Sairam, N., Schröter, K., Lüdtke, S., Merz, B., Kreibich, H. (2019): Quantifying Flood Vulnerability Reduction via Private Precaution. - *Earth's Future*, *7*, 3, pp. 235—249.

DOI: http://doi.org/10.1029/2018EF000994

Earth's Future

RESEARCH ARTICLE
10.1029/2018EF000994

Key Points:
- Private precaution significantly reduces the flood vulnerability of private households as shown by robust empirical matching methods
- State-of-the-art flood damage models differ strongly based on their ability to capture differences in vulnerability of private households
- Methodology applied and validated using an extensive object-level flood damage data set from Germany

Supporting Information:
- Supporting Information S1

Correspondence to:
N. Sairam,
nivedita@gfz-potsdam.de

Citation:
Sairam, N., Schröter, K., Lüdtke, S., Merz, B., & Kreibich, H. (2019). Quantifying flood vulnerability reduction via private precaution. *Earth's Future*, *7*, 235–249. https://doi.org/10.1029/2018EF000994

Received 23 JUL 2018
Accepted 30 JAN 2019
Accepted article online 4 FEB 2019
Published online 4 MAR 2019

Quantifying Flood Vulnerability Reduction via Private Precaution

Nivedita Sairam[1,2], Kai Schröter[1], Stefan Lüdtke[1], Bruno Merz[1,3], and Heidi Kreibich[1]

[1]Section Hydrology GFZ German Research Centre for Geosciences, Potsdam, Germany, [2]Geography Department, Humboldt University, Berlin, Germany, [3]Institute of Earth and Environmental Science, University of Potsdam, Potsdam, Germany

Abstract Private precaution is an important component in contemporary flood risk management and climate adaptation. However, quantitative knowledge about vulnerability reduction via private precautionary measures is scarce and their effects are hardly considered in loss modeling and risk assessments. However, this is a prerequisite to enable temporally dynamic flood damage and risk modeling, and thus the evaluation of risk management and adaptation strategies. To quantify the average reduction in vulnerability of residential buildings via private precaution empirical vulnerability data ($n = 948$) is used. Households with and without precautionary measures undertaken before the flood event are classified into treatment and nontreatment groups and matched. Postmatching regression is used to quantify the treatment effect. Additionally, we test state-of-the-art flood loss models regarding their capability to capture this difference in vulnerability. The estimated average treatment effect of implementing private precaution is between 11 and 15 thousand EUR per household, confirming the significant effectiveness of private precautionary measures in reducing flood vulnerability. From all tested flood loss models, the expert Bayesian network-based model BN-FLEMOps and the rule-based loss model FLEMOps perform best in capturing the difference in vulnerability due to private precaution. Thus, the use of such loss models is suggested for flood risk assessments to effectively support evaluations and decision making for adaptable flood risk management.

Plain Language Summary Private precautionary measures such as adapted building use, sealing basements and purchasing flood barriers reduce flood damage to residential buildings. Using an empirical dataset consisting of 948 flooded households in Germany, we estimate that the average loss reducing effect of implementing private precautionary measures is 11-15 thousand EUR per household. This is approximately equal to 27% of the average building loss suffered by the flooded households (48000 EUR). Despite this significant risk mitigation effect, these precautionary measures are hardly considered in flood risk assessment modelling. This results in biased flood loss predictions being used for evaluating risk management strategies. Hence, we compare state-of-the-art flood loss models in respect to their ability to account for building loss reduction due to private precaution. From all tested flood loss models, the expert Bayesian Network based model BN-FLEMOps and the rule-based loss model FLEMOps are best able to capture the damage reducing effect of private precaution. These models can be valuable for evaluating adaptable flood risk management strategies.

1. Introduction

An integrated approach toward flood risk management is conceptualized and accepted in many countries worldwide (Merz, Kreibich¸ et al., 2010). These concepts consider that flood defenses might fail and thus complement flood protection with nonstructural solutions, for example, private precaution, land use planning, and insurance (Bubeck et al., 2017; Kreibich et al., 2015; Kunreuther et al., 2009). Burby et al. (1988) revealed that floodplain management is able to divert development away from floodplains and reduce potential flood damage. According to section 5 of the German Federal Water Resource Act, it is the obligation of every person who is prone to flood risk to undertake appropriate actions that are reasonable and within one's means (Rolfsen, 2009). Reliable flood risk and cost-benefit analyses are essential for efficient risk management, since they support optimum investments in adaption measures. Cost-benefit analyses need to consider all suitable risk mitigation measures, associated costs, and expected flood losses, since incomplete accounting of costs and benefits, for example, only structural measures considered, will lead to a deviation from the global optimum in the analyses (Kreibich et al., 2014). The economic damage from floods has been increasing over the last decades, mostly due to societal factors such as increased standard of living, real per capita wealth, and population increase (Barredo, 2009; Mechler & Bouwer, 2015), and this trend is likely to continue (Intergovernmental Panel on Climate Change (IPCC), 2012; Jongman et al., 2014). Assessments need to account for this dynamic nature of risk to be able to detect relevant changes in risk and initiate appropriate adaptation to changes (Kreibich et al., 2014). Thus, there is a need to accurately estimate flood risks over long time periods. To be able to capture temporal dynamics in flood loss and risk, which is also a prerequisite to enable evaluations of risk management and climate adaptation strategies, loss models that are able to account for differences in vulnerability, for example, due to private precaution, are necessary.

Flood risk is influenced by a broad range of characteristics and processes, which can be categorized into hazard, exposure, and vulnerability (IPCC, 2012). Understanding the role of these components for changes in risk is essential for effective adaptation. Few studies are available, which investigate the role of vulnerability, using modeling (Jongman et al., 2015; Mechler & Bouwer, 2015) or empirical (Kreibich, Botto, et al., 2017) approaches. There are various definitions of *vulnerability*, and many vulnerability concepts consider a quite broad context (e.g., Brooks et al., 2005; Kelly & Adger, 2000; Turner et al., 2003). Additionally, there are suggestions to complement the concept of vulnerability with resilience, which adds considerations of recovery (Bruneau et al., 2003; de Bruijn, 2004; Fekete et al., 2014). For our study, we follow the natural sciences oriented approach, which defines vulnerability as the characteristic of a system that describes its potential to be harmed (Gouldby et al., 2005; IPCC, 2012). Thus, vulnerability is the susceptibility of a household to flooding, which is altered by precautionary measures as well as by changes in household or building characteristics (Few, 2003).

Precautionary measures that are commonly implemented among private households to reduce residential building loss include waterproof sealing, flood adapted use, and flood adapted interior fitting (Kreibich et al., 2005, 2015). It is generally assumed that precautionary measures are effective in mitigating flood losses (De Moel et al., 2014; Dutta et al., 2003; Holub & Fuchs, 2008), and also some empirically based quantitative information is available: The positive effect of private precautionary measures was revealed by loss reductions of 35% and up to 50% between two similar flood events in 1993 and 1995 at the Meuse and the Rhine Rivers, respectively, where many households had undertaken precautionary measures after the flood in 1993 (Bubeck et al., 2012; Wind et al., 1999). Some studies quantified the damage-reducing effect of individual precautionary measures and identified the most effective ones: These include scientific studies based on empirical damage data (e.g., Hudson et al., 2014; Kreibich et al., 2005; Poussin et al., 2015) as well as practical studies based on expert judgment and/or a rather not transparent database (e.g., ABI, 2003; Defra, 2008; ICPR, 2002). A study in France identified *elevating the ground floor* to be the most effective measure in reducing the damage to buildings by up to 5,500€ and to home contents by up to 6,500€ (Poussin et al., 2015). Studies in Germany identified the measures *flood adapted use* and *flood adapted interior fitting* as the most effective precautionary measures with building loss reductions of about 50% or in terms of absolute loss reductions of over 10,000€ (Hudson et al., 2014; Kreibich et al., 2005). On the other hand, cost-benefit analyses revealed low-priced measures like *elevating the boiler* and *securing the oil tank* as the most cost-effective ones (Kreibich et al., 2011, 2012; Poussin et al., 2015). Depth-damage curves were developed for different types of flood proofing adaptations through flood and exposure simulations (Dawson et al., 2011).

