THESIS

in partial fulfilment of the

Master of Public Policy – International Program

Economic Policy, Finance and Development

Graduate School of Public Policy

The University of Tokyo

(Dual Degree with Hertie School of Governance)

STABILITY OF RISK PREFERENCES IN THE FIELD
–
a Meta-Analysis

Student:	Valentin BAUMANN
Student No:	590411
Degree:	Master of Public Policy
Date:	15th June 2023
Word Count:	14 634 (excl. references)
Supervisor:	Hitoshi SHIGEOKA
Secondary Reviewers:	Yasutora WATANABE
	Yasuyuki SAWADA

List of Contents

Abstract

While research on the relationship between various shocks and risk preferences has intensified in recent years, there has been little convergence regarding the direction and magnitude of the effects. In an attempt to assess the quantitative status quo of the literature, I gathered a sample of 58 studies which investigate the effects of different kinds of shocks on risk preferences. I found that about half of the studies find a decrease in risk seeking in response to a shock compared to a quarter which find an increase and another quarter which find either no effect or mixed effects. Keeping in mind that comparisons of effect sizes across studies is somewhat problematic due to substantial variation in research designs and statistical methodologies, I conducted a meta-analysis of the effect sizes as well as a number of subgroup-comparisons, meta-regressions, and tests for publication bias. Consistent with the large share of studies finding negative effects on risk seeking, the overall effect appears to be negative. However, it is very small (average *Cohen's d = –0.043*) and seems to be driven by subgroups of medium-term effects and effects of economic shocks. Also, between-study-heterogeneity is very large. Publication bias remains a possibility, but I find no evidence of directional bias. Future meta-analytic work could further investigate the influence of specific methodologies and research designs. Future primary research could facilitate meta-analyses by standardising research designs, statistical methodologies, and reporting practices.

List of figures in main text

List of tables in main text

List of statistical outputs in main text

I. Introduction

Modern life has become incredibly complex. One of the central features of this complexity is an escalating number of risks and uncertainties individuals and societies are faced with at any moment. Unsurprisingly, there has been substantial academic interest in various facets of how people, organisations and societies approach risks and uncertainties. In fact, some of the most highly cited publications in the social sciences focus on risk and uncertainty, with "*Prospect Theory"* (Kahneman & Tversky, 2013), *"Knightian Uncertainty"* (Knight, 1921) or "*Risikogesellschaft"* (Beck, 2020) being household names in psychology and sociology. Due to its practical and conceptual importance for many aspects of modern economies and financial markets, there has been intense interest from empirically minded economists in the study of people's attitudes toward risk.

This research is also relevant for public policy and social science more broadly because people's relationship to risk, i.e. their willingness to take risks, their fears when it comes to uncertain futures, and their needs to find a balance between predictability on the one hand and potentially greater rewards when taking risks on the other hand, influence a wide range of decisions. On top of the literature on risk preferences in laboratory settings, research has become increasingly interested in relationships between risk preferences and real-world variables which has spurred a flurry of field research on risk preferences in many different contexts, in a wide range of locations, and with a wide range of participants. Furthermore, risk preferences have been studied both as independent and as dependent variables.

Some of the areas where risk preferences as an independent or mediating variable become a concern for stakeholders in the private as well as in the public sectors are decisions relating to insurance buying (Ali et al., 2021; Barsky et al., 1997b), occupational and career choices (Bellemare & Shearer, 2010), entrepreneurship (Cramer et al., 2002), investment decisions (Bryan et al., 2014; Meister & Schulze, 2022), corporate and executive behaviour (Witte, 2012; Yao et al., 2020), household's financial decisions (Khan et al., 2022; Kimball et al., 2008; Qiu et al., 2014; Volland, 2017), perceptions and attitudes regarding novel and potentially risky technologies (Gao et al., 2020; Huang et al., 2013), migration (Batista & Umblijs, 2014; Bauernschuster et al., 2014; Jaeger et al., 2010), and health-related behaviour behaviours (Anderson & Mellor, 2008; Barsky et al., 1997a; Dupas, 2014; Khwaja et al., 2006) just to name a few.

At the same time, substantial effort has been made to investigate the *stability of risk preferences as well as the factors influencing them*. For example, there has been research on the effects of emotions (Baillon et al., 2016; Campos-Vazquez & Cuilty, 2014; Cobb-Clark et al., 2022; Conte et al., 2018; Habib et al., 2015; J. S. Lerner & Keltner, 2001; Jennifer S. Lerner et al., 2003; Raghunathan & Pham, 1999; She et al., 2017; Treffers et al., 2016), other mental states (Castillo et al., 2017), health (Banks et al., 2020; Decker & Schmitz, 2016), social and familial factors (Alan et al., 2017; H. Brown & van der Pol, 2015; Gardner & Steinberg, 2005; Hermann, 2017), decision contexts (Barseghyan et al., 2018), lived experiences (Deole & Rieger, 2023), and individual characteristics such as gender (Charness & Gneezy, 2012). Furthermore, basic intertemporal stability has been investigated in a number of studies (Harrison et al., 2005; Josef et al., 2016).

Additionally, especially in the past decade, substantial effort has been made to learn more about the effects of exogenous shocks such as natural disasters, financial crises, conflicts and most recently the COVID-19 pandemic on risk preferences. It is this last topic that I am focusing on in this thesis. The main motivation for this research endeavour is that despite many studies having been conducted to investigate this question, there appears to be *a lack of convergence regarding the direction and magnitude of the effects*. Recent publications on the topic usually acknowledge that the existing evidence is mixed and this fact is also one of the factors that motivates new studies in the first place (Abatayo &

Lynham, 2020; Kuroishi & Sawada, 2019). As things stand today, the issue is less a lack of evidence but rather a lack of convergence when it comes to even the most fundamental questions such as whether shocks increase or decrease people's willingness to take risks. There have been laudable attempts to consciously improve methodologies (Kuroishi & Sawada, 2019; Reynaud & Aubert, 2020) and to better capture heterogeneity of estimates of the stability of risk preferences and the impacts of shocks (Rockmore & Barrett, 2022). However, even if those studies manage to avoid some weaknesses of prior work, they will at most come to valid conclusions regarding the stability on risk preferences in the specific context meaning that their results will still not necessarily be generalisable far beyond the these contexts, e.g. Japan after an earthquake or the Philippines after a tsunami (Kuroishi & Sawada, 2019). In the field, this context dependence appears to be common knowledge (Kettlewell et al., 2023).

All publications in the field do contain an element of literature review. However, because they focus on primary research, those reviews are necessarily selective and cannot paint a complete picture of the work that has been done on the topic. Studies usually mention only a small subset of the existing literature on the topic in the introductory section of their papers. While this may be largely adequate for explaining the motivation for a research endeavour, one should not read those parts as in any way complete accounts of the state of knowledge.

In addition to maintaining that the literature on the stability of risk preferences has so far delivered rather contradictory results, Schildberg-Hörisch, 2018, concludes in her broad narrative review of the literature on the stability of risk preferences that the "*literature is growing quickly, and in the future, it may become possible to do a meta-analysis that could shed light on the reasons behind the divergent findings*". Considering that in recent years, the literature has indeed grown, it appears reasonable to believe that a first meta-analysis on the topic may be viable now – with the caveat being that a number of papers on the effects of the Covid-19 pandemic is likely to be published relatively shortly after this thesis has been written. Given this context, the contribution of this thesis would be to provide both a methodological framework (selection criteria, weighting approach, code etc.) and an analysis of the status quo which can then be relatively easily expanded and updated with more incoming research.

Apart from the review by Schildberg-Hörisch, I am only aware of one other systematic review on the stability of risk preferences, namely that by Chuang & Schechter, 2015. It concludes that "*no systematic evidence that real world shocks influence play in games*". However, while methodologically useful, there are some important limitations to their review. Most of them are either a direct or indirect result of the fact that by 2015, there had only been a very limited number of publications on the stability of risk preferences. In fact, as I will touch on later, the majority of studies on the stability of risk preferences in response to shocks have been published between 2015 and now. Consequently, the review only contains 19 studies on risk preferences. Some of them are very old (Love & Robison, 1983) and some do not include quantitative estimates of stability. Furthermore, the review does not focus exclusively on the relationship between *shocks* and risk preferences but takes a broader look at *stability over time*. While review might have been able to expand the sample size somewhat (I did find a number of relevant pre-2015 studies that had not been included in the review) but the underlying problem of a rather small number of relevant studies was certainly the main constraint.

Since then, however, the situation has changed dramatically with a total of around 100 studies tackling the question of stability of risk preferences. With a more stringent focus on the consequences of shocks I was still able to gather a dataset of 66 studies (48 of which provide quantitative estimates of the effect size). Major accelerators appear to be the Covid-19 pandemic (Adema et al., 2022; Cicerale et al., 2022; Frondel et al., 2021; Graeber et al., 2020; Ikeda et al., 2020; K. K. Li et al., 2020; Lohmann et al., 2020; Meunier & Ohadi, 2021; Zhang & Palma, 2021) and the emergence of studies investigating long-term effects of the 2008 global financial crisis (Dohmen et al., 2016; Gerrans et al., 2015; Jetter et al., 2020)

With this amount of literature now available, it seems worthwhile to approach the relationship between shocks and risk preferences quantitatively by using the tools of meta-analysis. Employing a meta-analytic approach allows us to reach a number of goals. Firstly, it enables an assessment of whether the effects found by individual studies point into a certain direction when aggregated (Card, 2016; Harrer et al., 2021). Secondly, as a by-product, it will deliver a potentially useful quantitative overview of the literature, i.e. how many studies have been published, what types of shocks have they studied, do different types of influences (economic, natural, man-made etc.) tend to lead to different effects, what methodologies were employed, do different methods tend to yield different effects and much more. Drilling down further into the effects that were found and the characteristics of the studies may uncover potential patterns in the literature which might have remained under the radar when focusing on individual studies. Lastly, a meta-analysis can also investigate whether there is bias in the literature, be it publication bias or bias related to author's or journal's characteristics.

The goal of this thesis is to provide the to this date most comprehensive, transparent, systematic, and *quantitative* review of the literature on the stability of risk preferences. The primary research objective is to determine the *direction and magnitude* of the effect of different types of shocks on risk preferences. However, as I will explain in more detail later, I believe that this objective may yield unsatisfactory results and should be complemented by further quantitative and qualitative analyses of the existing literature. This way, I hope to paint a nuanced picture of the existing literature which includes statements about aggregated tendencies as well as heterogeneity and its possible causes.

As I mentioned before, the literature on shocks and risk preferences has reached a quantitative and qualitative level which allows for meaningful aggregate review and even some subgroup analyses that are hopelessly underpowered. With this meta-analysis, I hope to provide a valuable service for researchers in the field as well as outsiders, i.e. researchers in adjacent fields, policy makers etc. by *(i)* providing a condensed overview of the available literature, *(ii)* delivering estimates regarding true effect sizes and stability of risk preferences, *(iii)* identifying patterns and potential gaps in the literature, and *(iv)* informing future research endeavours.

While both meta-analyses and narrative reviews have their respective advantages, meta-analyses allow for an overall quantification of the direction in which the literature is pointing. Pure meta-analysis certainly risks losing some nuance, but, as for instance Card, 2016, mentions, using a meta-analytical approach is not incompatible with a potentially more nuanced narrative analysis in the very same article. I will approach the research question in this spirit even if my focus will be on quantitative analysis.

The specific methodology in this thesis is primarily based on the textbooks on meta-analysis by Harrer et al., 2021, and meta-analysis in social science by Card, 2016, with some details being informed by further contributions such as Davis et al., 2014. Results are reported according to the reporting guidelines for meta-analysis in economics by Havránek et al., 2020, for the Journal of Economic Surveys.

Employing a combination of search and selection criteria, I compiled a sample of 58 articles on the effects of different kinds of shocks and risk preferences. I found that about half of them find a decrease in risk seeking in response to a shock compared to a quarter finding an increase and another quarter finding either no effect or mixed effects. Comparing effect sizes across studies is somewhat problematic (more on that in (*II.vi)* but given this caveat I nevertheless conducted a general meta-analysis of the effect sizes as well as a number of subgroup-comparisons, meta-regressions, and tests for publication bias. Consistent with the large share of studies finding negative effects on risk seeking, the overall effect appears to be negative as well. However, it is very small and seems to be driven by medium-term effects and effects of economic shocks. Also, between-study-heterogeneity is very large. Publication bias remains a possibility, but funnel plots may be misleading due to the substantial heterogeneity.

Regarding the implications for policy, I would like to add that if risk preferences and assumptions about their (in)stability are ever going to be part of policy maker's considerations, there will need to be more work which systematic and ideally regularly updates the aggregate view of the topic in addition to careful review of smaller sets of studies that are specifically relevant to the respective question at hand.

