Natural Hazards
and Earth System
Sciences
Discussions

Weigh(t)ing the dimensions of social vulnerability based on a regression analysis of disaster damages

Vincent David Corvin Heß[1]

[1]Wegener Center for Climate and Global Change, University of Graz, Graz, Austria

Correspondence to: Vincent David Corvin Heß (vincent.hess@uni-graz.at)

Abstract. Social vulnerability determines how natural hazards can turn into social disasters. However, large uncertainties remain on how to quantify social vulnerability in a compact and comprehensive way. A principle component analysis (PCA) combines many social indicators, such as population density or the level of education, into a single vulnerability index. Whether the individual indicators increase or decrease vulnerability is usually based on consequential reasoning or derived from previous

5   studies. These assumptions are rarely tested for their applicability to the study area. In a case study for the Austrian federal state of Styria we observe no correlation between disaster damages and an initial vulnerability index which was based on literature-derived assumptions about the influence of each indicator and following best practices for the PCA. We show that a regression analysis of past damages can improve the interpretation of social indicators by better representing the local situation. With the results from the regression-based analysis we can improve the PCA approach and only the updated vulnerability

10 index correlates with observed damages. This indicates that the use of a PCA-based vulnerability index requires a thorough understanding about the regionally specific influence of each indicator on social vulnerability to explain differences in disaster damages.

1   Introduction

Current risk assessments recognize the importance of vulnerability as one of the key determining factors for natural disaster

15 risk. The fifth assessment report of the International Panel on Climate Change (IPCC) defines vulnerability as the predisposition or propensity to be adversely affected (Agard et al., 2014). This broad definition leaves space for multidimensional concepts of vulnerability and includes the impairment of human beings, their livelihoods, and assets (Olsson et al., 2014). The disaster risk community originally defined vulnerability as the physical susceptibility of exposed assets to sustain damages (Thomalla et al., 2006). Since then, the social dimension of vulnerability has experienced increased attention, not only in the disaster risk

20 community, but especially in the climate change adaptation (Birkmann et al., 2013) and risk management literature (Fuchs et al., 2012). Although social vulnerability became an essential part of vulnerability research, large conceptual differences of social vulnerability remain.

Social vulnerability is seen as a multidimensional phenomenon encompassing a wide variety of concepts and approaches. Various attempts have been proposed to define and assess social vulnerability. It is generally accepted that it is rooted in and

25 caused by a lack of adapting and coping capacity, a lack of resilience and/or a higher susceptibility towards natural hazards

(Cardona et al., 2012; Birkmann et al., 2013). This can manifest on different scales, from individuals and households to communities, municipalities and whole nations (Tapsell et al., 2010). This multi-dimensionality of social vulnerability, a lack of comparable data across regions and imprecise definitions of the term social vulnerability prevent a direct measurement of social vulnerability. We can only approximate social vulnerability through the use of proxy variables. A popular approach is

5 the use of social indicators to capture different aspects of vulnerability and combine them into a single vulnerability index.

The vulnerability index approach calculates vulnerability scores from social indicators, such as the share of elderly citizens, the population density or the level of education. The relationship between the indicators and vulnerability is typically derived from previous studies or assumed based on logical reasoning (Cutter et al., 2003). There are mainly two different ways to select and combine indicators. In a deductive approach, a small set of well-defined indicators are selected, which should represent

10 the main dimensions of social vulnerability. In an inductive approach, one uses a factor analysis to calculate the statistical relationship between a large number of social indicators and select and weight the most important ones (Tapsell et al., 2010).

Because the relationship between indicators and vulnerability is usually unknown, a large amount of studies focus on the inductive approach (e.g., Cutter et al., 2003, 2008; Armaș and Gavriș, 2013; Fekete, 2009; Khan, 2012; Koks et al., 2015; Rogelis et al., 2016; Rygel et al., 2006; Yoon, 2012; Willis and Fitton, 2016). Within the inductive approach, a principle

15 component analysis (PCA) is the most common factor analysis technique. The composition of a PCA index includes certain design choices. Several studies have conducted sensitivity analysis to estimate which step in the construction of a social vulnerability index based on PCA is associated with the largest uncertainties (Schmidtlein et al., 2008; Tate, 2012). Although every step of the composition influences the results and needs to be justified, the weighting scheme of the individual indicators and the principle components cause the largest deviations.

20 In this paper we construct a PCA-based social vulnerability index on a municipality level for the Austrian federal state of Styria. Initially, we assume an equal weight of each indicator and derive the sign of the influence from previous studies. The access to a large database of per capita damages on municipality basis allows us to alternatively retrieve indicator weights through the means of a regression analysis. For each municipality we regress the social indicators on per capita damages to represent the place-specific circumstances. While previous studies used regression analyses to externally validate vulnerability

25 scores, to the author's knowledge, this is the first time it is used to estimate individual indicator weights directly. We compare the results with the PCA index and subsequently use the regression coefficients to calculate an updated PCA index. This allows us to explain a large share of the variability in disaster damages.

We structure the paper as follows: Section 2 describes the design choices for our calculation of a vulnerability index based on a PCA. As an alternative vulnerability index, we present a regression analysis of social indicators against natural disaster

30 damages to calculate a regression analysis index (RA index). Both indices are applied in a case study. Section 3 provides an overview of the data used in our analysis. First, we describe the study area of the federal state of Styria consisting of 495 municipalities. Following is a description of all the available social indicators and their assumed influence on vulnerability. Lastly, we describe the damage database. Section 4 compares the results of both vulnerability indices and presents an updated PCA index that incorporates the results from the regression analysis. It also describes the sensitivity of the regression results

35 in relation to different hazard types. Section 5 discusses the results in comparison to other studies of social vulnerability and

details the limitations and advantages of our approach. Section 6 concludes that a PCA approach can only deliver meaningful results if we know the place-specific influence of all indicators.

