How Low is the Tail Insurance Demand?

Selim Mankaï[†] Sébastien Marchand[∓] Ngoc Ha Le[‡]

University of Clermont Auvergne

October 7, 2022

Abstract:

Evidence suggests that the demand for voluntary insurance against low-probability, high-impact (LPHI) losses is unexpectedly low. To assess the overall willingness to pay (WTP) for insurance against this class of risks, we conduct a meta-analysis of empirical contingent valuation studies. We also explore potential sources of heterogeneity, with an emphasis on certain theoretically grounded determinants. From a dataset of 65 outcomes spanning 2005 to 2021, we find that the average stated WTP is around 87% of the actuarially fair premium. After accounting for observed heterogeneity, this estimate is no longer significantly different from one. The meta-regression reveals high sensitivities of the WTP to survey mode, elicitation design and time period. WTP also appears to negatively depend on provided loss probability level and sample average income and age. Our results are robust to model specifications and publication bias and provide methodological recommendations to future research in this area.

Keywords: Low probability risks; Contingent valuation; Insurance demand; Stated preferences; Meta-analysis; Economic experiments.

JEL classifications: O12, O13.

† Department of Finance, IAE Clermont Auvergne-School of Management, CleRMa, France (Corresponding author: selim.mankai@uca.fr);

∓ Department of Economics, CERDI, University of Clermont Auvergne, France

‡ Ph.D. candidate, CERDI, University of Clermont Auvergne, France.

1. Introduction

The low demand for insurance against low-probability, high-impact (LPHI) events raises many questions about the relationship between risk perception and protection behavior (Kunreuther et al., 2013, Wagner, 2022). The severity of losses induced by this category of risks exerts significant economic pressure on people and communities, especially when their impact is uneven (Klomp and Valckx, 2014, Botzen et al., 2020). As a post-loss financing instrument, insurance enables to alleviate these unexpected economic shocks. Specific incentives could be created to encourage insureds to take more steps to reduce risks.[2]

A large body of research attempts to unravel the dynamics of insurance demand against LPHI events and particularly the underinsurance puzzle (see Jaspersen, 2016; Robinson and Botzen, 2019). The expected utility theory (EUT) has dominated the neoclassical economic literature, with agents' attitudes toward risk being a fundamental driver of insurance demand (Arrow, 1971). Despite its normative appeal, several systematic violations challenge the EUT model, which fails to explain low insurance demand against LPHI risks. Prospect Theory (PT), Rank-Dependent Utility (RDU), and Cumulative Prospect Theory (CPT) formulated by Kahneman and Tversky (1979), Quiggin (1982) and Tversky and Kahneman (1992), respectively, offered the most promising alternatives to the EUT model. These descriptive models incorporate additional components on risk preferences, such as reference-dependent behavior, probability weighting, and loss aversion. Similarly, behavioral biases and heuristics underlying LPHI risks are increasingly recognized in the literature as local drivers of underinsurance. Pitthan and De Witte (2021) emphasize their moderating effects on the relationship between extreme risk attributes, individuals' characteristics and insurance uptake.[3]

In addition to frictions in insurance demand driven by misunderstanding of risk or by biased discounting of severe losses, empirical research investigates more closely the drivers of insurance demand under the premise that households are price-takers. Based on insurance market data, a negative correlation is found between the premium costs and insurance uptake even for actuarially favorable rates. The level of public subsidies that intends to decrease the final insurance cost may also be an important factor influencing the decision to insure. However, the participation in insurance programs with subsidies remains low (Kousky, 2018; Landry and Jahan-Parvar, 2011). Kunreuther (1984) and Cai, De Janvry and Sadoulet (2015) focus on individual drivers and show that past loss experience and peers' effects increase demand. Furthermore, mitigation actions reduce the expected loss and, therefore, could reduce the perceived need for insurance. Individuals may also be liquidity constrained, especially in the absence of credit markets, and thus unable to afford insurance premiums (Cole et al., 2013).

The tightness or the absence of insurance markets for extreme risks is, however, a major limitation to the analysis of insurance determinants from transaction data. Additional methodological biases may also arise when estimating the impact of the premium cost on demand through market data due to endogeneity. Given these limitations, stated preference methods are increasingly used in the literature to measure the WTP as a proxy for the premium paid to avoid risk.[4] Contingent valuation is one of the most well-developed stated preference methods, which uses surveys and experiments to elicit willingness to pay for risk protection. This method allows for a more flexible measurement of the demand curve under various

[2] Hudson et al. (2017) show that insured populations are more likely to undertake disaster preparations and preventive actions than the non-insured.

[3] Jaspersen and Ragin (2021) propose a new descriptive model for decisions under risk based on anchoring and adjustment mechanisms.

[4] In many research areas, stated preference methods are the only available valuation approach. It is worth noting that contingent valuation is applied even where a revealed preference option is available.

conditions.[5] When compared to revealed preference results for LPHI insurance, contingent valuation surveys often confirm earlier results of low demand (Jaspersen, 2016; Robinson and Botzen, 2019). Contingent valuation methods are however sensitive to several methodological considerations. For instance, results from experience-based studies tend to show relatively high average WTP compared to actuarially fair prices, pointing to significant upward bias attributable to the data collection method (Leblois et al., 2020).[6] This discrepancy is fueled by a high degree of heterogeneity across studies conducted over different time periods, in different areas, for different types of risks and assets, and with different elicitation formats etc. All this limit the ability to answer some key questions about LPHI insurance demand, including: 1) How can all the stated average WTP information be aggregated? 2) Is there publication bias in the literature? 3) How low is the demand for LPHI risk insurance? While the focus has been on local drivers of insurance demand for target specific populations, little attention has been paid to measure the sensitivity of fair premium adjusted WTP estimates to risks attributes, socio-economic characteristics and elicitation methods specificities. As the number of studies on LPHI risks are expected to keep growing in the coming years, disentangling systematic and noise effects on WTP estimates is highly needed.

The purpose of this paper is to perform a meta-analysis to explain the overall WTP for LPHI insurance after adjusting for publication bias and heterogeneity. From a sample of 37 primary studies (65 observations), we conduct a meta-regression to examine whether observed variations across studies are due to structural determinants or to methodological and statistical artifacts. This meta-analysis is the first to focus on this topic spanning 17 years of research and extends, with a new quantitative perspective, the previous surveys of literature (e.g. Jaspersen, 2016; Robinson and Botzen, 2019; Harrison et al., 2019; Lucas, 2021). It provides estimates of the effect of several fundamental and methodological factors on stated WTP. It also allows for testing and correcting the literature for potential publication bias, a goal that cannot be achieved at the individual study level (Stanley et al., 2012). Finally, this meta-analysis contributes to improving the understanding of the empirical effect of hypothetical bias for various WTP elicitation designs (Lusk and Schroeder, 2004; Miller et al., 2011; Schmidt and Bijmolt, 2020).

Empirical analyses are performed according to meta-analysis guidelines (e.g. Havranek et al., 2020; Steel et al., 2021). We first show that the average stated WTP in our dataset is around 87% of the actuarially fair insurance premium. After heterogeneity is accounted for, this ratio is not significantly different from one suggesting that there is no a priori underestimation of tail risks in an international perspective. To explore the sources of heterogeneity, we conduct a meta-regression analysis using different moderators. To account for model uncertainty and moderator selection, we use the Bayesian Model Averaging approach (BMA). The meta-regression analysis reveals that stated WTP it is highly sensitive to risk characteristics and elicitation methodology, with a steady downward trend over time. WTP turns out to be higher for losses with extremely small descriptive probability levels. Laboratory experiments, within-subjects design, and participation fees also inflate average WTP. Conversely, sample average income and age negatively affect WTP. We find little evidence in support of the acknowledged vulnerability of stated WTP to hypothetical bias traditionally associated with contingent valuation methods. All these findings are robust to different model specifications and publication bias.

The rest of the paper is structured as follows. Section 2 presents the general theoretical framework underpinning the willingness to pay for insurance and formulates hypotheses regarding the relevant factors

[5] Contingent valuation methods suffer from several limitations attributed mainly to the hypothetical nature of the survey that tends to overestimate the true willingness to pay. For example, individuals might not be able to judge the value of the goods they have to evaluate, due to a lack of understanding; or they might have difficulty envisioning their income constraints in the proposed hypothetical setting (Diamond and Hausman, 1994). These limitations dramatically reduce external validity.

[6] Leblois et al. (2020) note a high take-up or willingness of insurance giving the following experimental studies as examples (Petraud et al., 2015; Norton et al., 2014; Serfilippi et al., 2018). They explain such discrepancy by the presence of seasonal liquidity constraint or distrust in the insurer that are perhaps ignored in experimental settings.

that may influence it. Section 3 describes the meta-dataset and discusses the metric construction. Section 4 examines publication bias. Section 5 investigates heterogeneity and presents the meta-regression results. In this section, we further check the robustness of the obtained results. The last section concludes.

2. Theory and hypotheses development

As a conventional measure of the change in welfare, compensating variation is defined as the maximum amount an individual would be willing to pay (WTP) to secure a change (i.e. restore the original welfare level) (Hanemann, 1991). For insurance coverage without deductibles or other cost-sharing limits, the willingness to pay (WTP) is implied from the following indifference condition between insurance and non-insurance decisions:

$$U(y - WTP, p, q_1; Z) = U(y, p, q_0; Z) \tag{1}$$

where U describes the agent indirect utility (value) function[7], y represents the individual's wealth (income), p is a vector of costs that the individual faces, q_i (with i=0, 1 and $q_0 < q_1$) reflects the safety value, and Z is a vector of personal characteristics (such as past loss experience, financial literacy age, gender, etc.). Parameters q_0 and q_1 describe different levels of the safety measure q_i. Parameter q_1 is associated with a measure that provides a higher level of safety compared with q_0 (Entorf and Jensen, 2020). Theoretical WTP for full insurance can be derived from different risk preference models. Based on the EUT model, WTP would be equal to the certainty equivalent of the expected utility without insurance. For this model, the difference between implied WTP and the actuarially fair premium corresponds to the standard risk premium for risk-averse agents.[8] In practice, the disparity between stated WTP and EU predictions calls for more comprehensive models with additional factors and extra risk premiums (such as risk attitudes, social preferences, sensitivity to heuristics and bias, etc.) (Baillon et al., 2022). Given that risk preferences are unobservable and difficult to measure objectively, it is of relevance to get insight into observable factors that moderate risk attitudes determinants and structurally impact demand. These factors fall into three categories: demand-side, supply-side and nature of extreme risk (Leblois et al., 2020).[9]

Based on observable factors, we formulate four hypotheses to investigate stated WTP variability. First, we consider the potential effect of small probability level on WTP normalized by actuarially fair prices. The significance of this relationship is essential to understand the extent to which small probability perception affects insurance decisions. Second, we evaluate the potential impact of income on insurance demand. Although this factor is expected to be positively related to insurance demand, empirical results are rather mixed. Third, we assess potential differences between small probability idiosyncratic and correlated losses. The last hypothesis refers to the influence of the elicitation method design and how it may introduce hypothetical bias in the estimation of the WTP. We predict that some specific elicitation details may exacerbate this bias. In addition to the above-mentioned factors, we examine other moderating variables described in section 5.

[7] See Barseghyan et al. (2018) for a summary of risk preference models.

[8] Under EU theory, a risk-averse agent will buy full insurance if and only if the premium is fair, i.e. equal to expected losses, (Mossin, 1968). For small probability risks, the demand for full insurance at unfair premiums or less than full insurance at fair premiums contradicts the EU theory (Schlesinger, 1997).

[9] While there are certainly others, these categories are the most frequently mentioned and examined in the literature.

2.1 Probability level

Economic theory recognizes the importance of risk attitude and probability of losses as two relevant inputs in decision-making under uncertainty (Michailova et al., 2020).[10] The perceived losses probability appears, however to have a nonlinear effect on insurance decision, a result highlighted by prior experimental studies (Robinson and Botzen, 2018). Two distinct behaviors are observed depending on the probability level. When people deem the probability of loss below their level of concern, they generally neglect risk and choose not to undertake protective actions (Kunreuther and Pauly, 2004; Kunreuther et al., 1978). In contrast, when they attribute a subjective likelihood of loss that is far higher than the actual probability, they become highly risk-aware and seek out mitigation strategies (e.g., Brouwer et al., 2014).

Two broad paradigms have dominated the literature on the impact of probability factor in choice under risk. The first one is related to the decision from description (DFD), where an explicit and precise description of the loss probability distribution is provided to subjects before decisions are made. The DFD framework has been widely used in laboratory, online and field experiments to elicit risk preferences. It gives rise to significant theoretical developments such as the probability weighting concept (Tversky and Kahneman, 1992). The second paradigm is related to the decision from experience (DFE) where losses and their associated probabilities are unknown at the outset. Individuals have to form a subjective representation of risk based on past observations or personal experience. Therefore, behavioral implications may potentially vary depending on whether uncertain choices are made from experience or description, a phenomenon often known as "description-experience gap" (Hertwig et al., 2009). For description-based prospects, people seem to be risk averse for gains and risk seeking for losses. However, for experience-based they become risk seeking for gains and risk averse for losses (Kudryavtsev and Pavlodsky, 2012).

In the same way, low-probability losses may also affect the description-experience gap. When small probability is only learned by observation or experience, it is underestimated as individuals form their own estimations through sampling (Hertwig et al., 2004). Due to small experienced samples, they may not experience losses, which might lead them to underweight these rare events. People acquire information throughout time and revise their beliefs about the probability of extreme risks occurrence. They tend generally to underweight small probability losses (Barron and Erev 2003; Hertwig et al., 2004; Weber et al. 2004, Bakkensen and Barrage (2021). They may also tend to focus on short-time horizons rather than acknowledging long-term exposure (Kahneman and Lovallo, 1993; Read et al., 1999; Redelmeier and Tversky, 1992). In this respect, Krawczyk et al. (2017) document a persistent underestimation of loss probability for LPHI risks, even when subjects learn about the risk over time.

On the other hand, when probability distribution is explicitly provided, it is used as inputs within "substantive" decision models such as expected-utility theory (EUT), rank-dependent utility (RDU) or Cumulative Prospect Theory (CPT).[11] A major theoretical development in this regard relates to the probability transformation model. Since the seminal work of Kahneman and Tversky (1979) it has been generally recognized that decision makers transform probabilities with respect to a subjective benchmark. The weighting function was found to be typically inverse S-shaped, i.e. overweighting small objective probabilities and underweighting large ones. More recently, Jaspersen et al. (2022) introduce a new local condition – the decreasing relative overweight (DRO) – on the probability weighting function. Under this model, extremely small probabilities are overweighed relative to their base value more than small probabilities. This new condition has been empirically and theoretically proven.

We summarize the potential role of small probability level on WTP through the following assumption:

[10] Several lines of evidence contradict this assertion showing that people may have difficulty understanding and using probability information (Tyszka and Sawicki, 2011) or may not be interested in receiving or searching for information about probabilities (Huber et al. 2001; Amelung and Funke, 2015).

