-indexes both show clusters of high value through northern Manhattan, the Bronx, northern Queens, and southern Brooklyn around Coney Island. Regionally, high values cluster in south eastern New Jersey, Treneton Patterson, Newark, Jersey City, and two transects running along the Hudson through Newburgh, Poughkeepsie, and Kingston, and another along the eastern Connecticut coast from New Haven through just south of Stamford.

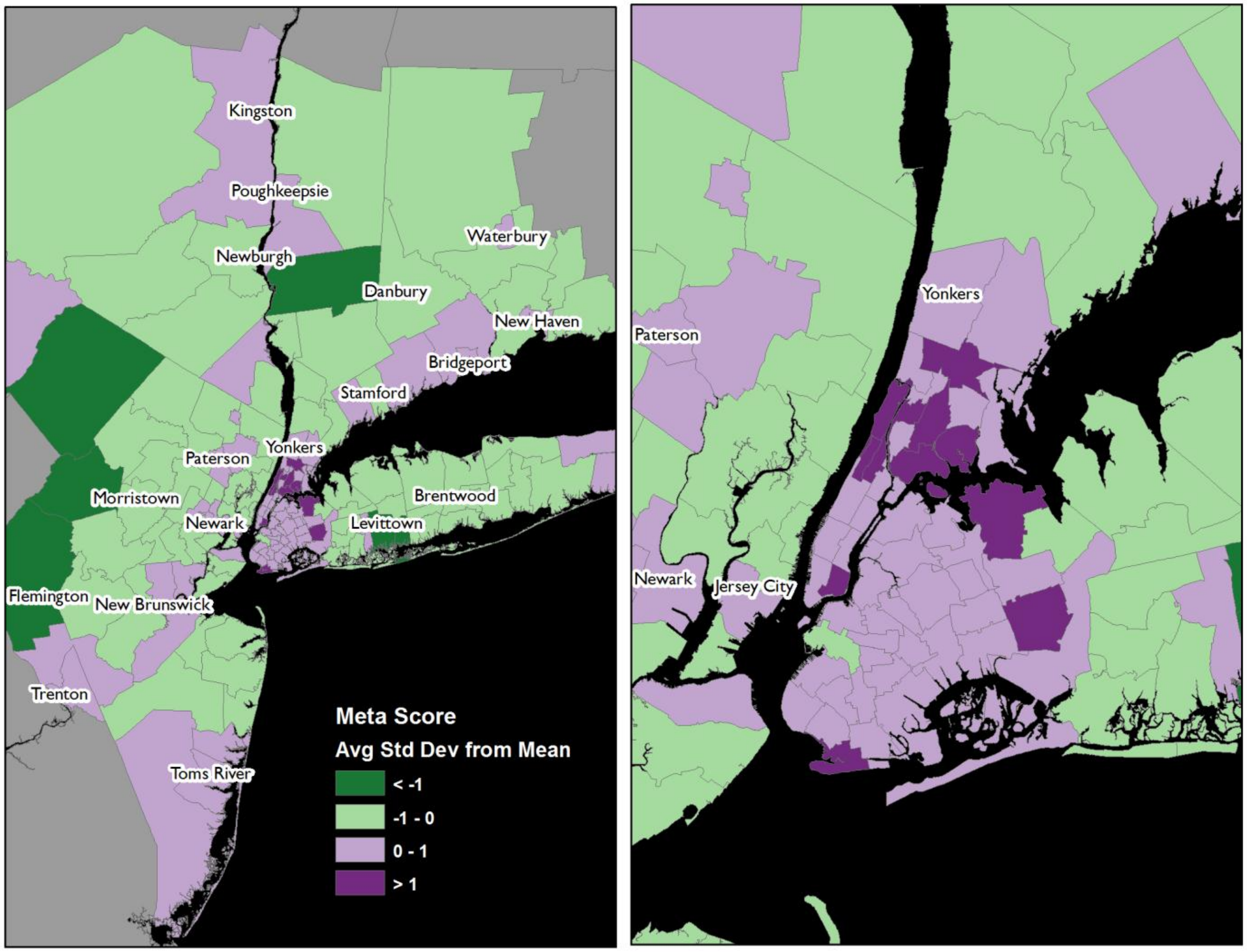


Figure A5.1a and A5.1b Meta Index based on the average of the standard deviations away from the mean for each of the four

Sources

Adger, W. N., N. Brooks, G. Bentham, M. Agnew, and S. Eriksen. 2004. New indicators of vulnerability and adaptive capacity. Norwich, England: Tyndall Centre for Climate Change Research, University of East Anglia.