II. Methodology and Data

Meta-analysis is not *per se* a "gold standard" or the "highest" form of evidence for any scientific question. Rather, it is only one of many different ways of advancing scientific knowledge. While aggregation allows for genuinely new insights, it necessitates abstraction and a loss of nuance and detail. Furthermore, the research question has to be adequate, carefully chosen, and well-defined with respect to scope and target question (Davis et al., 2014) in order to avoid applying selection criteria that are, for instance, too loose and which could lead to unwarranted claims – both false-positive and false-negative. This is especially true in the social sciences. In contrast to general practice in the medical sciences, meta-analysis in the social sciences sometimes combines evidence from studies with different levels of randomisation and different research contexts. Instead of excluding primary studies on the basis that they do not use active randomisation but rather other quasi-experimental methods or even non-causal empirical strategies, the standard approach is to use "best available" as the benchmark for an inclusion strategy. This is often necessary due to the still rather limited number of RCTs and in some fields even quasi-experimental methodologies in the social sciences (Davis et al., 2014).

Additionally, the social sciences tend to employ a wide range of different measurement methodologies for a variable such as risk preference. Additional between-study variation may arise from different sample compositions, locations and time lags between treatments and measurements just to name a few.

In the following chapters, I will describe the procedures and methodologies employed for this study and I explain why I believe the specific combination of decisions I made along the way represent a reasonable balance between scientific rigour, comprehensiveness, efficiency, and practicality.

i. Selection of studies

In order to enable other researchers to replicate the sample that was chosen in the present study, I will describe the process and criteria in some detail in this section. For meta-analyses, the goal is usually to include *all* relevant studies, or more precisely, an exhaustive and unbiased sample of studies. However, as there is no guarantee that *all* studies were found, the selected studies should be treated as a sample which reflects the overall "population" (Card, 2016).

The choice and in particular the strictness of search criteria will inevitably have an influence on the completeness of the sample. To avoid excluding relevant studies, I decided broadly on a two-step process. Firstly, I searched for as many *potentially* relevant publications as possible using relatively broad *search criteria* with the goal of *high* recall. Secondly, I selected the studies for the final sample based on the more stringent *selection criteria,* the goal here being *high precision*. For a number of cases, this two-step-process did not apply because I found them later on through references in the pre-selected publications. I documented the results of both processes (*Table 2 and 3, Appendix*).

For the initial literature search, I employed a number of sources such as Google Scholar, Web of Science, and PsycINFO. For all of them, I used the following search terms in different combinations:

> For dependent variables: *risk, risk, uncertainty, ambiguity, preferences, seeking, aversion, behaviour, perception*

> For independent variables: *shocks, natural disasters, earthquake, flood, conflict, violence, (global) financial crisis, economic crisis, recession, stability*

Based on the results from those searches, I picked all articles that appeared relevant to the research topic based on the respective titles and/or abstracts. Additionally, I employed "forward search" functions provided by the search engines to identify articles citing the articles in the primary results lists.

The second major source for studies were the review articles by Schildberg-Hörisch, 2018, and Chuang & Schechter, 2015. Additionally, while reading the already identified articles I engaged in "backward search", which means adding articles that are cited in those publications.[1] Through this process, I compiled a list of 94 articles.

Even though the goal of the *selection criteria* is high precision, I still chose relatively loose criteria in order to *(i)* capture as many relevant studies as possible[2], *(ii)* maximise sample size and thus statistical power, and *(iii)* allow for a large number of subgroup analyses and study of between-study-heterogeneity. Furthermore, applying relatively lax criteria in the selection stage still leaves the door open to be more restrictive later by including adherence to stricter criteria as a variable in the eventual dataset.

Even with relatively broad selection criteria, some subgroups contained only a very small number of studies which made meaningful analyses and meta-regressions difficult. Given that only few years ago, there existed barely enough work to conduct a meta-analysis let alone subgroup analyses, this should be regarded as progress nevertheless (Chuang & Schechter, 2015; Schildberg-Hörisch, 2018).

The central selection criterion was whether a study investigated the quantitative relationship between a clearly identified *real-world shock* and subsequent measurements of *risk preferences*. This means that studies with a focus on intertemporal stability of risk preferences in the absence of a clearly identified shock were excluded (Akesaka et al., 2021; Andersen et al., 2008; Brunnermeier & Nagel, 2008; Bucciol & Miniaci, 2018; Chetty & Szeidl, 2007; Ert & Haruvy, 2017; Harrison et al., 2020; Josef et al., 2016; Krčál et al., 2019; l'Haridon & Vieider, 2019). I exclude them mainly because the coefficients in the respective statistical analyses do not reflect the effects of shocks. Other publications that examine the effects of societal factors (Schmidt et al., 2019), professional experience (Krčál et al., 2019), and events with largely positive valence (Angerer et al., 2021; Haile et al., 2020) were also excluded. Whilst those studies are surely worthwhile and very interesting in their own right, they differ in too many ways from the studies I am focusing on.

In order allow at least for some conceptual comparisons, the shock had to be a predominantly negatively coded event such as natural disasters, economic crises, conflict & violence, or a pandemic. I will consider some edge cases in the following chapter. As a time frame I set the year 2000 as the starting date. The initial literature search showed that the overwhelming majority of studies had been published after 2000 and the only earlier studies on the stability of risk preferences were much older and did not have the specific focus on the relationship between a shock and risk preferences (Love & Robison, 1983; Wehrung et al., 1984).

[1] Both forward and backward search can be risky as they might introduce bias as articles may tend to cite and be cited by articles that find similar results. Therefore, they should not be the only approach. In combination with keyword searches in databases, I believe that employing those techniques is justified. Additionally, I found that many articles did indeed cite both articles that found similar results and such that found no results or completely opposing ones.

[2] After all, an important goal of this thesis is to provide a maximally comprehensive overview of the current state of the literature.

In contrast to the independent variable (i.e. the shock), which had to be a real-world event, the dependent variable (i.e. risk preferences) could be measured in a number of different ways. What the measurements had to have in common was minimum level of construct validity. This means that I included studies based on survey measures with self-reports and hypothetical decisions, experiments with incentivised as well as non-incentivised risk measures, and even some risk measures by proxy such as financial decision making of stock traders and corporate executives if the authors plausibly defended the validity of such measures as reflecting people's risk preferences. On the flipside, I excluded articles with a focus on real world behavioural outcomes which might be linked to risk preferences but are not explicitly introduced as *measures* of risk preferences (Adjei-Mantey & Horioka, 2022; Choi & Kim, 2022; Filipski et al., 2019; Schaller & Stevens, 2015). The full list of excluded articles, each with a brief justification can be found in *Table 3 (Appendix)*.

Regarding causal identification, I chose to first apply loose criteria as well in order to include as many relevant articles as possible. This means that I include studies that used a range of quasi-experimental, correlational, longitudinal, and cross-sectional methodologies. I did, however, include this information as variables in the dataset which allowed for later analysis of their influence on mean effect sizes.

The same applies for publication status and peer review. The sample includes both peer-reviewed journal articles as well as non-peer-reviewed working papers. Furthermore, studies were not required to report a single representative effect size in order to qualify for the dataset. However, only those that reported the effects could be included in the determination of an average effect size.

Another important decision was whether to exclude the publications that focused on the effects of the COVID-19 pandemic on risk attitudes. Possible reasons for doing so would have been that *firstly*, at the time I conducted this meta-analysis, the pandemic was still a very recent event. It did in fact spur significant research activity in the field of risk attitudes but as the research and publication process takes some time, there had only been a relatively small number of papers published on the topic. Consequently, including the research that had come through by the time of my research period may have represented a very incomplete and possible skewed picture of the eventual evidence on the topic. *Secondly*, the COVID-19 pandemic has been such a global and all-encompassing event and process that it is more difficult to classify than most other events. For different people in different circumstances, it likely presented itself in very different ways; for some, it was a natural disaster, for some a health crisis and for other a mostly economically or socially impactful event.

I believe that the relationships between the pandemic and risk attitudes warrants a further and potentially specifically dedicated analysis. However, for the present purpose I decided to nevertheless include publications dealing with the effects of Covid-19 in order to present the most up to date account of the state of the literature.

Ideally, the full set of inclusion criteria should have been determined *before* the start of the search process. However, due to my limited knowledge of the available literature prior to the process, I chose a somewhat *iterative approach* with slight adjustments of the criteria along the way. This enabled me to provide what I believe to be a more complete sample of studies which can share a common scientific focus while providing ample heterogeneity along a range of variables thus enabling interesting and relevant subgroup comparisons.

ii. Coding

After compiling the final list of publications, I completed the dataset by reading the articles in detail and coding a range of different variables. I should mention that even though I did define the majority of variables to code before starting the coding process, there were a number of minor adjustments that had to be made during the process in response to increasing knowledge of certain patterns in the literature.

For the simplest version of a meta-analysis, Card, 2016, recommends to simply code effect size and sample size for each study. However, in order to present a more comprehensive overview and to analyse the sources of heterogeneity, more characteristics need to be coded. In that sense coding study characteristics beyond effect size and sample size is both of descriptive and of explanatory value.

Table 1: Variables in final dataset

Category	***Variable***
Authors	Names, Share of male authors, Main academic discipline
Source characteristics	Publication status, Peer review, Citations on Google Scholar and Web of Science, Journal, Journal's impact factor
Article results	Title, Year of publication, Focus
	Reaction to positive events (if applicable), Reaction to negative events, (Main) Sources of within-study-heterogeneity, Statistical tools, Reported coefficient, Standardised effect size, Categorical classification of effect size, p-value, Significant effect, Reported standard error, Standard deviation, Standardised standard error, Persistence
Additional weights	Factor for double-counting, Quality index
Sample information	Location, WEIRDness, Size, Year, Sampling lag, Share of male participants, Mean age
Design	Experiment/Survey, Longitudinal/Cross-sectional, Source of variation, Causal identification
Dependent variable	Elicitation method, Incentivised
Independent variable	Shock type, GFC, COVID-19

In this section, I will outline the coding process and explain the major decisions I made during this process. I will give special consideration to "high inference" decisions, i.e. those decisions that required the highest degree of judgement on my side. The goal is to equip the reader with a good understanding of the process and reasonings for the sake of transparency as well as to enable future replications and expansions of my work. The variables included in the final dataset are displayed in *Table 1*. The full dataset also includes comments clarifying and explaining individual coding decisions.

I will start by explaining some of the key coding decisions. Due to the considerable variation in study designs and methodologies, many of the variables include some entries that were not entirely trivial to complete.

General information on the authors, articles and journals was the most straightforward to find and to code so there is no need to go into further detail. Article citation numbers as well as journal impact factors reflect the status quo on 26th May 2023. Furthermore, some of the more recent working papers may complete peer review in the coming months or years so their publication status (as well as their results in some cases) may change.

Regarding the *categorical* main results of the studies, i.e. in which direction (if at all) risk preferences changed in response to a shock, I followed the assessments of the authors. In most cases, the overall assessment was easy to confirm based on the reported statistical tests. However, in some cases, I was not able to identify one single representative coefficient that directly related to the general assessment. In those cases, I had to make the coding decision based on whether and to what extent I agreed with the author's reading of their results. In most cases, this was the case. However, in some cases, I had to amend the initial decision. For instance, Bchir & Willinger, 2013 report an increase in risk seeking in response to natural hazards. However, this result only holds for low-income participants while for the whole sample, the effect is small and not significant. In this case, I chose to include the study as two separate data-points with corresponding factors to avoid too much weight on the sample of this study. Similarly, in cases with two clearly distinct effects depending for instance on the elicitation method or subsample characteristics, the paper features twice in the dataset and its weight will be adjusted accordingly (Adema et al., 2022; Guiso et al., 2018; Hanaoka et al., 2018). Other publications such as Eckel et al., 2009; Holden & Tilahun, 2021, report *"substantial shock effects"* even though a closer reading of the results reveals that those effects go in very different directions. I therefore classified the result as "*mixed*". Publications with mixed effects were also excluded from the variable for exact coefficients and effect sizes as they did not provide a single measure. Furthermore, cases in which an effect was found only on loss aversion but not on overall on risk seeking (Meunier & Ohadi, 2021) as "*no effect*".

In rare cases which did not report sufficient data to calculate the effect size but which did mention a qualitative estimation of the magnitude of the effect (Malmendier & Nagel, 2011), I take the author's judgement at face value and include their effect size classification in the analysis.

In order to make the classifications comparable, they also had to be coded with respect to the same scale. In my coding, I work with *risk seeking* as the target metric, even though a significant part of the studies measure risk attitudes with respect to *risk aversion*. Accordingly, in those cases, both categorical assessments by the authors as well as the reported results had to be inverted so that *"increase"* or a positive effect size imply an increase in *risk seeking* whereas *"decrease"* or a negative effect size imply a decrease in risk seeking.