[11] Forgas (1995) defines "substantive" decision models where individuals execute complex mathematical operations with a high cognitive load.

H$_{1a}$: *The level of objective loss probability affects WTP: lower descriptive loss probability result in higher relative WTP.*

H$_{1b}$: *The level of objective loss probability affects WTP: lower unobserved loss probability result in lower relative WTP.*

2.2 Income and liquidity constraints

The liquidity constraint created by income may affect the WTP (Giné et al., 2008; Cole et al., 2013; Liu and Myers, 2016). Higher income households can purchase more insurance against LPHI risks. For instance, Yiannakoulias et al. (2018) observe that insurance demand is higher in high-income and high-hazard areas. Ali et al. (2021) show that farmers with liquidity constraints are less likely to join crop insurance programs, highlighting that financial restrictions inhibit insurance adoption. Atreya et al. (2015) show that income and price have a significant influence on the decision to purchase flood insurance. They identify a positive relationship between income and flood insurance purchases. Prior research of Kriesel and Landry (2004), which focuses on coastal residence insurance in the United States, has grouped participants' annual income factor into eight categories ranging from $30,000 to $250,000. They find that higher-income respondents were more likely to have flood insurance than those with lower incomes. Similarly, Kousky (2010) finds that people with higher incomes purchase more coverage and that an increase in income improves the level of coverage. An exception to this positive relationship is observed for the highest income level category.

Karlan et al. (2014) show, however, that liquidity constraints are not generally binding i.e. they are not prohibitive for insurance purchase. If smallholders are provided with the opportunity to purchase insurance, they will be able to finance it. This view is in line with the standard economic theory suggesting, based on plausible assumptions, that insurance is considered as an inferior good.[12] In other words, the perceived value of insurance declines as household wealth increases. A positive justification for this conjecture is that decreasing absolute risk aversion with wealth leads individuals to consider insurance as an inferior good. Additionally, wealthier households face lower costs to self-insure, a practice that may replace market insurance (Ehrlich and Becker, 1972). Self-insurance typically refers in the literature to preventive measures or post-loss actions to finance losses via assets liquidation or from saving.

Based on market data, Landry and Jahan-Parvar (2011) establish that higher income is associated with a larger demand for insurance, but that this relationship is not monotonous. The fact that households in the highest income bracket do not have a much larger demand than those in the lowest income bracket suggests that the former are able to compensate for losses out of pocket. De Nicola and Hill (2012), also, predict that demand for index-insurance will be hump-shaped in wealth. Wealthier households purchase more insurance than poor-households, up to the point where they can self-insure their assets. The prediction of De Nicola and Hill (2012) is in line with the findings of Clarke and Kalani (2011) who find that the demand for wealth is inverse U-shaped for wealth. We capture the effect of average income on WTP through the following hypotheses:

H$_{2a}$: *Average income levels affect WTP: higher income levels result in lower WTP.*

H$_{2b}$: *Average income levels affect WTP: higher income levels result in higher WTP.*

2.3 Idiosyncratic vs correlated risks

Evidence shows that people are more inclined to buy insurance for idiosyncratic risks than for correlated risks, Pitthan and De Witte (2021). One possible explanation of this stylized fact may lie in the difficulty of indemnifying concomitant and spatially correlated losses, resulting in fat tails loss distribution. High-cost correlated losses might thus threaten the solvency of insurers and their ability to fulfill their contractual commitments to policyholders (Biener et al., 2019). In addition to this non-performance

[12] For a recent discussion of the criteria that determine whether the demand for insurance is a normal or inferior good, see Peter (2022).

premise, a growing body of research considers that individuals' decisions related to risks are socially embedded decisions. Social comparison may introduce discrepancies into the agents' decisions compared to self-interested rational utility maximizer. Bault et al. (2008) find, for instance, that lottery results influence subjects' perceptions in different ways depending on the lottery outcomes of the other. Furthermore, subjects seems to respond more strongly to social benefits than to social losses. Linde and Sonnemans (2012) find that subjects take less risk when they can win at most as much as a certain payoff of a reference subject (social loss situation) compared to the case when they can win at least as much as a reference subject (social gain situation). Schwerter (2013) finds, however, that subjects take more risk when they are able to outperform a peer rather than remain ahead of a peer. Consequently, Schwerter (2013) interprets his results in favor of social loss aversion while Linde and Sonnemans (2012) argue that their findings suggest that loss aversion does not easily extend to the case of social comparisons.[13]

Related to the insurance market, Friedl et al. (2014) show that social comparisons make insurance less attractive when risks are correlated. They justify this result based on subjects' aversion to unequal payoffs. Fehr and Schmidt (1999) and Bolton and Ockenfels (2000) establish inequality aversion as an important dimension of social comparison. Bolton et al. (2005) find evidence in favor of inequality aversion. People are more affected by being worse off than others than being better off than others. There is a clear evidence that disadvantageous inequality (being worse off) reduces utility.

A growing body of literature shows how social reference points influence risk-related decisions in general and insurance underwriting in particular.[14] If inequality aversion is present, the existence of a social reference point makes insurance less appealing for correlated risks than for idiosyncratic risks, Fehr and Schmidt (1999). While the findings of Linde and Sonnemans (2012) and Vendrik and Woltjer (2007) reject decreasing sensitivity in the presence of social reference points, they report evidence in favor of loss aversion. In contrast, Bault et al. (2008) find the opposite of loss aversion (i.e., gain seeking) in the presence of social reference points. Their result, however, relies on subjects' evaluation of emotions and a model that does not precisely resemble that of the prospect theory. Another related experimental study conducted by Rohde and Rohde (2011), examines how participants' lottery decisions were affected by the risk incurred by a peer group. They find that subjects prefer risk to be distributed idiosyncratically rather than correlated.

Overall, these studies show that social comparison can have a strong impact on decisions under risk. Because the social reference point for insurance problems differs significantly for correlated and idiosyncratic risks, the demand for insurance can be strongly influenced by losses correlation. We formulate the following hypothesis:

H3: *Risk type affects WTP: correlated risks result in lower WTP.*

2.4 Elicitation methodology

The analysis would be incomplete without discussing the impact of elicitation design. Since the agent's true WTP is unobservable and there is no widely acknowledged elicitation method (Völckner, 2006), methodological choices and assumptions on valuation estimates may influence the results obtained. Carson et al. (2001) give some guidance in this area and outline two important aspects of the elicitation methodology. The first consideration is whether the elicitation process is based on price generation (i.e., an open-ended question format) or price selection tasks (i.e., a dichotomous choice question format) (Hofstetter et al., 2021). The second aspect is whether an actual economic commitment is required with the declared price. For example in a hypothetical setting, agents who declare their WTP have no obligation to purchase insurance.

---

[13] Rohde and Rohde (2011) do not find evidence of social comparisons affecting risk taking.

[14] Tversky and Kahneman (1992) suggest that a reference point may not depend solely on the status quo of the decision maker but may also be influenced by the social comparison that generates social reference points.

Whatever the chosen format, an important requirement of the elicitation methodology is that of incentive compatibility i.e. to accurately reflect the perceived economic value. The main drawback of contingent valuation methods is the lack of incentive compatibility in the sense that there is not a dominant strategy to bid truthfully (Wertenbroch and Skiera, 2002). Becker, DeGroot and Marschak (1964) formalize a first incentive-compatible elicitation procedure, known as BDM, where agents are asked to provide an offer for the contract being valued. This price is then compared to a randomly drawn fixed price, which is used as the trading price. A participant's dominant strategy is to offer exactly their value. A second incentive-compatible valuation method is the multiple price list (MPL) in which participants are presented with an ordered list of prices that they are asked to accept or reject. One price is randomly selected and the respondents' choice for that price is implemented (Andersen et al., 2006; Kahneman, Knetsch, and Thaler, 1990). Thus, MPL method is incentive-compatible because it is in the respondent's interest to accept if and only if the price is lower than the true WTP (Alfnes and Rickertsen, 2011).[15]

Empirical results have provided evidence of systematic bias in the estimation of the WTP referred to in the literature as "hypothetical bias". This bias tends to overestimate actual WTP versus a similar actual purchasing decision (Murphy et al., 2005). Hypothetical bias is generally attributed to the hypothetical nature of stated preferences surveys, which provide different incentives than those encountered in real-world situations (Hensher et al., 2015). In addition, previous studies have documented potential interplay between the hypothetical bias and the question format (e.g., Balistreri et al., 2001; Harrison and Rutström, 2008). Some specific question formats may encourage strategic behavior (Carson and Groves, 2014). It is well established that dichotomous choice typically overstates WTP relative to open-ended (Balistreri et al., 2001) and payment card formats (Ready et al., 1996; Welsh and Poe, 1998).[16] A number of calibration techniques have been proposed to de-bias contingent valuation surveys. Ex ante techniques attempt to improve hypothetical methods at the data collection stage by priming survey subjects, while ex post approaches attempt to calibrate the data after measuring WTP. The use of a certainty follow-up question is among the most popular ex post corrections measure. Consequentiality is an alternative technique used to mitigate hypothetical bias (Carson and Groves, 2014; Vossler et al., 2012).[17]

Related to insurance studies, Robinson and Botzen (2019) find some evidence for hypothetical bias in their results. Kesterniche et al. (2013) concluded that the significance and signs of the estimated coefficients for insurance demand are similar between market data and hypothetical choice experiments. The empirical results of Cole et al. (2020) provide no support for hypothetical bias in estimated demand for different elicitation mechanisms. All these results are in line with the earlier Loomis (2011) findings where hypothetical bias is less severe for WTP elicited for private goods. Despite all the empirical attempts to mitigate hypothetical bias, there is no widely accepted procedure to deal with this issue since its underlying causes are multi-dimensional and not yet sufficiently understood. We capture the potential effect of elicitation method formats on WTP through the following hypotheses:

H_{4a}: *Incentive-aligned elicitation mechanisms are associated with lower WTP.*

H_{4b}: *Price selection tasks are associated with a higher hypothetical bias and thus with a higher WTP.*

[15] Smith (1982) emphasizes the importance of salient payoffs, in which participants should get incentives (most typically money) for their involvement in the experiment, and these rewards should depend on the experiment's results.

[16] When asking agents directly about their WTP, they are more likely to dwell on the price or they may attempt to respond strategically if they believe their responses will affect future retail pricing (Breidert et al., 2006, Jedidi and Jagpal, 2009). Moreover, t simulating real purchasing experiences demand less mental effort than those asking respondents to indicate their WTP (Brown et al., 1996). There are also limitations to indirect approaches that might be possibly affect the hypothetical bias. Smith et al. (2019) argue that generally respondents are uncertain about their preferences and that this uncertainty leads to systematically different responses depending on the question format.

[17] Haghani et al. (2021) provide a comprehensive list of ex-ante and ex post bias mitigation.

3. Methodology: The Meta-Data Set

The first stage in the meta-analysis method consists in collecting and selecting the primary studies over the period from 2005 to 2021. To this end, we follow the reporting guidelines outlined in the PRISMA statement, Havranek et al. (2020). We search for empirical studies published in the Web of Science and Google Scholar databases using a combination of the following keywords: "Insurance", "willingness to pay", "low probability", "contingent valuation", "climate risk" and "natural disasters". We also reviewed the bibliographies of the retrieved papers for more empirical research.

In order to determine which primary studies to include, a list of selection criteria is established. These criteria are necessary to ensure that the final dataset contains studies with a reasonable degree of heterogeneity while still allowing for meaningful comparison. For an illustration of the PRISMA steps and results, see Figure 1. The identification stage included the 15,664 articles identified by the database search.[18] We remove duplicates and screen articles based on title (1,057 articles). Four main exclusion criteria were applied during the selection phase: (a) the nature of the risk considered (i.e. we exclude studies dealing with health or life risks), (b) Non contingent valuation based empirical research (i.e. WTPs obtained from theoretical modelling or discrete choice experiment articles are excluded), (c) Multiple protection mechanisms (studies considering other mechanisms in conjunction with insurance are excluded), and (d) the loss probability threshold (when it is given for idiosyncratic losses) is less than 5%.[19]

For the eligibility step, we include studies that report information allowing to directly or indirectly measure actuarially fair premium or, equivalently, the average historical cost of annual losses. In addition, two crucial statistics should be reported in studies: the mean WTP estimates and their corresponding standard errors or variance estimates. We included peer-reviewed articles published in English or in Chinese.[20] To identify additional studies, we reviewed the reference list of retrieved articles. We included 38 studies in our meta-analysis, which led to 74 data points because some studies included multiple observations (see Appendix C).[21, 22]

[18] These values are provided according to Google Scholar results, which covers more papers than Web of Science.

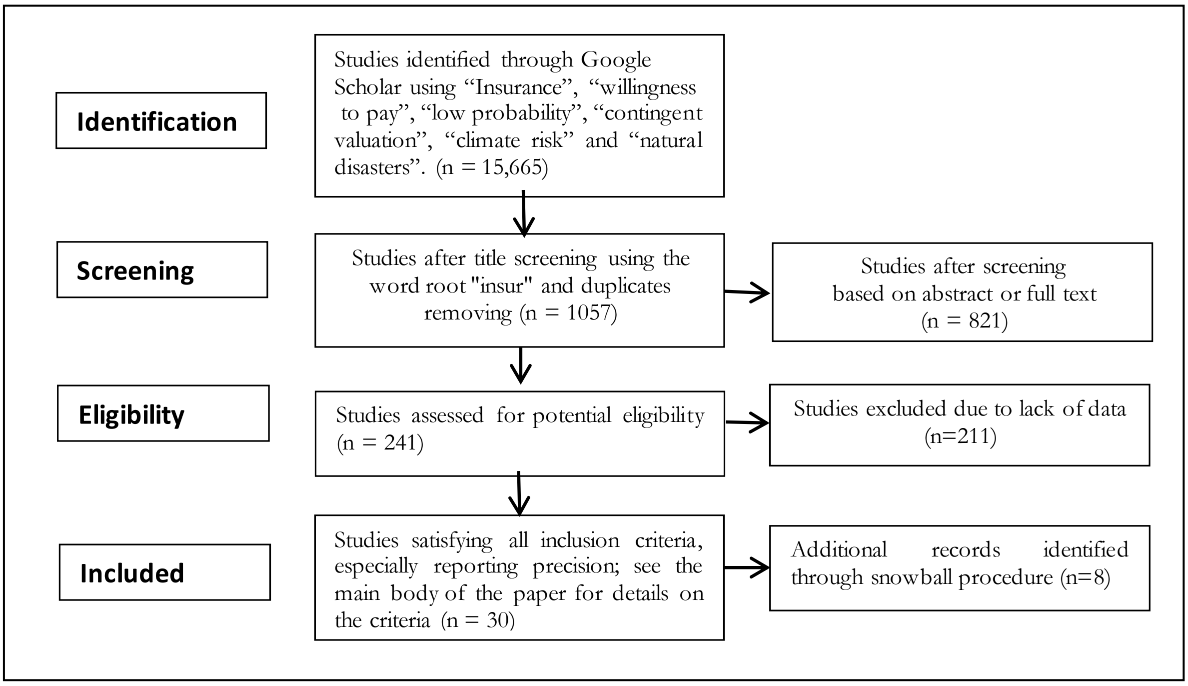
[19] For three experimental studies examining correlated simulated risks, the probability of loss is more than 5%.