2   Methods: Quantifying social vulnerability

We calculate two different social vulnerability indices. First, we conduct a PCA to calculate a PCA vulnerability index with
5   equal component weights based on prevalent literature. Second, we perform a regression analysis and use the coefficients as indicator weights for the calculation of a RA vulnerability index with explicit weighting.

2.1   Classic Approach: The PCA index

A literature review revealed that most studies which quantify social vulnerability on the basis of social indicators use a PCA to aggregate indicators into a social vulnerability index. A PCA reduces the dimensionality of the underlying data by transforming
10   the input data into a set of orthogonal vectors (the principal components). Each of the original indicators loads onto the principal components with a different weight. There are different possibilities to determine the number of components to retain for the analysis. Most studies use the Kaiser criterion, i.e. the number of eigenvalues of the covariance matrix that are greater than one (Kaiser, 1958). An improved criterion is based on a parallel analysis of a randomized matrix with the same dimensionality as the original data. To remove spurious components, principle components whose eigenvalues are smaller than that of the random
15   components are removed (Franklin et al., 1995; Tate, 2012; Ledesma and Valero-Mora, 2007). In our case, both criteria advise to keep three principal components. For the further analysis one needs to estimate the influence of the principal components on vulnerability, based on which original indicators load on the principle components. We use a varimax rotation to achieve higher loadings on individual components, which helps to interpret the resulting components (Kaiser, 1958). Based on the loadings, we assign each component a sign that reflects if it decreases or increases vulnerability.
20   The PCA vulnerability index is then calculated as the weighted sum of the scores of the rotated principal components for every municipality. This is the most crucial step in the calculation of a vulnerability index (Schmidtlein et al., 2008; Tate, 2013). Proposed weighting schemes include equal weighting (Cutter et al., 2003), expert judgments (Vincent, 2004), explainable variance weighted sum (Schmidtlein et al., 2008) or pareto ranking (Rygel et al., 2006). Since each method gives different weight to the principle components, the resulting vulnerability scores can differ from each other. As there is no justification to
25   use one approach over the others, we opted for the simplest and most widely used, equal weighting.

2.2   Regression Analysis: The RA index

Although the PCA index is based on the statistical relationship between social indicators, it still requires a priori knowledge about the sign of the influence of individual indicators on vulnerability to interpret the resulting components. Typically, the assumed influence is derived from other studies and study areas, based on expert judgments or based on logical reasoning. As
30   vulnerability is a dynamic notion, it is important to assess any indicator-based approach within the political, environmental, and socioeconomic landscape that it is being applied (Willis and Fitton, 2016). To weight the indicators appropriately, we conduct

Natural Hazards
and Earth System
Sciences
Discussions

a regression analysis of each social indicator against per capita damages to infer their influence on social vulnerability in our study area. We derive the per capita damages from the Austrian disaster fund, which covers damage from natural hazards and is described in further details in Sect. 3.3.

Rather than analyzing the intrinsic characteristics of the social indicator variance, we relate the social indicators directly to
5 damages per capita. This ensures a direct measurement of an individual indicator's influence on social vulnerability. The aim of the regression analysis is to see which indicators have a statistically significant influence and if they increase or decrease vulnerability.

Initially, a multiple linear regression against per capita damages is done according to equation (1), where damage per capita (DPC_i) of municipality i is dependent on the standardized regression coefficients (β) times the P social indicators (SI) and
10 the error term ϵ.

$$DPC_i = \beta_0 + \sum_{p=1}^{P} (\beta_{ip} \cdot SI_{ip}) + \epsilon_i \tag{1}$$

This simple regression does not take into account differences in disaster intensities and exposure of the individual municipalities. The resulting coefficients can be strongly biased and overestimate the importance of individual indicators. To account for differences in the number of affected households per natural disaster, we introduce the variable A. A is a proxy variable for
15 the disaster intensity, implemented as the number of disaster fund applications for a damage refund. A high number of applications can indicate either large-scale events, which affected multiple households simultaneously, municipalities that have been affected multiple times or both. Since the social indicators and the damage data are available for the years 2011 to 2014, we can calculate a fixed-effects model, where we include the individual years as dummy variables. This captures all the differences between municipalities that are not included in the social indicators and that do not change over time. Most importantly, this
20 includes differences in exposure, which we expect to play a major role but we do not expect to have changed over the short period of four years. The complete regression formula is

$$DPC_i = \beta_0 + \beta_1 A_i + \beta_2 C(year) + \sum_{p=1}^{P} (\beta_{ip} \cdot SI_{ip}) + \epsilon_i, \tag{2}$$

where $C()$ denotes the categorized fixed effect dummy variables for the years 2011 to 2014. The resulting standardized β coefficients are multiplied by the social indicators per municipality and summed up to derive a single RA score. The RA
25 vulnerability index is a vector of all individual municipality scores.

As an additional robustness check we split the dependent variable into different hazard types. The available categories are floods, landslides, storms and avalanches. The goal is to analyze which social indicators are hazard dependent and which are hazard independent and if it is justifiable to aggregate over all hazard types.

3 Data description

This section first describes the study area, for which the social indicators and the damage database were available. Following is a description of the eleven social indicators used in this study and their assumed influence on social vulnerability. This is the basis for the PCA approach. We then describe the damage database we used for the regression analysis.

3.1 Study area

We conduct this study on the municipality level for the Austrian federal state of Styria. The analysis includes 495 municipalities ranging from small and rural Alpine settlements with less than 200 inhabitants to the capital city of Graz with about 269,000 inhabitants (as of 2014). Because Styria's landscape ranges from Alpine foothills to the steep topographic reliefs of the Alps, it is exposed to multiple natural hazards. Fuchs et al. (2015) analyzed the exposure of Styrian buildings. 13.2% (50,419) of all buildings are exposed to natural hazards. About half of these buildings are exposed to river floods and half of them to flash floods and torrential flows. Only 0.1% (460) of the buildings are exposed to avalanches, while 0.9% (3,712) buildings are exposed to more than one hazard.