Somewhat surprisingly, coding the effect sizes in a way that enables aggregation and comparison across studies was not always straightforward as not every study provided a single representative coefficient and, in most cases, the coefficients were not standardised. Furthermore, due to the substantial variation both in independent variables and in especially the measurements of the dependent variable, the coefficients as they are reported are hardly comparable across studies. For the same reason, it is also problematic to convert them according to a common and meaningful scale. This may be possible in other meta analyses – for instance, Jackson & Mackevicius, 2023, standardise improvements in test scores in response to specific increases in spending on schools – but in the case of shocks and risk preferences, neither of the variables lend themselves to a conversion into such a scale. For that reason, the only plausible procedure I could find was to standardise the regression coefficients by dividing them by the respective standard deviations (Card, 2016; Kadlec et al., 2023).[3] This calculation yields *Cohen's d* as an effect size that can be compared across studies. For the studies which did not run any kind of

[3] Kadlec et al. (2023), find that meta-analyses regularly fail to consistently perform the correct standardisation procedure (dividing by the standard deviation and *not* by the standard error).

regression but instead based their analysis on t-tests and Chi-squared tests, I manually calculated Cohen's *d* using the meta-analysis effect size calculator provided by the *Campbell Collaboration.*[4]

Even in the cases where standardised effect sizes were reported, any comparison across studies should be taken with caution as statistical approaches and in particular controls differ widely between studies. I discuss the implications in *II.iii*.

Another case where coding was not always completely straightforward was sample size. Most cases were simple but for some, the sample size as reported in the main text deviated from the *N* for the relevant statistical tests in the results tables. Here, I chose the *N*s from the results tables. For longitudinal studies, I selected the number of participants which completed all of the relevant surveys or experimental sessions. For repeated cross sections with different participants, I calculated the total number of participants.

Generally, coding for meta-analyses should be reliable (Card, 2016). To ensure reliability or at least to provide transparency regarding coding reliability, one can empirically measure either inter- or intra-coder reliability. *Inter*-coder reliability, as the name suggests, requires two different coders whose results can be compared with each other. *Intra*-coder reliability requires only one person coding the same studies twice. As I conducted this study alone, I could only realistically consider intra-coder reliability. The disadvantage there is that there is no way to prevent knowledge of the first coding round to spill over to the second one even though this effect may be reduced through randomly choosing the second sample of studies so that at least there is no awareness of the sample that will be coded twice at the time of the first coding. Nevertheless, checking for intra-coder reliability can be used to identify "drift" due to learning over the course of the coding process. Due to time constraints, however, I chose not to conduct a dedicated coding reliability check. As mentioned in *II.i,* the coding process was somewhat iterative and after noticing patterns in the literature that were worth coding in the dataset, I did indeed re-read all of the publications at least partly and many of them even a third time. While doing this, I double-checked and, if necessary, corrected the previously coded variables. Given that at the time of re-reading the articles, I had already undergone a learning process, I believe that any existent drift should have been significantly reduced.

iii. Levels of confidence

As mentioned in the previous chapter, comparisons of effect sizes across studies are problematic when it comes to the literature on shocks and risk preferences. In this chapter, I will explain this problem in more detail and discuss the implications regarding the further analysis and the weight one should put on the different results.

The first issue is that for this study, I rely on the results as they are reported in the published articles and working papers. I do not have access to the raw data. Many of the studies employ more or less sophisticated statistical techniques to identify the effects of shocks on risk preferences. Only very few do not control for covariates. As (Card, 2016) notes, this is a problem and he even goes so far as to actively caution against comparing coefficients in regression analyses with very different sets of controls (*"it makes no sense"*).

In the, to this date, only systematic review on the topic, Chuang & Schechter, 2015, report correlation coefficients for some of the studies in the analysis but they do not use this data to calculate averages.

[4] Wilson, D. B., Ph.D. (n.d.). Practical Meta-Analysis Effect Size Calculator [Online calculator]. Retrieved June 9, 2023, from https://campbellcollaboration.org/research-resources/effect-size-calculator.html

Instead, they cautiously report the range of coefficients, and they mention the qualitative assessment that the coefficients to not seem to differ between longer and shorter periods of time nor between incentivised and hypothetical measures. However, they do not calculate average effect sizes and do not comment on whether this may be an adequate approach.

Furthermore, most studies try to paint a picture of the stability of risk preferences that is nuanced and reflects real life heterogeneity. This means that those studies cannot easily be reduced to one headline results. For example, (Guiso et al., 2018) study the effects of the 2007/2008 global financial crisis (GFC) on risk preferences. Besides general issues regarding generalisability beyond the specific sample (in this case Italian customers of a certain bank with at least 10.000€ in assets at that bank and the willingness to take part in a survey on several occasions), it may also be problematic to boil the study down to one simple result.

For instance, the authors employ both a categorical self-report measure as well as a hypothetical financial decision and then analyse the resulting data in a number of different ways such as simple comparisons of means as well as more complex methods involving regression models with different sets of control variables. Most of the resulting coefficients turn out to indicate a tendency towards a reduced willingness to take risks (increased risk aversion) but not all of them are statistically significant. Furthermore the qualitative self-assessment and the more quantitative financial decision may not necessarily reflect the same underlying processes. I will discuss this in more detail in the discussion section but for now it suffices to say that this means that the results section of the publication is both more tentative *and* more interesting than the headline result *"financial crisis tends to increase risk aversion"* may appear like.

This does not mean that reducing the publication to a single number for the purposes of a systematic review is not warranted but it does in my view strengthen the case for a hierarchy of confidence in the different levels of abstraction. The broad headline result (*"financial crisis tends to increase risk aversion"*), probably reflects the findings in the publication relatively well. Even though it is certainly simplification, it is consistent with most (but not all) of the statistical results.

A more precise effect size (e.g. an increase of .42 in the mean of the qualitative risk aversion measure in 2009 compared to 2007) is consistent with the headline result but meaningfully differs from the other measures. In the case of Guiso et al., 2018, the more quantitative measure points in the same direction and also correlates significantly with the qualitative one but the precise effect of the crisis is different.

Other studies, however, find that different measurement methodologies yield results that point in different directions (Adema et al., 2022). In this case, as well as the one of Guiso et al., 2018, it seems sensible to include the effect sizes of the different measures as separate data points and account for the double counting by reducing the respective weights. Most studies indeed rely on one main method for measuring risk preferences. In those cases, it appears to be more sensible to select the effects that are deemed most representative by the respective authors. The caveat that this is usually a stark simplification of the findings, however, remains.

Yet another reason why caution is warranted when interpreting effect sizes as representing the effect of a certain shock on risk preferences is that measurement and operationalisation of the dependent variable (i.e. the shock) varies substantially between studies. Some studies (Kettlewell, 2019; Kettlewell et al., 2023) explicitly analyse the impact of different levels or intensities of exposure to a shock within one paper. Kettlewell et al., 2023, for instance, find an overall increase in risk tolerance after a tsunami. However, further analysis revealed that this increase did only occur for subjects who experienced severe consequences such as injuries, property damage or even displacement but not for those who simply experienced the tsunami.

If, for the purpose of this meta-analysis, we need to choose one headline effect, I would argue that the statement "tsunami exposure tends to increase risk tolerance" better represents the finings than the statement "*tsunami exposure increases risk preferences by one third of a standard deviation*". On top of that, studies vary greatly in their reporting of standardised effect sizes and in the cases, where no standardised effect sizes are reported not all studies report sufficient information for manual calculations. There are ways to estimate comparable effect sizes using reported coefficients, p-values, sample sizes and standard deviations, but given that I could not work with the raw data, manual calculations may contain some errors or deviations from the true effect sizes. Nevertheless, I was able to calculate effect sizes for the majority of studies. The full dataset includes comments for the cases that required manual calculations.

As mentioned in *II.i.*, notwithstanding very few exceptions, the resulting effect sizes are mostly plausible and consistent with the qualitative descriptions in the respective publications. Against this background, I chose to split the analysis of the effects in three parts. *Firstly*, I classify the effects broadly depending on whether risk preferences increase, decrease, do not change ("no effect") or whether the results are mixed. *Secondly*, I classify the effects categorically into mixed, no effect, small, medium, and large. *Thirdly*, I analyse the effect sizes (when available) quantitatively.

Regarding the interpretation of those three parts, I suggest that there is an inverse relationship between granularity and warranted confidence. The broader the classification of the effect, the more confidence we should put in the findings. This implies some level of trust in the author's judgement of the general direction of the effect. It also avoids problematic comparisons of seemingly exact numerical estimates which look at different influences, rely on different specifications and control for different variables.

iv. Statistical analysis and weights

Even though in my estimation, the categorical assessments of the direction of effects are the most reliable judgements about the relationships between shocks and risk preferences, there is still value in analysing the quantitative effect sizes. Even though we should likely not rely too much on comparisons between different effect sizes, looking at them on an aggregate level and also considering the respective standard errors may tell us more about the distribution of effects and about relationships between study characteristics and effects than we could learn on the basis of the categorical classifications.

For the statistical analysis, I will mostly rely on Card, 2016, "*Applied Meta-Analysis for Social Sciences*" and Harrer et al., 2021, "*Doing Meta-Analysis in R: A Hands-on Guide*". For the analysis in R, I mainly rely on the packages *metafor* (Viechtbauer, 2010), tidyverse (Wickham et al. 2019), ggplot2 (Wickham, 2011), dplyr (Yarberry, 2021) as well as *meta* (Schwarzer, 2007), *psych* (Revelle, 2015) and *forestplot* (Gordon et al., 2019).

A crucial issue of any meta-analysis arises from the fact that relevant primary studies usually vary significantly in ways that have an impact on their informational value. This is where weights enter the picture. Treating every study as one simple datapoint would implicitly assign the same weight to studies that may differ substantially regarding the precision of their estimations of the true effects. As (Card, 2016) notes, there are many ("*virtually limitless*") ways of introducing differential weighting. Criteria may include sample size, p-values, and even more qualitative criteria such as methodology, robustness of the effects or journal quality.

In Card's view, however, the only "*statistically defensible*" option is weighting by standard errors as this increases confidence that the results will best reflect the true effects of the population. To be more precise, he suggests using the following formula for determining w_i (the weight of a study i) based on SE_i (standard error of study i):

$$w_i = \frac{1}{SE_i^2}$$

This implies an inverse relationship between the weight of a particular study and the variance of its point estimate, i.e. the lower the variance and thus the standard error, the greater the weight and vice versa. The weights that I calculated this way could be used for calculating the average effect sizes. For the analysis of the qualitative categorisations, however, this approach was not sufficient. As mentioned above, some studies in the sample unfortunately did not report the relevant standard errors or did not have results that could be reduced to one single (average) effect size which also meant that there was no single representative standard error that could be used for calculating the respective weight. Nevertheless, it does seem warranted to include some information about the sample size even in this kind of analysis. I will focus on means as the main index as it allows to incorporate weights. In addition, in some instances, I will report the medians which should, however, be understood as an additional but not the main index (Card, 2016). The weighted effect sizes are calculated by multiplying the effect size with the according weight of

$$w_i ES_i = w_i * ES_i$$

with ES_i being the effect size for study i. On this basis, we can calculate the overall average effect size for the sample of studies using

$$\overline{ES} = \frac{\sum(w_i ES_i)}{\sum w_i}$$

This simple approach will form the basis of my analysis. In a few cases where a publication reports the results for different sub-groups or different measures, it will be supplemented by a simple additional factor which prevents the samples in those studies from being counted more than once.[5] Practically, however, the *metafor* package in R (Viechtbauer, 2010) does these calculations and several other tests by default.

Additionally, a major choice that has to be made in any meta-analysis is the one between fixed and random effects model. The studies in the sample vary in many regards, including but not limited to measurement type, shock type, time between shock and measurement, location, and sample composition. We should therefore expect at least some heterogeneity between studies which implies that we should not assume that there is a true fixed effect for all of them. Consequently, for pooling effect sizes, the random-effects model is the adequate choice.

[5] The cases in question are:

- Adema et al. (2022): *factor 0.5 for each entry (two different measurements for the same sample)*
- Bchir and Willinger (2013): *factor 0.67 for overall sample and factor 0.33 for low-income sample*
- Guiso et al. (2018): *factor 0.5 for each entry (one qualitative and one quantitative risk measure)*
- Hanaoka et al. (2018): *factor 0.67 for overall sample and factor 0.33 for male sub-sample*
- Ingwersen et al. (2023): *factor 0.5 for each entry (two different types of exposure for the same sample)*
- Necker and Ziegelmeyer (2016): *factor 0.5 for each sample (different results for different sub-samples)*
- Thamarapani and Rockmore (2022): *factor 0.5 for each sample (different results for sub-samples)*
- Zhang and Palma (2021): *factor 0.25 for measures that find effects (3 out of 12), 0.75 for the rest*

III. Results

In this chapter, I will describe the findings of my research in some detail. In the first part, I will provide an overview of the literature in the final sample. I will do so while balancing comprehensiveness and focus on importance and relevance. This means that I will not go into every possible observation regarding the sample of studies but will focus more on the aspects that could be sorted into a relatively limited set of categories. This way, I hope to provide a useful quantitative overview of the literature on shocks and risk preferences. I will *not* be able consider in much detail the more nuanced differences on many particular aspects. In the second part I will focus on the analysis of effect sizes including subgroup analyses. Finally, I will comment on the issue of publication bias.