[20] Because the vast bulk of research on developing countries is conducted in China, we include Chinese published papers (see Appendix C).

[21] Attempts have been made to get variance estimates from authors.

[22] The final dataset includes WTP estimated either from observational or experimental studies. Three separate instances are covered by observational studies. WTP may be elicited using either an hypothetical/actual scenario that includes a comprehensive description of probability and loss levels or from a basic scenario with no probability information, or no scenario at all. On the other hand, experimental studies encompass two cases. The first one estimates the WTP for different levels of probabilities and/or losses. The context description is neutral and the insurance contract is without default risk. For the second type of studies, the WTP is elicited by the manipulation of additional factors (e.g. framing, default probability, etc.) other than probability or losses that are fixed and known throughout the experiment. For these studies, we select only WTP elicited from the control group, where the scenario description is neutral and the insurance contract is without default risk.

Figure 1: PRISMA flow diagram of the empirical studies included in the meta-analysis



3.2 Effect Size

As a natural measure of LPHI insurance perceived value, average WTP would be the most tractable. To ensure comparability, this metric expressed in different currencies across studies should be standardized. When a steady conversion factor is available such as purchasing power parity-corrected exchange rates, a potential issue is the lack of consistency caused by conversion. Various sources of heterogeneity are likely to exist across countries (e.g., risk characteristics, relative cost of insurance, cultural risk perception, etc.). Thus, expressing average WTP in a single currency unit would be inaccurate. Standardizing average WTP by its standard deviation eliminates the original units issue by expressing estimates in relative terms. However, standard deviations are highly dependent to the variable scale. In addition, studies using different experimental designs will have different standard deviation values, which will reduce comparability (Morris and DeShon 2002). Standard deviations may also be more subject to publication bias in that studies with large standard errors produce estimates with large confidence intervals and would be more difficult to publish. As an alternative to statistical rescaling, we index the average WTP by the actuarially insurance fair price to get more comparable values. The ratio is defined as the relative willingness to pay (RWTP):

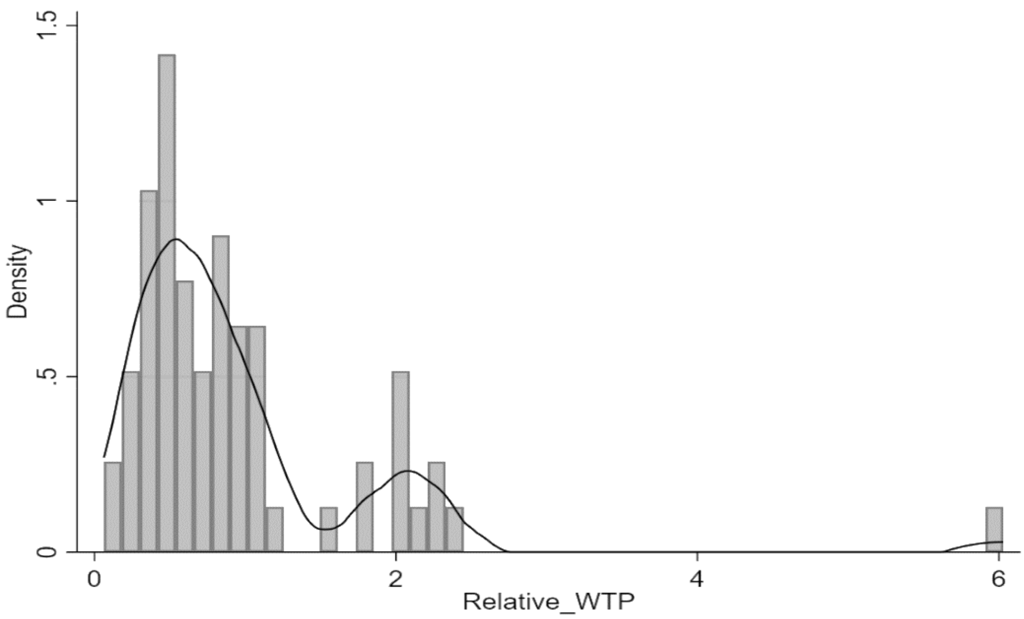
$$RWTP_i = \frac{\overline{WTP_i}}{FP_i} \tag{2}$$

where FP_i denotes the actuarially fair insurance price for study (i) approximated from historical average losses or calculated from the theoretical loss distribution. Rescaling with fair prices would also be a good indicator of value perception. If WTPs are marked up substantially below-actuarially fair levels, this indicates a negative attitude towards insurance, and vice versa.

Figure 1 illustrates the relative WTP distribution across all studies included in our meta-analysis. The distribution is bi-modal right-skewed. We note that 70.6 % of the dataset observations are less than one, and 13.8% are more than two. The weighted mean value of RWTP is 0.875, with minimum and maximum values of 0.062 and 6.029. The standard error varies greatly, which raises concerns about outliers that may distort the validity and robustness of the conclusions from the meta-analysis (Viechtbauer and Cheung, 2010). To alleviate this problem, we trim the RWTP and the standard error at the top of the 5% level, and then our final dataset has 65 observations.[23] Table 1 provides the average RWTP values for all estimates and for different groups of studies. The first column reports the unweighted means and the second column reports the weighted means. The overall mean of the RWTP weighted by the inverse of the number of reported observations is about 0.875. At this level, the overall average value should be interpreted with caution because of potential publication bias and heterogeneity. In the next section, we will examine whether publication bias exists and how it might affect estimates. Experimental studies appear to report higher values for RWTP than observational survey studies. Similarly, studies dealing with idiosyncratic risks appear to report higher values for RWTP than correlated risks. The difference between China on the one hand and Germany and the Netherlands on the other is significant. Finally, RWTP generated from laboratory experiments show the highest values.

The RWTP metric shares many common features with response ratio (RR) defined as the ratio of average outcomes between an experimental and a control groups (Schmidt and Bijmolt, 2020).[24] The response ratio is mainly employed in the ecological field to assess effect size (Koricheva and Gurevitch, 2014). To the best our knowledge, it has not yet been adopted in the insurance context. We perform statistical analyzes using the natural logarithm of the RWTP as the dependent variable. First, using the logarithm linearizes the metric, so that deviations in the numerator and denominator have the same impact (Hedges et al., 1999). Second, the moderating variables coefficients in the meta-regression would be easier to interpret. Third, the distribution of the logarithm of response ratios is approximately normally distributed (Hedges et al., 1999).

Figure 2: Histogram of the relative willingness to pay (RWTP) for LPHI insurance



[23] Winsorization is an alternative method to apply in the meta-analysis (Lipsey and Wilson, 2001). It substitutes the extreme values with the highest values in given percentiles.

[24] Without loss of generality, we may assume an artificial control group exposed to the same LPHI events than the treatment group and composed of identical synthetic individuals (perfectly rational, risk-neutral, maximizing a Neumann & Morgenstern type utility function and not subject to behavioral or psychological bias). Under these conditions, individual's nominal WTP would be equal to the certainty equivalent of a lottery of finite set of possible losses i.e. the fair actuarially premium. In other terms, the WTP for this group will be the same for all respondents – with a degenerate distribution– equals to the actuarially fair premium.

We define the effect size by the standardized willingness to pay (Henceforth SWPT) as follows:

$$SWTP_i = \ln\left(\frac{\overline{WTP_i}}{\mathrm{FP}_i}\right) \tag{3}$$

As in Lajeunesse (2015), we define the sampling variability of the SWTP as:

$$\operatorname{var}(RR_i) = \frac{std_i^2}{N_i \overline{WTP_i}^2} \tag{4}$$

where std_i denotes the standard deviation of the WTP for study (i) and N_i is the sample size.

Table 1. Full sample and subsample average RWTP

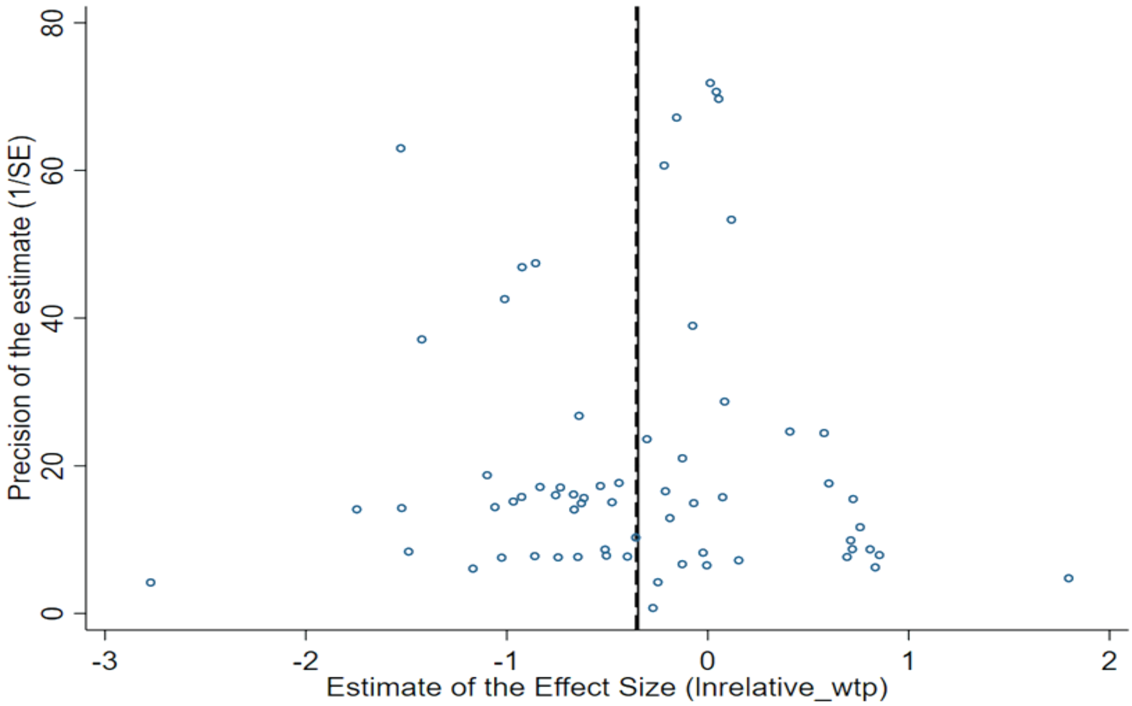
	Unweighted mean	**Weighted mean**	**Observations**
Full sample	0.944	0.875	65
Subsample observations			
Survey	0.613	0.641	40
Experiment	1.473	1.455	25
Idiosyncratic risk	1.534	1.672	14
Correlated risk	0.782	0.707	51
China	0.433	0.447	21
Germany	1.463	1.592	13
Netherlands	1.358	1.022	12
Laboratory	1.787	1.783	8

Note: In weighted means, SWTP values are weighted by the inverse of the number of estimates per study.

4. Publication Bias

Publication bias is a perennial concern in meta-analysis that arises when authors choose to hide their insignificant results, or when editors or reviewers subjectively favor studies with more significant results. Large publication bias may distort the estimation of the mean overall effect and the conclusions drawn. A starting point to get some insight into the presence of publication bias is by a visual inspection of the funnel plot. This graphic plots the effect size on the horizontal-axis and the precision of the estimates (1/standard errors) on the vertical-axis. If the distribution of the standard error is symmetrically distributed around the mean line, there is no publication bias. Figure 3 displays the funnel plot, which does not allow to make conclusive statements. The funnel plot is slightly skewed to the left –as the standard error increases– indicating a potential upward publication bias towards negative relative WTP.

Figure 3. Funnel Plot of the SWTP



Notes: The horizontal axis represents the SWTP values. The vertical represents the inverse of the standard errors

A more formal and accurate way to detect publication bias is the "Funnel Asymmetry Test"-"Precision Effect Test" (FAT-PET) proposed by Stanley (2008). This test assumes that publication selection induces a correlation between the estimated effect size and their standard errors. The FAT-PET is implemented through testing the slope of the regression of the effect size, i.e., SWTP, on its standard error:

$$SWTP_i = \alpha_0 + \alpha_1 SE(SWTP_i) + \varepsilon_i \tag{5}$$

where SWTP is the i^{th} standardized WTP estimated in study *s* and *SE(SWTP)* is the corresponding standard error, α_0 is the true effect after correcting for publication effect and α_1 is a measure of the importance of publication bias. Testing for $\alpha_0 = 0$ is a precision effect test (PET) for a genuine empirical effect net of publication bias, whilst testing for $\alpha_1 = 0$ is the funnel asymmetry test (FAT) for publication selection. The statistical significance of the estimate of α_1 is an indicator of the presence of publication bias. Since the empirical studies in our dataset use different methods, models and sample sizes, ε_i are likely to be heteroscedastic. Equation 5 is thus estimated using the weighted least square (WLS) method[25]. We use Fixed effects (FE) and Random effects (RE) methods. WLS-FE estimator assumes that there is a single underlying true effect and explain difference in estimates across studies as only due to sampling error. As a consequence, the weight is the inverse of the variance of the estimated effect, $1/SE(SWTP_i)$. WLS-RE estimator assumes, however, that true effects can differ across studies so that the variation in

[25] We further discuss the heteroscedasticity issue in section 5 when we produce heterogeneity analysis.

estimated effects is composed of two parts: heterogeneity (between studies) in the true effect and sampling error. The weight is thus $1/SE(RWTP_{I,s}) + \tau^2$ where τ^2 measures the variance of the true effect.[26]

Table 2: FAT-PET and PEESE of publication bias (WTP)

	OLS	WLS-RE	WLS-FE	OLS	WLS-RE	WLS-FE
	FAT-PET	FAT-PET	FAT-PET	PEESE	PEESE	PEESE
	(1)	(2)	(3)	(4)	(5)	(6)
Weight 1: Equal weight to each estimate						
Publication bias (α_1)	0.14	0.34	-0.87	0.044	0.056	0.25
	(0.32)	(1.03)	(4.09)	(0.065)	(0.16)	(5.27)
Precision effect (α_0)	**-0.36****	**-0.38*****	**-0.33**	**-0.35****	**-0.35****	**-0.36***
	(0.13)	(0.13)	(0.30)	(0.14)	(0.14)	(0.21)
Observations	65	65	65	65	65	65
R-squared	0.001	0.176	0.127	0.0002	0.175	0.275
Number of studies	36	36	36	36	36	36
Weight 2: Equal weight to each study						
Publication bias (α_1)	0.14	0.22	3.32	0.073	0.072	2.47
	(0.15)	(0.46)	(5.22)	(0.064)	(0.082)	(3.98)
Precision effect (α_0)	**-0.42*****	**-0.43*****	**-0.65***	**-0.41*****	**-0.41*****	**-0.55****
	(0.12)	(0.12)	(0.37)	(0.12)	(0.12)	(0.23)
Observations	65	65	65	65	65	65
R-squared	0.002	0.241	0.280	0.001	0.240	0.404
Number of studies	36	36	36	36	36	36

Notes: Robust standard errors clustered at study level are shown in parentheses. OLS = ordinary least squares, WLS-FE = weighted least square fixed effects; WLS-RE = weighted least square-random effects.*** $p<0.01$, **$p<0.05$,* $p<0.1$

Columns 1 to 3 of Table 2 present the results of three specifications based on equation (5): simple OLS, WLS-RE and WLS-FE. To accommodate within-study correlation of estimates for each specification, we report cluster-robust standard errors with a clustering by study. Moreover, two weighting schemes are used for each specification: equal weights for each estimate (weight 1) and equal weights for each study (weight 2). The second weighting scheme allows to consider multiple effect-size estimates reported by primary studies.[27] For all specifications, we do not reject the null hypothesis for α_1, which indicates the absence of publication bias: ($\alpha_1 = 0$ at the 10% significance level). Moreover, Stanley (2008) argues that the publication bias-corrected estimates of the mean true effect (α_0 in Equation (1)) may be biased downward when the null hypothesis is rejected. While the null hypothesis is not rejected in, we follow the procedure proposed by Stanley and Doucouliagos (2014) that consists in replacing the standard error with its squared term (quadratic specification), i.e., the variance. The meta-regression is called in this case the Precision Effect Estimate with Standard Error (PEESE). Columns 4 to 6 of Table 2 displays the PEESE results. We find the same results than columns 1 to 3, that is there is no publication bias in all specifications (OLS and WLS-RE in the two-weighting scheme)[28].