3.2 Social Indicators

The selection of social indicators determines how well a social vulnerability index can describe the actual social vulnerability. However, the availability of suitable social indicators on a municipality level constraints the selection. Table 1 summarizes the eleven social indicators we use, their numerical range, and the expected influence on vulnerability. We collected the majority of the indicators from Statistik Austria (2016). Although eleven indicators is at the lower spectrum of number of indicators used for the creation of social vulnerability indices, they represent some of the most commonly used indicators across social vulnerability studies (e.g. Armaș and Gavriș, 2013; Cutter et al., 2003; Frigerio and De Amicis, 2016). With our indicators we can account for differences in municipality characteristics, the composition of different social groups and the social status.

3.2.1 Municipality Characteristics

Our first three indicators depict basic municipality characteristics. *Population density* potentially increases a municipality's vulnerability, since densely populated areas are more difficult to evacuate and decrease the available space for adaptive measures (Cutter et al., 2003; Cutter and Finch, 2008; Yoon, 2012). Linked to the overall population density is the *average household size*, which states how many individuals share a household. Higher values are associated with larger families while lower values depict a larger share of single households, the latter of which are associated with a higher vulnerability due to fewer available funds (Fekete, 2009; Rogelis et al., 2016). The number of *fire station per citizens* reflects a municipality's ability to respond to natural disasters. A higher ratio decreases a municipality's vulnerability (Rogelis et al., 2016).

3.2.2 Social Groups

We use four indicators that reflect the share of different social groups. The *share of children* (below 15 years) and the *share of senior citizens* (above 65 years) both increase the vulnerability of a municipality (Cutter et al., 2003; Koks et al., 2015; Rogelis et al., 2016). Not only are those groups less able to cope with disasters themselves, they also require the attention from other people, who then themselves lack the ability to cope with the disaster. The *share of woman* increases vulnerability, since woman have on average less available funds and are less represented in voluntary disaster relief units (Morrow and Phillips, 1999). A high *share of foreigners* is generally linked to a higher vulnerability, since they are often hindered by cultural and/or language barriers to respond to emergency situations (Cutter et al., 2003; Koks et al., 2015).

3.2.3 Social Status

The social status of municipalities inhabitants is reflected in another four social indicators. We include *per capita tax payments* as a proxy for the wealth of a municipalities inhabitants. Higher wealth allows for self-provisional protection measures and also allows to more easily respond to natural disasters (Cutter et al., 2003; Fekete, 2009). Higher wealth therefore decreases vulnerability, although the monetary damages can increase due to a higher number and value of exposed assets. On the contrary, a high *unemployment quota* is associated with a lack of funds to prepare for natural disasters and cope with them and thus is considered as a vulnerability increasing factor (Armaș and Gavriș, 2013). Better *education* decreases vulnerability, because higher educational levels are associated with higher income and also with a higher level of risk awareness (Cutter et al., 2003; Rogelis et al., 2016). We measure *education* as the share of people having at least secondary education. A higher *share of commuters* per municipality is associated with a higher vulnerability to natural disasters. Their occupational absence make it difficult to respond to disasters in time and they are more dependent on a functioning infrastructure (Fekete, 2009).

3.3 Damage data

For the evaluation of our vulnerability indices and for the regression analysis we worked with a damage database from the Austrian Disaster Fund. In Austria, every damage from natural disasters above 1,000 €, which is not covered by insurance, is eligible to be partially covered by the disaster fund. The dataset includes the number of applications, reported damages and compensation payments for each municipality for the years 2011 to 2014. Compensation payments for our study area were usually between 30% and 50% (average: 34% of damages). We worked with the reported losses from private damages, which include applications from households and private companies, but not communal losses. The latter are assessed separately and are not available for this study. Considering the low insurance penetration rates across Austria (Holub and Fuchs, 2009), we assume a constant share of private, uninsured to total losses across municipalities. Since we analyze relative differences between municipalities, covering only a fraction of the total costs does not change our results.

Natural Hazards
and Earth System
Sciences

Discussions

Table 1. Social indicators used for the analysis and their assumed influence on social vulnerability

Indicator	Range (std.)	Assumed influence
Population density[a]	30–2620 km^{-1} (247)	increase
Average household size	1.81–3.77 (0.34)	decrease
Fire stations per citizens[a]	0–7.14 (1.21)	decrease
Share of children	6.9–19.7 % (1.98)	increase
Share of elderly citizen	10.1–38.7 % (3.5)	increase
Share of women	38.3–56.0 % (1.7)	increase
Share of foreigners	0.0–36.1 % (2.9)	increase
Per capita tax payments	651–2,883 €(232)	decrease
Unemployment rate	0.0–13.2 % (1.8)	increase
Education	4.7–84.3 % (4.7)	increase
Share of commuters	25.5–91.9 % (10.0)	increase

[a] The numerical calculations use the reciprocal values, namely population sparsity and citizens per fire stations

3.4 Data Preparation

Since our goal is to conduct a PCA and a regression analysis, we need to make sure that our variables approximate a normal distribution. This is already true for most indicators except *population density* and *fire stations per citizens*, which we need to transform. For the numerical analysis, we took the reciprocal values of these originally right-skewed indicators, namely population sparsity (area per inhabitants) and the number of citizens per fire station, which approximate a normal distribution. In the discussion of the results, however, we will continue to use the labels of *population density* and *fire stations per citizens*, as those more typical terms. The per capita damages and the number of disaster fund applications were also highly skewed. Distribution fitting revealed that they follow a heavy-tailed distribution. By taking the logarithm of the per capita damages and the disaster fund applications, we can approximate a normal distribution.

Since the indicators are measured in different units, the PCA would give more weight to indicators with larger values and higher variance. We need to normalize the indicators to compare them scale independently. Yoon (2012) analyzes the impact of different linear transformations of social indicators for the use of a PCA and finds little importance of the chosen transformation. We choose the z-score transformation, which subtracts the mean and divides by the standard deviation. We apply this rescaling for the social indicators and the log-transformed per capita damages.