Table 2: List of articles in the final sample

Article	*Status*	*Effect Direction*	*Sources of within-study-heterogeneity*	*Effect Size*	*Location*	*N*	*Sampling lag*	*Shock Type*
Abatayo & Lynham, 2020	Published	Increase	gender	0.219	Philippines	100	1-3 years	Natural disaster
Adema et al., 2022	Published	Decrease	measurement	- 0.148	India, Mexico, EU	303	1 year	Pandemic
Adema et al., 2022	Published	Increase	measurement	0.144	India, Mexico, EU	303	1 year	Pandemic
Ahsan, 2014	Published	Decrease		- 0.337	Bangladesh	250	3 years	Natural disaster
Angrisani et al., 2020	Working Paper	No effect		0.112	USA	108	< 3 months	Pandemic
Bchir & Willinger, 2013	Working Paper	Increase	income	0.314	Peru	162	< 1 year	Natural disaster
Bchir & Willinger, 2013	Working Paper	No effect	income	0.056	Peru	309	< 1 year	Natural disaster
Beine et al., 2020	Working Paper	Decrease	exposure intensity	- 0.179	Albania	1502	< 3 months	Natural disaster
Bernile et al., 2017	Published	Increase		0.082	USA	1508	> 10 years	Natural disaster
Bourdeau-Brien et al. 2020	Published	Decrease		-0.058	USA	7750	< 3 months	Natural disaster
P. Brown et al., 2018	Published	Decrease	ethnicity	- 0.158	Fiji	295	< 3 months	Natural disaster
R. Brown et al., 2019	Published	Decrease		- 0.017	Mexico	35000	< 3 months	Conflict & Violence
Callen, 2015	Published	Decrease		- 0.125	Afghanistan	977	< 3 months	Conflict & Violence
Cameron & Shah, 2015	Published	Decrease		- 0.059	Indonesia	1550	3 years	Natural disaster
Cassar et al., 2017	Published	Decrease		- 1.943	Thailand	334	3 years	Natural disaster
Chantarat et al., 2019	Published	Decrease		- 0.143	Cambodia	256	3 years	Natural disaster
Cicerale et al., 2022	Published	Decrease		- 0.377	Italy	350	< 3 months	Pandemic
Cohn et al., 2015	Published	Decrease	certainty	- 0.199	Switzerland	162	< 3 months	Economic
de Blasio et al., 2021	Published	Decrease		- 0.202	Italy	8000	2 years	Natural disaster
Di Falco & Vieider, 2022	Published	Decrease		- 0.156	Ethiopia	906	< 1 year	Natural disaster
Dohmen et al., 2016	Published	Decrease			Ukraine & Germ.	26056	1-3 years	Economic
Eckel et al., 2009	Published	Mixed	gend.		USA	352	1 year	Natural disaster
Enrique Fatas et al., 2021	Published	Increase	exposure intensity		Colombia	207	> 10 years	Conflict & Violence
Finger et al., 2023	Published	Increase	exp. type		Switzerland	1530	< 3 months	Natural disaster
Frondel et al., 2021	Working Paper	Decrease	exp. intensity	- 0.244	Germany	5500	< 1 year	Pandemic
Gassmann et al., 2022	Published	Increase		0.310	France	406	< 3 months	Pandemic
Gerrans et al., 2015	Published	Decrease		- 0.275	A, NZ, NA, UK	3368	1-3 years	Economic
Graeber et al., 2020	Working Paper	Decrease	gend., income	- 0.060	Germany	6393	< 3 months	Pandemic
Guiso et al., 2018	Published	Decrease	gend., age, inc., edu.	- 0.356	Italy	666	1-3 years	Economic
Guiso et al. 2018	Published	Decrease	age, education	- 0.518	Italy	666	1-3 years	Economic
Hanaoka et al., 2018	Published	No effect	gend., time		Japan	3352	1-5 years	Natural disaster
Hanaoka et al. 2018	Published	Increase	gend., time	0.076	Japan	1575	1-5 years	Natural disaster
Holden & Tilahun, 2021	Working Paper	Mixed	measurement		Ethiopia	830	1-3 years	Natural disaster
Ikeda et al., 2020	Working Paper	Mixed	stakes		Japan	3495	< 1 year	Pandemic
Ingwersen et al., 2023	Working Paper	Increase	exp. type & intensity		Indonesia	9860	< 1 year	Natural disaster
Ingwersen et al. 2023	Working Paper	Increase	exp. type & intensity		Indonesia	9860	5 years	Natural disaster
Jakiela & Ozier, 2019	Published	Decrease		- 0.118	Kenya	5047	1-3 years	Conflict & Violence
Jetter et al., 2020	Published	No effect	gender		Australia	22579	5 years	Economic
Kahsay & Osberghaus, 2018	Published	Increase	exp. intensity	0.028	Germany	6431	1-3 years	Natural disaster
Kettlewell, 2019	Published	Decrease	exp. type, time		Australia	4810	< 1 year	Economic
Kettlewell et al., 2023	Published	Increase		0.310	Sri Lanka	2946	1-3 years	Natural disaster
Kim & Lee, 2014	Published	Decrease	exp. intensity	- 0.039	Korea	7047	> 10 years	Conflict & Violence
Kuroishi & Sawada, 2019	Working Paper	Increase			Japan & Philipp.	344	6 years	Natural disaster
J.-Z. Li et al., 2011	Published	Increase		0.219	China	1072	< 3 months	Natural disaster
K. K. Li et al., 2020	Working Paper	Decrease	age	- 0.154	China	1040	< 3 months	Pandemic
Lohmann et al., 2020	Working Paper	No effect			China	539	< 3 months	Pandemic
Malmendier & Nagel, 2011	Published	Decrease			USA	28571	> 10 years	Economic
Meunier & Ohadi, 2021	Published	No effect	domain		USA, UK, A, EU	72	< 3 months	Pandemic
Moya, 2018	Published	Decrease	time, domain, risk	- 0.278	Colombia	284	1-5 years	Conflict & Violence
Necker & Ziegelmeyer, 2016	Published	Decrease	exp. type	- 0.048	Germany	2047	1-3 years	Economic
Necker & Ziegelmeyer 2016	Published	No effect	exp. type		Germany	2047	1-3 years	Economic
Page et al., 2012	Published	Increase		0.281	Australia	202	< 3 months	Natural disaster
Reynaud & Aubert, 2020	Published	Decrease	domain	- 0.163	Vietnam	448	1-5 years	Natural disaster
Rockmore & Barrett, 2022	Published	No effect	exposure type		Uganda	442	> 10 years	Conflict & Violence
Said et al., 2015	Published	Mixed	time, exp. intensity		Pakistan	384	1-3 years	Natural disaster
Shachat et al., 2021	Published	Mixed	domain		China	602	< 3 months	Pandemic
Shigeoka, 2019	Working Paper	Decrease		- 0.063	Japan	4165	> 10 years	Economic
Shupp et al., 2017	Published	Mixed	exp. type		USA	295	< 3 months	Natural disaster
Thamarapani & Rockmore, 2022	Published	Increase		0.090	Indonesia	2966	< 1 year	Natural disaster
Thamarapani et al. 2022	Published	No effect			Indonesia	2966	1-3 years	Natural disaster
Tsutsui & Tsutsui-Kimura, 2022	Published	Increase	time	0.278	Japan	3495	< 1 year	Pandemic
van den Berg et al., 2009	Working Paper	Decrease		- 0.188	Nicaragua & Peru	84	1-3 years	Natural disaster
Voors et al., 2012	Published	Increase	domain	0.056	Burundi	220	> 10 years	Conflict & Violence
Willinger et al., 2013	Working Paper	No effect			Indonesia	131	< 1 year	Natural disaster
Zhang & Palma, 2021	Published	Decrease	gend., measurement	- 0.180	USA	331	< 1 year	Pandemic
Zhang & Palma 2022	Published	No effect	gend., measurement		USA	331	< 1 year	Pandemic

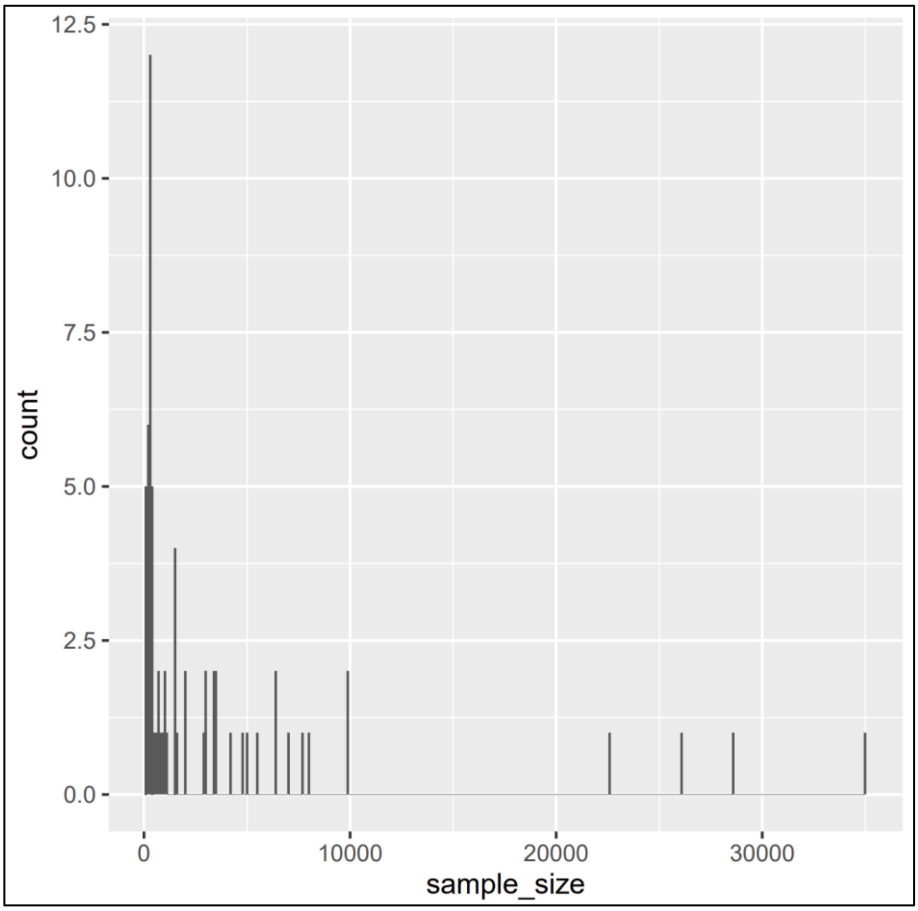
i. Sample and descriptive statistics

Applying the search and selection procedures mentioned in chapter *II.i* delivered a sample of 58 articles (*Table 2*).

Eight of them contained sufficiently distinct results within the same articles that they were entered as two datapoints with an according factor to prevent the respective samples from disproportionately influencing the results. As a subset of articles produced two sufficiently distinct results, I will from here on treat them as separate studies, which increases the number of studies in the sample to 66.

Of the 66 studies, 32 investigate the effects of natural disasters (earthquakes, floods, natural hazards etc.), 15 investigate effects of the COVID-19 pandemic, 11 investigate economic shocks (household finance shocks, 2007/2008 global financial crisis) and 8 look at the effects of conflict and violence (*Table 8*, *Appendix*). 32 studies measure short run effects (i.e. risk preferences less than one year after the respective shock), 23 studies measure medium run effects (one-to-five-year lag), and 11 studies measure long run effects of more than five or in seven cases even more than ten years (*Figure 1b, Appendix; Table 9 and 10, Appendix*). Sample locations are distributed across 5 continents with the most studies being conducted in the United States (10), Indonesia (6), Germany (6) and Japan (5). The respective frequency table (*Table 12*) can be found in the *Appendix*.

