מרכז הידע והמחקר הלאומי בתחום ההיערכות למצבי חירום
National Knowledge and Research Center for Emergency Readiness

Natural Hazards and Moral Hazards: Understanding the Insurance Coverage Limit

Daniel Felsenstein and Masha Vernick

Department of Geography,

Hebrew University of Jerusalem

Final Report

December 2020

Abstract: This research tests the moral hazard hypothesis in the insurance market for natural hazards. This states that insurance coverage does not reflect the distribution of natural hazards due to households tending to under-insure as the liability of risk is likely to be borne by others. We use insurance portfolio data (n=~12,000) from a large private insurance company linked to asset (dwelling unit) data from Tax Authority records for the Haifa. We control for housing attributes including price, size, year built, distance to hazards (natural and anthropogenic) and local socio-econ attributes include income and crime. We use spatial econometrics estimating a SAR (spatial autoregression) model to understand the effect of exposure to hazards on maximum insurance coverage (structure and content). Our estimation strategy accounts for selection bias in data (using Heckman procedure), spurious spatial relationships (residuals testing) and issues of identification (using SUR- seemingly unrelated regression). The findings differentiate between structure and content insurance. The former is directly related to dwelling unit attributes such as size, price, number of floors and other house prices in the vicinity. In terms of hazards we find a positive relationship to local crime rates and distance to industry and an inverse relationship to distance to forests. No relationship is established with distance to the centers of simulated earthquakes of different magnitudes. These findings support the hypothesis with respect to the existence to a moral hazard in relation to earthquakes. Policy implications are suggested.

תקציר מחקר זה בוחן את השערת קיומו של סיכון מוסרי בשוק הביטוח של סיכונים טבעיים ואנתרופוגניים. במידה והכיסוי הביטוחי אינו משקף את ההתפלגות המרחבית של הסיכונים אנו מפרשים תוצאה זו כמשקפת סיכון מוסרי שבו משקי בית נוטים לא לבטח את נכסיהם. אנו משתמשים בקובץ ייחודי של פוליסות ביטוח של אחד החברות המרכזיות בשוק באזור חיפה (12000 תיקים) ומחברים מידע זה עם מידע של נכסים מתוך מאגר כרמ״ן של רשות המיסים. אנו משתמשים באמידה אקונומטרית כדי להבין את הקשר בין כיסוי ביטוחי לבין החשיפה לסיכונים (טבעיים אנתרופונגניים). אסטרטגיית האמידה מתחשבת בהטעיית סלקציה במידע, קשרים מרחביים מטעים ושאלות של זיהוי (סיבתיות). הממצאים מבחינים בין כיסוי מבנה לבין כיסוי תכולה. ביחס לכיסוי מבנה נמצא קשר חיובי בין גובה הכיסוי לבין תכונות הנכס (כגון מחיר, שטח, מספר קומות, מחירים בסביבה וגם קשר שירי עם קרבה לתעשיה ולפשיעה מקומית. נמצא קשר הפוך עם קרבה ליערות ולא נמצא קשר מובהק למוקד רעידות אדמה מסומלצים בכל המגניטודות. ממצא זה מתפרש כתומך בהשערה המרכזית. מספר צעדי מדיניות אופרטיביים מוצעים.

1 .Introduction

Moral hazard is an endemic feature of insurance markets. In a competitive market, insuring against risk involves lump-sum transfers when an observable risk materializes. A moral hazard occurs when a change in the risk or the actions of an insured party takes effect and this change is unobservable to the insuring party. For example, faced with a natural (observable) hazard, households may take fewer measures to limit their own exposure. This is an action unobservable to the insurer. Households may reduce their own mitigation efforts because they expect insurers to compensate their damage irrespective of their self-protection measures. Individuals and households will thus tend to take larger than normal risks when insuring their own houses (structure and contents) as the liability from that risk is likely to be borne by others. As Arnott and Stiglitz (1988) note, "as more insurance is provided, the marginal private benefit to the individual of expending a given level of effort on accident prevention falls; as a result, he will tend to expend less effort which will increase the probability of his having an accident "(p385) . This increases household exposure to risk and results in an inverse relationship between risk-reduction measures and insurance. The outcome is that the more coverage the household purchases, the less likely it is to engage in risk-reducing activities (Ehrlich and Becker 1972).

Along with the tendency to underestimate probabilities of risk and to discount them highly, moral hazard incentivizes the tendency to under-insure (Kunreuther and Pauly 2004). The result is that insured households bear greater losses in the event of a natural disaster. Under such conditions, the insurance coverage limit represents the upper bound of the moral hazard. Understanding the insurance coverage limit is thus a key component in understanding behavior of households in the insurance market. This market is different to other markets that provide services as it involves an element of redistribution. The moral hazard serves to underscore this: there is little incentive to limit risk or take precautions as insurance providers are expected to pay on the basis of premia collected from unaffected parties. Because of asymmetric information, moral hazards are not always observable by insurance providers and therefore risk-taking behavior is not necessarily captured by higher insurance premia. On the other hand, moral hazards will impact on policy holders spatial behavior. As the latter are less likely to engage in risk mitigation, their location choices will not necessarily reflect a distance preference from low probability/high impact hazards.

This paper investigates the existence of a moral hazard (MH) with respect to insurance coverage for two natural hazards, forest fires and earthquakes. While the existence of a moral hazard has been addressed in the context of various insurance markets such as health, long-term care and car insurance (Sloan and Norton 1997, Finkelstein, McGarry, and Sufi 2005, Cohen and Siegelman 2010) it has only rarely been tested in relation to natural hazards (Carson et al 2013, Hudson et al 2017). The market for natural hazard insurance also differs to other insurance markets as its operation is invariably distorted by the existence of adverse selection (AS). This arises when only those faced by the hazard purchase insurance coverage. For example in terms of natural hazards, this would mean that only households residing close to potential hazard locations would buy insurance (akin to only sick people purchasing health insurance). Many studies exist with the aim of trying to predict which of these two distorters (MH or AS) are prevalent in insurance markets (Cohen and Siegelman 2010, Kreibich et al 2011). When MH is not observed, for example when households with natural hazard coverage take greater self-protection measures than uninsured households, this is commonly ascribed to behavioral attributes such as risk aversion (RA) (Dionne and Eeckhoudt 1985). These unobservables have a large effect on determining the insurance coverage limit.

In this study we test for the existence of a moral hazard in the market for natural hazard insurance with respect to earthquakes and forest fires. In both cases we hypothesize an expected inverse relationship between insurance coverage limit and the risk involved. We estimate the effect of probable earthquake damage (of differing intensities) at the building level on insurance coverage and the effect of proximity to forests (fire risk) on the insurance coverage limit. For both hazards, we investigate whether natural disaster coverage limits are consistent with the implications arising from the existence of a moral hazard. Using unique household insurance portfolio data provided by a commercial insurer for the Haifa metropolitan area, we estimate a model of insurance coverage limits. We are cognizant of the selection bias effect in such an exercise as we only observe households and housing units that purchase insurance. Consequently, we use the Heckman two-step estimation procedure to address this issue. Additionally we are aware of a potential identification threat in our empirical strategy given that structure insurance (the variable of interest here) is jointly dependent on content insurance. We use seemingly unrelated regression (SUR) estimation and test the hypotheses of zero covariance between structure and content, to diffuse this threat.