[26] There are several estimators of τ^2. We use the restricted maximum–likelihood (RELM) estimator and the Knapp-Hartung standard-error adjustment. Our results are robust to other estimators such as the Hedges estimator, the Šidák–Jonkman estimator and the DerSimonian–Laird estimator).

[27] To mitigate the domination effect of studies with large number of estimates, we estimate equation (5) with frequency weights, specified as the inverse of the number of estimates reported in each study.

[28] Because we do not find evidence for publication bias in our results, we decide not to conduct further available tests to measure the bias-adjusted true effect (α_0) such as the weighted average of adequately powered (WAAP) estimator by Ioannidis et al. (2017), the endogenous Kink (EK) estimator by Bom and Rachinger (2019) or the AK estimator by Andrews and Kasy (2019).

5. Heterogeneity Analysis

5.1 Variables Description

As shown in Figure1, the average RWTP varies considerably across studies. The null hypothesis of Cochran's Q test presented in appendix D indicates a large level of heterogeneity between studies (I^2 > 75%). To deal with heterogeneity and to identify the most effective factors that would explain differences between RWTP, we define 23 moderating variables (binary, multinomial and numeric) as covariates in the meta-regression. As a second objective, we define two synthetic study profiles that simulate an average RWTP using all estimates, but overweighting those that are better identified.

We separate the moderators into the following categories: WTP elicitation design, risk specificities, exposed assets and insurance features, sample characteristics, spatial-temporal variations and publication characteristics. Table 4 presents the definition and summary statistics of all variables included for heterogeneity analysis. In the first category, we consider moderators that focus on the survey design and measurement characteristics of WTP. As reported in table 2, this first category represents an important source of heterogeneity. We distinguish observational surveys 62% of our dataset from controlled experiments 38%, which breaks down to 25% for online experiments, 12% for laboratory experiments and 1% for field experiments. For some studies, there are several scenarios where authors use between-subjects or within-subjects designs.

For the WTP measurement methods, we describe the prevalence of compatible incentive mechanisms by a binary variable with an average value of 18%. We also consider the fact that WTP is measured using a price generation approach (e.g., an open-ended question) as opposed to a price selection approach (simple or double dichotomy method). Hypothetical bias mitigation correction is modelled by a binary variable that indicates whether researchers employ bias mitigation strategies (e.g. cheap talk, consequential script, follow-up question, etc.). The participation fees binary variable indicates whether participants received monetary compensation for their participation in the study.

We encoded the variability of the risks by considering the difference between idiosyncratic 21.5% and correlated risks 78.5%. For the second category, we specify different subclasses (flood 61%, various climatic risks 9%, and earthquake 1%). For the risk characteristics, we define a first numeric variable equal to the descriptive probability provided and a second binary variable that describes the studies in which the implicit probability of loss is below the 5% threshold. When the probability is not provided (68% of the cases), we estimate it from the fair premium or the average loss. Regarding the third category, we note that assets exposed to small probability risks are disparate. We define two main classes of assets: crops and property (house and contents), which account for 40% and 32% of our sample, respectively. Regarding insurance contracts, 60% (40%) are indemnity-based (index-based), where 29% have a subsidized premium. Sample and data characteristics include a set of dummy variables to indicate whether the estimates are related to the entire population or from targeting populations at risk (51% of the dataset). We also code two binary moderators for studies that distinguish for protest WTP and zero WTP. We create a set of variables related to the main countries of study, which are China, Germany and the Netherlands with the presence frequencies of 40%, 20% and 18% respectively. We include the year of data collection as well as the region. We introduce two variables related to average age and to average annual income converted to US dollars using the corresponding exchange rates.

The last category of moderators contains publication characteristics and relies on four variables. The first one is the number of citations to account for study quality. A second variable indicates whether the study was published in an international academic journal recognized by the French National Research Center (CNRS). We also denote by a binary variable studies with low citations (less or equal to one). Finally, we perform a diagnostic test for multi-collinearity on all variables. The values of the variance-inflation factors for all variables are lower than 9, with an average VIF less than 5.

Table 3. Description and summary statistics of variables

Variable	Description	Mean	Std. deviation	Weighted Mean
SWTP	Logarithm of average WTP divided by the actuarially fair insurance price	-0.346	0.763	-0.385
Standard error	Standard error of SWTP	0.101	0.165	0.112
Measurement of WTP design				
Observational survey	=1 if estimate is from observational survey data, 0 otherwise	0.615	0.50	0.713
Experience	=1 if estimate is from experience data, 0 otherwise	0.384	0.49	0.286
Lab	=1 if estimate is from laboratory experience, 0 otherwise	0.123	0.33	0.137
Field	=1 if estimate is from field experience, 0 otherwise	0.015	0.124	.0274
Online	=1 if estimate is from online experience, 0 otherwise	0.246	0.434	0.122
Within design	=1 if estimate is from within subjects design, 0 otherwise	0.353	0.481	0.181
Incentive compatible	=1 if the transaction is incentive compatible, 0 otherwise	0.184	0.391	0.126
HB mitigation	=1 if there is hypothetical bias mitigation, 0 otherwise	0.169	0.388	0.132
Elicitation method	=0 if WTP is generated through direct elicitation methods =1 if WTP is generated through hybrid elicitation methods =2 if WTP is generated through indirect elicitation methods	0.415	0.496	0.521
Participant fee/show-up	=1 if participants received participation fee or show-up, 0 otherwise	0.23	0.424	0.209
Exposure risks types and characteristics				
Idiosyncratic risk	=1 if study considers no correlated losses, 0 otherwise	0.221	0.414	0.173
Climate risk	=1 if study examine coverage against climate risk , 0 otherwise	0.092	0.291	0.137
Earthquake risk	=1 if study examine coverage against Earthquake risk, 0 otherwise	0.015	0.124	0.027
Flood risk	=1 if study examine coverage against Flood risk, 0 otherwise	0.615	0.49	0.593
Descriptive prob information	=1 if the probability of loss is provided, 0 otherwise	0.323	0.471	0.291
Descriptive prob value	Value of the provided probability of loss less than 5%	0.006	0.015	0.004
Probability estimated	Value of the estimated probability of loss	0.05	0.088	0.07
Probability estimated alpha	=1 if the probability is estimated and less than a specific threshold, 0 otherwise	0.415	0.496	0.324
Exposure assets and insurance characteristics				
House	=1 if exposed asset is a property (house/contents), 0 otherwise	0.323	0.471	0.232
Crop	=1 if exposed asset is a crop, 0 otherwise	0.4	0.493	0.438
Indemnity insurance	=1 if study considers indemnity insurance, 0 otherwise	0.6	0.493	0.561
Presence of subsidy	=1 if there is insurance premium subsidy, 0 otherwise	0.292	0.458	0.246

Variable	Description	Mean	Std. dev.	Weighted Mean
Sample characteristics				
Subject pool	=1 if WTP is estimated from general population, 0 otherwise	0.293	0.458	0.256
Random sample	=1 if WTP is estimated from random sample, 0 otherwise	0.923	0.268	0.89
Sample size	Number of observations of the study	348.18	369.4	382.82
Protest zeros WTP	=1 if study accounts for protest zeros WTP, 0 otherwise	0.415	0.496	0.401
Regions and study year				
China	=1 if study realized in China, 0 otherwise	0.323	0.471	0.301
Germany	=1 if study is realized in Germany, 0 otherwise	0.200	0.400	0.132
Netherlands	=1 if study is realized in Netherlands, 0 otherwise	0.180	0.391	0.113
Year	Year the study was conducted	2011.7	4.6	2012.67
Control variables				
Annual income	Logarithm of sample annual income in U.S. dollars (inflation-adjusted)	0.311	0.54	0.357
Average age	Sample average age in years	43.72	8.357	44.69
Publication characteristics				
Article type	=1 if the study is published , 0 if it is a working paper	0.953	0.211	0.972
Top ranked academic journal	=1 if the study is published in an acknowledged academic journal, 0 otherwise	0.323	0.471	0.324
Low number of citations	=1 if the average number of citations per year is less than one, 0 otherwise	0.261	0.442	0.246

Notes: The third column corresponds to the mean weighted by the inverse of the number of estimates per study.

5.2 Meta Regression Model

To investigate potential sources heterogeneity, we complete the model provided in equation (5) with additional study-level characteristics. We intend to estimate the "True" SWTP level after accounting for the potential effect of moderating variables. The baseline meta-regression model is then formulated as follows:

$$SWTP_{i,s} = \alpha_0 + \alpha_1 SE(SWTP_{i,s}) + \boldsymbol{X}_{i,s}\boldsymbol{\beta} + \varepsilon_{i,s} \tag{6}$$

With $SWTP_i$ the logarithm of the WTP divided by the actuarially fair price, $\boldsymbol{X}$ a vector of variables (moderating variables) to capture study-specific characteristics associated with the estimate *s* from study *i*, $\boldsymbol{\beta}$ a vector of coefficients, and $\boldsymbol{\varepsilon}_i$ the sampling error of the regression. The intercept term of the meta-regression, $\boldsymbol{\alpha}_0$, represents the true level of SWTP after controlling for publication bias and heterogeneity. A statistically insignificant intercept indicates that the observed effects are driven mainly by the characteristics of the primary studies. Conversely, a statistically significant negative (positive) intercept suggests an intrinsic perceived value of insurance robust to the presence of moderating variables.

Unlike conventional econometric models, we cannot assume that the estimation errors of the meta-regression model are independent and identically distributed. First, dependence is likely to arise, especially when there are multiple estimates from a unique study (Stanley and Doucouliagos, 2012). In such a case, this study's results might dominate the overall effect. Second, heteroscedasticity, i.e. non-constant variances of effect size estimates, could also be present due to primary studies using different sample sizes, sample randomness, and sampling method (Nelson and Kennedy, 2009). Therefore, estimating a meta-regression with OLS method might lead to inconsistent estimates, though unbiased. For these reasons, we estimate meta-regression with weighted-least squares (WLS) method:

$$\frac{SWTP_{i,s}}{SE(SWTP_{i,s})} = \frac{\alpha_0}{SE(SWTP_{i,s})} + \alpha_1 + \frac{\boldsymbol{X}_{i,s}}{SE(SWTP_{i,s})}\boldsymbol{\beta} + \frac{\varepsilon_{i,s}}{SE(SWTP_{i,s})} \tag{7}$$

We perform the meta regression using three estimators: (1) a cluster-robust ordinary least squares (OLS) estimator; (2) a cluster-robust random effects model (Unweighted RE) (3) a weighted random effects model by the inverse of the standard error (Weighted RE).[29]

As pointed out by Brada et al. (2021) and Kocenda and Iwasaki (2022), meta-regression faces the so-called "model uncertainty" problem, which implies that the true model cannot be identified in advance. Having the wrong variables in the regression model leads to misspecification bias and invalid inference. To address this problem, we estimated our models with moderators selected by Bayesian model averaging (BMA). The objective of this method is to define the best possible approximation of the distribution of regression parameters. BMA analysis provides three basic statistics for each parameter: the posterior mean, the posterior variance, and the posterior inclusion probability. The likelihood of each model is reflected by the model's posterior probabilities. The posterior means are then calculated as the estimated coefficients weighted across all models by their posterior model probability. We follow Jeffreys (1961) to interpret the posterior inclusion probabilities (PIPs) of BMA means, who characterizes evidence of an effect as "weak" for a PIP between 0.5 and 0.75, "substantial" for a PIP between 0.75 and 0.95, "strong" for a PIP between 0.95 and 0.99, and "decisive" for a PIP above 0.99.[30]

---

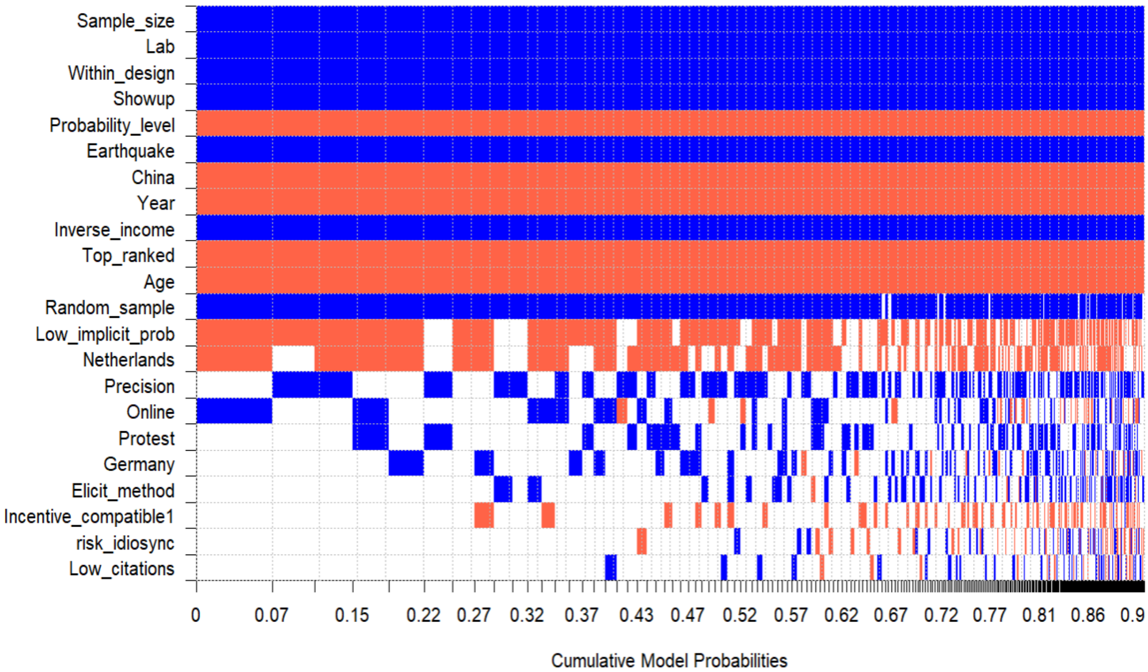
[29] The random-effect (RE) model is more appropriate in the presence of high heterogeneity. RE model weights correspond to $1/(\tau^2+v_i)$. When heterogeneity is high, v_i would be negligible compared to τ^2 so that all data points would have the same weight$\approx 1/\tau^2$. The model becomes almost unweighted, which can be problematic in the presence of publication bias. We weight the dataset by 1/vi before using the RE model.