Figure 1: Histogram for sample sizes (bucket size =100)

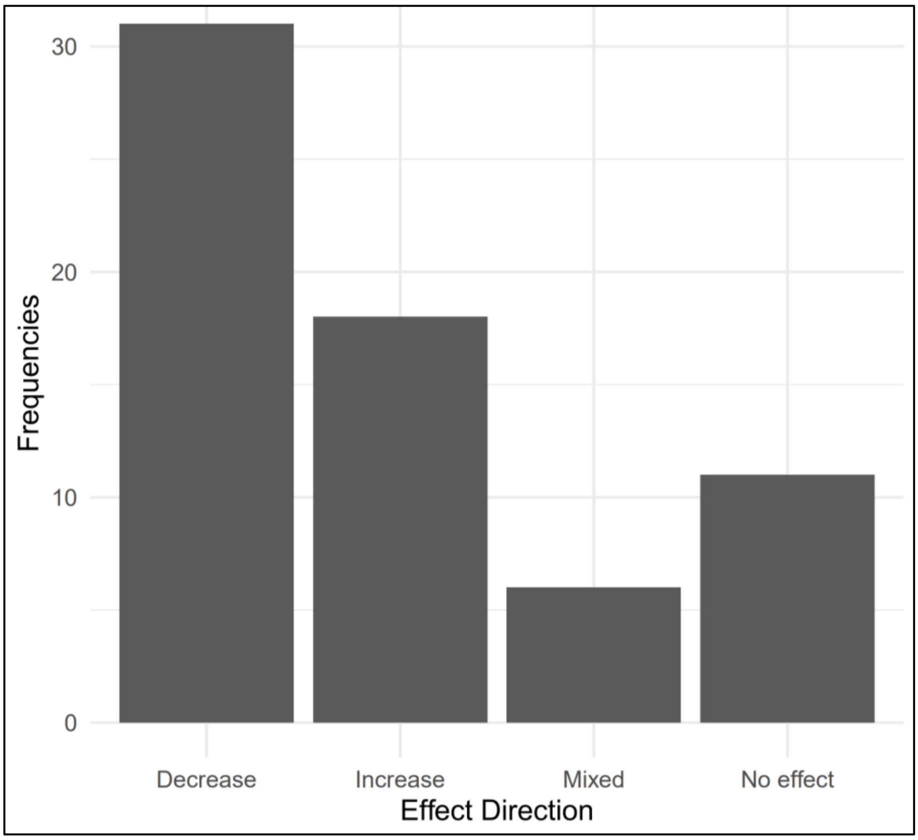


The sample sizes vary from 72 to 35 000 (mean = 3 578, SD = 6331.4). Most studies feature sample sizes below 1000 participants with the Median being substantially lower than the mean at just 868 (*Table 13, Appendix; Figure 1 & d, Appendix*).

One major goal of this thesis was to give an overview of the effect directions that were found in the literature. As mentioned in *II.iii*, of all the results in the dataset, we can probably trust the overall directions found in the individual papers the most. As so many articles mention, there is significant variation in the research as to whether shocks increase or decrease risk seeking. My literature search and selection now allow us to finally quantify this impression. A majority of 31 studies find a decrease in risk seeking while 18 find an increase, 6 find mixed effects and 11 find no effects (*Table 4, Appendix; Figure 2*). This means that while there certainly is substantial between-study-heterogeneity (i.e. the impression that estimated effects point in very different directions) it is also true that there is a pronounced tendency towards a decrease in risk seeking in response to shocks.

Additionally, a substantial share of the publications reports heterogeneous results (*Table 2, column 4*) with a wide range of sources for heterogeneity including gender, income, education, exposure type and intensity, timing, domain (gain vs loss) and measurement type.

Figure 2: Frequencies of effect directions



This *within*-study heterogeneity may even be severely under-estimated as most studies employ aggregate measures of their independent variables meaning that they likely collapse a substantial variety of ways shocks are experienced into one measurement (Rockmore & Barrett, 2022). There is good reason for doing so in many cases (practicality chiefly among them), but it should nevertheless be noted.

ii. Meta-analysis of effect sizes

Keeping in mind the caveats outlined in *II.iii*, I will now move on to the meta-analysis of effect sizes. While 85% of the studies in the sample record significant effects of some kind (*Table 7, Appendix*), a majority of 58% (38 studies) shows only very small (< 0.1) or small (0.1 – 0.3) effects . Larger effects become increasingly rare, with five studies finding small-to medium (0.3 – 0.5) effects and medium (> 0.5, one study) and large (> 0.8, one study)[6] effects (*Table 7, Appendix*).

Figure 3 shows the forest plot for the studies for which quantitative effect sizes were available.[7] The forest plot shows effect sizes and confidence intervals for the studies in the sample. The last line indicates the simple average effect size of –0.06, graphically represented by the light blue diamond and the dotted line, and the corresponding confidence interval ranging from –0.11 to 0. Leaving aside Cassar et al. (2017) with an effect size of –1.94, there all effects remain in a relatively modest range from –0.52 (Guiso et al., 2018) to 0.31 (Bchir & Willinger, 2013).

While *Figure 3* shows the raw standardised effect sizes for all of the studies, the meta-analysis uses the adjusted effect sizes for the studies that have been entered twice. Consequently, the results differ slightly. *Output 1* shows the results of the random-effects meta-analysis for all 48 studies for which effect sizes where available. The analysis delivers an average effect, measured as standardised mean difference (SMD), of –0.0887 (CI: –0.1805 – 0.0031). This effect is slightly significant on the 10% significance level ($t =$ –1.94, $p =$ 0.0578).

In addition to within-study-heterogeneity, there is substantial heterogeneity regarding research methodology, design, location, sample size and sample composition. This is usually referred to as *between*-study-heterogeneity. I mentioned this earlier while explaining why a random effects model should be employed instead of a fixed effects model because it could not be assumed that there is only one true effect.

The results appear to vindicate the previously stated assumption that the heterogeneity regarding research designs implies that there may not be a single common true effect. Following Harrer et al., 2021, I employ a restricted maximum likelihood estimator to estimate the between-study heterogeneity variance of $\hat{\tau}^2 =$ 0.0968 (95%, CI: 0.0652 – 0.1515) and $I^2 =$ 98% (95%, CI: 97.7% - 98.2%).

The prediction interval ranged from $g =$ –0.7218 to $g =$ 0.5443. The test for heterogeneity is highly significant ($p <$ 0.0001). This indicates *substantial or even very high heterogeneity* (classification according to Higgins & Thompson, 2002) and thus a generally high likelihood of future studies finding effects in both the positive (more risk seeking) and the negative (less risk seeking) direction.

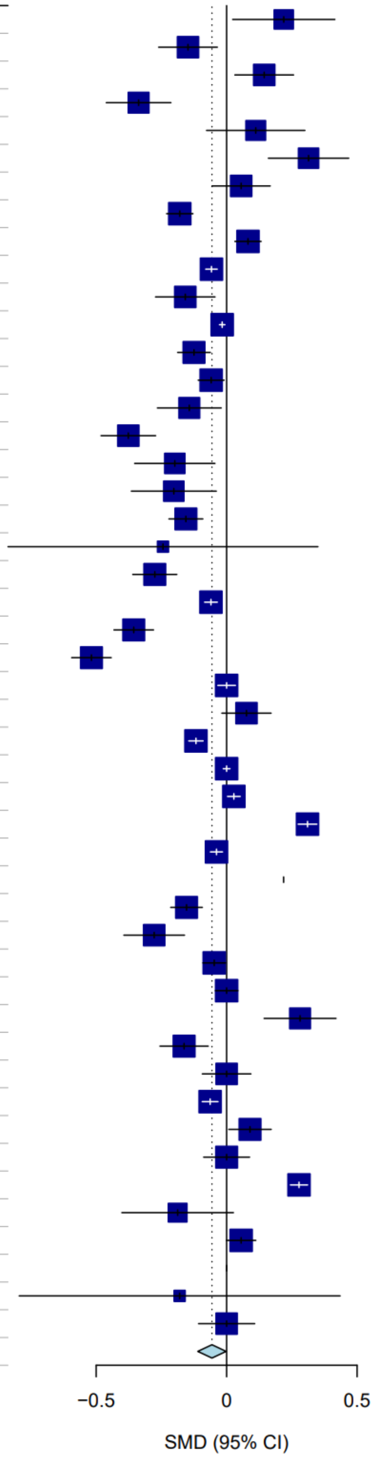
As a first robustness check, *Output 2* estimates the same model but without Cassar et al., 2017. The results are rather similar to the original model. Despite the omission of a very large negative effect and the consequently slightly lower SMD of now –0.0471, the p-value ($t =$ –2.08, $p =$ 0.0427) in this model becomes even smaller, making it significant on the 5% significance level. Unsurprisingly, heterogeneity decreases markedly but remains *substantial* and highly significant ($p <$ 0.0001).

[6] The only study with an effect classified as "large" (Malmendier and Nagel (2011) does not report one easily interpretable effect size but the authors consider their effects as large.

[7] In order to fit the forest plot in one page, I omitted the studies which did not provide quantitative effect sizes. The forest plot which includes all studies can be found in the *Appendix (Figure 2a&b)*. I also omitted Cassar (2017) because its exceptionally large effect size (of which I am doubtful anyways) would distort the image too much. The forest plot which includes Cassar (2017) can be found in the *Appendix (Figure 2a)*.

Figure 3: Forest plot for studies that find effects (no Cassar et al., 2017)

Source	SMD (95% CI)
Abatayo & Lynham 2020	0.22 [0.02; 0.41]
Adema et al. 2022	−0.15 [−0.26; −0.04]
Adema et al. 2022	0.14 [0.03; 0.26]
Ahsan 2014	−0.34 [−0.46; −0.21]
Angrisani et al. 2016	0.11 [−0.08; 0.30]
Bchir & Willinger 2013	0.31 [0.16; 0.47]
Bchir & Willinger 2013	0.06 [−0.06; 0.17]
Beine et al. 2020	−0.18 [−0.23; −0.13]
Bernile et al. 2017	0.08 [0.03; 0.13]
Bourdeau−Brien & Kryzanowski 2020	−0.06 [−0.08; −0.04]
Brown et al. 2018	−0.16 [−0.27; −0.04]
Brown et al. 2019	−0.02 [−0.03; −0.01]
Callen et al. 2014	−0.13 [−0.19; −0.06]
Cameron & Shah 2015	−0.06 [−0.11; −0.01]
Chantarat et al. 2019	−0.14 [−0.27; −0.02]
Cicerale et al. 2021	−0.38 [−0.48; −0.27]
Cohn et al. 2015	−0.20 [−0.35; −0.04]
de Blasio et al. 2020	−0.20 [−0.37; −0.04]
Di Falco & Vieider 2022	−0.16 [−0.22; −0.09]
Frondel et al. 2021	−0.24 [−0.84; 0.35]
Gerrans et al. 2015	−0.28 [−0.36; −0.19]
Graeber et al 2020	−0.06 [−0.08; −0.04]
Guiso et al. 2018	−0.36 [−0.43; −0.28]
Guiso et al. 2018	−0.52 [−0.59; −0.44]
Hanaoka et al. 2018	0.00 [−0.03; 0.03]
Hanaoka et al. 2018	0.08 [−0.02; 0.17]
Jakiela & Ozier 2019	−0.12 [−0.15; −0.09]
Jetter et al. 2020	0.00 [−0.01; 0.01]
Kahsay & Osberghaus 2018	0.03 [0.00; 0.05]
Kettlewell et al. 2023	0.31 [0.27; 0.35]
Kim & Lee 2014	−0.04 [−0.06; −0.02]
Li et al. 2011	0.22
Li et al. 2020	−0.15 [−0.21; −0.09]
Moya 2018	−0.28 [−0.39; −0.16]
Necker & Ziegelmeyer 2015	−0.05 [−0.09; −0.00]
Necker & Ziegelmeyer 2015	0.00 [−0.04; 0.04]
Page et al. 2012	0.28 [0.14; 0.42]
Reynaud & Aubert 2020	−0.16 [−0.26; −0.07]
Rockmore & Barrett 2022	0.00 [−0.09; 0.09]
Shigeoka 2019	−0.06 [−0.09; −0.03]
Thamarapani & Rockmore 2022	0.09 [0.01; 0.17]
Thamarapani & Rockmore 2022	0.00 [−0.09; 0.09]
Tsutsui & Tsutsui−Kimura 2022	0.28 [0.24; 0.31]
van den Berg et al. 2009	−0.19 [−0.40; 0.03]
Voors et al. 2012	0.06 [−0.00; 0.11]
Willinger et al. 2013	0.00
Zhang & Palma 2022	−0.18 [−0.80; 0.44]
Zhang & Palma 2022	0.00 [−0.11; 0.11]
Total	−0.06 [−0.11; −0.00]



Heterogeneity: χ^2_{45} = 1313.02 (P < .001), I^2 = 97%

Output 1: Random effects meta-analysis (all studies)

```
Number of studies combined: k = 48

                              SMD             95%-CI      t p-value
Random effects model (HK) -0.0887 [-0.1805; 0.0031] -1.94  0.0578

Quantifying heterogeneity:
 tau^2 = 0.0968 [0.0652; 0.1515]; tau = 0.3111 [0.2553; 0.3893]
 I^2 = 98.0% [97.7%; 98.2%]; H = 7.05 [6.60; 7.52]

Test of heterogeneity:
       Q d.f. p-value
 2333.93   47       0
```