The paper proceeds as follows. Section 2 outlines the utilities to be gained from natural hazard insurance and the conditions under which different levels of coverage arise. This is followed by a description in Section 3 of the data generation, processing and study area that underpin the empirical analysis of the paper. The empirical strategy adopted is presented in Section 4 emphasizing the use of suitable estimation procedures. The empirical results are discussed in Section 5 and their implications highlighted in the conclusions.

2. Defining the Insurance Coverage Limit for Natural Hazards

Our approach builds on the foundations articulated in by Ehrlich and Becker (1972) who claim that comprehensive insurance decision making involves three inter-related activities. The first involves acquiring market insurance. The second relates to self-insurance through the adoption of mitigating measures that reduce the probability of an event. For example, building protective defenses against flooding in order to mitigate damage. The third is self-protection which leads to reducing the impact of an event. At the extreme, this involves relocating out of a hazardous area. Ehrlich and Becker (1972) show that in the absence of a clear linkage between the level of insurance premia and risk reducing measures, market insurance and self-insurance are substitutes. This assumes, along with models predicting MH, that insurance purchasers make rational decisions when faced with natural hazards. However if insurance purchasers are driven by RA, substitution may not occur as latent risk convictions become divorced from insurance purchasing. Furthermore, a whole string of behavioral factors can lead to divergence from assumed rationality. Insurers may tend to under-estimate risk probabilities (Pahl et al. 2005), over-prioritize perceived risk (Kunreuther and Pauly 2004) and the likelihood of positive outcomes (Kunreuther et al 2013) and resist processing new information (Botzen and van den Bergh 2012).

While market insurance and self insurance are substitutes, market insurance and self-protection are complements (Ehrlich and Becker 1972). One strategy in self-protection is locational choice. As noted above, the choice to remain proximate to a natural hazard may reflect the existence of a moral hazard in the insurance market. Households take larger than normal risks expecting insurers to compensate their damage irrespective of their risk reduction efforts. While some claim that relocation as a mitigation is 'not a viable option' (Carson et al 2017, p309), we consider this behavior (or its absence) as a key variable in understanding the insurance coverage limit for natural hazards.

We argue that in the case of natural hazards, the MH problem associated with household insurance is likely to be mitigated. Natural hazard insurance is a useful mechanism in disaster risk management as it limits the costs of natural disasters spreading risk over space and time and over a diverse group of policy holders. It also facilitates recovery by providing financial compensation after a disaster attempting to restore pre-disaster levels of wealth. The insurance market for natural hazards differs from other markets (such as health or accidents) as natural hazards are place or point oriented with known probabilities and return times. Therefore they are less random in both time and space than other insurable hazards such as car accidents or ill-health. Their coverage limit is thus expected to be bounded both in terms of both time and space.

We define this coverage limit with respect to the loss and the premium associated with the natural hazard. Denoting α as coverage, π as premium, H as value of the property, W as value of the policy-holders wealth, L as loss, p as the probability of loss ($1-p$ = no-loss probability), θ as the rate of loss, X as the magnitude of the event, δ as the probability of having coverage and r as the interest rate, we can define the following terms:

Loss (L): $= X \theta W_0$ (initial wealth prior to the hazard)

Premium (π): this is related to W_0, such that $\pi = \pi' W_0$, where $\pi' W_0 > 0$

Coverage (α): coverage limit (α^*) is $\alpha^* = \alpha\pi$

In the event of a disaster, if the insurance pays the limit α^*, then the actual loss covered by insurance is equal to either coverage (α) or the value of the loss itself (L), whichever is smaller,

Following Ehrlich and Becker (1972) and Kelly and Kleffner (2003) the individual will choose the insurance coverage limit that maximizes expected utility (EU):

$$maxEU_{\alpha} = pU[W_0\, \alpha\pi H - (1-p)] + (1-p)U[W_0 - \alpha\pi H] + \sum_{t=1}^{n} \frac{1}{1+r^{t-1}} \qquad (1)$$

This simple utility is comprised of two components: a cost function expressed as $pU[W_0\, \alpha\pi H - (1-p)]$ and a loss function denoted by $(1-p)U[W_0 - \alpha\pi H]$. The pre-hazard situation of the household (W_0), which determines the premium and coverage, serves as the base for future time periods. If the probability of having coverage is δ, then the actual recovery (R) associated with insurance coverage is:

$R = \delta W_0 \min\{\pi, \alpha\}$ where π' W_0>0

Substituting for L and R, the post-disaster wealth (W_1) of the individual or household with coverage (α) will be:

$$W_1(X, W_0)_{\alpha} = W_0 - \theta(X)W_0 + p(W_0)min\{\alpha(p(W_{0)}, \theta(X)W_0\}$$

Equation (1) can be interpreted as the coverage limit for self-insurance (Briys and Schlesinger 1990). This includes all forms of self-mitigation against natural hazards such costs of defenses and protection measures. As noted earlier, an alternative strategy of self-protection that lowers the probability of a loss includes locational choice, ie maximizing utility from insurance coverage given a particular location (which at the extreme, would include relocating). This is expressed as:

$$maxEU_{\alpha} = p(x)U[W_0 - L - c(x) + [1 - p(x)]U[\,W_0 - c(x)] \qquad (2)$$

where x denotes the level of self-protection and c is the cost associated with self-protection.

The theoretical implication arising from equations (1) and (2) is that measures taken to enhance self-insurance tend to reduce risk whereas measures for self-protection do not (Briys and Schlesinger 1990). Therefore, risk-reducing behavior will result in increasing self- insurance but not necessarily in increasing self-protection. In fact, risk-averse behavior does not contradict reduced levels of self-protection. This may provide an explanation for both the behavioral immobility of households in the face of natural hazards (why do people continue to live in low probability but high risk areas?) and the perceived suboptimal insurance coverage of households in natural hazard zones (why are households in high risk areas under-insured?).

3. Data

We use a unique data set relating to the level of insurance coverage for assets (structure and content) provided by a commercial insurance company with a share of roughly 9% of the Israeli national market for house insurance. The data contain insurance portfolios for close to 100,000 housing units for a geographic area that comprises the continuous built up area of the city of Haifa and a cluster of small towns and suburbs to its north and east. While the study area overlaps much of the Haifa metropolitan area the two areas do not correspond totally (Fig 1). The area chosen includes two primary natural hazards. The first is the seismic hazard zone defined by the Yagur fault. This 20km transform comprises a system of faults that was last active in August 1984 when seismic shock of magnitude ~M5.3 was recorded at a depth of 15-20 km (Salamon et al 2013, Levi et al 2018). As this fault runs through major population centers in the north of Israel, it thus considered potentially active and one of the likely sources of seismic damage in the country.