[30] BMA requires explicit priors on the model (model prior) and regression coefficients (g-prior). As suggested by Eicher et al.(2011), we use the uniform model prior and the unitary g -prior information.

5.3 Results

Table 4 presents the empirical results obtained from different estimation techniques after controlling for multi-collinearity. The meta-regression results show that the publication bias term is not statistically significant in all of models. This finding corroborates the results presented in Table 2 and proves that our dataset does not suffer from this issue. Figure 3 depicts the outcomes of the BMA analysis, where the vertical axis lists all our moderating variables sorted by the posterior inclusion probability (PIP) in descending order and the horizontal axis refers to the posterior model probability (PMP) of each model sorted in ascending order.[31] From figure 4, we note that 14 moderating variables have PIP higher than 0.5 and allow to explain heterogeneity of the SWTP metric. As shown in Table 5, the estimates are not sensitive to the choice of estimator. There are small differences in estimation results between the weighted (1 & 2) and unweighted (3) methods.

***Figure* 4**: Model inclusion in BMA



Note: On the vertical axis, the explanatory variables are ranked according to their posterior inclusion probabilities from the highest at the top to the lowest at the bottom. The horizontal axis shows the values of cumulative posterior model probability. Blue color (darker in grayscale) = the estimated parameter of a corresponding explanatory variable is positive. Red color (lighter in grayscale) = the estimated parameter of a corresponding explanatory variable is negative. No color = the corresponding explanatory variable is not included in the model. Numerical results are reported in Table 5. All variables are described in Table 3.

For the different estimations methods (1), (2), and (3) reported in Table 5, the intercept of the meta-regression, α_0, is not statistically significant except for BMA method (4). To reconcile this contradictory result from the BMA analysis, further analysis is undertaken using general-to-specific stepwise regression. After accounting progressively for key moderators from the BMA, the intercept is not substantially different from zero, as shown in Table 5. An average effect size, which is statistically indistinguishable from zero, suggests no a priori global underestimation of tail expected losses. This result contradicts a

[31] The blue color indicates the positive sign of the variable in the model, and the red color denotes the negative sign of the variable. The blank cell suggests that the parameters associated to these variables are not significantly different from zero for most models.

common explanation of low local demand for insurance against tail risks related of risk underestimation (Wagner, 2022; Dionne et al., 2022). This result may also be interpreted as no systematic rejection of insurance that might be considered by households under a narrow framing context as a poor financial investment (Gottlieb and Smetters, 2020). This result also provides evidence in favor of extraneous factors' impact on low insurance demand leaving less to be attributed to risk preferences.

The meta-regression results show that laboratory experiments are positively related to SWTP, whereas online experiments report no significant effect. A possible explanation of this finding could be related to the social desirability bias that may encourage individuals to adjust their reaction or overstate their WTP to enhance their self-image and be socially acceptable (Lusk and Norwood, 2009; Carlsson et al., 2010; Paulhus, 1991). For instance, Leggett et al. (2003) show that WTP values derived from face-to-face interviews can be as much as 23%–29% higher relative to self-administered surveys. Similarly, the presence of participation (show-up) fees is positively related to SWTP regardless the amount given. Multiple WTP estimates based on within subject design also report significant positive effect. Such design is more subject to potential correlation of treatment effects and to other dynamic behavioral patterns (e.g. learning, subject fatigue, wealth effects, etc.) (Landry, 2017).

Our results have not shown a significant effect of elicitation design and do not confirm hypotheses 4a and 4b. None of the two moderators describing this factor (incentive-based contract and method of election) is robustly significant. We find nonetheless some evidence in support of compatible-incentive (CI) designs that have minor negative effect on SWTP. These findings may suggest that the effect of CI settings may vary with context. The characteristics of the sample do not allow for more detailed analysis.

Regarding risk features, very small descriptive probability tends to impact SWTP, a finding in line with (Schade et al., 2012), where WTP for insurance is much more than the expected loss when the probability of loss is known. All else being equal, a 1% downward variation of the provided probability tend to increase SWTP by 0.26. This result supports hypothesis 1 and confirms the inverse S-shaped probability weighting. It also gives support to the DRO setting developed by Jaspersen et al. (2022) where extreme small probabilities are overweighed relative to their baseline value more than large ones.

Average sample incomes are negatively associated with SWTP. This result verifies hypothesis 2a and supports the view of insurance as an inferior good. Decreasing absolute risk aversion in wealth for samples with high average income implies a decrease in the value of insurance. Moreover, wealthier agents face lower costs to self-insure, an activity that substitutes for market insurance (Ehrlich and Becker, 1972). Age is negatively related to the dependent variable in the sense that samples with high average age exhibited lowest SWTP. This result confirms the finding of Browne et al. (2015) related to revealed preference analysis for purchasing flood versus bicycle theft insurances. Their results show that demand for both types of coverage decreases with age.[32]

Our results show that the SWTP are geographically dependent. We find that SWTP is smaller in China than in Europe mainly in Germany and the Netherlands. This finding is consistent with several empirical studies conducted in developing countries. A plausible explanation is that for collectivity-oriented society as in China, post loss financing may rely on informal family and community solidarity more than on formal market insurance products. Therefore, private insurance may play a smaller role as a risk management mechanism. The results show that there is a notable difference between Germany and the Netherlands. Seifert et al. (2013) compared insurance demand in Europe using data from these two countries and find that WTP is higher in Germany. The charity hazard resulting from the disparity in disaster insurance systems is a possible explanation for this gap (Browne and Hoyt, 2000). Post-disaster public funding as in the Netherlands encourages individuals to expect receiving contributions from public relief money in the event of a major catastrophe (Yan and Faure, 2021).

Our results also show that the SWTP tends to decrease over time, a result that coincides with the persistent trend of underinsurance puzzle discussed in the literature. Related to the effect of research quality, we show that top ranked journals seem to report smaller levels of SWTP. Our empirical findings do not confirm hypothesis 3, which predicts a negative effect between correlated risks WTP.

[32] For stated preference surveys, age may have a significant influence on protest responses. It has been argued that mental abilities decline with age. Cognitive effort need for making decisions in hypothetical scenarios leads to protest responses. In addition, younger individuals may be more likely to accept hypothetical scenarios than older individuals may.

Table 4 Estimation of the multivariate general-to-specific stepwise meta-regression

	(1)Unweighted OLS (Clustered)			(2) Unweighted RE (Clustred)			(3) Weighted RE			(4) Weighted BMA		
	Coef.	**SE**	**pval**	Coef.	**SE**	**pval**	Coef.	**SE**	**pval**	**Post Mean**	**Post SE**	**PIP**
~~Constant~~ (true effect α_0)	-0.1957	0.6779	0.7742	-0.1054	0.3942	0.7934	-0.2304	0.2613	0.3828	**-1.2171**	NA	**1.0000**
~~Precision~~ (pub. bias α_1)	0.127	1.2851	0.9217	-0.6998	1.3425	0.6109	0.0877	0.1381	0.5288	0.3248	0.4103	0.4905
Lab	**1.423*****	0.2135	**0.0001**	**1.4585*****	0.1375	**0.0001**	**1.3451*****	0.1080	**0.0000**	**1.3168**	0.1533	**1.0000**
Online	0.2322	0.2072	0.2688	0.1844	0.137	0.2014	0.1441	0.1486	0.3376	0.0870	0.1882	0.3522
Within_design	**0.5433****	0.1741	**0.0033**	**0.5396****	0.1481	**0.003**	**0.4682*****	0.1130	**0.0002**	**0.4306**	0.0829	**1.0000**
Showup	**1.437*****	0.2509	**0.0001**	**1.5404*****	0.2149	**0.0001**	**1.5017*****	0.1753	**0.0000**	**1.1290**	0.1600	**1.0000**
Incentive_compatible	-0.3602	0.2796	0.2047	**-0.3723***	0.1975	**0.082**	**-0.4694****	0.1628	**0.0062**	-0.0344	0.1028	0.1967
Elicit_method	0.0797	0.1067	0.4595	0.0892	0.0841	0.3085	**0.1403***	0.0598	**0.0239**	0.0112	0.0358	0.2050
Probability_level	**-27.966*****	5.4187	**0.0001**	**-33.225*****	2.6467	**0.0001**	**-27.930*****	3.7179	**0.0000**	**-25.7667**	5.4836	**0.9993**
Low_implicit_prob	-0.209	0.1645	0.211	**-0.2827***	0.1376	**0.0606**	**-0.2384****	0.0860	**0.0083**	**-0.1255**	0.1042	**0.7102**
Risk_idiosync	0.0395	0.2154	0.8553	0.1302	0.1393	0.3669	0.1198	0.0945	0.2118	0.0029	0.0682	0.1402
Earthquake	**1.9783*****	0.5103	**0.0004**	**2.0919*****	0.3638	**0.0001**	**2.0016*****	0.2621	**0.0000**	**1.6628**	0.3207	**0.9999**
Sample size	**0.0005***	0.0002	**0.0154**	**0.0003***	0.0001	**0.0283**	**0.0002****	0.0001	**0.0024**	**0.0004**	0.0001	**0.9991**
China	**-1.3885*****	0.1659	**0.0001**	**-1.2488*****	0.2285	**0.0001**	**-1.1231*****	0.1646	**0.0000**	**-1.3690**	0.1133	**1.0000**
Year	**-0.0761*****	0.0132	**0.0001**	**-0.0663*****	0.0104	**0.0001**	**-0.0658*****	0.0063	**0.0000**	**-0.0770**	0.0078	**1.0000**
Germany	0.3067	0.3127	0.3324	0.442	0.2595	0.1123	0.4737	0.2443	0.0592	0.0580	0.1455	0.2748
Netherlands	**-0.6131***	0.2794	**0.0338**	**-0.5004***	0.1764	**0.014**	**-0.5173*****	0.1350	**0.0004**	**-0.3087**	0.3141	**0.6226**
Protest	0.0735	0.1795	0.6842	0.1865	0.1796	0.318	0.1602	0.1306	0.2267	0.0348	0.0687	0.2944
Random_sample	**0.5957****	0.1808	**0.002**	**0.7408*****	0.1237	**0.0001**	**0.6823*****	0.1133	**0.0000**	**0.5309**	0.1941	**0.9602**
Inverse_income	**0.3394*****	0.0688	**0.0001**	**0.2799*****	0.0576	**0.0003**	**0.2671*****	0.0396	**0.0000**	**0.3186**	0.0644	**0.9999**
Age	-0.0187	0.0131	0.1597	**-0.0206***	0.0073	**0.0146**	**-0.0194****	0.0057	**0.0014**	**-0.0246**	0.0071	**0.9990**
Top_ranked	**-0.6102****	0.1782	**0.0014**	**-0.8085*****	0.1245	**0.0001**	**-0.6974*****	0.0965	**0.0000**	**-0.5202**	0.1264	**0.9976**
Low_citations	-0.0735	0.1405	0.6035	**-0.2988***	0.1648	**0.0929**	**-0.3011***	0.1244	**0.0199**	0.0020	0.0307	0.1237
R^2		94.02%			91.53%			91.53%				-
H^2		-			32.66			32.66				-
Q stat. (p.value)		-			449.67 (.0001)			449.67 (.0001)				-
Tau^2 (SE)		-			0.047 (0.0121)			0.047 (0.0121)				-
I^2		-			96.94%			96.94%				-
N-obs		65			65			65				65

Notes: The results are from three estimators: (1) a cluster-robust ordinary least squares (OLS) estimator; (2) a cluster-random-effects model (Unweighted RE) (3) a weighted random effects model by the inverse of the standard error (Weighted RE). All variables are described in Table 3. Variables with PIP above 0.5 or significant are emphasized in bold. SD = standard deviation. SE = standard error. PIP= posterior inclusion probability. N.A. = not available. *, **, and *** denote statistical significance at 10, 1%, and .1%, respectively.

Table 5 Estimation of the multivariate meta-regression

	Model (1)			Model (2)			Model (3)		
	Coef.	SE	pval	Coef.	SE	pval	Coef.	SE	pval
~~Constant~~ (true effect α_0)	-0.347*	0.1375	0.0161	0.1425	0.5502	0.7978	0.4688	0.3297	0.1698
~~Precision~~ (pub. bias α_1)	-	-	-	-0.1527	1.0873	0.8895	-1.3379	1.2567	0.2991
Lab	-	-	-	0.870***	0.2083	0.0003	1.1886***	0.1036	0.0001
Online	-	-	-	-	-	-	-	-	-
Within_design	-	-	-	0.3901*	0.1645	0.0257	0.4817***	0.1098	0.0003
Showup	-	-	-	1.254***	0.3285	0.0008	1 .3387***	0.1958	0.0001
Incentive_compatible	-	-	-	-	-	-	-	-	-
Elicit_method	-	-	-	-	-	-	-	-	-
Probability_level	-	-	-	-23.09**	7.5271	0.0051	-31.581***	3.7465	0.0001
Low_implicit_prob	-	-	-	-	-	-	-0.3182*	0.1404	0.0341
Risk_idiosync	-	-	-	-	-	-	-	-	-
Earthquake	-	-	-	-	-	-	1.8898***	0.2432	0.0001
Sample size	-	-	-	-	-	-	-	-	-
China	-	-	-	-0.917***	0.2081	0.0002	-1.3709***	0.1692	0.0001
Year	-	-	-	-0.044**	0.0135	0.0028	-0.0777***	0.0118	0.0001
Germany	-	-	-	-	-	-	-	-	-
Netherlands	-	-	-	-0.488*	0.2075	0.0268	-0.4425**	0.1402	0.0048
Protest	-	-	-	-	-	-	-	-	-
Random_sample	-	-	-	0.3013	0.3594	0.4097	0.716***	0.1474	0.0001
Inverse_income	-	-	-	-	-	-	0.2116**	0.0733	0.0088
Age	-	-	-	-0.0173**	0.006	0.0078	-0.0235***	0.0035	0.0001
Top_ranked	-	-	-	-	-	-	-0.6104***	0.1441	0.0004
Low_citations	-	-	-	-	-	-	-	-	-
R^2		0.00%			71.20%			88.49%	
H^2		410.42			108.63			43.27	
Q stat (p.value)		16616.53 (.0001)			1847.98 (.0001)			858 (.0001)	
Tau^2 (SE)		0.560 (0.1014)			0.1613 (0.0329)			0.0645 (0.0145)	
I^2		99.76%			99.08%			97.69%	
N-obs		65			65			65	

Notes: The results are from a cluster-random-effects model (3) a weighted random effects model by the inverse of the standard error (Weighted RE). All variables are described in Table 3. *, **, and *** denote statistical significance at 10, 1%, and .1%, respectively

5.4. Robustness checks

We estimate several additional regressions to check the robustness of estimation results obtained in Table 4. First, we cheek the sensitivity of our results to weighting factors. We perform BMA analysis without weighting, and then with a weight equal to the inverse of the number of data points per study. The results given in Table 6 are robust and corroborate previous findings, where the magnitude and the sign of the main variables exhibit little variance.