Estimating the damage-reducing effect of precautionary measures from observed flood loss data should consider the possible bias due to confounding variables. One approach was to estimate the difference in average flood loss experienced by households with precaution and households with no precaution, while controlling for similar inundation depth (Kreibich & Thieken, 2009) or inundation depth and building characteristics (Kreibich et al., 2011). This approach to analyze controlled household groups faces two challenges: (1) controlling for hazard and building variables results in small samples that can be used for further analysis and (2) controlling the influence of a large number of variables is not feasible. In order to overcome these challenges, Poussin et al. (2015) developed a regression-based method to determine the effectiveness of individual precautionary measures by controlling for the effects of potential flood risk variables. Another suitable approaches to control for confounding variables are matching techniques, since they tests causal inference with fewer assumptions than typical regression models, using a smaller, preprocessed data set (Rosenbaum & Ruby, 1983). For instance, Aldrich (2012) used five different methods of matching on propensity scores, that is, kernel, radius, nearest neighbor, nearest neighbor with replacement, and Mahalanobis matching, to investigate the influence of social capital on the pace of population recovery following the 1923 Tokyo earthquake. Allaire (2016) tested the effectiveness of online information and social media in enabling households to reduce disaster losses using propensity score matching (PSM); that is, nearest neighbor and kernel matching was undertaken followed by a postmatching regression analyses. That is, the average treatment effect (ATE) was estimated using the matched sample to run postmatching regression of the outcome on covariates that are associated with flood losses, but not necessarily the likelihood of using social media. Hudson et al. (2014) implemented expert-selected PSM to quantify the treatment effect of different precautionary measures on building and content losses. This method is able to control for an extensive set of variables, that is, all variables that are likely to introduce selection bias. Our study builds on these approaches to determine the ATE of private precaution in general (not focused on individual measures) by matching based on confounders of private precaution and applying postmatching regression controlling for variables describing flood hazard, warning, and emergency measures.

Various flood loss models have been developed for estimating direct economic loss to buildings (Carisi et al., 2018; Merz, Hall, et al., 2010; Schröter et al., 2014; Smith, 1994). Many models represent the loss in terms of relative loss, which is the ratio between costs of loss and the value of asset at the time of the event. A standard approach is depth-damage functions that model the loss as a function of one variable, that is, inundation depth, commonly differentiated according to the building type or use (Grigg & Helweg, 1975; Penning-Rowsell et al., 2005; Smith, 1994; White, 1964). Recently, multivariable flood loss models have been developed. For instance, FLEMOps+r (Elmer et al., 2010) is a rule-based model to estimate flood loss to residential buildings based on five different classes of water depth, three individual building types, two classes of building quality, three classes of flood frequency, three classes of contamination, and three classes of private precaution. Further, more complex models, based on machine learning algorithms and covering various aspects of flood damage processes, are being developed. Examples are multivariable tree-based models (Hasanzadeh Nafari et al., 2016; Kreibich, Di Baldassarre, et al., 2017; Merz et al., 2013). They do not require any special treatment for discrete and continuous variables and no specific prior assumptions about the distributions of variables. Bagging decision tree is an ensemble approach with a number of individual trees. The loss estimate is then determined using the mean as the prediction of the ensemble of trees. Also, Bayesian networks are used in flood loss estimation (Schröter et al., 2014; Vogel et al., 2014; Wagenaar et al., 2018). Bayesian networks are Directed Acyclic Graphs (DAG) constructed from assertion of dependencies and principle of conditional independence (Heckerman, 1998). They have the advantage of inherently quantifying uncertainty associated with the loss estimation. Thus, a variety of models with varying complexities and working concepts are currently available, and it is not trivial to decide which one to use for a specific application (Apel et al., 2009; de Moel et al., 2015; Figueiredo et al., 2018). Several studies have tested and compared various flood loss models in respect to their predictive accuracy and reliability (e.g., Cammerer et al., 2013; Gerl et al., 2016; Hasanzadeh Nafari et al., 2016; Jongman et al., 2012). In contrast, to the best of our knowledge, no study so far has examined the ability of loss models in capturing differences in vulnerability due to private precaution. However, loss models with this ability are necessary to enable temporally dynamic flood damage and risk modeling and thus the evaluation of risk management and adaptation strategies.

Hence, our study aims at quantifying the average loss-reducing effect of private precaution, by taking into account possible biases due to confounding variables, and to assess how well different types of flood loss estimation models are able to represent this difference in vulnerability.

2. Data and Methods

2.1. Description of Data Set

The data set contains flood loss data collected via cross-sectional telephone surveys of private households that had suffered from losses due to floods in 2002, 2005, 2006, 2010, 2011, or 2013 mainly in the Elbe and Danube catchments in Germany (Table 1). On basis of flood reports, press releases, and flood masks derived from satellite data (www.zki.dlr.de), lists of inundated streets were compiled separately after one or two flood events. On basis of these lists, property-specific random samples of potentially affected households; that is, their telephone numbers were selected from the public telephone directory. Property-specific means that only one household was interviewed per address. Computer-aided telephone interviews were undertaken in independent campaigns in April/May 2003, in November/December 2006, in February/March 2012,

10.1029/2018EF000994

Table 1
Flood Surveys: Computer-aided Telephone Interviews With Private Households Who Suffered Flood Loss

Characteristics	Surveys			
Date of survey:	April/May 2003	November/December 2006	February/March 2012	February/March 2014
Flood(s):	August 2002	August 2005, April 2006	August 2010, January 2011	June 2013
Affected regions	Elbe and Danube catchments	Elbe and Danube catchments	Elbe, Oder, and Rhine catchments	Elbe, Danube, Rhine, and Weser catchments
Number of households interviewed:	1697	461	658	1652
Response rate	15%	18%	16%	17%
Sampling type	random	comprehensive	comprehensive	comprehensive
References	Thieken et al. (2007)	Kreibich et al. (2011)	Kienzler et al. (2015)	Kreibich, Di Baldassarre, et al. (2017)

and in February/March 2014 (Table 1). In 2003, households from the list of telephone numbers were sampled randomly. In the following campaigns, comprehensive surveys were conducted; that is, all the researched telephone numbers were contacted. Each interview was focused on one specific flood event. At the beginning of the interview, it was asked if the household had suffered losses due to the specific flood event; the interview was only continued if this was the case. Thus, the data set does not contain cases, where the precautionary measures fully prevented loss. This limitation of the data set is considered when interpreting the results. At the beginning of the telephone call, the person on the phone was asked who in the household has the best knowledge about the flood event and the incurred economic losses. Then the interview was undertaken with this person. The questionnaires for all the survey campaigns contained about 180 questions including aspects of hazard (e.g., inundation depth, duration, and velocity), flood experience and awareness, early warning, emergency and precautionary measures, building and socio-economic characteristics, and building and content losses. Building loss includes all costs (e.g., costs of wages and material) that are associated with repairing the damage caused by floods to the building structure. Damage may be due to moisture penetration as well as cracks, pushed in doors and windows, subsidence, or deformation of walls and ceilings, etc. Repair works may include, for instance, plastering, laying bricks, replacing construction components, or broken windows. Building losses are adjusted to prices as of 2013 (inflation) by adjusting the reported loss estimates given at the time of the events by the building cost index (DESTATIS, 2013). The losses reported by the surveyed households were believed to be reliable, since most people had restored their building by the time of the survey (except for after the 2002 flood; Kienzler et al., 2015) and had claimed their losses either from government funds or from their insurers. The responses from the survey after the 2002 flood was confirmed by comparing it with official loss data provided by the Saxon Bank of Reconstruction, which looked after providing governmental disaster assistance after the 2002 flood in the federal state of Saxony (Thieken et al., 2005). Nevertheless, data collected via surveys is associated with uncertainty, which is however difficult to quantify since hardly alternative means to measure these variables exist. The building loss ratio was calculated consistently for all surveys as follows: the absolute losses reported by the surveyed households are divided by the building values as at the time of the flood event. Actuarial valuation method VdS guideline 772 1988-10 (Dietz, 1999), which is commonly used in the insurance sector for Germany was used to estimate absolute values of residential building in terms of replacement costs (in contrast to depreciated values). In order to apply this valuation, some attributes from the survey responses such as total floor space, basement area, number of storeys, and roof type are used. In respect to precautionary measures people were asked about the kind of measure (check list and additional open answers possible and multiple answers possible) and the time of realization (check list: *undertaken before the flood*, *after the flood*, *planned within the next 6 months*, and *not intended*). The check list contained among others the following building precautionary measures: adapt interior fitting, adapt use, and adapt building structure. Adapting interior fitting involves using less expensive fittings that are easily replaceable or preferably water proof fittings in lower floors; adapting usage to floods means for instance to use the flood endangered floors in a low-value way; Adapting building structure to floods include structural measures like sealing the basement. These measures are also sometimes referred to as passive preparedness measures (Cumiskey et al., 2018) undertaken often after flood events during the reconstruction phase, however, always much before an event. Thus,