The second natural hazard relates to the threat of forest fires. The study area is bounded by the Carmel forest to its south. Over a 5-day period in November 2010 this was the location of the largest forest fire in Israeli history. This event resulted in 44 fatalities, the evacuation of 17,000 residents and the destruction of 25 sq km of forest and vegetation. More recently (Nov 2016) the city of Haifa suffered extensive fire damage that penetrated the urban area via the vegetated valleys and pine forest corridors that characterize city topography. These fires resulted in direct damage estimated as $180m, the evacuation of nearly 70,000 city residents and the destruction of nearly 600 homes (mainly apartments) in 77 buildings leaving 1,600 people homeless (State Comptroller 2018).

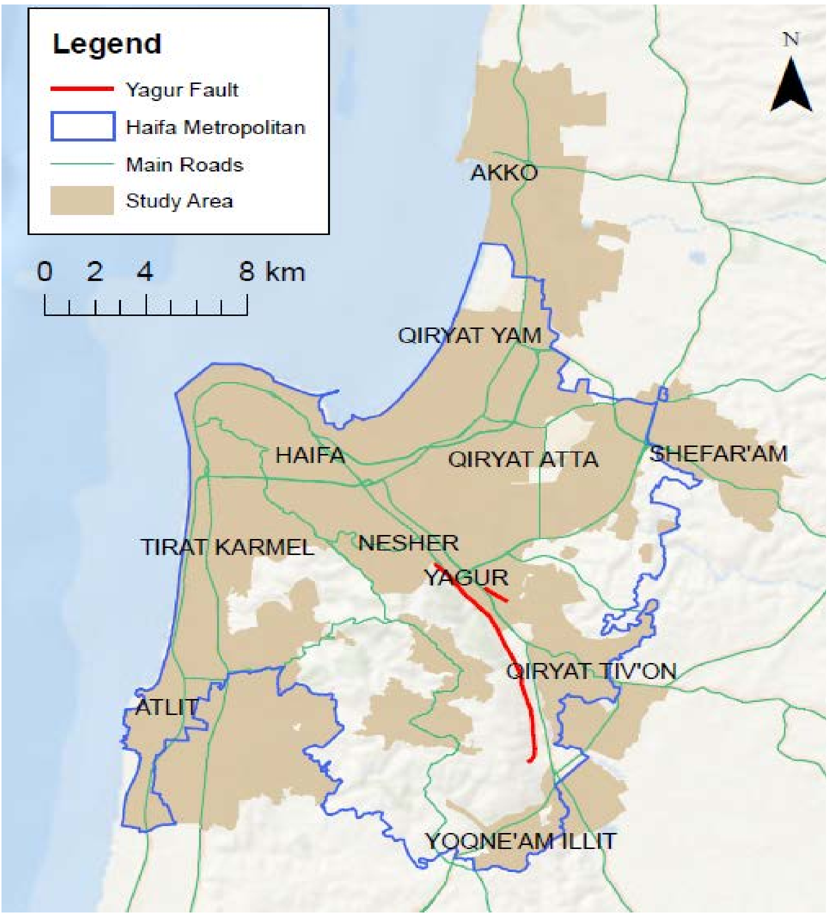


Fig 1: The Study Area

3.1 The Allocation Process

The core data comprise insurance portfolios by housing unit. These contain details on maximal insurance coverage for structure and/or content as well as attributes of the dwelling unit (floor number, type), the building in which it is located (year built, number of floors in building, type ie single family home, apartment block etc) and information on the policy owner (age, gender, marital status). The data are not fully geo-coded but provide city and street names. Instead of randomly assigning each asset (housing unit) to an address on its relevant street, we create an allocation algorithm that considers the building attributes and allocates the portfolios to a national GIS buildings (assets) data layer received from the Survey of Israel. The script randomly allocates each portfolio to matching assets based on city name, street name, floor number and building type (private, duplex or multi-unit). Additionally, this allocation process serves as a weight matrix for assigned streets, representing the density of the associated potential assets. Allocation variables are described in Figure 2 and the data flow in Figure 3. Ultimately, 11,926 portfolios in the study area are matched this way to an asset and receive *x,y* co-ordinates.

Fields used for the allocation:	
Portfolio data:	**GIS Assets layer:**
City	Full address
Street	Number of floors in the building
Asset's floor	Floor (string field)
Building type (private, duplex or multi-unit)	

Table 1: Variables used in the data allocation process

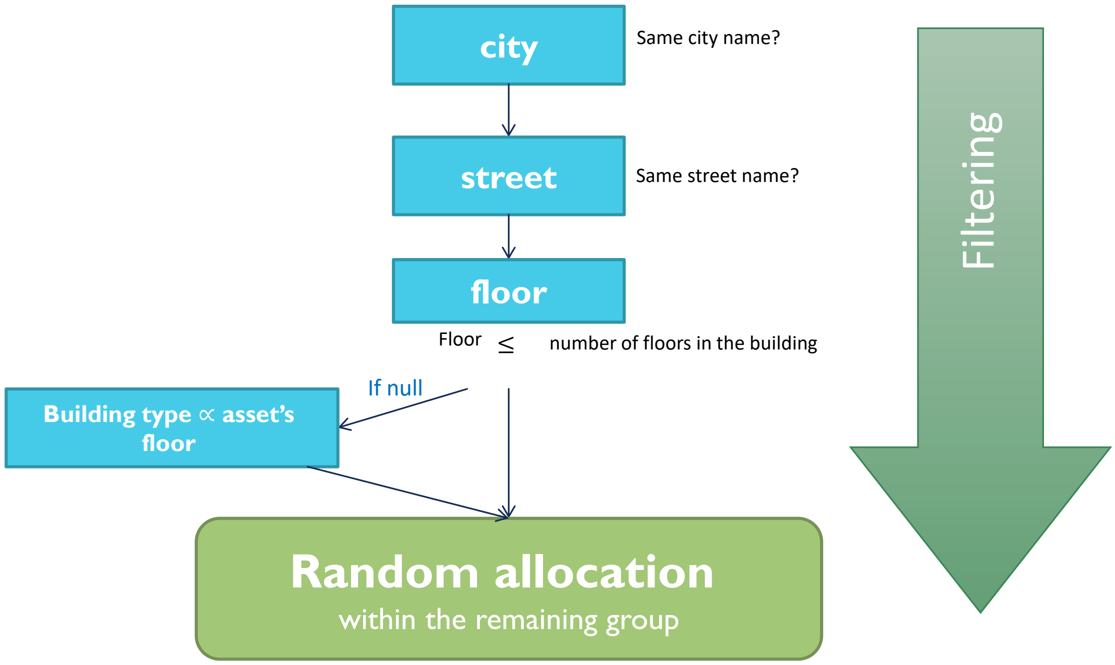


Fig 2: Data Allocation Flow Chart

3.2 Creating Synthetic Data:

Our data on portfolios only includes households purchasing insurance coverage. To adjust for systematic differences that may exist between households that purchase insurance coverage and those that do not, we employ the Heckman two-step selection model to deal with the unobserved households. This adjustment is described below.