Output 2: Random effects meta-analysis (no Cassar et al., 2017)

```
Number of studies combined: k = 47

                              SMD              95%-CI      t p-value
Random effects model (HK) -0.0471 [-0.0926; -0.0016] -2.08  0.0427

Quantifying heterogeneity:
 tau^2 = 0.0211 [0.0136; 0.0347]; tau = 0.1453 [0.1165; 0.1864]
 I^2 = 95.8% [95.0%; 96.4%]; H = 4.88 [4.49; 5.30]

Test of heterogeneity:
       Q d.f.  p-value
 1095.70   46 < 0.0001
```

iii. Subgroup analyses & Meta-regressions

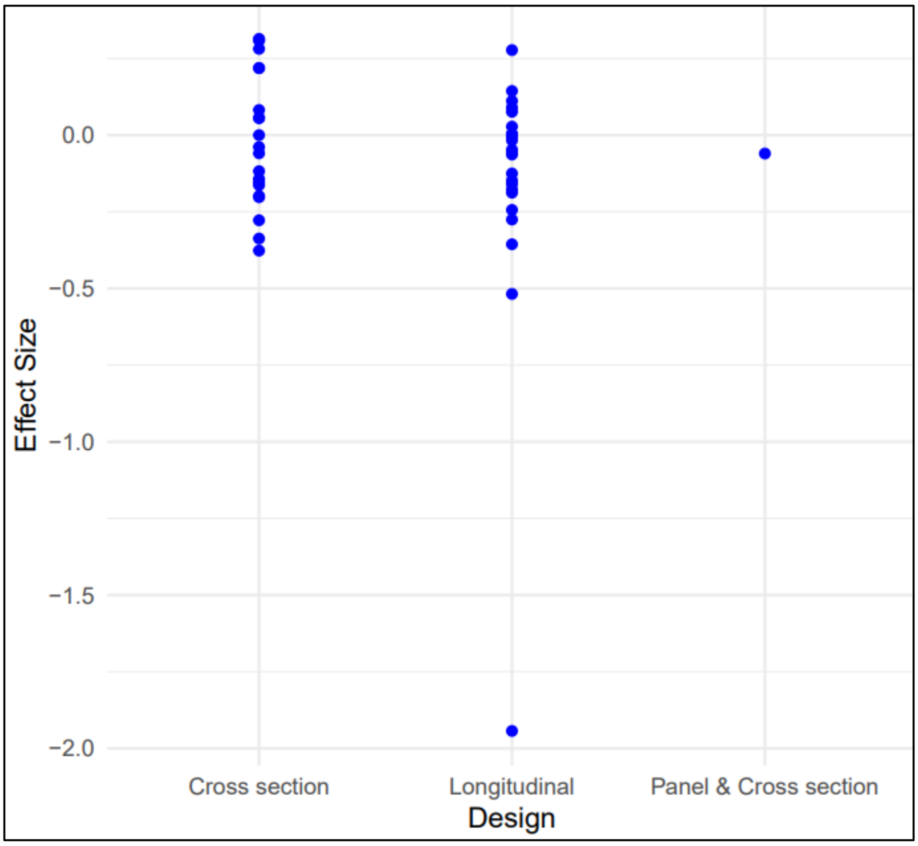
Beyond this general quantitative analysis of heterogeneity, analysing between-study-heterogeneity may yield additional informational value as there may be relationships between the some of the study characteristics and the study results. In the following section, I will therefore report the main findings of a number of subgroup analyses and meta-regressions. For that, it is important, however, to note that none of the findings should be interpreted as causal. Furthermore, as the sample includes only 48 studies with effect sizes, any sub-group analysis will necessarily be relatively low-powered. This is especially true for cases with very unequal sample sizes. Nevertheless, I believe that it is worth performing these analyses and to create the corresponding graphs in order to paint a more detailed picture of where the literature stands at this moment in time. I will begin with subgroup analyses for the whole sample (*III.iii.a-b*). After that, I will go into more detail on the effects of the different types of shocks (*III.iii.c-f*).

a. Study design and sample characteristics

Research on the effects of shocks on risk preferences has to grapple with the fundamental problem that shocks cannot be planned or foreseen. Given that fact, it is somewhat surprising to see that the majority of studies (41) on the topic have employed longitudinal designs which allow for comparisons of risk preferences before and after the respective shock. 25 studies employed cross-sectional designs which measure risk preferences only after the shock *(Table 11, Appendix)*. In those cases, the comparison is usually between participants with more vs less exposure to the shock. Cross-sectional designs may be at a higher risk of falling victim to selection effects which they may be unable to control for. However, longitudinal studies may suffer from attrition effects which they also may not be able to sufficiently control for as they might lack the relevant data due to difficulties of adequately planning pre-post comparisons of inherently unforeseeable events. Comparison of the effects estimated with longitudinal vs cross-sectional approaches shows no significant difference ($p = 0.623$, *Statistical output 2f, Appendix; Figure 3m, Appendix, Figure 3h*).

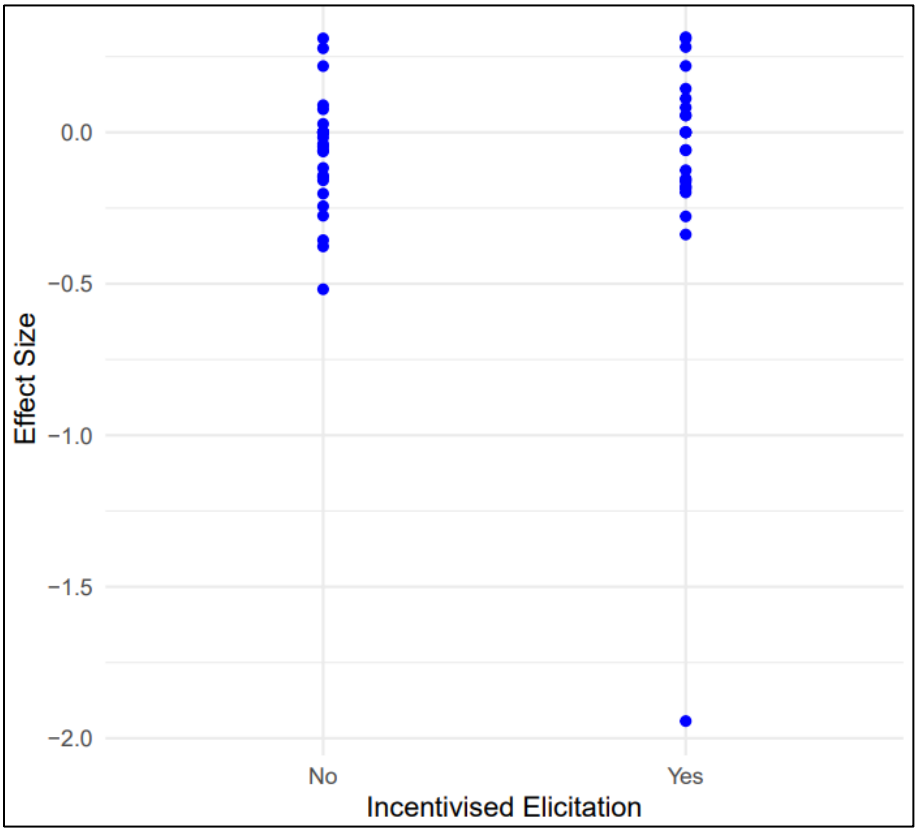
The results nicely illustrate the difference aggregation can make when compared to individual studies. Adema et al., 2022, find that as a reaction to the same shock (in this case COVID-19 exposure), self-reported risk seeking declines (small effect) while participants are more risk seeking in incentivised lotteries (small effect). On aggregate, however, incentivisation does not have a significant effect on effect direction. Both incentivised (–0.1227) and non-incentivised (–0.0512) tend to measure on average small reductions in risk seeking that do not differ significantly from each other ($Q = 0.57$, $df = 1$, $p = 0.4521$; *Output 2h, Appendix; Output 3n, Appendix*). Despite negative averages, however, both groups are very heterogeneous with plenty of studies measuring positive effects on risk seeking (*Figure 4*). Results are similar for comparisons between experimental elicitations, which often but not always are incentivised, and surveys, which are often but not always non-incentivised (*Output 3d, Appendix*).

Figure 4: Cross-sectional vs longitudinal design



Regarding sample composition, there is not much to report, as reporting of sample characteristics in the studies was relatively unreliable. The dataset therefore includes only very few variables that lend themselves as plausible covariates to test the influence of sample characteristics on effect sizes. One of them is the mean age of the respective participants. However, when conducting subgroup analyses or meta-regression, there is severe danger of ecological bias (Harrer et al., 2021), i.e. the unwarranted inference from the aggregate to the individual or from the macro to the micro. This is especially problematic in cases where we do not know the underlying distribution of a factor (e.g. age) in the different primary studies and therefore also do not know what is driving the respective effects. (Harrer et al., 2021) therefore caution against subgroup analyses on the basis of aggregate information (*"never, ever use aggregate information in subgroup analyses and meta regressions"*). Nevertheless, I did produce a graph showing that there appears to be no systematic relationship between mean sample age and effect sizes (*Figure 3a, Appendix*).

Figure 5: Incentivised vs non-incentivised risk preference elicitation



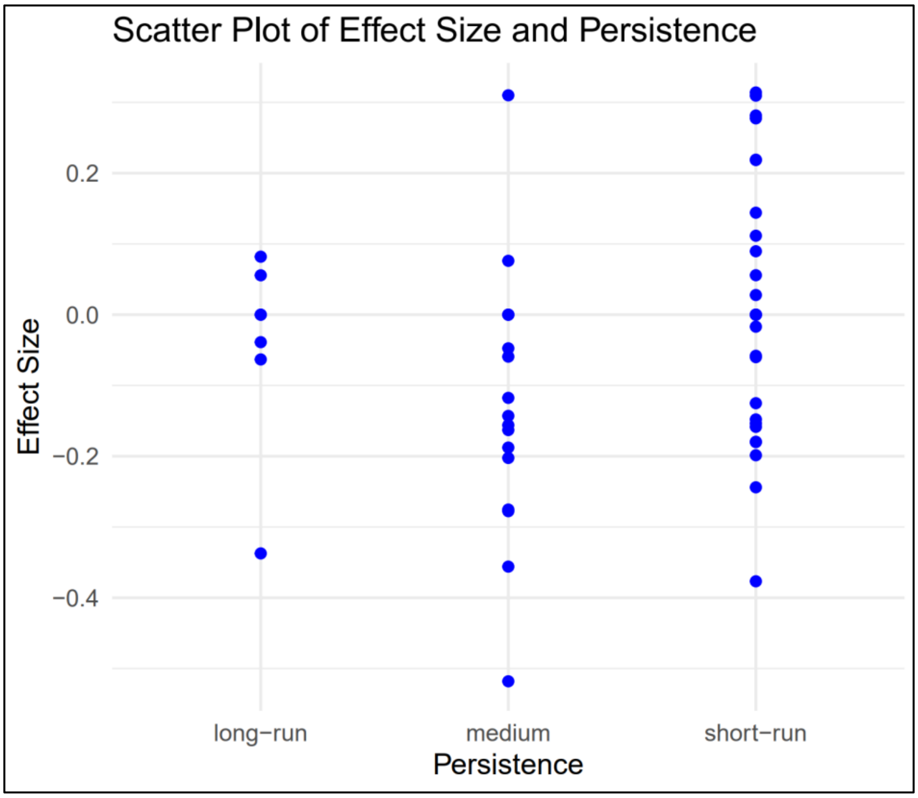
Another characteristic of interest could be the location of a sample. In order to increase the subgroup sample sizes, I classified the sample locations in an ad-hoc way according to their WEIRDness (WEIRD being Western, Educated, Industrialised, Rich and Democratic, Henrich et al., 2010). I find that WEIRD samples tend to exhibit slightly more risk aversion in response to shocks, but the difference is far from significant ($Q = 1.25$, $df = 2$, $p = 0.4844$, *Output 2l, Appendix, Figure 5 & 4e, Appendix*).

Output 3: Medium-run effects

```
Number of studies combined: k = 17

                              SMD             95%-CI      t p-value
Random effects model (HK) -0.1212 [-0.2180; -0.0243] -2.65  0.0174

Quantifying heterogeneity:
 tau^2 = 0.0342 [0.0179; 0.0788]; tau = 0.1850 [0.1338; 0.2807]
 I^2 = 97.7% [97.1%; 98.2%]; H = 6.61 [5.88; 7.42]

Test of heterogeneity:
      Q d.f.  p-value
 698.93   16 < 0.0001
```

Similarly, the lag between the shock and the risk preference elicitation, i.e. whether the study measures short-, medium-, or long-run effects, does not appear to have any systematic relationship with effect sizes at least when analysed ($Q = 2.72$, $df = 3$, $p = 0.2564$, *Output 2c, Appendix*). Also, heterogeneity is very large for all groups.