Table 2
List of Potential Confounders of Private Precaution

Categories	Attributes	Type	Attribute explanation-range, unit
Building characteristics Building characteristics may induce limitations or technical feasibility to be able to undertake some precautionary measures (Cumiskey et al., 2018).	Building quality (bq)	ordinal	1 – very good; 6 – very bad
	Building area (ba)	continuous	[24,299997] sq. meters
	Single-family house (bt1)	dichotomous	0 – no, 1 - yes
	Multi-family house (bt2)	dichotomous	0 – no, 1 - yes
	Building value corrected for inflation 2013 (bv)	continuous	[98496, 10411183] EUR
	Number of flats in the building (nfb)	continuous	(1,45) flats
Socio-economic attributes People from varying socio-economic groups vary in aspects like sense of responsibility, willingness to respond and ability to invest in mitigation measures (Bubeck et al., 2012; Cumiskey et al., 2018).	Ownership – Apartment (own_1)	dichotomous	0 - tenant, 1 – apartment owner
	Ownership – building (own_2)	dichotomous	0 – not building owner, 1 – building owner
	Age of the interviewee (age)	continuous	[16,99] years
	Household size (hs)	continuous	[1,20] persons
	Household monthly net income indicator (inc)	ordinal	11 = below 500 EUR to 16 = 3,000 EUR and more
Flood experience and awareness • Flood experience and strong social networks improve awareness about hazard and coping appraisal (Atreya et al., 2017; Bubeck et al., 2013; Kreibich et al., 2005; Parker et al., 2007).	Knowledge about flood hazard (kh)	dichotomous	1 – Has sufficient knowledge, 0 – Has no knowledge
	Flood experience (fe)	ordinal	0 – no flood experience, 9 – recent flood experience
	Neighborhood preparedness programs (neigh_ind)	dichotomous	1 – participated in neighborhood programs, 0 – not participated in neighborhood programs
	Flood insurance (ins_ind)	dichotomous	1 – Has flood insurance, 0 – Has no flood insurance
	Perceived effectiveness of private precaution (epre)	ordinal	1 = very efficient to 6 = not efficient at all
	Event	nominal	Flood events in 2002, 2005, 2006, 2011, and 2013

precautionary measures are not dependent on event forecast and early warning information, in contrast to emergency measures. The questionnaire included also questions that reveal the perception of the interviewee regarding aspects like effectiveness of precautionary measures, usefulness of early warning information, and the quality of their building. People were asked to assess these qualitative variables on a scale from 1 to 6; the meanings of the end points of the scales were given to the interviewee. Indicators were developed for some variables such as flood experience, emergency measures, and warning information. Variables used in this study are described in Tables 2 and 3. The corresponding questions, possible options for answers, and score computation for indicators are included in supporting information S.I.1 (S.I.1.1 - S.I.1.3). Further details about the development and calculation of indicators are given in Thieken et al. (2005) and Elmer et al. (2010). More information about the individual flood events, the surveys, and their results were published in Thieken et al. (2007), Kreibich et al. (2011), Kreibich, Botto, et al. (2017), and Kienzler et al. (2015). A total of 4,468 interviews were completed, of which 2,671 interviews furnished building loss in EUR. If one or more of the precautionary measures are not practically applicable for a particular household, this data set is not included in the analysis. For example, households with no basement/cellar are not potential candidates for all structural adaptation measures (e.g., sealing the basement). Hence, the households with no basement are removed from the analysis. Since the methodology does not deal with missing variables, households with incomplete data are removed. Thus, data consisting of 974 households with complete observations are available for the analyses.

2.2. Difference in Vulnerability Between Households With Respect to Private Precaution

2.2.1. ATE Considering Selection Bias

A dichotomous indicator (0/1) is used to distinguish private households into low/high vulnerability with respect to implementation of precautionary measures (pre). Private precautionary measures considered are adapt interior fitting, adapt use, and adapt building structure. The indicator for private precaution takes a value of 1 for households with one or more of these precautionary measures implemented before the flood (treatment group) and 0 for households with none of these measures implemented before the flood (control group). Actually, many of the households have undertaken several precautionary measures, which differ in their way how they mitigate flood damage to the building structure and function jointly in the case of a flood event.

The average effect of private precaution in reducing building structure losses in EUR, referred to as the ATE, contributes to the differences in vulnerability between the two groups. ATE is estimated using the Roy-Rubin model (Roy, 1951; Rubin, 1974):

$$\mathrm{ATE}(T) = E[Y(T=1) - Y(T=0)] \tag{1}$$

where T is the treatment—implementation of one or more private precautionary measures (1/0). Y is the outcome that is influenced by the treatment, that is, the reported building loss in EUR.

Considering the heterogeneity among the households with respect to building characteristics, socio-economic attributes, flood experience, and awareness, the observed difference in losses between the two groups may not be necessarily only due to the effect of private precaution. This is due to the fact that the households from treatment and control groups have different probabilities of undertaking private precaution. The attributes that influence a household to undertake private precaution are the confounding variables or confounders of private precaution (Table 2). The bias in ATE caused due to the effect of confounding is called selection bias. Matching households from treatment and control groups based on the sufficient set of confounders provides an appropriate solution to get rid of selection bias. It is important to only include pretreatment variables to the list of confounders, whose measurement is not altered by the implementation of private precaution (Pearl, 2009). Equation (1) is altered to equation (2), where the building loss estimate is conditioned on the treatment variable, that is, private precaution, as well as the set of confounding variables.

$$\mathrm{ATE}(\mathrm{pre}) = E(\mathrm{loss}|\mathrm{pre}>0, X) - E(\mathrm{loss}|\mathrm{pre}=0, X) \tag{2}$$

where ATE(pre) is the treatment effect of implementing private precaution; loss is the reported building loss of households (EUR), and X is the set of confounding variables.

From direct answers to interview questions and derived indicators described in section 2.1, we choose 16 attributes that potentially influence whether a household undertakes private precaution (Table 2). These attributes are potential pretreatment confounders. They are categorized into building characteristics, socio-economic attributes, flood experience, and awareness. In order to remove hidden bias due to unaccounted variation in the characteristics of different flood events that lead to selection bias, we include *event* as a nominal covariate in the set of confounders.

2.2.2. Matching Distances and Methods

There are a number of matching methods and distance estimates that can be used to eliminate selection bias and obtain a matched data set. We test the suitability of two distance estimates: (1) PSM and (2) Mahalanobis distance matching (MDM).

PSM has been used widely in socio-economics and medical studies (Dehejia & Wahba, 1999; Vincent et al., 2002). Propensity score is the probability that a particular household will undertake precautionary measures, conditioned on the set of confounding variables (equation (3)).

$$p_i \equiv P(T_i = 1|X) = \frac{1}{(1 + \mathrm{e}^{X_i \beta})} \tag{3}$$

where p_i is the propensity score of ith household in the data set obtained through linear logistic regression, T is the private precaution indicator (Treatment),X is the set of confounding variables, β is the set of regression coefficients, and base e denotes the exponential function. The distance between matched households from the two groups is estimated as the scalar difference between their propensity scores. The common support for propensity scores is determined using equation (4). Only households that lie in the range of common support are considered for matching.

$$\text{Common Support} = [\max(\min(P_t, P_c)), \min(\max(P_t, P_c))] \tag{4}$$

where P_t and P_c are propensity scores of households with private precaution and with no precaution, respectively.

MDM is a covariate matching method. It uses the Mahalanobis distance as the distance estimate. The Mahalanobis distance matrix is furnished using the distance estimates between pairs of households from the two groups, with the set of confounders as covariates (equation (5)).

$$M(X_i, X_j) = \sqrt{\left[(X_i - X_j)^T \mathscr{E}^{-1} (X_i - X_j)\right]} \tag{5}$$

where $M(X_i, X_j)$ is the Mahalanobis distance estimate between two households i and j based on the set of confounders X, X_i and X_j represent column matrices of values of confounders from treatment and control households, $(X_i - X_j)^T$ denotes the transpose of the matrix $(X_i - X_j)$ resulting in a row matrix, and $\mathscr{E}^{-1}$ is the inverse covariance matrix. This results in $M \times N$ Mahalanobis distance matrix (where M and N are the numbers of households in treatment and control groups from the original population).