In order to create a credible first stage selection in the Heckman adjustment process, we create 'semi-synthetic' observations for all other households in the study area who do not have insurance coverage. To achieve this, we use information on the share of households with insurance coverage in every statistical area (SA)[1] in order to assign an insurance status to the remaining households (assets) in the GIS layer. This status indicates whether they purchased content/structure coverage or not, as well as the age of the head of the household (portfolio owner). The insurance status allocation (purchaser vs non-purchaser) and the age allocation are both random, independent of each other and do not consider any other properties of the asset.

The synthetic data is thus created by random assignment of SA-level data to assets based on the GIS data layer of all residential buildings in the study area described above. The SA-level data on insurance expenditure comes from pooled survey data from the CBS Household Expenditure Survey (2009-2015) where respondents indicate if they have household insurance and the monetary value of their coverage. This survey also includes average socio-demographics for each SA.

[1] An SA is a roughly homogenous administrative division defined by the Israel Central Bureau of Statistics (CBS) akin to a census tract. Most SA's are uniform in population size with an average of 3000 people. Large cities can have hundreds of SA's and the study area comprises around 120.

3.3 Description of Data

Table 2 presents the data used in this study. The insurance portfolio data and the GIS building layer are the primary sources for asset-level information. Variables relating to income, crime rates, house prices and educational level are all aggregate level (SA) controls. Distance to hazardous industry serves the same purpose and is derived from GIS measurement.

The key variables of interest relate to the risk associated with natural hazards. Distance to fire hazards is a building-level variable measured using GIS land use data. Seismic risk is derived from HAZUS damage estimates generated in conjunction with the Israel Geological Survey (IGS). The IGS uses a HAZUS MH 2.1 model to estimate damage scenarios for seven active seismic faults in Israel (Levi 2015). IGS earthquake scenarios calculate ground motion resulting from a specific earthquake and in HAZUS, damage assessment relate to SA centroids. Thus for a given SA a single set of ground motion parameters is applied to all buildings irrespective of within tract variation in ground motion and soil conditions. HAZUS outputs compute site-specific loss estimations, based on ground acceleration and inventories of SA's, buildings and infrastructure.

Based on these outputs, we allocate damage probabilities (DPs) to each building within each SA. In this way, DP's are considered properties of individual buildings. As HAZUS damage calculations differ across scenarios, we consider three plausible options, all of them located along the potentially-active Yagur fault. The parameters of each scenario are given in Table 3. Scenarios A and B are calibrated according to the Geophysical Institute of Israel's (GII) standard peak ground acceleration (PGA) maps and their recurrence interval is known. The recurrence interval of scenario C is not calibrated by a standard map, but is estimated to be less than 475 years (probability of reoccurrence over 50years >0.10). Based on the DP's, each building in a SA is assigned a probability of 1. no damage, 2. slight damage, 3.moderate damage, 4. extreme damage and 5. complete damage for each of the three scenarios. This results in 15 different HAZUS-related values for each observation. To generate a unified variable representing building vulnerability (*I*) we create the following index:

$$i_X = 0 * DP_{none}(x) + 2 * DP_{slight}(x) + 3 * DP_{moderate}(x) + 8 * DP_{extensive}(x) + 10 * DP_{complete}(x)$$

and $I = \sum p_x * i_X = \sum \frac{i_X}{return\ period_x}$

therefore: $I = \frac{i_{5.5}}{200} + \frac{i_6}{475} + \frac{i_{6.7}}{975}$

As a given building is expected to reflect different levels of resilience under changing magnitudes of hazards, the index is designed to reflect the changing probabilities of damage at higher levels of exposure. Index weights are meant to capture the marginal cost of damage for each level of damage exposure.

Table 2: Data Description and Sources

Variable	**Measurement unit**		**Source**
Structure coverage	ILS	Maximal coverage limit for structure	Portfolio data
Content coverage	ILS	Maximal coverage limit for content	Portfolio data
Communal (dummy)		A dummy for a communal building	Portfolio data
distance to industry	Meters	Aerial distance between assigned location and nearest petrochemical industrial zone	GIS layer - Survey of Israel
floor number		Floor number of asset	Portfolio data
crime_per1k in SA		Number of break-ins per 1000 residents, by SA	Israel Police,
Sqm	sqm	Area of the apartment	Portfolio data
bld_age	years	The age of the building	Portfolio data
$DP_i(x)$	Probability [0,1]	The probability for damage scenario i given an earthquake of X Moment magnitude. $i\begin{cases} none \\ slight \\ moderate \\ extensive \\ complete \end{cases}$	HAZUS, Israel Geological Survey (IGS)
I		Combined vulnerability index $I = \sum p_x * i_X$	HAZUS, IGS
$i_{5.5}$		Vulnerability index for a 5.5 Moment magnitude earthquake	HAZUS, IGS
i_6		Vulnerability index for a 6 Moment magnitude earthquake	
$i_{6.7}$		Vulnerability index for a 6.7 Moment magnitude earthquake	
Age	Years	Age of the portfolio owner; imputed age for the rest	Portfolio data, CBS
n_floors in the building		Number of floors in the building	Portfolio data
Proximity to forest (dummy)		=1 if the asset is within 50 meters from a forest	GIS layer-Survey of Israel
income in SA	ILS	Median annual earned income 2015	CBS, 2015
h_prices in SA	ILS	Average prices per sqm	Carmen data base (Israel Tax Authority)
pec_edu_13y in SA	%	Share of people with >13 years of education	CBS, 2013
pec_ownership in SA	%	Share of households owning the apartment in which they reside	CBS Household Expenditure Survey (HES) 2009-2015
%insured HH in SA	%	Share of HH with insurance portfolio (structure or content)	

Table 3: Three earthquake scenarios associated with the Yagur Fault

Scenario	Magnitude (moment magnate scale)	Recurrence Interval
A	6.0	475yr (10% @50yr)
B	6.7	975yr (5% @50yr)
C	5.5	Less than 475yr

4. Estimation Issues

4.1 Selection Bias

The observed data in this study suffer from selection bias in that they only relate to assets (housing units) for which insurance coverage is purchased. Not all households purchase insurance and there may be a latent selection variable $z*$ that governs the insurance decision for a particular dwelling unit. Generally $z*$ is not observable and its sign can only be inferred. If a household purchases insurance then $z*$ is assumed to be positive and $z=1$. The opposite holds if insurance is not purchased ($z=0$). This decision choice can be estimated using standard probit estimation. However, estimating the extent of insurance coverage (the nonlimit sample) using OLS will produce errors that are heteroscedastic by construction (Heckman 1979).

The solution to this estimation issue is the Heckman two-step selection procedure. This involves first estimating the decision (selection) equation and obtaining estimates of λ_i (ie the inverse Mills ratio) for each observation in the nonlimit sample. The second step calls for the OLS estimation of y on x and $\hat{\lambda}$ for the nonlimit (outcome) equation, where $\hat{\lambda}$ serves as a control for selection bias. If $\hat{\lambda}$ is statistically significant, its omission produces selection bias. If $\hat{\lambda}$ is not significant its omission does not lead to biased errors and OLS estimation of the outcome model will generate consistent estimators.