Natural Hazards
and Earth System
Sciences

Open Access

Discussions

EGU

4   Results

In this section we first describe the results of the PCA and how we interpret the components to calculate our initial PCA index.
We then describe the results of the regression analysis that lead to the RA index. We subsequently compare both indices with
each other and per capita damages. Finally, we make use of the results from the regression analysis to calculate an updated
5   PCA index that correlates well with per capita damages.

4.1   Initial PCA index

Our PCA of eleven social indicators results in three principle components, which can explain 62% of the original variance.
Table 2 shows the main factor loadings for each components. To derive a single vulnerability score we need to add the principle
components in a way that reflects an increase in vulnerability with increasing scores. The first principle component is mainly
10   influenced by *population sparsity* and *education*. Both indicators load onto it with opposite signs. The negative absolute value
of the first principle component can indicate either a badly educated or a densely populated municipality, both increasing
vulnerability. We interpret this component as an indicator of "urbanity". We label the second component as "age", which is
influenced by the *share of children* and the *share of senior citizens*. We add the absolute value, because a high share of either
group increases vulnerability. We interpret the third principle component as "wealth", because the *share of foreigners*, the *per
15   capita tax payments*, the *unemployment rate*, and the *share of commuters*, have the highest loadings. Since all of indicators
increase vulnerability and the majority load positively on "wealth", no transformation was necessary for the last principle
component. Following the prevalent literature we applied no weights to the components and added them as in equation (3)

$$\text{PCA index} = -abs(\text{PC1} \times \text{SI}) + abs(\text{PC2} \times \text{SI}) + \text{PC3} \times \text{SI}, \tag{3}$$

where PC are the varimax-rotated principle component vectors and SI is a matrix of all indicators and all municipalities.

20   ### 4.2   RA index

We conduct a regression analysis as an alternative measurement of a vulnerability score and to estimate the influence of
individual social indicators on per capita damages. Instead of using the statistical relationship between indicators as represented
by the principle components, we use the social indicators directly. The fixed effects regression also includes the number of
disaster fund application as a proxy for disaster intensities. This approach allows us to explain 63% of the variance in per
25   capita damages. The resulting coefficients are then used as weights for adding the social indicators into a composite RA index.

We find that only few indicators had a statistically significant ($p < .01$) influence on per capita damages (see Tab. A1 in
the appendix for the complete regression table). The only indicators that significantly increase vulnerability are the *share
of commuters*, the *share of elderly citizens*, and *population sparsity*. The *citizens per fire station* and the *average household
size* decrease vulnerability for per capita damages. Overall, *population sparsity* and the *citizens per fire station* are the most
30   important indicators, while the *share of women, foreigners and children*, the *per capita tax payments*, the *unemployment rate*,

Natural Hazards
and Earth System
Sciences

Open Access

EGU

Discussions

Table 2. Main factor loadings of the social indicators onto the first three principle components. Interpretation is based on the indicators and transformed to reflect increasing vulnerability

Name	Social indicator	Factor loading	Interpretation	Variance explained
PC1	Population sparsity	−0.53	Urbanity	33%
	Education	+0.49	(−abs)	
PC2	Share of children	−0.60	Age	18%
	Share of elderly citizen	+0.62	(+abs)	
PC3	Share of foreigners	+0.54	Wealth	12%
	Per capita tax payments	+0.47	(+)	
	Unemployment rate	+0.47		
	Share of commuters	−0.45		

and *education* show no significant influence (see Fig. 1). These results indicate that few of the chosen indicators actually determine the differences in social vulnerability.

Although the effect (i.e. vulnerability increasing or decreasing) is as expected for most of the indicators, we can observe notable and important differences. The most influential indicator, *population density*, is associated with lower per capita damages, contrary to our expectations. The opposite is true for the *fire stations per capita*, which, to our surprise, increase vulnerability. We will discuss the reasons in Sect. 5. The results also hold true when we differentiate between hazard types. While *population density* and the *fire stations per capita* remain the most important indicators for all hazard types, the *share of elderly citizens* is only important for landslides. Floods are the only hazard for which the *share of foreigners* decrease vulnerability statistically significant and the *share of commuters* increase vulnerability. This is likely caused by the fact that flood damages can be avoided if people are at home or nearby. This is not true for landslides and avalanches which happen rapidly. The complete regression table for the hazard-specific regression can be found in Tab. A1 in the appendix.

4.3 An updated PCA approach

When we correlate the RA index with the PCA index we see a small negative correlation (spearmans $r = −.12, p < .001$), but a moderate positive correlation with per capita damages ($r = .44, p < .001$). To analyze the discrepancy between the PCA index and the RA index, we review the role of the individual indicators in the PCA approach. Originally, we attribute the first principle component ("urbanity") a negative absolute influence on social vulnerability. Because we find *population sparsity* to be the most important indicator and *education* insignificant, we update the contribution of "urbanity" to increase vulnerability with *population sparsity*. We also subtract rather than add the last principle component "wealth". The *share of commuters* is the only significant indicator of it and it increases vulnerability, but decreases the value of the "wealth" component.
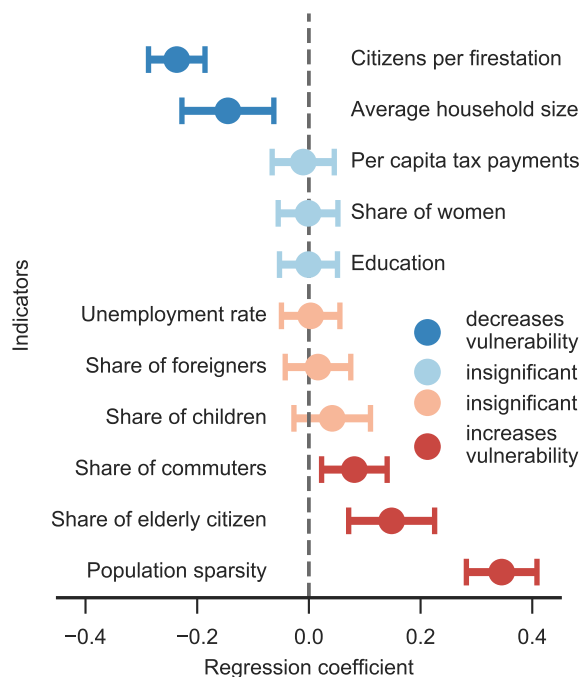
Figure 1. Relative parameter weights calculated from a regression analysis of eleven social indicators against per capita damages. The 95% confidence interval is shown and significance is based on p-values $< .01$. Negative weights decrease vulnerability, while positive weights increase it. Fixed effects and disaster fund applications are also part of the regression, but are not shown.