Once the distance estimates are obtained, different methods of matching (Ho et al., 2007) are tested: (1) nearest neighborhood (NN) with/without replacement, with/without caliper; (2) inverse probability treatment weighting; and (3) genetic matching algorithm (Diamond & Sekhon, 2006).

Small pruning threshold/caliper is required to reduce bias while matching. We consider 1/4th standard deviation of the PSM and Mahalanobis distances as the caliper to remove unsuitable matches, since it reduces the imbalance by at least 90% (Rosenbaum & Rubin, 1985). From the two distance estimates and six matching methods, 12 potential matched data sets are obtained.

2.2.3. Quality of Matching

Two tests are performed to assess imbalance in individual confounders after matching: the two sample weighted t test and the standardized differences test. The potential matched data sets that pass the two tests for all confounders are chosen for the estimation of ATE.

The two sample weighted t test evaluates whether the distributions of confounders belonging to treatment and control households are significantly different (Rosenbaum & Rubin, 1985). In the standardized differences test (equations [(6a)] and [(6b)]), an absolute value of standardized difference less than 10% for each of the confounders belonging to treatment and control households is considered to be an accurate match (Austin & Mamdani 2006). It is a point estimate with no significance limits attached to it.

$$\text{Standardised difference (continuous)} = \frac{(\bar{x}_T - \bar{x}_C)}{\sqrt{\frac{(s_T^2 + s_C^2)}{2}}} \tag{6a}$$

where $\bar{x}$ is the covariate mean of treatment (T) and control (C) groups; s is the covariate standard deviation of treatment (T) and control (C) groups.

$$\text{Standardised difference (dichotomous)} = \frac{(p_T - p_C)}{\sqrt{\frac{[p_T(1-p_T) + p_C(1-p_C)]}{2}}} \tag{6b}$$

p is the sample prevalence (proportion of TRUE (value = 1) in the sample of a dichotomous variable) of the covariate in treatment (T) and control (C) groups.

2.2.4. Postmatching Regression and Sensitivity Analysis

Postmatching regression/model fitting is performed in order to control for the bias in treatment effect introduced by aspects that influence the outcome (building loss) but do not potentially influence the treatment

Table 3
Attributes for Postmatching Regression

Attributes	Attribute Type	Attribute explanation - range, unit
Inundation depth (wst)	continuous	[−245,674] cm
Duration of inundation (d)	continuous	[0,1440] hr
Contamination (con)	ordinal	0 – no contamination to 2– heavy contamination
Velocity of water (v)	dichotomous	0: $v = 0$, 1: $v > 0$
Emergency measures (em)	ordinal	1 = no measures undertaken to 17 = many measures undertaken
Warning information (wi)	ordinal	0 = no helpful information to 12 = many helpful information

(implementation of private precaution). Varying flooding intensities across different households influence the degree of damage experienced. Further, emergency response measures such as pump out water, use sandbags/barriers, and switch-off electricity and gas also potentially reduce flood damage. Unlike precautionary measures, the implementation of emergency measures are highly dependent on event forecast and early warning. Hence, in addition to the matching procedure, which controls for the pretreatment variables (Table 2), potential bias due to flooding situation, emergency measures, and warning information (Table 3) is removed via postmatching regression. The choice of postmatching regression model depends on the ability of the model to account for the influence of flooding scenario, early warning and emergency measures on incurred loss. A standard linear regression model is commonly used to remove bias in postmatching. In addition to linear regression, bagging decision trees (ensemble of 1,000 regression trees) are used as the postmatching regression model due to its ability to predict losses with least errors compared to standard linear regression models (Merz et al., 2013). Bagging decision trees are an ensemble of regression trees built on bootstrapped samples of the data such that model dependency and overfitting are reduced. Bagging decision trees approximate nonlinear regression to heterogeneous data. Using the matched samples that pass the quality check, regression models (linear and bagging decision trees), is built for predicting building loss (in EUR) using the treatment variable (private precaution - pre) and predictors from Table 3. Two intervention scenarios—treatment (pre = 1) and control (pre = 0)—are applied to the model and the loss estimates (in EUR) are determined for each scenario. The difference between the two groups of model estimates result in ATE of private precaution. The survey questions and score computations corresponding to the variables for postmatching regression are included in supporting information S.I.1 (S.I.1.4).

The ATE estimate may still be sensitive to the choice of confounders (Caliendo & Kopeinig, 2008). Since the list of confounders is chosen through expert knowledge, there is a possibility that some aspects of confounding may be missing or unmeasured. The matching methodology cannot eliminate potential bias due to unobserved or missing confounders. Potential unobserved or missing confounders that the analysis does not control for may be specific building or contents characteristics, which may favor or hamper certain building precautionary measures or differences in the ability of households to undertake measures. Rosenbaum's sensitivity analysis (Rosenbaum, 2002) using Hodges-Lehmann point estimate quantifies the robustness of the causal relationship between treatment (precaution) and outcome (building loss) to the presence of bias introduced by missing confounders (DiPrete & Gangl, 2004).

Two households, which are matched based on the set of confounders, may vary in the probability of undertaking private precaution by at most a factor of Γ (sensitivity parameter). If $\Gamma = 1$, the two groups of matched households have the same probability of undertaking precaution (no hidden bias). If $\Gamma = 2$, the matched households in the treatment group may have at most twice the probability of undertaking precautionary measures when compared to the households in the control group. When Γ is increased from 1.0, the bounds of Hodges-Lehmann point estimate (Rosenbaum's bounds) widen and the certainty with which we estimate the treatment effect decreases. The robustness of the estimate is represented by the value of Γ,

10.1029/2018EF000994

Table 4
Summary of the Flood Loss Estimation Models

Model	Variables	Type
FLEMOps+r (Elmer et al., 2010)	Inundation depth, return period, building value, building type, building quality, precautionary, and contamination indicators	Point estimate
Regression Trees: RT-FLEMOps (Merz et al., 2013)	inundation depth, return period, duration of inundation, flood experience, precautionary measure, building area, and building type	Point estimate
Bagging Decision Trees: BT-FLEMOps (Kreibich, Botto, et al., 2017)	inundation depth, return period, duration of inundation, flood experience, precautionary measure, building area, and building type	Point estimate (ensemble approach)
Bayesian Networks: BN-FLEMOps (Wagenaar et al., 2018)	inundation depth, return period, duration of inundation, flood experience, precautionary measure, building area, and building type	Distribution function

when the Rosenbaum's bounds extend further from the positive effect of treatment and bracket to zero (Keele, 2010).

2.3. Ability of Flood Loss Estimation Models to Capture Differences in Vulnerability Due to Private Precaution

2.3.1. Flood Loss Estimation Models

A range of flood loss estimation models are applied to the matched data set to test to which extend the models are able to capture differences in vulnerability due to private precaution. The models are of varying complexities from deterministic rule-based models to probabilistic Bayesian network-based models. All these models estimate the relative building loss (brloss) for private household buildings using multivariable predictors from the surveys. The brloss values range between 0 (no loss) and 1 (total loss). From the brloss estimates, the absolute losses are computed by multiplying with the building value of the respective private buildings (in EUR) corrected to 2013 inflation.

FLEMOps+r (Elmer et al., 2010; Thieken et al., 2008) estimates relative building losses based on defined rules associating seven input variables (Table 4) to relative building loss. FLEMOps+r works in two steps: first, relative flood loss is estimated on basis of water level and building characteristics (i.e., building type and building quality); second, the estimate is refined by a scaling factor which considers contamination (in three classes, i.e., no, medium, and heavy), precaution (in three classes, i.e., little, medium, and strong), and recurrence interval (in three classes, i.e., 1–9 years, 10–99 years, and from 100 years onward). In this model, private precaution takes a value of 0 for little precaution, 1 for medium precaution, and 2 for strong precaution.