Alwang, J., P. Siegel, and S. Jorgensen. 2001. Vulnerability: A View From Different Disciplines: The World Bank.

Americans with Disabiliies Act Amendments Act. Public Law 110-325. 2008.

Association of State and Territorial Health Officials (ASTHO). 2008. At Risk Populations and Pandemic Influenza: Planning Guidance for State, Territorial, Tribal, and Local Health Departments.

Barnett, J., S. Lambert, and I. Fry. 2008. The Hazards of Indicators: Insights from the Environmental Vulnerability Index. *Annals of the Association of American Geographers* 98 (1):102-102.

Birkmann. 2006. *Measuring Vulnerability To Natural Hazards*: TERI Press.

Brooks, N., N. Adger, and M. Kelly. 2005. The determinants of vulnerability and adaptive capacity at the national level and the implications for adaptation. *Global Environmental Change Part A* 15 (2):151-163.

Buckle, P., G. Mars, and S. Smale. 2000. New Approaches to Assessing Vulnerability and Resilience. *Australian Journal of Emergency Management* 15 (2):8-14.

Chandra, A., J. Acosta, L. Meredith, K. Sanches, S. Stern, L. Uscher-Pines, M. Williams, and D. Yeung. 2010. Understanding Community Resilience in the Context of National Health Security, A Literature Review. Santa Monica, CA: Rand Corporation.

Clark, G., S. Moser, S. Ratick, K. Dow, W. Meyer, S. Emani, W. Jin, J. Kasperson, R. Kasperson, and H. Schwarz. 1998. Assessing the Vulnerability of Coastal Communities to Extreme Storms: The Case of Revere, MA., USA. *Mitigation and Adaptation Strategies for Global Change* 3:59-82.

Cutter, S., B. Boruff, and W. L. Shirley. 2003. Social Vulnerability to Environmental Hazards. *Social Science Quarterly* 84 (2):242-261.

Cutter, S. L. 1996. Vulnerability to environmental hazards. *Progress in Human Geography* 20 (4):529-539.

Cutter, S. L., and C. T. Emrich. 2006. Moral Hazard, Social Catastrophe: The Changing Face of Vulnerability along the Hurricane Coasts. *The ANNALS of the American Academy of Political and Social Science* 604 (1):102-112.

Cutter, S. L., and C. Finch. 2008. Temporal and spatial changes in social vulnerability to natural hazards. *Proceedings of the National Academy of Sciences* 105 (7):2301-2306.

Cutter, S. L., J. T. Mitchell, and M. S. Scott. 2000. Revealing the Vulnerability of People and Places: A Case Study of Georgetown County, South Carolina. *Annals of the Association of American Geographers* 90 (4):713-713.

Dwyer, A. 2004. *Quantifying social vulnerability : a methodology for identifying those at risk to natural hazards*. Canberra Australia: Geoscience Australia.

Ebert, A., N. Kerle, and A. Stein. 2008. Urban social vulnerability assessment with physical proxies and spatial metrics derived from air- and spaceborne imagery and GIS data. *Natural Hazards* 48 (2):275-294.

Federal Emergency Management Agency (FEMA). 2008. National Response Framework. Washington, DC: US Department of Homeland Security.

Flanagan, B. E., E. Gregory, E. Hallisey, J. Heitgerd, and B. Lewis. 2011. A Social Vulnerability Index for Disaster Management. *Journal of Homeland Security and Emergency Management* 8 (1).

Fonseca, C. M., and P. J. Fleming. 1995. An overview of evolutionary algorithms in multiobjective optimization. *Evolutionary computation* 3 (1):1-16.

Gall, M. 2007. Indices of social vulnerability to natural hazards: A comparative evaluation, University of South Carolina, United States -- South Carolina.

Haines, V. A., J. S. Hurlbert, and J. J. Beggs. 1996. Exploring the determinants of support provision: provider characteristics, personal networks, community contexts, and support following life events. *Journal of Health and Social Behavior* 37 (3):252-264.