The updated interpretation is captured in equation 4

$$\text{updated PCA index} = +\text{PC1} \times \text{SI} + abs(\text{PC2} \times \text{SI}) - \text{PC3} \times \text{SI}. \tag{4}$$

As a result, we observe a much better correlation between the updated PCA index and the RA index ($r = .53, p < .001$) and per capita damages ($r = .23, p < .001$), see also Fig. 2. To account for the number of disaster fund applications as a proxy for disaster intensity and fixed effects, we conduct a further regression analysis of all vulnerability indices against per capita damages. The results are summarized in Tab. 3. The initial PCA index does not improve the regression in comparison to an approach without any vulnerability index. The updated PCA index shows an increase in the explained variance (R^2 increases 22% from 0.41 to 0.50) compared to the initial PCA index and to the regression without any social indicators. But the updated PCA index can not only explain even more of the variance, it is only the updated PCA index that has a highly significant positive coefficient.

Natural Hazards
and Earth System
Sciences

Open Access

Discussions

Table 3. Comparison of regression results from the different vulnerability indices against per capita damages. Applications are disaster fund applications and year fixed effects are dummy variables for the individual years

	no SI	initial PCA	RA	updated PCA
Vulnerability index		-0.06**	0.47***	0.31***
(PCA or RA)		(0.03)	(0.02)	(0.03)
Applications	0.62***	0.62***	0.64***	0.64***
	(0.93)	(0.03)	(0.02)	(0.03)
Fixed effects				
Year 2011	-0.30***	-0.32***	-0.20***	-0.39***
	(0.09)	(0.09)	(0.08)	(0.09)
Year 2012	0.09	0.09	-0.01	0.06
	(0.06)	(0.06)	(0.05)	(0.06)
Year 2013	0.05	0.06	0.06	0.07*
	(0.05)	(0.05)	(0.04)	(0.04)
Year 2014	-0.03	-0.03	-0.00	-0.01
	(0.04)	(0.04)	(0.04)	(0.04)
Observations	807	807	807	807
adjusted R^2	0.41	0.41	0.62	0.50

Standard errors in parentheses.
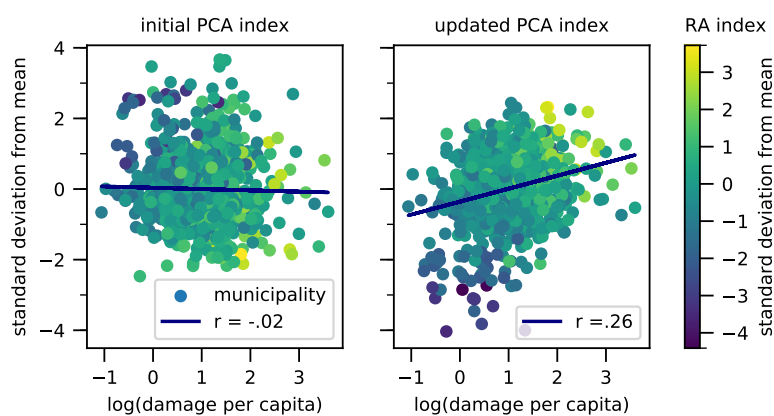
$* p < .1, ** p < .05, *** p < .01$

Figure 2. Correlation between per capita damages and vulnerability indices. While the initial PCA index shows no correlation, the updated PCA index shows a clear correlation and is more closely related to the RA index.

5 Discussion

Although we have followed best practices in calculating a social vulnerability index via a PCA of social indicators, our initial index performs poorly when we correlate it against per capita damages. Validation of vulnerability indices is a critical issue, but is not always considered. While some studies interpret the indices as they are, other studies try to externally validate their

5　vulnerability index with independent proxy data (Tate, 2013). Mortality (Adger et al., 2004), economic losses (Schmidtlein et al., 2011) or the number of disaster declarations (Cutter et al., 2003) have been proposed as validation variables for social vulnerability. Another approach is the internal validation of the index via a sensitivity study (Tate, 2012; Schmidtlein et al., 2008). We used a mixed approach, where we used an external proxy (disaster damages) to calculate indicator weights and internally compare the results to the results of a PCA. We find the initial correspondence between both indices is small due to

10　false assumptions about the influence of individual indicators in the initial PCA approach.

　　Contrary to our expectations, we identify the most influential indicator, *population density*, as a vulnerability decreasing factor. A considerable amount of literature classifies population density as a factor that increases vulnerability (e.g. Mazumdar and Paul, 2016). We identify three reasons why we observe the opposite. First, densely populated areas are better protected, since implementing adaptation measures is more cost-effective in such areas. Second, we measure damage as damage per

15　capita. While this makes it possible to compare large and small municipalities, it also decreases the importance of individual damages in large (and typically more dense) municipalities. If a disaster affects only few citizens in a large municipality, even if individual households experience large losses, per capita damages in the whole municipality remain low. Third, we measure social vulnerability only as monetary damages. This neglects the fact that other, non-monetary effects can be more severe. For instance, densely populated areas are considered more vulnerable, because they are more difficult to evacuate. This

20　influences individual stress levels, but does not necessarily result in higher monetary damages. Our initial assumption that densely populated areas increase vulnerability was based on individual and a priori factors, but neglected the aggregate effects on a municipality level that already addresses some of the issues.