Tree-based models (Kreibich, Botto, et al., 2017; Merz et al., 2013), that is, regression trees (RT-FLEMOps) and bagging decision trees (BT-FLEMOps), are grown with seven variables (Table 4). RT-FLEMOps is grown with a minimum of 60 households in each leaf, resulting in 25 leaves (Figure 1a). In RT-FLEMOps, the precautionary measure indicator appears only once in the bottom of the tree, and hence, the variable does not hold an important role in estimating relative building losses. BT-FLEMOps is an ensemble approach consisting of 1,000 trees. The variable importance plot (Figure 1b) shows that private precaution has a relatively low importance. The tree based algorithms are developed using Statistics and Machine Learning toolbox.

BN-FLEMOps (Wagenaar et al., 2018) is a discrete Bayesian network model, which is constructed with seven variables (Table 4). The continuous variables in the model were discretized on the basis of bins with equal frequency with inundation depth (wst) and relative building loss (brloss) in 10 classes, return period (rp) and inundation duration (d) in five classes, and building area (ba) in three class. The network structure (Figure 2) describing the conditional dependencies between the variables is learnt using 500 iterations of score-based local search algorithms—Fast-IAMB (Tsamardinos et al., 2003) and a hill-climbing approach using the Bayesian Dirichlet Equivalent (Heckerman et al., 1995). The set of network structures and all arcs that occurred at least in 80% of all iterations

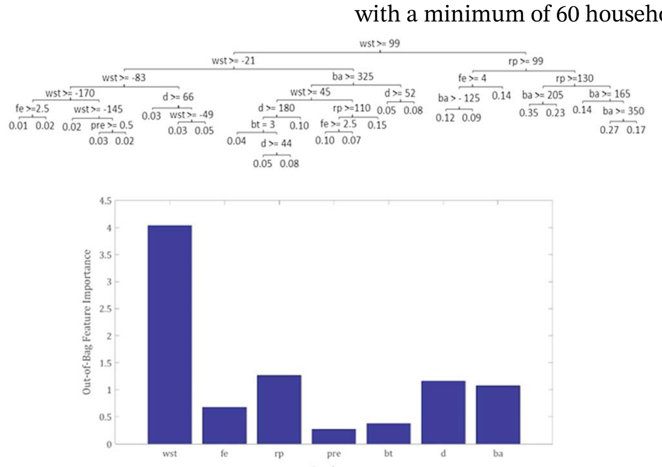

Figure 1. (a) Regression tree with seven variables and 25 leaves (RT1). (b) Feature importance of flood loss predictors for Bagging decision trees BT with seven variables and an ensemble of 1,000 trees.

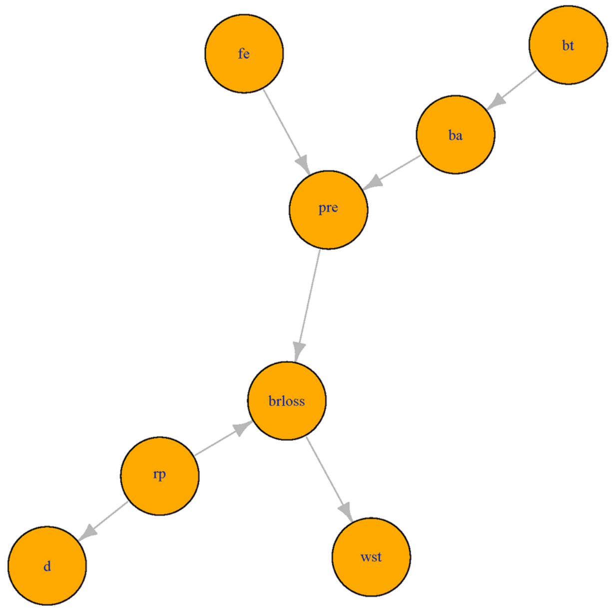

Figure 2. Structure of the Bayesian network: BN-FLEMOps (Wagenaar et al., 2018).

provided the basis to define the network used. Relative building losses are estimated as the medians of the conditional probability distributions of the *brloss* node in the network. The discrete Bayesian network is derived using *bnlearn package, R version 3.3.1* (R Core team, 2016; Scutari, 2009).

2.3.2. Model Performance-Loss Estimation and Vulnerability Differences

The performance of the tested flood loss estimation models are evaluated using (1) accuracy of the models in estimating flood losses to buildings and (2) vulnerability differences due to private precaution accounted by the models.

Two point estimate accuracy indictors—root-mean-square error and mean bias error) from 1,000 bootstrap iterations of the overall sample of households are used for assessing accuracy in loss estimation. The influence of vulnerability differences due to private precaution on the model outcome is captured by introducing an intervention for private precaution, that is, forcing the model to consider two scenarios: (1) pre > 0: all households have implemented one or more private precaution measures (treatment), and (2) pre = 0: all households have no precaution (control). The scenarios are applied to determine the model loss estimates for the matched households. The differences between the averages of the loss estimates obtained from the two scenarios are the differences in vulnerability due to private precaution, captured by the models.

$$\begin{aligned}&\text{Difference in vulnerability accounted by loss models (pre)}\\&\quad = E(\text{loss estimate }|\text{pre}>0) - E(\text{loss estimate}|\text{pre} = 0)\end{aligned} \tag{7}$$

where pre represents the precautionary measure indicator of respective models.

3. Results and Discussion

3.1. Matching Households With and Without Private Precaution

In order to determine the effectiveness of private precaution in mitigating building loss, the data are controlled for heterogeneity due to potential confounding variables. We use pretreatment variables pertaining to the households (Table 2) for removing selection bias from the survey data set and then perform post-matching regression using variables pertaining to the flooding and response scenarios (Table 3). 948 households with no missing confounding variables undergo the matching procedure. Households with propensity scores in the common support region (equation (4)) [0.07, 0.93] between treatment and control groups are

Table 5
PS - Summary of Overall and Matched Data Sets

Data Set	Criterion (Precaution)	Sample size	Min	Median	Mean	Max
Overall	Treatment	454	0.07	0.64	0.61	0.99
	Control	494	0.05	0.31	0.36	0.93
Households in common support	Treatment	425	0.07	0.61	0.58	0.92
	Control	491	0.08	0.32	0.36	0.93
PSM-NN with caliper (households in matched data set)	Treatment	248	0.07	0.46	0.46	0.92
	Control	248	0.09	0.45	0.46	0.93
PSM-NN with caliper and with replacement	Treatment	352	0.07	0.56	0.56	0.92
(households in matched data set)[a]	Control	203	0.08	0.55	0.56	0.93
PSM-genetic matching (households in matched data set)[a]	Treatment	425	0.07	0.61	0.58	0.92
	Control	197	0.08	0.59	0.56	0.90

[a]Estimates are adjusted for weights created during matching.

Table 6
ATE Estimates From Matched Data Sets

	ATE estimate from matched data sets in EUR		
Postmatching regression models	PSM-NN with caliper	PSM-NN with caliper and with replacement	PSM-genetic matching algorithm
No postmatching regression model	−26097 (6372)	−29305 (6639)	−16474 (5304)
Linear regression	−17025 (5713)	−21850 (5750)	−14330 (4541)
Bagging decision trees	**−12217 (2608)**	**−15053 (2947)**	**−11238 (2348)**

considered for matching. This results in 32 households outside the common support and 916 households within the common support, which are considered for PSM.

PSM and MDM distance estimates for selection bias combined with six different methods of matching (section 2.2.2) result in 12 potential matched data sets. The following three data sets pass the quality checks for suitable matches (<10% standard error and insignificant bias, as described in section 2.2.3): (1) PSM-NN with caliper and no replacement, (2) PSM-NN with caliper and with replacement, and (3) PSM-genetic matching.

The summary statistics of propensity scores of households from the overall data set, common support, and suitable matched data sets are provided in Table 5. In supporting information S.I.2 (Table S.I.2.1), the imbalance in covariates before and after matching is summarized.