In our case, the selection (probability of having insurance coverage) equation is:

$$z_i^* = \gamma' w_i + +u_i \quad \text{where } u_i \;\; is\; N[0,1] \qquad (3)$$

$$z_i = 1 \;\; if \;\; z_i^* > 0 \;\text{ and }\; z_i = 0 \;\; if \;\; z_i^* \le 0$$

$$Prob\;(z_i = 1) \; = \Phi(\gamma' w_i) \text{ and } Prob\;(z_i = 0) \; = 1 - \Phi(\gamma' w_i)$$

where Φ denotes the cumulative normal distribution.

The nonlimit (insurance coverage) equation in reduced form is:

$$p_i = \mathbf{B}'x_i + \varepsilon_i \quad \text{where } \epsilon_i \;\; is\; N[0,1] \qquad (4)$$

where: p_i is the insurance coverage limit, $\mathbf{B}'$ is a vector of covariates including dummies and controls and (ϵ_i, u_i) are N[0,1 , σ^2 ρ] with ρ denoting the correlation between p and z. Equation (4) is only observed if $z_i=1$. This implies that estimating the coefficient vector **B** in (4) without Heckman's correction will result in omitted variable bias. Estimates of **B** will be inconsistent and oblivious to the fact that housing units that purchase insurance may differ systematically from those that do not. Equation (4) can be reformulated to account for the fact that z_i and w_i are observed for a random sample of insurance purchasers but p_i is observed only when $z_i=1$, as follows:

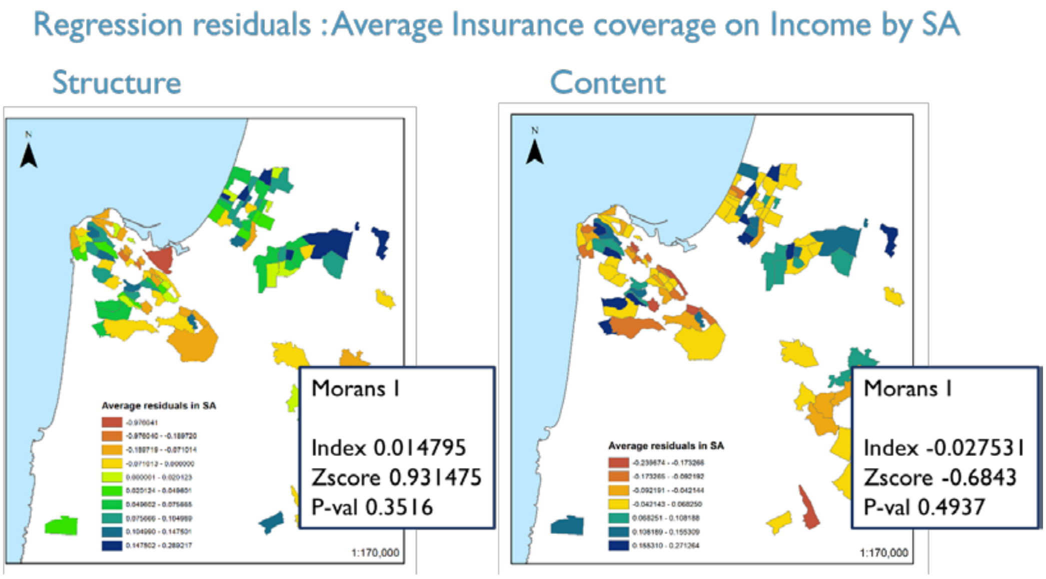
$$E[p_i|z_i = 1] = \mathbf{B}'x_i + \rho\sigma, \lambda(\gamma' w_i) \qquad (5)$$

where $\lambda(\gamma' \mathrm{w_i})$ is the inverse Mills ratio given by $\phi(\gamma' \mathrm{w_i})/[1 - \Phi(\gamma' \mathrm{w_i})]$. The normal density and distribution function are denoted by and ϕ and Φ respectively.

4.2 Spurious Relationships

To diffuse claims that the spatial pattern of insurance coverage does not reflect natural hazards but some other confounding relationship, we test for the existence of a spatial association between insurance cover and income. We utilize aggregate data at the SA level for both insurance coverage and income variables. Fig 3 maps the residuals of the relationship between these two variables. There is no visually apparent picture of spatial concentration nor is there any evidence of any local spatial association with small and insignificant Moran's I values in both cases.

Fig 3: Testing for Spatial Association



4.3 Identification Threats

The identification threat to estimating structure coverage comes from the fact that some households are 'natural insurers' and therefore structure (S) and content (C) insurance are jointly dependent. Our data covers household insurance policies in all forms of combination as follows:

Structure	X	X	
Content	X		X

We assume the existence of latent variables C^*_i and S*i:

The utility from purchasing C_i: $C_i^* = \alpha x_i + \varepsilon_{iC}$

If $C_i^* < 0 \quad \rightarrow \quad C_i = 0$ (i.e. will not purchase)

Utility from purchasing S_i: $S_i^* = \beta z_i + \varepsilon_{iS}$

If $S_i^* < 0 \quad \rightarrow \quad S_i = 0$

To identify 'natural insurers' ie those people who have both S and C coverage we need to test for covariance , ie. $\rho_{\varepsilon_C,\varepsilon_S} > 0$

Our data includes only those for whom $C_i^* > 0 \quad or \;\; S_i^* > 0$

As noted, theory posits various unobservables that will result in sub-optimal coverage, for example risk aversion (RA), moral hazard (MH) and adverse selection (AS). These will affect the covariance between the errors of the factors influencing both C and S, such that :

$$\varepsilon_S = \lambda_s RA + MH$$

+ - -

$$\varepsilon_c = \lambda_c RA + MH + AS$$

The covariance (hitherto ignored) between $\varepsilon_S, \varepsilon_c$ is expressed as

$$Cov(\varepsilon_S, \varepsilon_c) = \lambda_s \cdot \lambda_c \cdot cov^2(RA) + \lambda_s \cdot cov(RA, AS) - \lambda_c \cdot cov(RA, MH)$$

If we assume that $Cov(\varepsilon_S, \varepsilon_c) = 0$, we can focus on either S or C. However more probably $Cov(\varepsilon_S, \varepsilon_c)$ is large and non-zero. The most efficient way of dealing with this dependence is to estimate models for S and C as a SUR (seemingly unrelated regression, Zellner 1962) and to obtain $Cov(\varepsilon_S, \varepsilon_c)$. SUR regression assumes that cross section dependence in the residuals is seemingly unrelated, ie it has no spatial or common factor sources. Essentially its is an extension of linear regression and allows for estimating coefficients in a system of multiple equations with parameter restrictions and correlated error terms. This yields estimates that are asymptotically more efficient than single equation models.

Tables A1 and A2 present the SUR models and covariance results respectively. As anticipated, the models for S and C and very similar results and most significantly, $Cov(\varepsilon_S, \varepsilon_c)$ is far from 0.