To further investigate heterogeneity, we apply a three-level structure to the meta-regression model which allows for examining differences in outcomes within studies (i.e., within-study heterogeneity) as well as differences between studies (i.e., between-study heterogeneity).[33] Unlike other models, we do not need to know correlations between outcomes reported within primary studies since the second level accounts for sampling covariation (Van den Noortgate et al., 2013). Note that the random effects hierarchical method that we use for estimation allows coefficients to vary randomly across studies (Ugur et al. 2016; Neves and Sequeira, 2018). We test different candidate variables as third level. The results do not change from previous results, confirming the absence of correlation between SWTP within studies and therefore the appropriateness of the two-level model used to analyze heterogeneity.[34] None of the third-level candidate variables is able to explain more of the variability between studies.

As a third robustness check, we conduct outlier analyses by first examining extreme SWTP with confidence intervals that did not overlap with the confidence interval of the pooled effect. We perform influence analyses via a "leave-one-out" method, in which effect size is recalculated when a single study is left out of the analysis (Viechtbauer and Cheung, 2010). We identified three observations as potential outliers or influential outcomes with the leave one out analysis and the Baujat plot displayed in figure 4. After progressively excluding these three data points, we further reduce heterogeneity while confirming the obtained estimation results.[35] Finally, as a last robustness check we estimate a meta-regression model using conditional SWTP (excluding zeros WTP) as effect size. The estimation results are quite similar to those reported in Table 4[36].

Because of the structural heterogeneity in our dataset related the data collection method (experiments vs surveys), we deem it more useful to present average SWTP value for two extreme research designs than for a unique "best study". We decide a priori on a set of characteristics that constitute profile 1 that are survey-based studies with a between design for outcome measurement, no participation fees, non-given probability of loss and a moderate size sample. Such studies are expected to provide low SWTP. On the country, the second extreme profile covers laboratory-based experiments with participation fees, within-design, young subjects with moderate-income and extremely small given probabilities of losses. For this profile, we expect to observe larges SWTP.

Related to the second profile, futures experimental studies should account for the sensitivity of the WTP estimation to the choice of loss probability level. Providing extremely low probability values are more difficult to process and more subject to decision heuristics and cognitive bias than higher values (e.g. Kunreuther and Slovic, 1978; Hertwig et al., 2004; Kunreuther et al., 2001). Future studies should also take heed of the sensitivity of stated WTP estimation to the qualitative characteristics of the subjects sample in terms of randomness and representatively. A non-random sample may yield biased WTP estimations since it can over represent some socioeconomic groups exacerbating or moderating the effect of age, income or other potential variables not captured in our study (e.g. loss experience, financial literacy, etc.).

[33] Three-level meta-analytic model assumes different variance components distributed over the levels of the model: sampling variance of the SWTP extracted at level 1; variance between effect sizes extracted from the same study at level 2; and variance between studies at level 3.

[34] Estimation results are reported in Appendix G.

[35] Estimation results are reported in Appendix H.

[36] Estimation results are reported in Appendix I.

***Figure* 5**: Baujat plot of SWTP

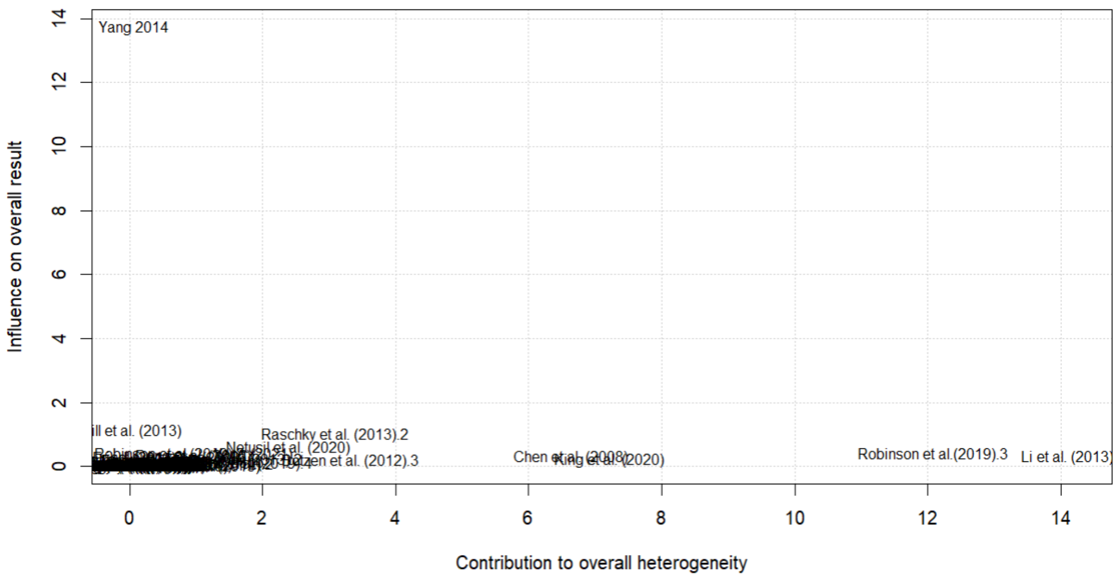


Table 6 Additional BMA meta-regression

	Unweighted BMA			Weighted BMA (number of data points per study)		
	Post Mean	**Post SE**	**PIP**	**Post Mean**	**Post SE**	**PIP**
Constant (true effect α_0)	**NA**	**0.1469**	1	**-0.0195**	**NA**	**1**
Precision (pub. bias α_1)	-	-	-	**-0.0004**	**0.1541**	0.1403
Std. error	-0.0160	0.1307	0.1528	-0.0167	0.0806	0.1492
Lab	**1.2467**	**0.2291**	**1**	**1.2161**	**0.1587**	**1**
Online	0.0752	0.1706	0.2788	0.0709	0.1827	0.2490
Within_design	**0.4964**	**0.1464**	**0.9870**	**0.5225**	**0.1590**	**0.9873**
Showup	**1.3927**	**0.1893**	**1**	**1.2912**	**0.1477**	**1**
Incentive_compatible1	-0.0370	0.1408	0.1834	-0.0838	0.1527	0.3263
Elicit_method	0.0078	0.0425	0.1615	0.0008	0.0271	0.1381
Probability_level	**-33.928**	**4.6663**	**1**	**-31.8202**	**4.8846**	**1**
Low_implicit_prob	**-0.3060**	**0.1923**	**0.8160**	**-0.4661**	**0.0982**	**0.9996**
risk_idiosync	0.0315	0.1245	0.1654	0.0103	0.0662	0.1285
Earthquake	**1.9184**	**0.4919**	**0.9925**	**2.0572**	**0.2786**	**1**
Sample_size	0.0001	0.0001	0.4193	0.0000	0.0001	0.2633
China	**-1.2403**	**0.1636**	**1**	**-1.2151**	**0.0995**	**1**
Year	**-0.0695**	**0.0147**	**0.9998**	**-0.0736**	**0.0105**	**1**
Germany	0.1960	0.2563	0.4812	**0.2546**	**0.2367**	**0.6421**
Netherlands	**-0.5309**	**0.2450**	**0.9306**	**-0.4813**	**0.2116**	**0.9498**
Protest	0.1178	0.1552	0.4699	0.0243	0.0637	0.2179
Random_sample	**0.7529**	**0.2169**	**0.9863**	**0.7594**	**0.1458**	**0.9998**
Inverse_income	**0.1916**	**0.1087**	**0.8539**	**0.2534**	**0.0518**	**0.9998**
Age	**-0.0226**	**0.0101**	**0.9283**	**-0.0164**	**0.0045**	**0.9859**
Top_ranked	**-0.6349**	**0.1480**	**0.9993**	**-0.6617**	**0.1208**	**0.9999**
Low_citations	**-0.1513**	**0.1619**	**0.5712**	**-0.1209**	**0.1233**	**0.5907**

Notes: The results are from specifications with no weight and with weight equal to the number of data points per study. All moderating variables are described in Table 3. Variables with PIP above 0.5 are shown in bold. SD = standard deviation. SE = standard error. PIP= posterior inclusion probability. N.A. = not available. N-obs: 65.

6. Conclusion

In this paper, we conduct a meta-analysis on the stated willingness to pay for insurance against LPHI risks. To the best of our knowledge, this work offers the most comprehensive analysis of this issue and extends previous literature reviews by addressing publication bias and heterogeneity considerations. After revising these two aspects, our main finding is that true willingness to pay to insure extreme risks is not significantly different from the actuarially fair price. From a global perspective, we can view this result as a new evidence of no a priori underestimation of tail losses or systematic rejection of insurance considered under some circumstances as a poor financial investment. The meta-regression analysis captures various systematic factors affecting the WTP estimates. Preferences elicitation setting and risk characteristics significantly influence standardized WTP. For instance, SWTP turns out to be higher for extremely small descriptive probabilities. Laboratory experiments, within-subjects design, and participation fees also appear to inflate SWTP. Conversely, income level and average sample age negatively affect SWTP. This latter result might highlight the importance of identifying sub-groups with distinct observable heterogeneity to provide decision-makers with central information for policy design. SWTP also shows a steady downward trend over time, a result that would be interesting to examine in future research. In general, we find little evidence in support of hypothetical bias for WTP measures, a finding that should renew interest for stated preference methods and alleviate concerns about its external validity. In the absence of market data on actual demand for LPHI risks, stated preferences elicitation remains a central option for understanding individuals' values of insurance. More effort is needed, however, to improve external validity through a unified procedure conjoining relevant instruments used to mitigate hypothetical bias.

To be viable, policymakers and insurance industry must be aware that LPHI insurance should be supported by prevention and resilience actions. Achieving this goal may be more difficult for certain social groups. Building on the results identified in our analysis, one way to ease the budget constraint is to combine insurance with other financial instruments such as credit. This will reduces the initial cash payment of premiums and reduces the liquidity constraint. For the other factors coupled with budget constraints, we still need more systematic evidence on the relationship between small-probability losses, risk perception, and insurance uptake. For instance, some simple steps could be taken to mitigate the problems of overweighting or ignoring small-probability risks. This could be done by improving public understanding and access to multi-period probability of LPHI events so that these risks become salient to the decision makers. Finally, our results call for more empirical analysis on how fundamental, behavioral and methodological factors interact and impact insurance demand. A better understanding of these interactions will provide valuable policy guidance, particularly improving affordability through more subsidy targeting and/or a progressive shift toward risk-based pricing systems.

Our study is not without limitations. Although passing most of the robustness checks, we first acknowledge the dataset size restriction and its potential impact on the stability of the meta-regression coefficient. This constraint also limits our ability to test non-linear and interaction effects between moderators to further explain heterogeneity. The second limitation is that we were unable to explore potentially significant latent drivers of WTP due to a lack of data. Some important factors, such as financial literacy, past loss experience, or perceived insurance providers' quality are not included in our meta-regression. Third, because protesters are not systematically identified and corrected across studies, false-zero WTP responses may occur, leading to a systematic downward bias in the results. Future meta-analyses should attempt to resolve some of these issues. In this respect, data-driven methods applying regularization techniques would present promising prospects. Combining flexible machine learning and simple linear models can provide powerful and interpretable results overcoming small sample and missing data problems.

References

Alfnes, F., Rickertsen, K. (2011). Non-market valuation: experimental methods. *The Oxford handbook of the economics of food consumption and policy*, *215*, 242.

Ali, W., Abdulai, A., Goetz, R., Owusu, V. (2021). Risk, ambiguity and willingness to participate in crop insurance programs: Evidence from a field experiment. *Australian Journal of Agricultural and Resource Economics*, *65*(3), 679-703.

Amelung, D., Funke, J. (2015). Laypeople's risky decisions in the climate change context: climate engineering as a risk-defusing strategy? *Human and ecological risk assessment: An International Journal*, *21*(2), 533-559.

Andersen, S., Harrison, G. W., Lau, M. I., Rutström, E. E. (2006). Elicitation using multiple price list formats. *Experimental Economics*, *9*(4), 383-405.

Andrews, I., Kasy, M. (2019). Identification of and correction for publication bias. *American Economic Review*, *109*(8), 2766-94.

Arrow, K. (1971). Essays in the Theory of Risk-Bearing (Amsterdam: North-Holland).

Atreya, A., Ferreira, S., Michel-Kerjan, E. (2015). What drives households to buy flood insurance? New evidence from Georgia. *Ecological Economics*, *117*, 153-161.

Baillon, A., Kraft, A., O'Donnell, O., Van Wilgenburg, K. (2022). A behavioral decomposition of willingness to pay for health insurance. *Journal of Risk and Uncertainty*, *64*(1), 43-87.

Bakkensen, L., and Barrage, L. (2021). Flood Risk Belief Heterogeneity and Coastal Home Price Dynamics: Going Under Water?, *NBER Working Paper 23854*

Balistreri, E., McClelland, G., Poe, G., Schulze, W. (2001). Can hypothetical questions reveal true values? A laboratory comparison of dichotomous choice and open-ended contingent values with auction values. *Environmental and Resource Economics*, 18(3), 275-292.

Barron, G., & Erev, I. (2003). Small feedback‐based decisions and their limited correspondence to description‐based decisions. *Journal of behavioral decision-making*, *16*(3), 215-233.

Barseghyan, L., Molinari, F., O'Donoghue, T., Teitelbaum, J. C. (2018). Estimating risk preferences in the field. *Journal of Economic Literature*, *56*(2), 501-64.

Bault, N., Coricelli, G., Rustichini, A. (2008). Interdependent utilities: How social ranking affects choice behavior. *PLoS One, 3*, 1–10.

Becker, G. M., DeGroot, M. H., Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science*, 9(3), 226–232.

Biener, C., Landmann, A., Santana, M. I. (2019). Contract nonperformance risk and uncertainty in insurance markets. *Journal of Public Economics*, *175*, 65-83.