Handmer, J. 2003. We are all vulnerable. *The Australian Journal of Emergency Management* 18 (3):55 - 60.

Hinkel, J. 2011. Indicators of vulnerability and adaptive capacity: Towards a clarification of the science-policy interface. *Global Environmental Change* 21 (1):198-208.

Hufschmidt, G. 2011. A comparative analysis of several vulnerability concepts. *Natural Hazards* 58 (2):621-643.

Kailes, J. I., and A. Enders. 2007. Moving Beyond "Special Needs.". *Journal of Disability Policy Studies* 17 (4):230-237.

Kleinosky, L. R., B. Yarnal, and A. Fisher. 2006. Vulnerability of Hampton Roads, Virginia to Storm-Surge Flooding and Sea-Level Rise. *Natural Hazards* 40 (1):43-70.

Klinenberg, E. 2003. *Heat Wave: A Social Autopsy of Disaster in Chicago*: University of Chicago Press.

Kumpulainen, S. 2006. Vulnerability Concepts in Hazard and Risk Assessment. In *Natural and Technological and Risks Affecting the Spatial Development of European Regions*, 65-74: Geological Survey of Finland.

Murphy, B. 2007. Locating social capital in resilient community-level emergency management. *Natural Hazards* 41 (2):297-315.

Pandemic and All-Hazards Preparedness Act. Public Law No. 109-417. 2006.

Rupasingha, A., S. J. Goetz, and D. Freshwater. 2006. The production of social capital in US counties. *Journal of Socio-Economics* 35 (1):83-101.

Rygel, L., D. O'Sullivan, and B. Yarnal. 2006. A Method for Constructing a Social Vulnerability Index: An Application to Hurricane Storm Surges in a Developed Country. *Mitigation and Adaptation Strategies for Global Change* 11 (3):741-764.

Scheffler, R. M., T. T. Brown, L. Syme, I. Kawachi, I. Tolstykh, and C. Iribarren. 2008. Community-level social capital and recurrence of acute coronary syndrome. *Social Science & Medicine* 66 (7):1603-1613.

Tapsell, S., S. McCarthy, H. Faulkner, and M. Alexander. 2010. Social Vulnerability to Natural Hazards. London: Flood Hazard Research Centre - FHRC, Middlesex University.

Thywissen, K. 2006. Components of Risk: A Comparative Glossary. In *Studies of the University Research Counsel (SOURCE)*. Bonn Germany: UNU - EHS.

United Nations Development Programme - Bureau for Crisis Prevention and Recovery (UNDP-BCP). 2004. Reducing Disaster Risk, A Challenge for Development. New York City, NY: United Nations.

US Bureau of the Census. 2009. A Compass for Understanding and Using American Community Survey Data, What PUMS Data Users Need to Know. Washington, DC: US Government Printing Office.

———. 2010. American Community Survey, Puerto Rico Community Survey, 2010 Subject Definitions. Washington, DC: US Bureau of the Census.

———. 2012. *2010 Census Mail Participation Rate* 2012a [cited June 4 2012]. Available from http://2010.census.gov/2010census/take10map/.

———. *County Business Patterns* 2012b [cited. Available from http://www.census.gov/econ/cbp/.

US Department of Health and Human Services. 2012. *Special Medical Needs: Definitions and Related Terms* 2012 [cited June 4 2012]. Available from http://www.phe.gov/Preparedness/planning/abc/Pages/medical-needs.aspx.

USA Today. 2012. *Presidential vote by county* 2008 [cited June 4 2012]. Available from http://content.usatoday.com/news/politics/election2008/PresidentialByCounty.aspx?oi=P&rti=G&sp=NY.

Villagran De Leon, J. C. 2006. Vulnerability: A Conceptual and Methodological Review. In *Studies of the University Research Counsel (SOURCE)*. Bonn, Germany: United Nations University - Institute for Environment and Human Security (UNU-EHS).

Wisner, B., P. Blaikie, T. Cannon, and I. Davis. 2004. *At Risk: Natural Hazards, People's Vulnerability and Disasters*. London: Routledge.

World Health Organizations (WHO). *International Classification of Functioning, Disability, and Health (ICF)* 2012 [cited. Available from http://www.who.int/classifications/icf/en/.