　　For most of the other indicators, we can not find an influence on social vulnerability as measured by per capita damages. This, however, does not have to mean that those indicators are unimportant for social vulnerability. For one, their effects

25　on vulnerability are well documented and might already be addressed by the municipalities. For example, we expect a high share of children to have an effect similar to a high share of elderly citizens. However, parents are aware of their children's vulnerability and take precautionary measures to resist natural disasters. We also observe a hidden influence for the share of women. First, the share of women is expected to increase vulnerability mainly through secondary effects like lower income levels. Second, the share of women does not substantially fluctuate between municipalities in Styria and differences are more

30　difficult to observe in a regression analysis. Again, our initial assumptions did not take into account that higher vulnerabilities might already be addressed or differences within some indicators are less distinct than for other indicators.

　　A disadvantage of using social indicators as proxies for social vulnerability also arises from the fact that they are aggregate measurements in specific units. For example, education has been identified as an important factor of social vulnerability, because it is linked to risk awareness, income and thus an increased ability to prepare for and respond to damages. Yet, we only

measured education as the share of citizens having at least secondary education. This neglects a lot of the different concepts of education. We observe another example for the importance of the measurement unit for the *fire stations per capita*. We measure *fire stations per capita* as a variable that reflects the coping capacity of a municipality. Although that should decrease vulnerability, we observe that a higher number of fire stations per citizens increases vulnerability. However, we only collected

5   the number of fire stations without any details about their size or their equipment. This shows that we need to be careful when we derive social vulnerability from indicators. Although their conceptual influence is in one direction, we might measure a different aspect. With the knowledge of these limitations and the true influence of our indicators on per capita damages, we can improve the results of the PCA approach considerably.

Some of the ambiguity of social vulnerability is also caused by the various definitions of vulnerability to what. Most studies

10   that quantify social vulnerability from social indicators do not specify the risks and negative outcomes associated with natural disasters. Negative outcomes are thought of as an aggregate from direct damages to non-economic losses such as the loss of life or traumatic disorders. This makes it difficult to analyze since the detailed processes leading to vulnerability differ substantially. We specifically focus our social vulnerability assessment on per capita damages, representing direct and monetary damages resulting from natural disasters. We did not include differences in the physical vulnerability, which can be substantial for direct

15   costs (Papathoma-Köhle et al., 2011; Merz et al., 2010). While we see the concrete definition of vulnerability as the propensity to suffer direct, monetary damages as one of the main advantages of our research, it is also the biggest limitation. Our study does neither address the indirect costs, nor intangible effects (direct and indirect) (Meyer et al., 2013). For these kinds of costs, differences in the social vulnerability might play a more important role than it does for monetary damages.

6   Conclusions

20   We have seen that a vulnerability index based on a PCA of social indicators can only give meaningful results if we consider the place and domain-specific circumstances. We have used a regression analysis to estimate the true influence of social indicators on per capita damages in a case study of Styrian municipalities. Not only is the influence of most indicators insignificant, but it can even have an effect contrary to general assumptions. The units of measurement, unobserved secondary effects and the considered outcome variable can all influence the importance of individual indicators.

25   Therefore, indicators need to be reviewed on a per case basis with keeping in mind how social vulnerability is defined and in what unit the indicators are measured. Although a damage-dependent regression analysis might seldom be possible, with the right choice of indicators and their influence, a PCA based vulnerability index can relate closely to disaster damages. This should encourage future research to focus attention on few, but decisive indicators for which their influence is well-defined rather than include an excessive number of indicators whose local influence is unknown.

Appendix A

Table A1. Complete regression table of log(damage per capita) against the number of disaster fund applications (log(Applications)), yearly fixed effects and eleven social indicators.

	All	Flood	Landslide	Storm	Avalanche
log(Applications)	0.6428***	0.6841***	0.4793***	0.6889***	0.4597***
	(0.0224)	(0.0267)	(0.0348)	(0.0511)	(0.0713)
Fixed effects					
Year 2011	-0.2044***	-0.1830*	-0.4073***	0.0259	
	(0.0776)	(0.0967)	(0.1239)	(0.1568)	
Year 2012	-0.0104	0.0416	0.0995	0.0535	-0.5823***
	(0.0494)	(0.0529)	(0.0739)	(0.1070)	(0.2048)
Year 2013	0.0597	-0.0047	-0.0157	0.1574**	0.0858
	(0.0377)	(0.0517)	(0.0549)	(0.0770)	(0.1291)
Year 2014	-0.0025	0.0125	-0.0453	-0.0470	-0.1544
	(0.0358)	(0.0426)	(0.0601)	(0.0920)	(0.1050)
Social indicators					
Share of children	0.0418	0.0165	0.0904	0.0368	0.1121
	(0.0349)	(0.0439)	(0.0587)	(0.0800)	(0.0977)
Share of senior citizens	0.1482***	0.0246	0.1741***	0.1348	0.1247
	(0.0392)	(0.0534)	(0.0625)	(0.0925)	(0.1370)
Share of foreigners	0.0163	0.0733**	0.0145	0.0040	-0.0538
	(0.0299)	(0.0369)	(0.0406)	(0.0818)	(0.1255)
Unemployment rate	0.0032	-0.0951***	0.0825**	-0.0087	0.0256
	(0.0267)	(0.0365)	(0.0396)	(0.0604)	(0.0724)
Share of women	-0.0015	0.0502	0.0144	-0.0277	0.1328*
	(0.0272)	(0.0348)	(0.0402)	(0.0657)	(0.0767)
Share of commuters	0.0814***	0.1390***	0.0229	0.0713	0.1745*
	(0.0299)	(0.0372)	(0.0478)	(0.0663)	(0.1020)
Average household size	-0.1451***	-0.2364***	-0.1593**	-0.1286	-0.1156
	(0.0419)	(0.0550)	(0.0699)	(0.0885)	(0.1140)
Citizens per fire station	-0.2366***	-0.2570***	-0.3717***	-0.1885***	-0.1901*
	(0.0258)	(0.0349)	(0.0413)	(0.0410)	(0.1028)
Education	-0.0007	-0.0377	-0.1166***	0.0299	-0.0846
	(0.0265)	(0.0351)	(0.0388)	(0.0625)	(0.0793)
Population sparsity	0.3452***	0.3907***	0.1844***	0.3323***	0.4466***
	(0.0322)	(0.0381)	(0.0552)	(0.0681)	(0.0891)
Per capita tax payments	-0.0102	0.0221	-0.0290	-0.0895	0.1252
	(0.0283)	(0.0338)	(0.0409)	(0.0766)	(0.1193)
Observations	807	456	431	181	125
adjusted R^2	0.62	0.70	0.53	0.61	0.47
F-statistics	87.96	71.57	33.03	19.86	8.78
Prop(F)	6e-157	1e-107	2e-61	1e-29	2e-12

Standard errors in parentheses.