Figure 6: Persistence and Effect Size

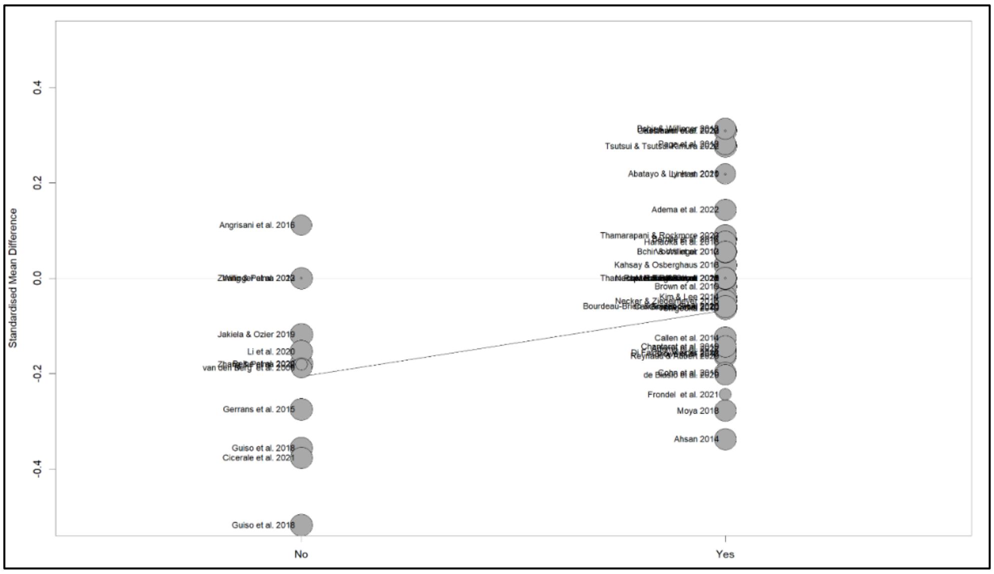


Looking at the groups separately, however, I find that medium-run effects differ significantly from zero ($SMD = –0.1212$, $t = –2.65$, $p = 0.0174$, *Output 3, Figure 6*) while the effects for both short-run ($SMD = –0.0079$) and long-run effects ($SMD = –0.0357$) are both much smaller and not significantly different from zero (*Output 2i & 2k, Appendix)*. This is a notable result as it implies that the overall effect reported in *III.ii* appears to be driven substantially by the medium-run effects (1-5 years after the shock).

b. Article quality and author characteristics

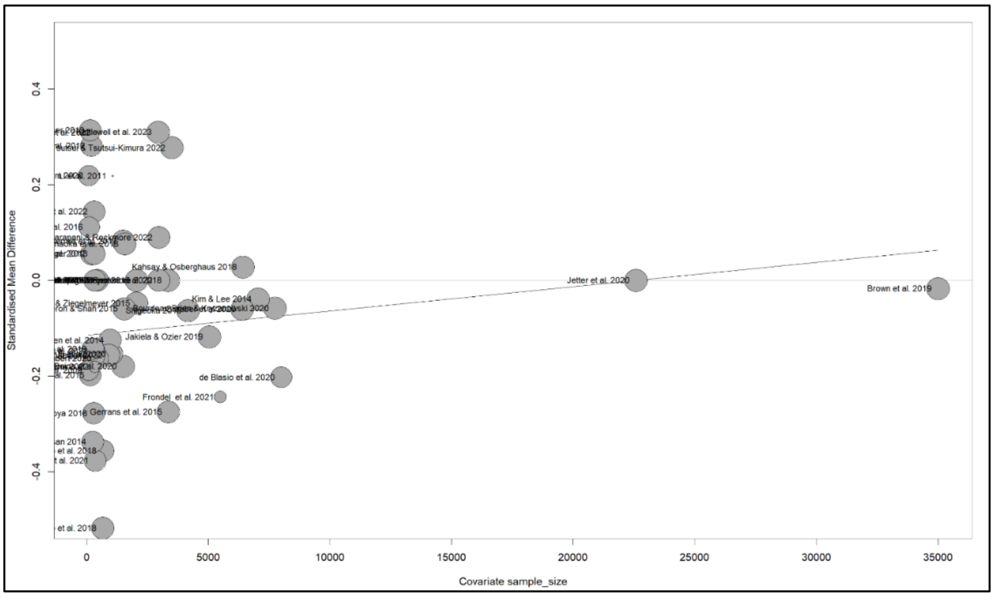
When weighting the effect sizes for the general meta-analysis I relied only on standard errors as a measure of precision. However, precision, i.e. the likelihood of measuring the true effect may also be influenced by the quality of a study. Within the scope of this thesis, I was not able to construct elaborate measures of study quality as doing so in a reliable and valid way would have required more subject matter expertise and time. As a second-best approach, I collected information on a range of indicators which on their own may be very flawed proxies of quality but taken together could serve as a useful approximation of study quality.

Figure 7: Causal Design and Effect Sizes



For instance, not all studies in the sample use a quasi-experimental approach. Of the 48 studies, 37 use some form of empirical strategy which allows for causal claims regarding the relationship between shocks and risk preferences (*Output 2g, Appendix; Figure 7*) . The mean effects differ slightly between the groups with a smaller effect for causal approaches (–0.0695) when compared to non-causal approaches (–0.1684). The effect is not significant, however ($Q = 2.03$, $df = 1$, $p = 0.1538$).

Figure 8: Sample size and effect size

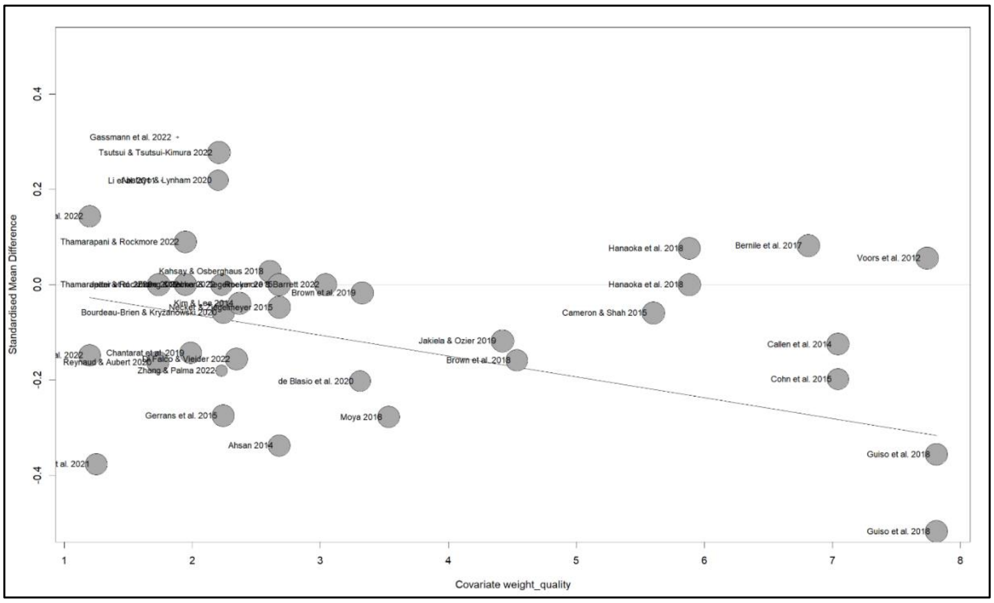


Another marker for study quality is sample size. Due to the standard errors in the dataset being standardised according to standard deviations, using sample size as a marker for quality is not entirely redundant. However, estimating a meta-regression shows that it has zero effect on effect sizes (*Output 3b, Appendix; Figure 8 & 3h, Appendix; Figure 4a & 4b, Appendix*).

I also consider article citations and journal impact factor as possible correlates of effect size and direction. On their own, they are non-significant and negligible (*Output 3c and 3d, Appendix*). Additionally, I constructed a simple ad-hoc index for publication quality based on citations per year, journal impact factor[8] and regressed it on the effect sizes. Effects tended slightly towards less risk seeking with higher journal quality (coefficient = –0.0351) but the effect was not significant ($df = 35$, $p = 0.0195$, *Output 3d, Appendix; Figure 9*)

The last variables I considered have little to do with study quality but might still be of interest. Firstly, the overwhelming majority of articles with quantitative effect sizes were written by authors from the field of economics (40 studies) with only 3 studies by psychologists and 5 studies by interdisciplinary teams.

Figure 9: Quality index and effect sizes



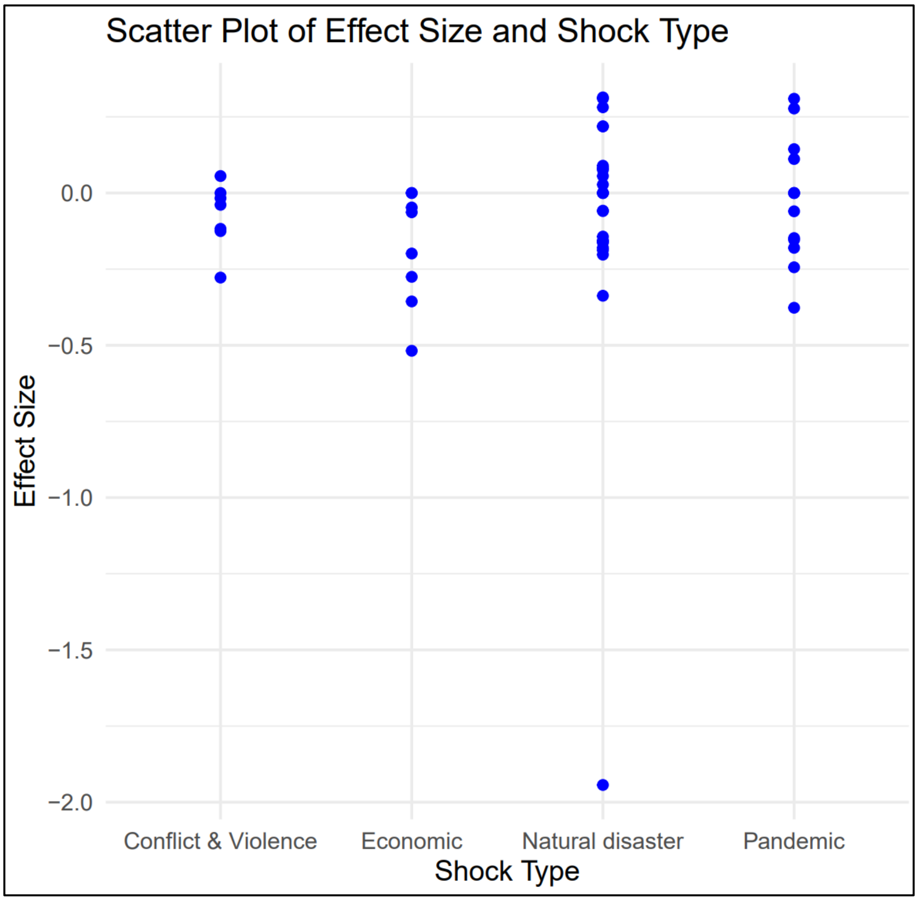
While means differ markedly between studies by economists (–0.1051) and studies by psychologists (0.0792), the comparison is under-powered and thus unsurprisingly non-significant ($Q = 3.89$, $df = 2$, $p = 0.1430$, *Output 2b, Appendix*). Secondly, while the year of publication is not correlated with effect sizes, *Figure 3i*, (*Appendix*) shows nicely how many studies have been published in the 5 years since the publication of the review paper by Schildberg-Hörisch, 2018. Thirdly, there is no relationship between the author's gender composition and effect sizes. However, as *Figure 3n* (*Appendix*) shows very clearly, the field appears to be heavily skewed towards male authors with majority female teams of authors being the rare exception.

[8] $quality\ index = (impact\ factor + \sqrt{citations/year}) / 2$

c. Shock types

The last major subgroup analysis focuses on the different types of shocks that have been investigated in the literature so far. Going by *Figure 10* (as well as *Figures 2fghj*, *Appendix*), there appear to be some differences between the different shock's influences on risk preferences. However, likely due to the relatively small respective sample sizes, the test for subgroup differences did not yield significant results (Q = 2.05, df = 3, p = 0.5619, *Output 2i*, *Appendix*). Selecting only the studies for the different subgroups, average effect sizes do not differ significantly from zero for natural disasters (SMD = –0.009, t = –0.23, p = 0.82, N = 21, *Output 2c*, *Appendix*), the COVID-19 pandemic (SMD = –0.0398, t = –0.62, p = 0.551, N = 10, *Output 2e, Appendix*) and Conflict & Violence (SMD = –0.063, t = –1.75, p = 0.131, N = 7, *Output 2f, Appendix*).