3.2. Differences in Vulnerability Due to Private Precaution-Empirical Estimate

Vulnerability reduction of households due to private precaution is estimated as the ATE of private precaution undertaken. The ATE estimates with standard deviation in brackets for all three suitable matched data sets are provided in Table 6. A scenario with no postmatching regression and a simple linear model are also included for reference. It is evident that Bagging decision trees provide a better estimate of ATE with least standard deviations than using a linear model for postmatching regression or no postmatching regression at all. Thus, our best estimate of ATE of private precaution is between 11,238 and 15,053 EUR. Detailed results of postmatching regression along with the regression tables and feature importance plots from Bagging decision trees are provided in supporting information S.I.3. (Table S.I.3.1 and Figures S.I.3.1, S.I.3.2). Despite the fact that building loss of households with and without private precaution and ATE estimates are based on empirical data controlled for pretreatment confounders and posttreatment loss influencing variables, there might still be alternative explanations for precautionary measures being associated with reduced building losses. Building loss may differ due to damage that a frequently affected household had not repaired after a previous flood. Also, bias may still be present due to specific building characteristics for which the approach has not controlled for. However, Rosenbaum's sensitivity bounds for robustness of the estimated ATEs confirm that ATE of private precaution is unlikely to be sensitive to unobserved confounders (Table 7). The monetary loss reduction of 11,238–15,053 EUR is equal to approximately 27% of the average losses across all the households (47,769 EUR). An average 27% loss reduction due to general private precaution is lower than the reported 50% reduction in median residential building loss comparing the 1993 and 1995 Rhine floods, which was attributed to a considerable general increase in the implementation of private precautionary measures (Bubeck et al., 2012). It is also in the lower range of loss reduction due to wet and dry flood proofing presented in the review of Kreibich et al. (2015). However, these studies hardly controlled for confounding variables. The generalized effectiveness of private precaution of 11,238–15,053 EUR is comparable with the average treatment effects of individual private precautionary measures reported by Hudson

Table 7
Rosenbaum's Bounds for ATE of Private Precaution

Matched Data Set	Γ where ATE becomes statistically insignificant (p value > 0.05)	Γ where ATE brackets to zero (treatment effect = 0)
PSM-NN with caliper	2.4	2.0
PSM-NN with caliper and with replacement	2.4	2.0
PSM-genetic matching algorithm	2.0	1.8

Table 8
Comparison of Flood Loss Models

Flood loss models	ATE estimate from matched data sets in EUR* (Vulnerability reduction due to private precaution): PSM-NN with caliper	PSM-NN with caliper; with replacement	PSM-genetic matching	Mean	Relative loss estimation accuracy: RMSE	MBE
FLEMOps+r	−13,497 (4,202)	−11,997 (4,188)	−11,060 (4,307)	−12,185**	0.122	0.001
RT-FLEMOps	−742 (3,806)	−763(3,486)	−657 (4,127)	−721	0.122	0.000
BT-FLEMOps	−3,816 (3,448)	−3,619 (3,149)	−3,557 (3,867)	−3,664	0.116	0.000
BN-FLEMOps	−14,502 (3,744)	−15,142 (3,609)	−14,416 (4,239)	−14,687**	0.130	0.002

*Standard errors corresponding to the ATE estimate are provided in brackets **Significant ATE estimate (p value ≤ 0.05).

et al. (2014): 14,385 EUR for flood adapted use and 11,302 EUR for flood adapted interior fitting. However, due to the survey design, we do not have households, where the precautionary measures fully prevented loss, for example, water barriers, which hindered water to reach the building. Hence, the contribution of flood barriers to the generalized effectiveness of private precaution is not quantified in the analysis.

3.3. Assessment of Flood Loss Models

The comparison of flood loss models described in section 2.3.1 is provided in Table 8. All tested models perform relatively similar in predicting building loss, with the Bagging decision tree model (BT-FLEMOps) showing the lowest root-mean-square error and mean bias error and the Bayesian network model BN-FLEMOps showing the highest errors. To test the ability of the models to capture differences in vulnerability, we evaluate how close the model-based ATE estimates are to the empirical ATE estimate by comparing the model results obtained for both the vulnerability groups using equation (7).

Only two of the models result in a significant ATE for the implementation of private precaution. The ATE estimates from FLEMOps+r and BN-FLEMOps are 12,185 EUR and 14,687 EUR, respectively (Table 8), which fall within the range of the empirical estimates (11,238–15,053 EUR). Both models have been developed through a combination of expert knowledge and analysis of empirical data and explicitly take into consideration the direct influence of private precaution (Elmer et al., 2010; Wagenaar et al., 2018). The rule-based model FLEMOps+r considers precaution in the second model step together with contamination and recurrence interval. BN-FLEMOps has private precaution indicator (pre) in the Markov blanket of relative building loss (brloss; Figure 2). This implies that in BN-FLEMOps, predictions of building relative loss are directly influenced by private precaution. The treatment and control interventions of private precaution in tree-based models do not result in significant differences in loss estimations between vulnerability scenarios for households. The tree-based models are developed exclusively based on association inferences from empirical data, not using expert knowledge. In RT-FLEMOps, the precautionary measure indicator appears only once in the bottom of the tree, and also, the variable importance plot of BT-FLEMOps reveals a low importance of precaution. Hence, the influence of private precaution on the estimation of building loss is superseded by the effect of other, more important variables. Thus, the building loss estimates of the tree-based models result in an insignificant difference in losses between the two groups of households.

4. Conclusions

We provide robust evidence from a rigorous statistical analysis of a large empirical data set that implementation of private precaution reduces residential building loss with an ATE of 11,238–15,053 EUR currently in Germany. More generally, this confirms previous results that undertaking private precaution is an effective means to reduce vulnerability of households against floods. Our methodology implements matching confounders of private precaution using two distance estimates and six matching methods. From this, three matched data sets are obtained with no significant bias between covariates of households with/without private precaution. Each step in the implemented methodology is customized and tested for its appropriateness for matching flood loss predictors influencing private precaution.

Dynamic risk assessments that account for the differences in vulnerability are necessary for efficient climate-based adaptation in flood risk management. Since flood loss estimation models are crucial to quantify risk, it is important that these models appropriately capture differences in vulnerability, including private precaution. Only two of the tested models are able to capture these differences; these are the rule-based FLEMOps and the expert Bayesian network-based BN-FLEMOps models. In comparison with the tree-based data mining models, the accuracy with which these models predict flood losses are lower. The estimate of ATE and model performances is limited to Germany. Hence, one direction for further research could be assessing data- and model-based quantification of vulnerability due to private precaution in a spatial transferability scenario. It is also evident from the assessment of flood loss models that further research to account for the aspects of dynamic risk without compromising on prediction accuracy is required. Possible other directions in research would include developing better graphical models based on expert knowledge complemented by machine learning algorithms (Chipman et al., 2010). These models should represent causal relationships among potential flood loss estimators and also provide model based scenarios of flood vulnerability.

Acknowledgments
This research has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement 676027 MSCA ETN System-Risk. Flood damage data of the 2005, 2006, 2010, 2011, and 2013 events along with instructions on how to access the data are available via the German flood damage database, HOWAS21 (http://howas21.gfz-potsdam.de/howas21/). Flood damage data of the 2002 event were partly funded by the reinsurance company Deutsche Rückversicherung (www.deutscherueck.de) and may be obtained upon request. The surveys were supported by the German Research Network Natural Disasters (German Ministry of Education and Research (BMBF), no. 01SFR9969/5), the MEDIS project (BMBF; 0330688), the project "Hochwasser 2013" (BMBF; 13N13017), and by a joint venture between the German Research Centre for Geosciences GFZ, the University of Potsdam, and the Deutsche Ruckversicherung AG, Dusseldorf.

References

ABI (2003). Assessment of the cost and effect on future claims of installing flood damage resistant measures. Association of British Insurers, London.

Aldrich, D. P. (2012). Social, not physical, infrastructure: The critical role of civil society after the 1923 Tokyo earthquake. *Disasters*, *36*(3), 398–419.

Allaire, M. C. (2016). Disaster loss and social media: Can online information increase flood resilience? *Water Resources Research*, *52*, 7408–7423. https://doi.org/10.1002/2016WR019243

Apel, H., Aronica, G. T., Kreibich, H., & Thieken, A. H. (2009). Flood risk analysis—How detailed do we need to be? *Natural Hazards*, *49*(1), 79–98.

Atreya, A., Czajkowski, J., Botzen, W., Bustamante, G., Campbell, K., Collier, B., et al. (2017). Adoption of flood preparedness actions: A household level study in rural communities in Tabasco, Mexico. *International Journal of Disaster Risk Reduction*, *24*, 428–438.

Austin, P. C., & Mamdani, M. M. (2006). A Comparison of prognosis score methods: A case-study estimating the effectiveness of post-AMI statin use. *Statistics in Medicine*, *25*(12), 2084–2106.