* $p < .1$, ** $p < .05$, ***$p < .01$

Competing interests. The author declares that he has no conflict of interest.

5

Natural Hazards
and Earth System
Sciences
Discussions

References

Adger, W. N., Brooks, N., Bentham, G., and Agnew, M.: New indicators of vulnerability and adaptive capacity, 2004.

Agard, J., Schipper, E., Birkmann, J., Campos, M., Dubeux, C., Nojiri, Y., Olsson, L., Osman-Elasha, B., Pelling, M., Prather, M., Rivera-Ferre, M., Ruppel, O., Sallenger, A., Smith, K., Clair, A. S., Mach, K., and Mastrandrea, M.: Annex II: Glossary, in: Clim. Chang. 2014
5   Impacts, Adapt. Vulnerability. Part B Reg. Asp. Contrib. Work. Gr. II to Fifth Assess. Rep. Intergov. Panel Clim. Chang., edited by Barros, V. R., Field, C. B., Dokken, D. J., Mastrandrea, M. D., Mach, K. J., Bilir, T. E., Chatterjee, M., Ebi, K. L., Estrada, Y. O., Genova, R. C., Girma, B., Kissel, E. S., Levy, A. N., MacCracken, S., Mastrandrea, P. R., and White, L. L., pp. 1757–1776, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2014.

Armaș, I. and Gavriș, A.: Social vulnerability assessment using spatial multi-criteria analysis (SEVI model) and the Social Vulnerability
10   Index (SoVI model) – a case study for Bucharest, Romania, Nat. Hazards Earth Syst. Sci., 13, 1481–1499, doi:10.5194/nhess-13-1481-2013, 2013.

Birkmann, J., Cardona, O. D., Carreño, M. L., Barbat, A. H., Pelling, M., Schneiderbauer, S., Kienberger, S., Keiler, M., Alexander, D., Zeil, P., and Welle, T.: Framing vulnerability, risk and societal responses: the MOVE framework, Nat. Hazards, 67, 193–211, doi:10.1007/s11069-013-0558-5, 2013.

15   Cardona, O., van Aalst, M., Birkmann, J., Fordham, M., McGregor, G., Perez, R., Pulwarty, R., Schipper, E., and Sinh, B.: Determinants of Risk: Exposure and Vulnerability, in: Manag. Risks Extrem. Events Disasters to Adv. Clim. Chang. Adapt., edited by Field, C., Barros, V., Stocker, T., Qin, D., Dokken, D., Ebi, K., Mastrandrea, M., Mach, K., Plattner, G.-K., Allen, S., Tignor, M., and Midgley, P., chap. A Special, pp. 65–108, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, doi:10.1017/CBO9781139177245.005, 2012.

20   Cutter, S. L. and Finch, C.: Temporal and spatial changes in social vulnerability to natural hazards., Proc. Natl. Acad. Sci. U. S. A., 105, 2301–2306, doi:10.1073/pnas.0710375105, 2008.

Cutter, S. L., Boruff, B. J., and Shirley, W. L.: Social Vulnerability to Environmental Hazards*, Soc. Sci. Q., 84, 242–261, doi:10.1111/1540-6237.8402002, 2003.

Cutter, S. L., Barnes, L., Berry, M., Burton, C., Evans, E., Tate, E., and Webb, J.: A place-based model for understanding community
25   resilience to natural disasters, Glob. Environ. Chang., 18, 598–606, doi:10.1016/j.gloenvcha.2008.07.013, 2008.

Fekete, A.: Validation of a social vulnerability index in context to river-floods in Germany, Nat. Hazards Earth Syst. Sci., 9, 393–403, doi:10.5194/nhess-9-393-2009, 2009.

Franklin, S. B., Gibson, D. J., Robertson, P. A., Pohlmann, J. T., Fralish, J. S., Pohlman, J. T., and Fralish, J. S.: Parallel analysis: a method for determining significant principal components, J. Veg. Sci., 6, 99–106, doi:10.2307/3236261, 1995.

30   Frigerio, I. and De Amicis, M.: Mapping social vulnerability to natural hazards in Italy: A suitable tool for risk mitigation strategies, Environ. Sci. Policy, 63, 187–196, doi:10.1016/j.envsci.2016.06.001, 2016.

Fuchs, S., Birkmann, J., and Glade, T.: Vulnerability assessment in natural hazard and risk analysis: current approaches and future challenges, Nat. Hazards, 64, 1969–1975, doi:10.1007/s11069-012-0352-9, 2012.

Fuchs, S., Keiler, M., and Zischg, A.: A spatiotemporal multi-hazard exposure assessment based on property data, Nat. Hazards Earth Syst.
35   Sci., 15, 2127–2142, doi:10.5194/nhess-15-2127-2015, 2015.

Holub, M. and Fuchs, S.: Mitigating mountain hazards in Austria - legislation, risk transfer, and awareness building, Nat. Hazards Earth Syst. Sci., 9, 523–537, doi:10.5194/nhess-9-523-2009, 2009.

Kaiser, H. F.: The varimax criterion for analytic rotation in factor analysis, Psychometrika, 23, 187–200, doi:10.1007/BF02289233, 1958.