Table A1: SUR regression results for Coverage Limits; Structure and Content Insurance

	SUR model	
VARIABLES	Structure	Content
Multi-unit *(dummy)*	426,454***	259,877***
	(141,116)	(77,917)
Private *(dummy)*		-36,806
		(105,517)
crime_per1k *in SA*	-7.339e+07*	-4.468e+06
	(3.826e+07)	(1.833e+07)
Distance: to industry	2.940	6.381
	(21.43)	(10.34)
-to forests (log)	22,418	-4,219
	(48,293)	(23,325)
Floor # of the asset	105,178***	15,009*
	(16,398)	(7,868)
Sqm	72,592***	20,636***
	(1,391)	(864.6)
i_	4.493e+06	5.790e+06
	(7.502e+06)	(3.554e+06)
Client’s age	30,748***	20,349***
	(2,641)	(1,270)
# rooms		40,413*
		(23,188)
bld_age	10,211***	
	(1,908)	
Income in SA	140.4***	102.2***
	(29.10)	(14.06)
Constant	-4.439e+06***	-2.735e+06***
	(476,540)	(235,744)
Observations	6,793	6,793
R-squared	0.361	0.222

*** p<0.01, ** p<0.05, * p<0.1
Note: Dependent variable non-log format. Standard errors in parentheses. Households without structure/content policy are flagged as 0

Table A2: Covariance results

Covariance:	U_s	U_c
U_s – *residuals from Structure reg.*	$1.0*10^{13}$	
U_c – *residuals from Content reg.*	$3.0*10^{12}$	$2.4*10^{12}$

Correlation	U_s – *residuals from Structure reg.*
U_c – *residuals from Content reg.*	0.6161, $p < 0.0000$

5. Empirical Results

Two-sample t test with equal variances

	Mean					N. Observations	
Structure coverage	No	Yes	St_Err	t_value	p_value	No	Yes
Age - head of HH	54.58	51.893	.285	9.4	0	57176	3321
# floors in the bld	4.239	4.898	.055	-12.15	0	57176	3321
log deal	12.996	13.22	.012	-19.1	0	57176	3321
share education >13y in SA	47.784	50.127	.255	-9.2	0	57176	3321
crime per 1k in SA	.002	.002	0	4.65	0	57176	3321
income in SA	6123.5	6538	26.549	-15.6	0	57176	3321
pec ownership in SA	66.862	68.222	.247	-5.5	0	57176	3321
DPs index (combined)	.018	.017	0	14.1	0	54433	3321
i 55	1.857	1.698	.011	14.55	0	54433	3321
i 6	2.078	1.891	.015	13	0	54433	3321
i 67	4.093	3.72	.028	13.45	0	54433	3321

1) Outcome Model

Variables	y=log(max. coverage) (1)	(2)	(3)	(4)
Multi-unit building (dummy)	0.0765	0.0766	0.0766	0.0767
	(0.0525)	(0.0515)	(0.0515)	(0.0520)
Asset floor #	0.00803	0.00800	0.00799	0.00801
	(0.00642)	(0.00643)	(0.00639)	(0.00642)
Sqm	0.0129***	0.0129***	0.0129***	0.0129***
	(0.000559)	(0.000562)	(0.000555)	(0.000559)
log(price)	0.0884***	0.0883***	0.0883***	0.0883***
	(0.0271)	(0.0269)	(0.0267)	(0.0270)
crime_per1k	-18.47	-18.69	-18.82	-18.58
	(14.59)	(14.60)	(14.74)	(14.59)
log(dist forest)	-0.0533**	-0.0529**	-0.0527**	-0.0530**
	(0.0231)	(0.0225)	(0.0227)	(0.0228)
i_55	0.00451			
	(0.0342)			
i_6		-0.000629		
		(0.0272)		
i_67			-0.00134	
			(0.0143)	
i_ - combined				0.0934
				(3.483)
Constant	13.32***	13.33***	13.33***	13.33***
	(0.497)	(0.492)	(0.487)	(0.493)
Lambda (inv. Mills ratio)	0.0558***	0.0592***	0.0611***	0.0579***
	(0.0754)	(0.0738)	(0.0734)	(0.0747)
Rho	0.1292	0.1371	0.1413	0.1341
	(0.1724)	(0.1682)	(0.1672)	(0.1705)
Sigma	0.4315***	0.4319***	0.4321***	0.4318***
	(0.0163)	(0.0166)	(0.0166)	(0.0164)
Observations	44,972	44,972	44,972	44,972

Robust standard errors in parentheses (adjusted for 96 clusters in SA)
*** p<0.01, ** p<0.05, * p<0.1

2) Selection Model: Probit - max likelihood

	y=insurance purchase (dummy)			
Variables	(1)	(2)	(3)	(4)
Age	-0.00590***	-0.00591***	-0.00592***	-0.00591***
	(0.00145)	(0.00145)	(0.00146)	(0.00145)
n_floors	0.0272**	0.0272**	0.0272**	0.0272**
	(0.0121)	(0.0121)	(0.0121)	(0.0121)
l_deal	0.243***	0.243***	0.243***	0.243***
	(0.0522)	(0.0522)	(0.0522)	(0.0522)
pec_edu_13y	0.00402	0.00402	0.00402	0.00402
	(0.00271)	(0.00271)	(0.00271)	(0.00271)
Constant	-4.627***	-4.626***	-4.626***	-4.627***
	(0.682)	(0.682)	(0.682)	(0.682)
athrho	0.130	0.138	0.142	0.135
	(0.175)	(0.171)	(0.171)	(0.174)
Slnsigma	-0.840***	-0.839***	-0.839***	-0.840***
	(0.0377)	(0.0383)	(0.0385)	(0.0381)
Observations	44,972	44,972	44,972	44,972
Wald chi^2	652.71	653.11	653.74	652.59
Log pseudo likelihood	-13424.4	-13424.46	-13424.42	-13424.46

Robust standard errors in parentheses (adjusted for 96 clusters in SA)
*** p<0.01, ** p<0.05, * p<0.1

1) Average marginal effects

*All models had similar marginal effects result, differing only in their 6th digit

Variables	dy/dx	Delta-method Std. Err.	P>|z|
age	-0.00092	0.00024	0.000
n_floors	0. 00422	0.00200	0.035
l_deal	0..03765	0.00962	0.000
pec_edu_13y	0. 00062	0.00043	0.148