Barredo, J. I. (2009). Normalized flood losses in Europe: 1970–2006. *Natural Hazards and Earth System Sciences*, *9*(1), 97–104.

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

Bruneau, M., Chang, S. E., Eguchi, R. T., Lee, G. C., O'Rourke, T. D., Reinhorn, A. M., et al. (2003). A framework to quantitatively assess and enhance the seismic resilience of communities. *Earthquake Spectra*, *19*(4), 733–752.

Bubeck, P., Botzen, W. J. W., Kreibich, H., & Aerts, J. C. J. H. (2012). Long-term development and effectiveness of private flood mitigation measures. *Natural Hazards and Earth System Sciences*, *12*(11), 3507–3518.

Bubeck, P., Botzen, W. J. W., Kreibich, H., & Aerts, J. C. J. H. (2013). Detailed insights into the influence of flood-coping appraisals on mitigation behaviour. *Global Environmental Change*, *23*(5), 1327–1338.

Bubeck, P., Botzen, W.J.W., Laudan, Aerts, J.C.J.H. Thieken, A.H. (2017): Insights into flood-coping appraisals of protection motivation theory: Empirical evidence from Germany and France. *Risk Analysis*, online first 17 Nov 2017, https://doi.org/10.1111/risa.12938

Burby, R J, Bollens, S A, Holloway, I M, Kaiser, E J, Mullan, D and Sheaffer, J R (1988) Cities under Water: A comparative evaluation of ten cities' efforts to manage Flood Plain Land Institute of Behavioral Science, University of Colorado, Boulder, CO.

Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, *22*(1), 31–72.

Cammerer, H., Thieken, A. H., & Lammel, J. (2013). Adaptability and transferability of flood loss functions in residential areas. *Natural Hazards and Earth System Sciences*, *13*(11), 3063–3081.

Carisi, F., Schröter, K., Domeneghetti, A., Kreibich, H., & Castellarin, A. (2018). Development and assessment of uni- and multi-variable flood loss models for Emilia-Romagna (Italy). *Natural Hazards and Earth System Sciences Discussion*, https://doi.org/10.5194/nhess-2017-342.

Chipman, H. A., George, E. I., & McCulloch, R. E. (2010). BART: Bayesian additive regression trees. *Annals of Applied Statistics*, *4*(1), 266–298.

R Core Team (2016). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. URL https://www.R-project.org/.

Cumiskey, L., Priest, S., Valchev, N., Viavattene, C., Costas, S., & Clarke, J. (2018). A framework to include the (inter) dependencies of disaster risk reduction measures in coastal risk assessment. *Coastal Engineering*, *134*, 81–92.

Dawson, R. J., Ball, T., Werritty, J., Werritty, A., Hall, J. W., & Roche, N. (2011). Assessing the effectiveness of non-structural flood management measures in the Thames Estuary under conditions of socio-economic and environmental change. *Global Environmental Change*, *21*(2), 628–646.

De Bruijn, K. M. (2004). Resilience and flood risk management. *Water Policy*, *6*(1), 53–66.

de Moel, H., Jongman, B., Kreibich, H., Merz, B., Penning-Rowsell, E., & Ward, P. J. (2015). Flood risk assessments at different spatial scales. *Mitigation and Adaptation Strategies for Global Change*, *20*(6), 865–890. https://doi.org/10.1007/s11027-015-9654-z

De Moel, H., van Vliet, M., & Aerts, J. C. (2014). Evaluating the effect of flood damage-reducing measures: A case study of the unembanked area of Rotterdam, the Netherlands. *Regional Environmental Change*, *14*(3), 895–908.

Defra, A. (2008). Framework for pro-environmental behaviors. Department for Environment, Food and Rural Affairs, London.

Dehejia, R. H., & Wahba, S. (1999). Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs. *Journal of the American Statistical Association*, *94*(448), 1053–1062.

DESTATIS (2013). Statistisches Jahrbuch – Deutschland und Internationales. Retrieved from https://www.destatis.de/DE/Publikationen/StatistischesJahrbuch/StatistischesJahrbuch2013.pdf (last access: 20 May 2018), 2013 (in German).

Diamond, A., & Sekhon, J. S. (2006). Genetic matching for estimating causal effects: A general multivariate matching method for achieving balance in observational studies. UC Berkeley: Institute of Governmental Studies. Retrieved from https://escholarship.org/uc/item/8gx4v5qt

Dietz, H. (1999). *Wohngebäaudeversicherung. Kommentar* (2nd ed.p. 756). Karlsruhe: VVW Verlag Versicherungswirtschaft GmbH.

DiPrete, T. A., & Gangl, M. (2004). 7. Assessing bias in the estimation of causal effects: Rosenbaum bounds on matching estimators and instrumental variables estimation with imperfect instruments. *Sociological Methodology*, *34*(1), 271–310.

Dutta, D., Herath, S., & Musiake, K. (2003). A mathematical model for flood loss estimation. *Journal of Hydrology*, *277*(1–2), 24–49.

Elmer, F., Thieken, A. H., Pech, I., & Kreibich, H. (2010). Influence of flood frequency on residential building losses. *Natural Hazards and Earth System Sciences*, *10*(10), 2145–2159.

Fekete, A., Hufschmidt, G., & Kruse, S. (2014). Benefits and challenges of resilience and vulnerability for disaster risk management. *International journal of disaster risk science*, *5*(1), 3–20.

Few, R. (2003). Flooding, vulnerability and coping strategies: Local responses to a global threat. *Progress in Development Studies*, *3*(1), 43–58.

Figueiredo, R., Schröter, K., Weiss-Motz, A., Martina, M. L., & Kreibich, H. (2018). Multi-model ensembles for assessment of flood losses and associated uncertainty. *Natural Hazards and Earth System Sciences*, *18*(5), 1297–1314.

Gerl, T., Kreibich, H., Franco, G., Marechal, D., & Schröter, K. (2016). A review of flood loss models as basis for harmonization and benchmarking. *PLoS ONE*, *11*(7), e0159791.

Gouldby, B., Samuels, P., Klijn, F., Messner, F., Van Os A., Sayers, P., & Schanze, J. (2005) Language of risk—Project definitions. *FLOODsite Report: T32–04-01*, access (29/06/2017).

Grigg, N. S., & Helweg, O. J. (1975). State-of-the-art of estimating flood damage in urban areas. *JAWRA Journal of the American Water Resources Association*, *11*(2), 379–390.

Hasanzadeh Nafari, R., Ngo, T., & Mendis, P. (2016). An assessment of the effectiveness of tree-based models for multi-variate flood damage assessment in Australia. *Water*, *8*(7), 282.

Heckerman, D. (1998). A tutorial on learning with Bayesian networks. In *Learning in graphical models* (pp. 301–354). Netherlands: Springer.

Heckerman, D., Geiger, D., & Chickering, D. M. (1995). D.M. Machine Learning, 20-197. https://doi.org/10.1023/A:1022623210503

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis*, *15*(3), 199–236.

Holub, M., & Fuchs, S. (2008). Benefits of local structural protection to mitigate torrent-related hazards. *WIT Transactions on Information and Communication Technologies*, *39*, 401–411.

Hudson, P., Botzen, W. J. W., Kreibich, H., Bubeck, P., & Aerts, J. C. J. H. (2014). Evaluating the effectiveness of flood damage mitigation measures by the application of propensity score matching. *Natural Hazards and Earth System Sciences*, *14*(7), 1731–1747.

ICPR, 2002. International Commission for the Protection of the Rhine: Nonstructural flood plain management—Measures and their effectiveness, Koblenz.

IPCC (2012). Managing the risks of extreme events and disasters to advance climate change adaptation. In C. B. Field, V. Barros, T. F. Stocker, D. Qin, D. J. Dokken, K. L. Ebi, M. D. Mastrandrea, K. J. Mach, G.-K. Plattner, S. K. Allen, M. Tignor, & P. M. Midgley (Eds.), *A special report of Working Groups I and II of the Intergovernmental Panel on Climate Change* (p. 582). Cambridge, UK, and New York, NY: Cambridge University Press.

Jongman, B., Kreibich, H., Apel, H., Barredo, J. I., Bates, P. D., Feyen, L., et al. (2012). Comparative flood damage model assessment: towards a European approach. *Natural Hazards and Earth System Sciences*, *12*(12), 3733–3752.