Khan, S.: Vulnerability assessments and their planning implications: a case study of the Hutt Valley, New Zealand, Nat. Hazards, 64, 1587–1607, doi:10.1007/s11069-012-0327-x, 2012.

Koks, E. E., Jongman, B., Husby, T. G., and Botzen, W. J. W.: Combining hazard, exposure and social vulnerability to provide lessons for
5   flood risk management, Environ. Sci. Policy, 47, 42–52, doi:10.1016/j.envsci.2014.10.013, 2015.

Ledesma, R. and Valero-Mora, P.: Determining the number of factors to retain in EFA: An easy-to-use computer program for carrying out parallel analysis, Pract. assessment, Res. Eval., 12, 2007.

Mazumdar, J. and Paul, S. K.: Socioeconomic and infrastructural vulnerability indices for cyclones in the eastern coastal states of India, Nat. Hazards, doi:10.1007/s11069-016-2261-9, 2016.

10 Merz, B., Kreibich, H., Schwarze, R., and Thieken, A.: Review article "Assessment of economic flood damage", Nat. Hazards Earth Syst. Sci., 10, 1697–1724, doi:10.5194/nhess-10-1697-2010, 2010.

Meyer, V., Becker, N., Markantonis, V., Schwarze, R., van den Bergh, J. C. J. M., Bouwer, L. M., Bubeck, P., Ciavola, P., Genovese, E., Green, C., Hallegatte, S., Kreibich, H., Lequeux, Q., Logar, I., Papyrakis, E., Pfurtscheller, C., Poussin, J., Przyluski, V., Thieken, a. H., and Viavattene, C.: Review article: Assessing the costs of natural hazards – state of the art and knowledge gaps, Nat. Hazards Earth Syst.
15   Sci., 13, 1351–1373, doi:10.5194/nhess-13-1351-2013, 2013.

Morrow, B. and Phillips, B.: What's gender'Got to do with it'?, Int. J. Mass Emerg. Disasters, 17, 5–13, 1999.

Olsson, L., Opondo, M., Tschakert, P., Agrawal, A., Eriksen, S. H., Ma, S., Perch, L. N., and Zakieldeen, S. A.: Livelihoods and poverty, in: Clim. Chang. 2014 Impacts, Adapt. Vulnerability. Part A Glob. Sect. Asp. Contrib. Work. Gr. II to Fifth Assess. Rep. Intergov. Panel Clim. Chang., edited by Field, C. B., Barros, V. R., Dokken, D. J., Mach, K. J., Mastrandrea, M. D., Bilir, T. E., Chatterjee, M., Ebi, K. L.,
20   Estrada, Y. O., Genova, R. C., Girma, B., Kissel, E. S., Levy, A. N., MacCracken, S., Mastrandrea, P. R., and White, L. L., pp. 793–832, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2014.

Papathoma-Köhle, M., Kappes, M., Keiler, M., and Glade, T.: Physical vulnerability assessment for alpine hazards: state of the art and future needs, doi:10.1007/s11069-010-9632-4, 2011.

Rogelis, M. C., Werner, M., Obregón, N., and Wright, N.: Regional prioritisation of flood risk in mountainous areas, Nat. Hazards Earth Syst.
25   Sci., 16, 833–853, doi:10.5194/nhess-16-833-2016, 2016.

Rygel, L., O'sullivan, D., and Yarnal, B.: A Method for Constructing a Social Vulnerability Index: An Application to Hurricane Storm Surges in a Developed Country, Mitig. Adapt. Strateg. Glob. Chang., 11, 741–764, doi:10.1007/s11027-006-0265-6, 2006.

Schmidtlein, M., Shafer, J., Berry, M., and Cutter, S.: Modeled earthquake losses and social vulnerability in Charleston, South Carolina, Appl. Geogr., 31, 269–281, doi:10.1016/j.apgeog.2010.06.001, 2011.

30 Schmidtlein, M. C., Deutsch, R. C., Piegorsch, W. W., and Cutter, S. L.: A sensitivity analysis of the social vulnerability index, Risk Anal., 28, 1099–1114, doi:10.1111/j.1539-6924.2008.01072.x, 2008.

Statistik Austria: Gemeindeergebnisse der Abgestimmten Erwerbsstatistik und Arbeitsstättenzählung ab 2011, 2016.

Tapsell, S., Mccarthy, S., Faulkner, H., Alexander, M., Kuhlicke, C., Brown, S., Walker, G., Scolobig, A., De, B., and Bianchizza, C.: Social vulnerability to natural hazards, CapHaz-Net WP4 Rep., p. 524, 2010.

35 Tate, E.: Social vulnerability indices: A comparative assessment using uncertainty and sensitivity analysis, Nat. Hazards, 63, 325–347, doi:10.1007/s11069-012-0152-2, 2012.

Tate, E.: Uncertainty Analysis for a Social Vulnerability Index, Ann. Assoc. Am. Geogr., 103, 526–543, doi:10.1080/00045608.2012.700616, 2013.

Natural Hazards
and Earth System
Sciences
Open Access
Discussions

EGU

Thomalla, F., Downing, T., Spanger-Siegfried, E., Han, G., and Rockström, J.: Reducing hazard vulnerability: Towards a common approach
between disaster risk reduction and climate adaptation, Disasters, 30, 39–48, doi:10.1111/j.1467-9523.2006.00305.x, 2006.

Vincent, K.: Creating an index of social vulnerability to climate change for Africa, Tyndall Cent. Clim. Chang. Res., 56, 41, 2004.

Willis, I. and Fitton, J.: A review of multivariate social vulnerability methodologies: a case study of the River Parrett catchment, UK, Nat.

5   Hazards Earth Syst. Sci., 16, 1387–1399, doi:10.5194/nhess-16-1387-2016, 2016.

Yoon, D. K.: Assessment of social vulnerability to natural disasters: a comparative study, Nat. Hazards, 63, 823–843, doi:10.1007/s11069-
012-0189-2, 2012.