VARIABLES	(3) log(structure)	(4) structure	(5) log(structure)	(6) structure	(7) log(structure)	(8) structure
Multi-unit dwelling (dummy)	0.0263	213,515	0.0253	207,690	0.0259	211,612
	(0.0227)	(159,396)	(0.0228)	(158,877)	(0.0228)	(159,196)
distance to industry	6.83e-07	-69.61***	1.07e-07	-76.24***	1.09e-07	-74.76***
	(3.80e-06)	(26.19)	(3.76e-06)	(26.13)	(3.86e-06)	(26.63)
floor number	0.0159***	62,472**	0.0157***	60,612**	0.0159***	61,969**
	(0.00302)	(27,529)	(0.00300)	(27,312)	(0.00302)	(27,426)
Next to forest (dummy)	-0.0492	-1.003e+06**	-0.0516	-1.025e+06**	-0.0514	-1.020e+06**
	(0.0604)	(428,176)	(0.0608)	(429,467)	(0.0606)	(430,786)
crime_per1k in SA	-4.836	-2.090e+07	-6.062	-3.067e+07	-5.497	-2.575e+07
	(6.873)	(4.544e+07)	(6.935)	(4.460e+07)	(6.899)	(4.520e+07)
sqm	0.0130***	63,249***	0.0130***	63,320***	0.0130***	63,236***
	(0.000349)	(2,354)	(0.000352)	(2,360)	(0.000347)	(2,352)
bld_age	-0.00396***	-71,180***	-0.00393***	-70,505***	-0.00395***	-70,842***
	(0.000766)	(8,940)	(0.000756)	(8,917)	(0.000761)	(8,896)
h_prices in SA	2.42e-05***	254.7***	2.38e-05***	249.9***	2.40e-05***	252.4***
	(4.00e-06)	(48.40)	(4.00e-06)	(48.34)	(4.01e-06)	(48.48)
i_						
i_55	-0.00914	-31,750				
	(0.0165)	(107,263)				
i_6			-0.0142	-96,788		
			(0.0136)	(73,661)		
i_67					-0.00492	-28,737
					(0.00682)	(40,812)
Constant	13.70***	-8.299e+06***	13.71***	-8.121e+06***	13.71***	-8.212e+06***
	(0.0586)	(830,123)	(0.0579)	(831,058)	(0.0580)	(834,356)
age	-0.00510***	-0.00536***	-0.00510***	-0.00536***	-0.00510***	-0.00536***

	(0.000475)	(0.000440)	(0.000475)	(0.000439)	(0.000475)	(0.000439)
n_floors in the building	-0.00721	-0.00730	-0.00718	-0.00717	-0.00720	-0.00725
	(0.00566)	(0.00469)	(0.00566)	(0.00469)	(0.00566)	(0.00469)
Next to forest (dummy)	-0.134*	-0.125*	-0.134*	-0.125*	-0.134*	-0.125*
	(0.0750)	(0.0751)	(0.0750)	(0.0750)	(0.0750)	(0.0750)
income in SA	-9.40e-05***	-5.94e-05***	-9.39e-05***	-5.93e-05***	-9.38e-05***	-5.92e-05***
	(2.72e-05)	(2.13e-05)	(2.71e-05)	(2.12e-05)	(2.72e-05)	(2.12e-05)
h_prices in SA	8.20e-05***	7.70e-05***	8.18e-05***	7.66e-05***	8.19e-05***	7.67e-05***
	(1.55e-05)	(1.27e-05)	(1.55e-05)	(1.27e-05)	(1.55e-05)	(1.27e-05)
bld_age	-0.0198***	-0.0200***	-0.0198***	-0.0200***	-0.0198***	-0.0200***
	(0.00115)	(0.00112)	(0.00115)	(0.00112)	(0.00115)	(0.00112)
pec_ownership in SA	-0.00576**	-0.00488**	-0.00577**	-0.00490**	-0.00577**	-0.00489**
	(0.00248)	(0.00195)	(0.00249)	(0.00197)	(0.00248)	(0.00196)
pec_edu_13y in SA	0.00413*	0.00222	0.00417*	0.00231	0.00415*	0.00226
	(0.00250)	(0.00200)	(0.00251)	(0.00202)	(0.00251)	(0.00201)
Constant	-0.384*	-0.475***	-0.384*	-0.475***	-0.384*	-0.475***
	(0.219)	(0.183)	(0.219)	(0.183)	(0.219)	(0.183)
Athrho	0.564***	1.387***	0.569***	1.389***	0.565***	1.388***
	(0.0539)	(0.0856)	(0.0526)	(0.0852)	(0.0534)	(0.0857)
Lnsigma	-0.713***	15.48***	-0.711***	15.48***	-0.713***	15.48***
	(0.0263)	(0.0570)	(0.0258)	(0.0571)	(0.0260)	(0.0572)
Observations	91,992	92,317	91,992	92,317	91,992	92,317

6. Conclusions and Policy Implications

The statistical estimations presented in this research test for the existence of the moral hazard hypothesis is insurance coverage of natural and anthropogenic hazards. The evidence presented is mixed. Controlling for dwelling unit attributes and area (neighborhood) characteristics in the case of both structure and content policy, does not yield unambiguous results. We find structure coverage directly related to proximity to industry and to level of neighborhood crime rates and inversely related to distance form forest areas. In contrast is not related to the simulated presence of earthquakes. This last finding is interpreted as evidence supporting the moral hazard hypothesis. In the case of content coverage we find similar but generally weaker results.

These results beg the question as to why insurers don't take more direct and spatially differentiated action in promoting hazard insurance given the moral hazard use and the tendency of households to under-insure. We offer two explanations. The first is the uncertainty and magnitude of the natural and anthropogenic risks involved. Given the history of government intervention in the case of unanticipated disasters (including warfare) and the existence of a national compensation mechanism, the private sector has tended to avoid this market. Second the fear of catastrophic disaster and the inability to use standard actuarial in its forecasting has caused this area of insurance to overlooked or ignored. This while the demand side of the market is subject o moral hazard , the supply side has been overshadowed by a tradition of heavy handed government intervention in the field of hazard insurance

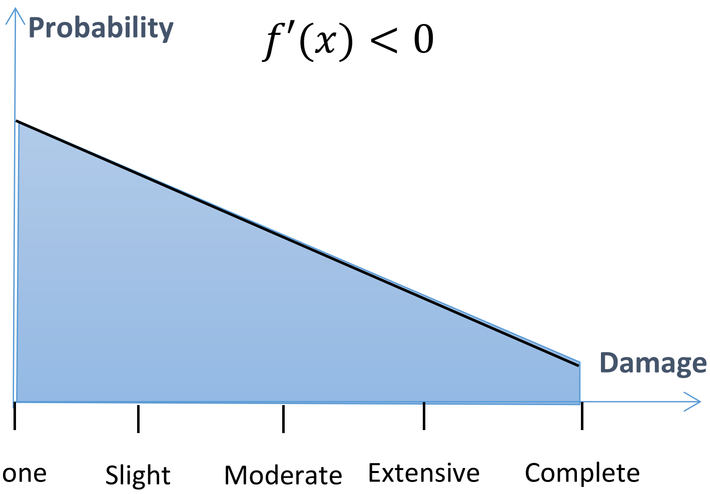
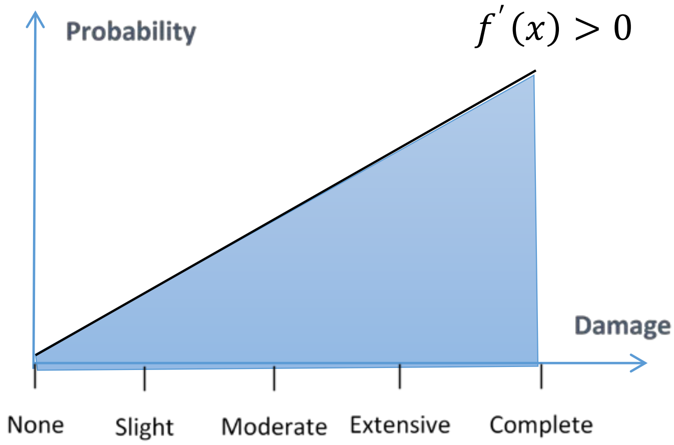
Given this situation we suggest the following policy measures to generate a better functioning market and to iron-out some of its' imperfections:

- a public awareness campaign of natural and anthropogenic hazards. Household behavior is governed by information which is not readily available in the case of hazards. An informed public needs to be aware of the environmental hazards it faces. Delivering this information through readily available platforms such as dashboards on websites, should become standard practice in conveying information to the public which has come to expect this kind of information delivery through the Covid-19 pandemic
- spatially differentiated insurance policies for different forms of hazards (proximity to active seismic zones, flood zones, areas prone to forest fires,

heavy industry and toxic waste concentrations etc), Blanket national premia for earthquakes for example, simply perpetuate the current situation. Devising such as pricing system is of course no trivial matter and is a research project in its own right. But it is critical for combatting the moral hazard inherent in insurance coverage.

- the devolution of disaster management to the local level. Again this another Corona-learned lesson where central government has begrudgingly been forced to diffuse authority to local authorities in those areas of responsibility such as shelter, health and welfare where intervention is inherently local.

Appendix 1

a. <u>**Strong structure**</u>	b. <u>**Weak structure**</u>
The probability of severe damage is low	Higher risk of severe damage
	

We would expect the curve to shift from A type towards B-like type when the magnitude increases. Hence, we designed the Index in a manner that will preserve this trait.

Bibliography

Arnott RJ and Stiglitz JE (1988) The Basic Analytics of Moral Hazard, *Scandinavian Journal of Economics* 90 (3), 383–413

Bin O, Kruse, JB and Landry CE (2008) Flood Hazards, Insurance Rates, and Amenities: Evidence from the Coastal Housing Market, *Journal of Risk and Insuranc*e 75(1), 63-82.

Bin O and Landry CE (2013) Changes in Implicit Flood Risk Premiums, Empirical evidence from the Housing market, *Journal of Environmental Economics and Management*, 65(3), 361-376.

Booth K and Tranter B (2018) When disaster strikes: Under-insurance in Australian households, *Urban Studies*, 55(14) 3135–3150

Botzen, WJW and van den Bergh JCJM (2012) Monetary Valuation of Insurance against Flood Risk under Climate Change, *International Economic Review* 53(3), 1005–26.

Briys E and Schlesinger H (1990) Risk Aversion and the Propensities for Self Insurance and Self Protection, *Southern Economic Journal*, 57, 458-467.

Carson JM, McCullough KA and Pooser DP (2013) Deciding Whether to Invest in Mitigation Measures: Evidence from Florida, *Journal of Risk and Insurance*, 80 (2), 309–327.

Cohen and Siegelman P (2010) Testing for Adverse Selection in Insurance Markets, *Journal of Risk and Insurance,* **77** (1), 39–84.

Dionne,G.,and.Eeckhoudt L (1985) Self Insurance, Self Protection and Increased Risk Aversion, *Economics Letters*, 17: 39-42.

Ehrlich I and Becker GS (1972) Market Insurance, Self-insurance, and Self-protection, *Journal of Political Econom*y 80 (4), 623-648.

Finkelstein A, McGarry K Sufi A (2005 Dynamic Inefficiencies in Insurance Markets: Evidence from Long-Term Care Insurance, *American Economic Review* 95(2): 224–28

Heckman, J (1979), Sample Selection Bias as a Specification Error, *Econometrica*, 47(1), 153-61.

Hudson P, Botzen WJW , Feyen L and Aerts JCJH (2016) Incentivising Flood Risk Adaptation through Risk Based Insurance Premiums: Trade-offs between Affordability and Risk Reduction, *Ecological Economics* 125, 1-13.

Hudson O, Botzen WJW., Czajkowski J and Kreiblich H (2017), Moral Hazard in Natural Disaster Insurance Markets: Empirical Evidence from Germany and the United States, *Land Economics*, 93(2), 179-208.

Kunreuther H and Pauly M (2004) Neglecting Disaster: Why Don’t People Insure against Large Losses? *Journal of Risk and Uncertainty* 28 (1): 5–21.

Kunreuther, H, Pauly M and McMorrow S (2013) *Insurance and Behavioral Economics: Improving Decisions in the Most Misunderstood Industry*. New York: Cambridge University Press.

Kreibich H, Thieken AH, Petrow T, Muller M and 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), 117–26.

Kreibich, H, Christenberger S, and Schwarze R (2011) Economic Motivation of Households to Undertake Private Precautionary Measures against Floods, *Natural Hazards and Earth System Sciences* 11(2), 309–321.

Levi T, Bausch D, Katz O, Rozelle J and Salamon A (2015)Insights from Hazus loss estimations in Israel for Dead Sea Transform earthquakes, *Natural Hazards* 75:365–388

Levi T, Salamon A, Bausch D, Rozelle J, Cutrell A, Hoyland S Hamiel Y, Katz O, Calvo R, Gvirtzman Z, Ackerman B and Gavrieli I (2018) Earthquake scenario in a national drill, the case of ‘‘Turning Point 6’’, 2012, Israel , *Natural Hazards* 92,113–132

Osberghaus D. (2015) The Determinants of Private Flood Mitigation Measures in Germany: Evidence from a Nationwide Survey. *Ecological Economics* 110, 36-50.

Pahl, S, Harris PR, Todd HA, and Rutter DR (2005) Comparative Optimism for Environmental Risks, *Journal of Environmental Psychology* 25(1), 1–11.

Pérez-Blanco CD (2015) Reflections on the Current Debate on How to Link Flood Insurance and Disaster Risk Reduction in the European Union, *Natural Hazards* 79(3), 1451–1479.

Petrolia DR, Hwang J, Landry CE, and Coble KH (2015) Wind Insurance and Mitigation in the Coastal Zone, *Land Economics* 91 (2), 272–295.

Salamon A, Aksinenko T. Zaslavsky Y (2013). *The location of the Carmel Fault between Yagur and Haifa Port as inferred from subsurface data*. Israel Geological Survey, Rep. GSI/01/2013, Jerusalem (Hebrew).

Sloan FA and Norton EC (1997) Adverse Risk Selection, Bequests, Crowding Out, and Private Demand for Insurance: Evidence from the Long-Term Care Insurance Market, *Journal of Risk and Uncertainty* 15(3): 201–19.

Thieken, A H., Petrow H, Kreibich H, and Merz B (2006) Insurability and Mitigation of Flood Losses in Private Households in Germany, *Risk Analysis* 26(2), 383–395.

Wang D., Davidson RA., Tainor JE., Nozick LK and Kruse J (2017) Homeowner purchase of insurance for hurricane-induced wind and flood damage, *Natural Hazards*, 88, 221-245.

Zellner, A (1962) An efficiency method of estimating seemingly unrelated regression equations and tests for aggregation bias, *Journal of the American Statistical Association,* 57, 348–368.