Bolton, G. E., Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *American economic review*, *90*(1), 166-193.

Bolton, G. E., Brandts, J., Ockenfels, A. (2005). Fair procedures: Evidence from games involving lotteries. The Economic Journal, 115(506), 1054-1076.

Bom, P. R., Rachinger, H. (2019). A kinked meta‐regression model for publication bias correction. *Research synthesis methods*, *10*(4), 497-514.

Botzen, W. W., Deschenes, O., Sanders, M. (2020). The economic impacts of natural disasters: A review of models and empirical studies. *Review of Environmental Economics and Policy*, 13(2), 167-188.

Brada, J. C., Drabek, Z., Iwasaki, I. (2021). Does investor protection increase foreign direct investment? A meta‐analysis. *Journal of Economic Surveys*, *35*(1), 34-70.

Breidert, C., Hahsler, M., Reutterer, T. (2006). A review of methods for measuring willingness-to-pay. *Innovative Marketing*, 2(4), 8–32.

Brouwer, R., Tinh, B. D., Tuan, T. H., Magnussen, K., Navrud, S. (2014). Modeling demand for catastrophic flood risk insurance in coastal zones in Vietnam using choice experiments. *Environment and Development Economics*, *19*(2), 228-249.

Brown, T. C., Champ, P. A., Bishop, R. C., McCollum, D. W. (1996). Which response format reveals the truth about donations to a public good? *Land Economics*, 72(2), 152–166.

Browne, M.J., Hoyt, R. (2000). The demand for flood insurance: empirical evidence. *Journal of Risk and Uncertainty*, 20(3):291–306.

Browne, M. J., Knoller, C., Richter, A. (2015). Behavioral bias and the demand for bicycle and flood insurance. *Journal of Risk and Uncertainty*, *50*(2), 141-160.

Cai, J., De Janvry, A., Sadoulet, E. (2015). Social networks and the decision to insure. American Economic Journal: Applied Economics, 7(2), 81-108.

Carlsson, F., García, J. H., Löfgren, Å. (2010). Conformity and the demand for environmental goods. *Environmental and resource economics*, *47*(3), 407-421.

Carson, D., Gilmore, A., Perry, C., Gronhaug, K. (2001). *Qualitative marketing research*. Sage.

Carson, R.T., Groves, T., List, J.A. (2014). Consequentiality: A theoretical and experimental exploration of a single binary choice. *Journal of the Association of Environmental and Resource Economists*, 1(1/2), 171-207.

Clarke, D., Kalani, G. (2011, November). Micro-insurance decisions: evidence from Ethiopia. In *CSAE 25th Anniversary Conference*.

Cole, S., Fernando, A. N., Stein, D., Tobacman, J. (2020). Field comparisons of incentive-compatible preference elicitation techniques. *Journal of Economic Behavior Organization*, 172, 33-56.

Cole, S, Giné, X., Tobacman, J., Topalova, T., Townsend, R., Vickery., J. (2013). Barriers to Household Risk Management: Evidence from India. *American Economic Journal: Applied Economics*, 5 (1): 104–35.

De Nicola, F., Hill, R. V. (2012). Interplay among credit, weather insurance and savings for farmers in Ethiopia. In *Presentation at the American Economic Association Meetings*.

Diamond, P. A., Hausman, J. A. (1994). Contingent valuation: is some number better than no number?. *Journal of economic perspectives*, *8*(4), 45-64.

Dionne, G., Desjardins, D. (2022). A re-examination of the US insurance market's capacity to pay catastrophe losses. *Available at SSRN 4113490*.

Ehrlich I., Becker G. (1972). Market insurance, self-insurance and self-protection. *Journal of Political Economy*, 623-648.

Eicher, T. S., C. Papageorgiou, A. E., Raftery, M. (2011). Default priors and predictive performance in Bayesian model averaging, with application to growth determinants. *Journal of Applied Econometrics* 26(1), 30–55.

Entorf, H., Jensen, A. (2020). Willingness-to-pay for hazard safety–A case study on the valuation of flood risk reduction in Germany. *Safety science*, *128*, 104657.

Fehr, E., Schmidt, K. M. (1999). A theory of fairness, competition and cooperation. *Quarterly Journal of Economics*, 114, 817–868.

Forgas, J. P. (1995). Mood and judgment: The affect infusion model (AIM). *Psychological Bulletin*, 117(1):39–66.

Friedl, A., de Miranda, K. L., Schmidt, U. (2014). Insurance demand and social comparison: An experimental analysis. *Journal of Risk and Uncertainty*, 48(2), 97–109.

Giné, X., Townsend, R., Vickery, J. (2008). Patterns of Rainfall Insurance Participation in Rural India. *World Bank Economic Review*, 22 (3), 539–66.

Gottlieb, D., Mitchell, O. S. (2020). Narrow framing and long-term care insurance. *Journal of Risk and Insurance*, *87*(4), 861-893.

Haghani, M., Bliemer, M. C., Rose, J. M., Oppewal, H., Lancsar, E. (2021). Hypothetical bias in stated choice experiments: Part II. Conceptualisation of external validity, sources and explanations of bias and effectiveness of mitigation methods. *Journal of choice modelling*, *41*, 100322.

Hanemann, W. M. (1991). Willingness to pay and willingness to accept: how much can they differ?. The American Economic Review, 81(3), 635-647.

Harrison, G. W., Ng, J. M. (2019). Behavioral insurance and economic theory: A literature review. *Risk Management and Insurance Review*, *22*(2), 133-182.

Harrison, G. W., Rutström, E. E. (2008). Experimental evidence on the existence of hypothetical bias in value elicitation methods. *Handbook of experimental economics results*, 1, 752-767.

Havránek, T., Stanley, T. D., Doucouliagos, H., Bom, P., Geyer-Klingeberg, J., Iwasaki, I., Reed, W. R., Rost, K., Van Aert, R. C. M. (2020). Reporting guidelines for meta-analysis in economics. *Journal of Economic Surveys*, 34(3), 469–475.

Hedges, L. V., J. Gurevitch, and P. S. Curtis. 1999. The meta- analysis of response ratios in experimental ecology. *Ecology*, 80, 1150 –1156.

Hensher, D.A. (2015). Data challenges: more behavioural and (relatively) less statistical–a think piece. *Transportation Research Procedia,* 11, 19-31.

Hertwig, R., Erev, I. (2009). The description–experience gap in risky choice. Trends in cognitive sciences, 13(12), 517-523.

Hertwig, R., Barron, G., Weber, E. U., Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological science*, *15*(8), 534-539.

Hofstetter, K.M. Miller, H. Krohmer, Z.H. Zhang. (2021). A de-biased direct question approach to measuring consumers' willingness to pay. *International Journal of Market Research*, 38 (1), 70-84

Huber, K. M., Roder, J. C., Bear, M. F. (2001). Chemical induction of mGluR5-and protein synthesis–dependent long-term depression in hippocampal area CA1. *Journal of neurophysiology*, *86*(1), 321-325.

Hudson, P., Botzen, W. J. W., Czajkowski, J., Kreibich, H. (2017). Moral hazard in natural disaster insurance markets: Empirical evidence from Germany and the United States. *Land Economics, 93*(2), 179–208.

Ioannidis, J. P., Stanley, T. D., Doucouliagos, H. (2017). The power of bias in economics research. *The economic Journal,* 127(605), 236-265.

Jaspersen, J.G. (2016). Hypothetical surveys and experimental studies of insurance demand: a review. *Journal of Risk and Insurance*, 83(1): 217–255.

Jaspersen, J. G., Ragin, M. A. (2021). A Model of Anchoring and Adjustment for Decision-Making under Risk. *Available at SSRN 3845633*.

Jaspersen, J. G., Peter, R., Ragin, M. A. (2022). Probability weighting and insurance demand in a unified framework. *The Geneva Risk and Insurance Review*, 1-47.

Jedidi, K., Jagpal, S. (2009). Willingness to pay: measurement and managerial implications. In *Handbook of pricing research in marketing*. Edward Elgar Publishing.

Jeffreys, H. (1961): Theory of Probability. Oxford Classic Texts in the Physical Sciences. *Oxford University Press, Third edition.*

Kahneman, D. Tversky, A. (1979). Prospect theory: An analysis of decision under risk, *Econometrica*, 47(2), 263–291.

Kahneman, D., Lovallo, D. (1993). Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Science*, 39(1), 17–31.

Kahneman, D., Knetsch, J. L., Thaler, R. H. (1990). Experimental tests of the endowment effect and the Coase theorem. *Journal of political Economy*, *98*(6), 1325-1348.

Karlan, D., Osei, R., Osei-Akoto, I., Udry, C. (2014). Agricultural decisions after relaxing credit and risk constraints. *The Quarterly Journal of Economics*, *129*(2), 597-652.

Kesternich, I., Heiss, F., McFadden, D., Winter, J. (2013). Suit the action to the word, the word to the action: Hypothetical choices and real decisions in Medicare Part D. *Journal of Health Economics*, 32(6), 1313-1324.

Klomp, J., K. Valckx. (2014). Natural disasters and economic growth: a meta-analysis. *Global Environmental Change* 26: 183–95.

Kočenda, E., Iwasaki, I. (2022). Bank survival around the world: A meta‐analytic review. *Journal of Economic Surveys*, *36*(1), 108-156.

Koricheva, J., Gurevitch, J. (2014). Uses and misuses of meta‐analysis in plant ecology. *Journal of Ecology*, *102*(4), 828-844.

Kousky, C. (2010). Learning from extreme events: Risk perceptions after the flood. *Land Economics*, *86*(3), 395-422.

Kousky, C. (2018). Financing flood losses: A discussion of the national flood insurance program. *Risk Management and Insurance Review*, *21*(1), 11-32.

Krawczyk, M. W., Trautmann, S. T., van de Kuilen, G. (2017). Catastrophic risk: social influences on insurance decisions. *Theory and Decision*, 82(3), 309–326.

Kriesel, W., Landry, C. (2004). Participation in the National Flood Insurance Program: An empirical analysis for coastal properties. *Journal of Risk and Insurance*, *71*(3), 405-420.

Kudryavtsev, A., Pavlodsky, J. (2012). Description-based and experience-based decisions: individual analysis. *Judgment Decision Making*, *7*(3).

Kunreuther,H. (1984). Causes of underinsurance against natural disasters. *Geneva Papers on Risk and Insurance*, 206–220.

Kunreuther, H., R. Ginsberg, L. Miller, P. Sagi, P. Slovic, B. Borkan, N. Katz. (1978). *Disaster Insurance Protection: Public Policy Lessons. Wiley, New York.*

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

Kunreuther, H., Pauly, M. (2004). Neglecting disaster: Why don't people insure against large losses?. *Journal of Risk and Uncertainty*, *28*(1), 5-21.

Kunreuther, H., Novemsky, N., Kahneman, D. (2001). Making low probabilities useful. *Journal of risk and uncertainty*, 23(2), 103-120.

Lajeunesse, M. J. (2015). Bias and correction for the log response ratio in ecological meta-analysis. *Ecology*, *96*(8), 2056-2063.

Landry, C. E. (2017). Experimental methods in valuation. In *A Primer on Nonmarket Valuation* (pp. 391-429). Springer, Dordrecht.

Landry, C. E., Jahan-Parvar, M. R. (2011). Flood insurance coverage in the coastal zone. *Journal of Risk and Insurance*, *78*(2), 361-388.

Leblois, A., Le Cotty, T., d'Hôtel, E. M. (2020). How might climate change influence farmers' demand for index-based insurance?. *Ecological economics*, *176*, 106716.

Leggett, C. G., N. S. Kleckner, K. J. Boyle, J. W. Duffield, and R. C. Mitchell. 2003. 'Social Desirability Bias in Contingent Valuation Surveys Administered through In-Person Interviews. *Land Economics*, 79 (4), 561–75.

Linde, J., Sonnemans, J. (2012). Social comparisons and risky choices. *Journal of Risk and Uncertainty*, 44(1), 45–72.

Lipsey, M. W., Wilson, D. B. (2001). *Practical meta-analysis*. SAGE publications, Inc.

Liu, Y., and R.J. Myers. (2016). The dynamics of micro-insurance demand in developing countries under liquidity constraints and insurer default risk. *Journal of Risk and Insurance,* 83 (1): 121–138.

Loomis, J. (2011). What's to know about hypothetical bias in stated preference valuation studies? *Journal of Economic Surveys*, 25(2), 363–370.

Lucas, C. H., Booth, K. I., Garcia, C. (2021). Insuring homes against extreme weather events: a systematic review of the research. *Climatic Change*, *165*(3), 1-21.

Lusk, J. L., Norwood, F. B. (2009). Bridging the gap between laboratory experiments and naturally occurring markets: an inferred valuation method. *Journal of Environmental Economics and Management*, *58*(2), 236-250.

Lusk, J. L., Schroeder, T. C. (2004). Are choice experiments incentive compatible? A test with quality differentiated beefsteaks. *American journal of agricultural economics*, 86(2), 467-482.

Michailova, J., Tyszka, T., Gawryluk, K. (2020). Are people interested in probabilities of natural disasters? *Large Risks With Low Probabilities*, 21.

Miller, K. M., Hofstetter, R., Krohmer, H., Zhang, Z. J. (2011). How should consumers' willingness to pay be measured? An empirical comparison of state-of-the-art approaches. *Journal of Marketing Research*, 48(1), 172–184.

Morris, S. B., DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological methods*, *7*(1), 105.

Mossin, J. (1968). Aspects of rational insurance purchasing. *Journal of Political Economy*, 76 (4): 553–568.

Murphy J, Allen G, Stevens T, Weatherhead D. (2005). A meta-analysis of hypothetical bias in stated preference valuation. *Environmental and Resource Economics*, 30, 313-325.

Nelson, J. P., Kennedy, P. E. (2009). The use (and abuse) of meta-analysis in environmental and natural resource economics: an assessment. *Environmental and resource economics*, *42*(3), 345-377.

Neves, P. C., Sequeira, T. N. (2018). Spillovers in the production of knowledge: A meta-regression analysis. *Research Policy*, *47*(4), 750-767.

Norton, M., Osgood, D., Madajewicz, M., Holthaus, E., Peterson, N., Diro, Gebremichael, M. (2014). Evidence of demand for index insurance: experimental games and commercial transactions in Ethiopia. *Journal of Development Studies*, *50*(5), 630-648.

Paulhus, D. L., Reid, D. B. (1991). Enhancement and denial in socially desirable responding. *Journal of Personality and Social Psychology*, 60, 307-317.

Peter, R. (2022). When is safety a normal good? *Available at SSRN 4147815*.

Petraud, J., Boucher, S., Carter, M. (2015). *Competing theories of risk preferences and the demand for crop insurance: Experimental evidence from Peru* (No. 1008-2016-80041).

Pitthan, F., De Witte, K. (2021). Puzzles of insurance demand and its biases: A survey on the role of behavioural biases and financial literacy on insurance demand. *Journal of Behavioral and Experimental Finance*, *30*, 100471.