Jongman, B., High-Stigler, S., Feyen, L., Aerts, J. C., Mechler, R., Botzen, W. W., Bouwer, L. M., et al. (2014). Increasing stress on disaster-risk finance due to large floods. *Nature Climate Change*, *4*(4), 264.

Jongman, B., Winsemius, H. C., Aerts, J. C., de Perez, E. C., van Aalst, M. K., Kron, W., & Ward, P. J. (2015). Declining vulnerability to river floods and the global benefits of adaptation. EGU General Assembly 2016, held 17-22 April, 2016 in Vienna Austria, id. EPSC2016-7630

Keele, L. (2010). An overview of rbounds: An R package for Rosenbaum bounds sensitivity analysis with matched data. White Paper. Columbus, OH, 1-15.

Kelly, P. M., & Adger, W. N. (2000). Theory and practice in assessing vulnerability to climate change and facilitating adaptation. *Climatic Change*, *47*(4), 325–352.

Kienzler, S., Pech, I., Kreibich, H., Müller, M., & Thieken, A. H. (2015). After the extreme flood in 2002: Changes in preparedness, response and recovery of flood-affected residents in Germany between 2005 and 2011. *Natural Hazards and Earth System Sciences*, *15*(3), 505–526.

Kreibich, H., Botto, A., Merz, B., & Schröter, K. (2017). Probabilistic, multivariable flood loss modeling on the mesoscale with BT-FLEMO. *Risk Analysis*, *37*(4), 774–787.

Kreibich, H., Bubeck, P., Van Vliet, M., & De Moel, H. (2015). A review of damage-reducing measures to manage fluvial flood risks in a changing climate. *Mitigation and Adaptation Strategies for Global Change*, *20*(6), 967–989. https://doi.org/10.1007/s11027-014-9629-5

Kreibich, H., Christenberger, S., & Schwarze, R. (2011). Economic motivation of households to undertake private precautionary measures against floods. *Natural Hazards and Earth System Sciences*, *11*(2), 309.

Kreibich, H., Christenberger, S., & Schwarze, R. (2012). Corrigendum to "Economic motivation of households to undertake private precautionary measures against floods" published in Nat. Hazards Earth Syst. Sci., 11, 309–321, 2011. *Natural Hazards and Earth System Sciences*, *12*(2), 391–392.

Kreibich, H., Di Baldassarre, G., Vorogushyn, S., Aerts, J. C., Apel, H., Aronica, G. T., Arnbjerg-Nielsen, K., et al. (2017). Adaptation to flood risk: Results of international paired flood event studies. *Earth's Future*, *5*, 953–965.

Kreibich, H., & Thieken, A. H. (2009). Coping with floods in the city of Dresden, Germany. *Natural Hazards*, *51*(3), 423.

Kreibich, H., Thieken, A. H., Petrow, T., Müller, M., & Merz, B. (2005). Flood loss reduction of private households due to building precautionary measures—Lessons learned from the Elbe flood in August 2002. *Natural Hazards and Earth System Sciences*, *5*(1), 117–126.

Kreibich, H., Van Den Bergh, J. C., Bouwer, L. M., Bubeck, P., Ciavola, P., Green, C., Hallegatte, S., et al. (2014). Costing natural hazards. *Nature Climate Change*, *4*(5), 303.

Kunreuther, H. C., Michel-Kerjan, E. O., Doherty, N. A., Grace, M. F., Klein, R. W., & Pauly, M. V. (2009). *At war with the weather: Managing large-scale risks in a new era of catastrophes*. Cambridge, MA: MIT Press.

Mechler, R., & Bouwer, L. M. (2015). Understanding trends and projections of disaster losses and climate change: is vulnerability the missing link? *Climatic Change*, *133*(1), 23–35.

Merz, B., Hall, J., Disse, M., & Schumann, A. (2010). Fluvial flood risk management in a changing world. *Natural Hazards and Earth System Sciences*, *10*(3), 509.

Merz, B., Kreibich, H., & Lall, U. (2013). Multi-variate flood damage assessment: a tree-based data-mining approach. *Natural Hazards and Earth System Sciences*, *13*(1), 53–64.

Merz, B., Kreibich, H., Schwarze, R., & Thieken, A. (2010). Review article "Assessment of economic flood damage". *Natural Hazards and Earth System Sciences*, *10*(8), 1697.

Parker, D., Tapsell, S., & McCarthy, S. (2007). Enhancing the human benefits of flood warnings. *Natural Hazards*, *43*(3), 397–414.

Pearl, J. (2009). Causal inference in statistics: An overview. *Statistics Surveys*, *3*, 96–146.

Penning-Rowsell, E., Floyd, P., Ramsbottom, D., & Surendran, S. (2005). Estimating injury and loss of life in floods: A deterministic framework. *Natural Hazards*, *36*(1), 43–64.

Poussin, J. K., Botzen, W. W., & Aerts, J. C. (2015). Effectiveness of flood damage mitigation measures: Empirical evidence from French flood disasters. *Global Environmental Change*, *31*, 74–84.

Rolfsen, M. (2009). Das neue wasserhaushaltsgesetz. *Natur und Recht*, *31*(11), 765.

Rosenbaum, P. R. (2002). Observational studies. In *Observational studies* (pp. 1–17). New York, NY: Springer.

Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, *39*(1), 33–38.

Rosenbaum, P. R., & Ruby, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrics*, *70*(1), 41–55.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers*, *3*(2), 135–146.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, *66*(5), 688.

Schröter, K., Kreibich, H., Vogel, K., Riggelsen, C., Scherbaum, F., & Merz, B. (2014). How useful are complex flood damage models? *Water Resources Research*, *50*, 3378–3395. https://doi.org/10.1002/2013WR014396

Scutari, M. (2009). Learning Bayesian networks with the bnlearn R package. *Journal of Statistical Software*, *35*(3), 1–22. URL http://www.jstatsoft.org/v35/i03/

Smith, D. I. (1994). Flood damage estimation—A review of urban stage-damage curves and loss functions. *Water SA*, *20*(3), 231–238.

Thieken, A. H., Kreibich, H., Müller, M., & Merz, B. (2007). Coping with floods: Preparedness, response and recovery of flood-affected residents in Germany in 2002. *Hydrological Sciences Journal*, *52*(5), 1016–1037.

Thieken, A. H., Müller, M., Kreibich, H., & Merz, B. (2005). Flood damage and influencing factors: New insights from the August 2002 flood in Germany. *Water Resources Research*, *41*, W12430. https://doi.org/10.1029/2005WR004177

Thieken, A. H., Olschewski, A., Kreibich, H., Kobsch, S., & Merz, B. (2008). Development and evaluation of FLEMOps—a new flood loss estimation model for the private sector. *WIT Transactions on Ecology and the Environment*, *118*, 315–324.

Tsamardinos, I., Aliferis, C. F., Statnikov, A. R., & Statnikov, E. (2003). Algorithms for large scale Markov blanket discovery. In *FLAIRS conference* (Vol. 2, pp. 376–380).

Turner, B. L., Kasperson, R. E., Matson, P. A., McCarthy, J. J., Corell, R. W., Christensen, L., et al. (2003). A framework for vulnerability analysis in sustainability science. *Proceedings of the National Academy of Sciences*, *100*(14), 8074–8079.

Vincent, J. L., Baron, J. F., Reinhart, K., Gattinoni, L., Thijs, L., Webb, A., & Investigators, A. B. C. (2002). Anemia and blood transfusion in critically ill patients. *JAMA*, *288*(12), 1499–1507.

Vogel, K., Riggelsen, C., Korup, O., & Scherbaum, F. (2014). Bayesian network learning for natural hazard analyses. *Natural Hazards and Earth System Sciences*, *14*(9), 2605.

Wagenaar, D., Lüdtke, S., Schröter, K., Bouwer, L. M., & Kreibich, H. (2018). Regional and temporal transferability of multivariable flood damage models. *Water Resources Research*, *54*, 3688–3703. https://doi.org/10.1029/2017WR022233

White, G. F. (1964). Choice of adjustment to floods. *CHICAGO, UNIVERSITY OF CHICAGO, 1964.150 P.*

Wind, H. G., Nierop, T. M., Blois, C. D., & Kok, J. D. (1999). Analysis of flood damages from the 1993 and 1995 Meuse floods. *Water Resources Research*, *35*, 3459–3465. https://doi.org/10.1029/1999WR900192