Quiggin, J. (1982). A Theory of Anticipated Utility, *Journal of Economic Behavior Organization*, 3(4), 323-343

Read, D., Loewenstein, G., Rabin, M. (1999). Choice bracketing. *Journal of Risk and Uncertainty*, 19, 171–197.

Ready, R.C., Buzby, J.C., Hu, D. (1996). Differences between continuous and discrete contingent value estimates. *Land Economics*, 72(3), 397–411

Redelmeier, D. A., Tversky, A. (1992). On the framing of multiple prospects. *Psychological Science*, 3(3), 191–193.

Robinson, P.J., Botzen, W.J. (2019). Economic experiments, hypothetical surveys and market data studies of insurance demand against low-probability/high-impact risks: A systematic review of designs, theoretical insights and determinants of demand. *Journal of economic Surveys*, 33(5), 1493–1530.

Robinson, P. J., Botzen, W. J. (2018). The impact of regret and worry on the threshold level of concern for flood insurance demand: Evidence from Dutch homeowners. *Judgment and Decision making*, *13*(3), 237-245.

Robinson, P. J., Botzen, W. J. (2019). Determinants of Probability Neglect and Risk Attitudes for Disaster Risk: An Online Experimental Study of Flood Insurance Demand among Homeowners. *Risk Analysis*, 39(11), 2514–2527.

Rohde, I. M., Rohde, K. I. (2011). Risk attitudes in a social context. *Journal of Risk and Uncertainty*, *43*(3), 205-225.

Schade, C., Kunreuther, H., Koellinger, P. (2012). Protecting against low-probability disasters: The role of worry. *Journal of Behavioral Decision Making*, *25*(5), 534-543.

Schlesinger, H. (1997). Insurance demand without the expected-utility paradigm. *Journal of Risk and Insurance*, 19-39.

Schmidt, J., Bijmolt, T. H. (2020). Accurately measuring willingness to pay for consumer goods: a meta-analysis of the hypothetical bias. *Journal of the Academy of Marketing Science*, *48*(3), 499-518.

Schwerter, F. (2013). *Social reference points and risk taking* (No. 11/2013). Bonn Econ Discussion Papers.

Seifert, I., Botzen, W. J. W., Kreibich, H., H. Aerts, J. C. J. (2013). Influence of flood risk characteristics on flood insurance demand: A comparison between Germany and the Netherlands. *Natural Hazards and Earth System Sciences*, 13(7), 1691–1705.

Serfilippi, E., Carter, M., Guirkinger, C. (2020). Insurance contracts when individuals "greatly value" certainty: Results from a field experiment in Burkina Faso. *Journal of Economic Behavior and Organization*, 180, 731–743.

Smith, G. S., Day, B. H., Bateman, I. J. (2019). Preference uncertainty as an explanation of anomalies in contingent valuation: coastal management in the UK. *Regional Environmental Change*, 19(8), 2203-2215.

Smith, V. L. (1982). Microeconomic systems as an experimental science. *The American economic review*, *72*(5), 923-955.

Stanley, T.D. (2008). Meta-regression methods for detecting and estimating empirical effect in the presence of publication selection. *Oxford Bulletin of Economics and Statistics*, 70(1), 103–127.

Stanley, T. D., Doucouliagos, H. (2012). *Meta-regression analysis in economics and business*. Routledge.

Stanley, T. D., Doucouliagos, H. (2014). Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods*, *5*(1), 60-78.

Steel, P., Beugelsdijk, S., Aguinis, H. (2021). The anatomy of an award-winning meta-analysis: Recommendations for authors, reviewers, and readers of meta-analytic reviews. *Journal of International Business Studies*, 52(1), 23-44.

Tversky, A., Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, 5(4), 297-323.

Tyszka, T., Sawicki, P. (2011). Affective and cognitive factors influencing sensitivity to probabilistic information. *Risk Analysis: An International Journal*, *31*(11), 1832-1845.

Ugur, M., Trushin, E., Solomon, E., Guidi, F. (2016). R&D and productivity in OECD firms and industries: A hierarchical meta-regression analysis. *Research Policy*, *45*(10), 2069-2086.

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., Sánchez-Meca, J. (2013). Three-level meta-analysis of dependent effect sizes. *Behavior research methods*, *45*(2), 576-594.

Vendrik, M., Woltjer, G. B. (2007). Happiness and loss aversion: Is utility concave or convex in relative income? *Journal of Public Economics, 91*, 1423–1448.

Viechtbauer, W., Cheung, M. W. L. (2010). Outlier and influence diagnostics for meta-analysis. *Research synthesis methods*, 1(2), 112-125.

Voelckner, F., (2006). An empirical comparison of methods for measuring consumers' willingness to pay. *Marketing Letter* 17: 137-149.

Vossler, C.A., Doyon, M., Rondeau, D., (2012). Truth in consequentiality: theory and field evidence on discrete choice experiments. *American Economic Journal: Microeconomics* 4(4), 145-171.

Wagner, K. R. (2022). Adaptation and adverse selection in markets for natural disaster insurance. *American Economic Journal: Economic Policy*, *14*(3), 380-421.

Weber, E. U., Shafir, S., Blais, A. R. (2004). Predicting risk sensitivity in humans and lower animals: risk as variance or coefficient of variation. *Psychological review*, *111*(2), 430.

Welsh, M., Poe, G.L. (1998). Elicitation effects in contingent valuation: comparisons to a multiple bounded discrete choice approach. *Journal of Environmental Economics and Management*, 36(2):170–185.

Wertenbroch, K, Skiera, B. (2002). Measuring consumers' willingness to pay at the point of purchase *Journal of Marketing Research*, 39 (2), pp. 228-241.

Yan, Y., Faure, M. (2021). Government interventions in microinsurance: evidence from China. *The Geneva Papers on Risk and Insurance-Issues and Practice*, *46*(3), 440-467.

Yan, Y., Faure, M. (2021). Government interventions in microinsurance: evidence from China. *The Geneva Papers on Risk and Insurance-Issues and Practice*, *46*(3), 440-467.

Yiannakoulias, N., Darlington, J. C., Elshorbagy, A., Raja, B. (2018). Meta-analysis based predictions of flood insurance and flood vulnerability patterns in Calgary, Alberta. *Applied geography*, *96*, 41-50.

Meta-Analysis References

1. Aditya, K. S., Kishore, A., Khan, M. T. (2020). Exploring farmers' willingness to pay for crop insurance products: A case of weather-based crop insurance in Punjab, India. *Agricultural Economics Research Review*, 33(2), 135–146.
2. Adzawla, W., Kudadze, S., Mohammed, A. R., Ibrahim, I. I. (2019). Climate perceptions, farmers' willingness-to-insure farms and resilience to climate change in Northern region, Ghana. *Environmental Development*, 32(October), 100466.
3. Botzen, W. J. W., van den Bergh, J. C. J. M. (2012). Risk attitudes to low-probability climate change risks: WTP for flood insurance. *Journal of Economic Behavior and Organization*, 82(1), 151–166.
4. Bradt, J. (2022). Comparing the effects of behaviorally informed interventions on flood insurance demand: an experimental analysis of 'boosts' and 'nudges'. *Behavioural Public Policy*, 6(3), 485-515.
5. Budhathoki, N. K., Lassa, J. A., Pun, S., Zander, K. K. (2019). Farmers' interest and willingness-to-pay for index-based crop insurance in the lowlands of Nepal. *Land Use Policy*, 85(March), 1–10.
6. Budiasa, I. W., Temaja, I. G. R. M., Ustriyana, I. N. G., Nuarsa, I. W., Wijaya, I. G. B. A. (2020). The willingness of farmers to pay insurance premiums for sustainable rice farming in Bali. *Journal of the International Society for Southeast Asian Agricultural Sciences*, 26(1), 63–72.
7. Chen, Z.Y., Ling, Y.Y.(2008). Analysis of influence factor and willingness to pay for farmer on agricultural insurance-a case of tobacco insurance in XingShan county, Hubei province.*South China Journal of Economics*,7,61-68.
8. Deng, Y., Munn, I. A., Coble, K., & Yao, H. (2015). Willingness to pay for potential standing timber insurance. *Journal of Agricultural and Applied Economics*, 47(4), 510-538
9. Ellis, E. (2017). Willingness to Pay for Index Based Crop Insurance in Ghana. *Asian Economic and Financial Review*, 7(7), 700–721.
10. Fonta, W. M., Sanfo, S., Kedir, A. M., & Thiam, D. R. (2018). Estimating farmers' willingness to pay for weather index-based crop insurance uptake in West Africa: Insight from a pilot initiative in Southwestern Burkina Faso. Agricultural and Food Economics, 6(1), 1-20.
11. Friedl, A., Lima de Miranda, K., Schmidt, U. (2014). Insurance demand and social comparison: An experimental analysis. *Journal of Risk and Uncertainty*, 48(2), 97–109.

12. Hill, R. V., Hoddinott, J., Kumar, N. (2013). Adoption of weather-index insurance: Learning from willingness to pay among a panel of households in rural Ethiopia. *Agricultural Economics*, 44(4–5), 385–398.
13. Hillebrandt, M.-A. (2020). What Drives the Willingness to Pay for Insurance Contracts with Nonperformance Risk? Experimental Evidence. *SSRN Electronic Journal*, https://doi.org/10.2139/ssrn.3720078
14. King, M., Singh, A. P. (2020). Understanding farmers' valuation of agricultural insurance: Evidence from Vietnam. *Food Policy*, 94(January), 101861.
15. Kong, R., Turvey, C. G., He, G., Ma, J., & Meagher, P. (2011). Factors influencing Shaanxi and Gansu farmers' willingness to purchase weather insurance. *China Agricultural Economic Review*, 3(4), 423–440.
16. Krawczyk, M. W., Trautmann, S. T., Van de Kuilen, G. (2017). Catastrophic risk: social influences on insurance decisions. *Theory and Decision*, 82(3), 309–326.
17. Li, Y., Lin, Y., Kong, X. (2013).Study on Farmers' Willingness to Pay for Policy Forest Insurance Based on Cox Model. *Journal of Nanjing Agricultural University*, 40, 2, 103-108.
18. Lo, A. Y. (2013). Household Preference and Financial Commitment to Flood Insurance in South-East Queensland. Australian Economic Review, 46(2), 160–175. https://doi.org/10.1111/j.1467-8462.2013.12009.x
19. Mcintosh, C., Sarris, A., Papadopoulos, F. (2013). Productivity, credit, risk, and the demand for weather index insurance in smallholder agriculture in Ethiopia. *Agricultural Economics*, 44(4–5), 399–417.
20. Mol, J. M., Botzen, W. J. W., Blasch, J. E. (2020). Behavioral motivations for self-insurance under different disaster risk insurance schemes. *Journal of Economic Behavior and Organization*, 180, 967–991.
21. Netusil, N. R., Kousky, C., Neupane, S., Daniel, W., Kunreuther, H. (2021). The willingness to pay for flood insurance. *Land Economics*, 97(1), 17-38.
22. Ning, M. X., Miao, Q., Xing, L. (2006). An empirical analysis of farmers' willingness to pay for agricultural insurance: A case study of Manas River Basin in Xinjiang. *Chinese Rural Economy*, 6,43-51.
23. Ozdemir, O., Morone, A. (2014). An experimental investigation of insurance decisions in low probability and high loss risk situations. *Journal of Economic Interaction and Coordination*, 9(1), 53–67.
24. Petrova, D. G., Van der Pligt, J., Garcia-Retamero, R. (2014). Feeling the numbers: On the interplay between risk, affect, and numeracy. *Journal of Behavioral Decision Making*, 27(3), 191–199.
25. Raschky, P. A., Schwarze, R., Schwindt, M., Zahn, F. (2013). Uncertainty of Governmental Relief and the Crowding out of Flood Insurance. *Environmental and Resource Economics*, 54(2), 179–200.
26. Ren, J., Holly Wang, H. (2016). Rural Homeowners Willingness to Buy Flood Insurance. *Emerging Markets Finance and Trade*, 52(5), 1156–1166.
27. Robinson, P. J., Botzen, W. J. W. (2019). Determinants of Probability Neglect and Risk Attitudes for Disaster Risk: An Online Experimental Study of Flood Insurance Demand among Homeowners. *Risk Analysis*, 39(11), 2514–2527.
28. Roder, G., Hudson, P., & Tarolli, P. (2019). Flood risk perceptions and the willingness to pay for flood insurance in the Veneto region of Italy. *International Journal of Disaster Risk Reduction*, 37, 101172.
29. Seifert, I., Botzen, W. J. W., Kreibich, H., H. Aerts, J. C. J. (2013). Influence of flood risk characteristics on flood insurance demand: A comparison between Germany and the Netherlands. *Natural Hazards and Earth System Sciences*, 13(7), 1691–1705.
30. Serfilippi, E., Carter, M., Guirkinger, C. (2020). Insurance contracts when individuals "greatly value" certainty: Results from a field experiment in Burkina Faso. *Journal of Economic Behavior and Organization*, 180, 731–743.
31. Sun, X. (2008). Crop Insurance Knowledge, Trust on Government and Demand for Crop Insurance- an Empirical Study of Peasant Households' Willingness-To-Pay in Huai'an, Jiangsu Province. *Journal of Nanjing Agricultural University*, 8(1),48-54.
32. Xiu, F., Xiu, F., & Bauer, S. (2012). Farmers' willingness to pay for cow insurance in Shaanxi province, China. Procedia Economics and Finance, 1, 431-440.
33. Yang, T., Chen, S., Mu, Q., Liu, X., Liu, W., Yee, T., Ye, T. *(2021).* Willingness-to-pay for yak snow disaster weather index insurance: A case study of Yushu Tibetan Autonomous Prefecture, Qinghai Province. *Research of Agricultural Modernization, 42(4), 745-754*
34. Yang, X., Liu, Y., Bai, W., Liu, B. (2014). Evaluation of the crop insurance management for soybean risk of natural disasters in Jilin Province, China. *Natural Hazards*, *76*(1), 587-599
35. Zhang, C. M. (2020). Seismic risk-coping behavior in rural ethnic minority communities in Dali, China. *Natural Hazards*, 103(3), 3499–3522.
36. Zheng, H., Mu, H., Zhao, X. (2018). Evaluating the demand for aquaculture insurance: An investigation of fish farmers' willingness to pay in central coastal areas in China. *Marine Policy*, 96(June), 152–162.
37. Zimmer, A., Gründl, H., Schade, C. D., Glenzer, F. (2018). An Incentive-Compatible Experiment on Probabilistic Insurance and Implications for an Insurer's Solvency Level. *Journal of Risk and Insurance*, 85(1), 245–273.

Zimmer, A., Schade, C., Gründl, H. (2009). Is default risk acceptable when purchasing insurance? Experimental evidence for different probability representations, reasons for default, and framings. *Journal of Economic Psychology*, 30(1), 11–23.