Figure 10: Shock types



Economic shocks, however, do appear to on average reduce risk seeking significantly (SMD = –0.1794, t = –2.67, p = 0.031, N = 8, *Output 2g*, *Appendix*) with an even slightly larger average effect size if we restrict the sub-sample exclusively to the global financial crisis (SMD = –0.1971, t = –2.23, p = 0.0762, N = 6, *Output 2h*, *Appendix*). Given the very small sample sizes and the number of statistical tests performed, we should likely not be too confident of these results. However, similar to the finding that medium-term effects appear to drive the overall effect towards a decrease in risk seeking, looking at different shock types reveals that research on economic crises may play a similar role.

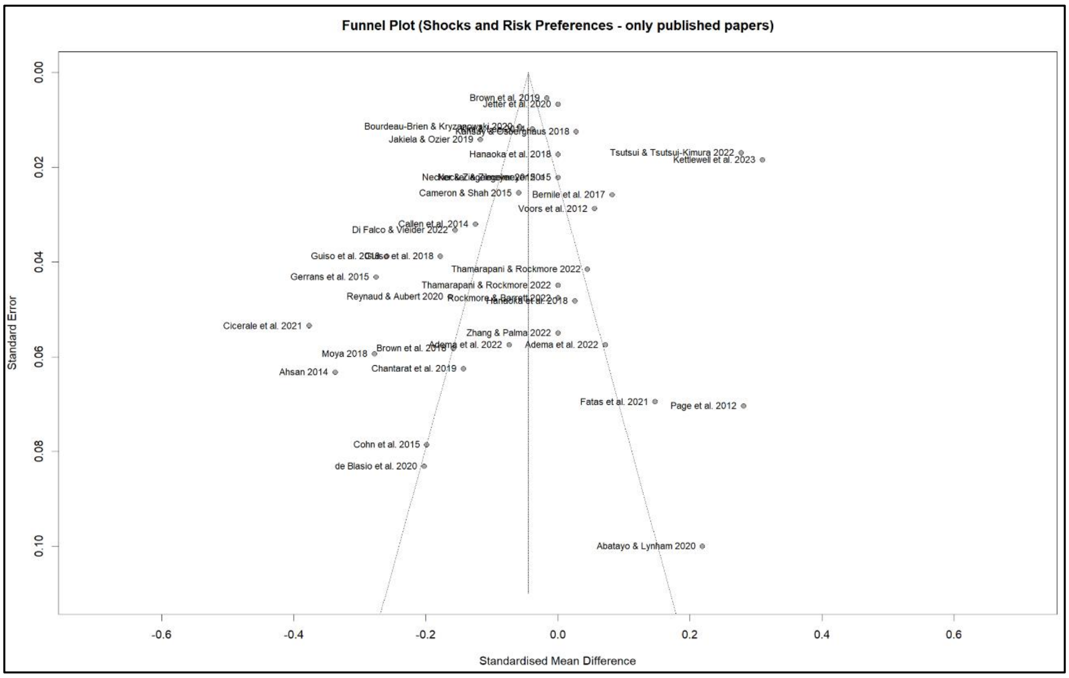
With this note, I conclude the main meta-analysis of effect sizes. In the next chapter, I will briefly investigate the issue of publication bias.

iv. Publication bias

As a last step of this meta-analysis, I investigate whether any signs of publication bias can be detected. While I have tried to compile a list of studies that approximates or at least represents the whole research that has been done on the effects of shocks on risk preferences, this approach relies on this research being published in some form, be it as a peer-reviewed journal article or as a working paper. Given the institutional design of academia and the incentive structures of the various stakeholders, there are many ways for research to be conducted without being published.

While publication bias is difficult to prove, there are multiple possible methods which can at least give some clues as to whether heightened attention to the issue is warranted (Harrer et al., 2021). A first approach can be to compare the results of peer-reviewed publications and working papers. While this does not account for research which gets discarded before even some form of article is written and submitted to a repository, it may tell us something about potential biases being introduced by the peer-review and publication process. Comparing average effects does not reveal notable differences. To the contrary, they are remarkably similar both in size ($SMD = –0.0448$ for journal publications vs $SMD = –0.0638$ for working papers) and in their inferential statistics ($t = –1.69$, $p = 0.0986$, $N = 38$ for journal publications vs $t = –1.72$, $p = 0.1232$, $N = 8$, *Output 4a & 4b, Appendix, Figure 3d & 3e, Appendix)*. Forest plots for journal publications and working papers can be found in the Appendix (*Figure 2d, Appendix and Figure 2e, Appendix)*.

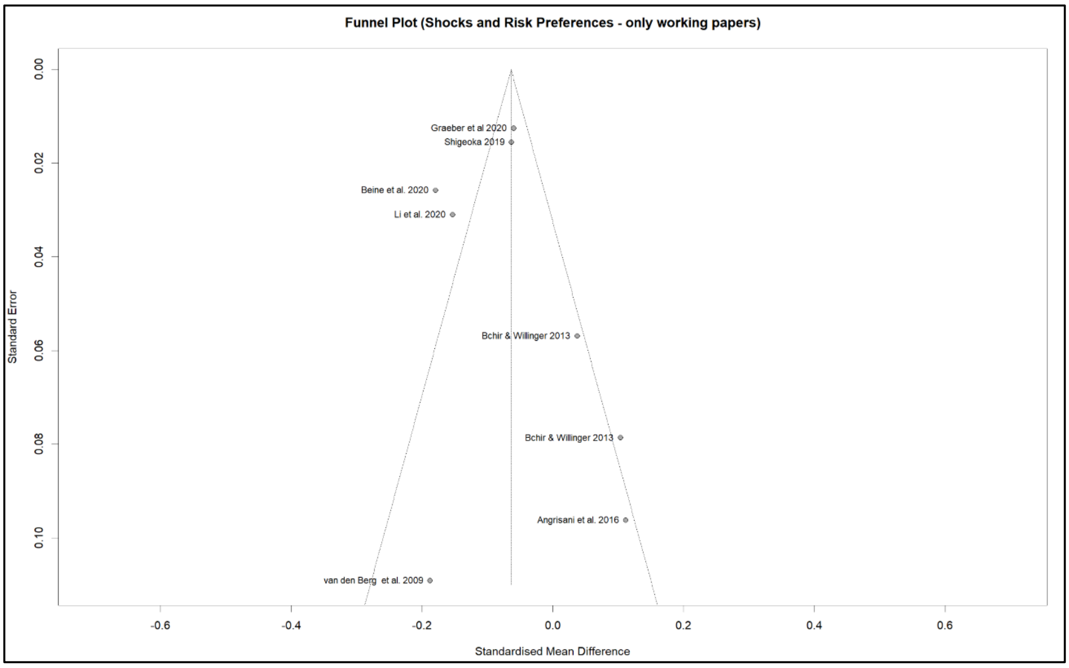
Figure 11: Funnel Plot (only published studies)



Meta-analysts, however, have developed more sophisticated ways to test for publication bias. Apart from p-curve analysis (which was not possible in this case due to a lack of precisely reported p-values), one central approach are so-called funnel plots (*Figure 11 & 5b, Appendix, Figure 5c &5d)*. Funnel plots display the standardised mean difference (effect size) on the x-axis and the standard error on the y-axis. Each dot represents one study. As standard errors are inversely related to sample size, the plot tends to display studies with smaller sample sizes at the bottom and studies with larger sample sizes at the top. The higher a dot, the more precise the estimate.

The vertical line in the pyramid represents the true effect based on the data. Sterne et al., 2011, explain that *"[in] the absence of bias and between study heterogeneity, the scatter will be due to sampling variation alone and the plot will resemble a symmetrical inverted funnel [...]. A triangle centred on a fixed effect summary estimate and extending 1.96 standard errors either side will include about 95% of studies if no bias is present and the fixed effect assumption (that the true treatment effect is the same in each study) is valid"*. According to this criterion, there appears to be some bias towards somewhat inflated effect sizes in the literature, as can be observed in *Figure 5c*. While there appears to be some convergence around the peak, around half of the published studies fall outside of the funnel. This seems to be the case for studies that find decreases in risk seeking slightly more so than for studies that find increases.

Figure 12: Funnel Plot (only working papers)



The situation looks slightly different for working papers (*Figure 12*) which for the most part appear to conform to the funnel shape with few large deviations from the expected structure. The number of observations here is very small, however.

Following the standard recommendations for interpreting funnel plots, we should likely conclude that some suspicion regarding bias in the literature is warranted. There are, however, two major problems with this standard interpretation of the comparisons and funnel plots when applied to the literature on shocks and risk preferences. Firstly, as the quote from Sterne et al., 2011, mentions, an important assumption when interpreting funnel plots is that the vertical line represents the *true treatment effect*. Given the enormous variation in treatment types and measures as well as sample compositions, locations, sampling lags and elicitation methods, there may not even be one true effect but rather multiple ones. In that case, even large deviations from the vertical line in the funnel for studies with small effect sizes would not be a strong indicator of publication bias. Funnel plot asymmetry may be an indicator of publication bias. However, it can also be the result of true heterogeneity within the sample (Sterne et al., 2011). The slight leftward tilt of the distribution in *Figure 5c* may, for instance, be due to the aforementioned subgroups for medium-run effects and economic crises which had comparatively large negative average effects. Secondly, my confidence in comparability of the specific effect sizes and standard errors is relatively low. Hence, I have to further decrease the confidence of publication bias being an important issue in this field.

IV. Discussion & Conclusion

In contrast to Chuang & Schechter, 2015, I find that the majority of the publications that investigated the impact of shocks find significant effects on risk preferences. Importantly though, those effects point in very different directions. On aggregate, there appears to be a tendency for studies to find reductions in risk seeking in response to a shock. Heterogeneity is substantial but it should be noted that of the many subgroups I analysed, only very few had average effects pointing towards more risk seeking (none of them significant). Subgroup comparisons and analyses deliver mostly null results. Highly significant result for medium-run effects as well as relatively large negative effects for economic factors imply that the small overall negative effect may be driven by few subgroups. However, due to the small sample sizes for subgroups, confidence in those results should be rather low. Finally, while I cannot discard the possibility of publication bias but there appears to be little to no evidence for bias in any one direction.

i. Contribution and Limitations

This study should be seen as a preliminary aggregate overview of the current state of the literature on the relationship between exogenous shocks and risk preferences. Its goal is to give a quantitative account of some of the core aspects of the existing body of work on the topic. Its goal was to be exhaustive in the sense that it covers as close to a complete set of publications as possible. However, it is by no means a complete review of the literature as there are many methodological and material issues that warrant further investigation. For instance, there are numerous important and complicated issues such as the mechanisms through which (different kinds of) shocks influence risk preferences, e.g. trust (Ayton et al., 2020), stress (Cahlíková & Cingl, 2017), cognitive resources (Castillo et al., 2017), depression and emotions (Cobb-Clark et al., 2022; Eckel et al., 2009; Moya, 2018) as well as economic conditions and expectations (Cohn et al., 2015) just to name a few.

Additionally, this study focuses on the effects of shocks on risk preferences but leaves out a number of other factors that may influence risk preferences such as inequality (Schmidt et al., 2019), less pronounced and less negatively valenced economic factors (Brunnermeier & Nagel, 2008; Chetty & Szeidl, 2007) as well as general stability over time (Einav et al., 2012) and contexts (Barseghyan et al., 2011).

Finally, I hope that this research endeavour will be a first conceptual and practical step towards a continuously expanding and updating repository of studies with information on effect sizes and a range of other study characteristics.

An important research area that remains largely untouched by my study is the role of measurement of risk preferences. I did compare incentivised to non-incentivised elicitation techniques and results were similar. As mentioned above, however, some of the studies in my sample deliver different results depending on the measurement method (Adema et al., 2022; Holden & Tilahun, 2021; Reynaud & Couture, 2012; Zhang & Palma, 2021). Further reviews and research could aim to explain this apparent disconnect between within-study differences and aggregate similarity by exploring in much more detail the relevant literature on measurement of risk preferences from a general psychometric perspective and in the context of different possible influences on risk preferences. In this regard, an especially interesting development appears to be an increasingly systematic research effort focusing on the validity and predictive power of incentivised preference elicitation (lotteries, gambles, MLPs) versus non-incentivised surveys with relatively broad self-assessments of risk preferences (Anderson & Mellor, 2009; Arslan et al., 2020; Festjens et al., 2015; Finger et al., 2023; Galizzi & Miniaci, 2016; Garagnani, 2020; Hackethal et al., 2023; Hertwig et al., 2019; Lönnqvist et al., 2015; Mata et al., 2018; Mudzingiri et al., 2021; Mudzingiri & Koumba, 2021). Paying attention to those kinds of developments which tend to happen in the field of psychology could be beneficial for researchers in all related disciplines as it may contribute to sharpening concepts and measurement approaches.

The considerable variation regarding research designs, measurement tools, geographical locations, participant characteristics, sample sizes and reporting of results, makes it very difficult to approach the topic with the standard tools of a meta-analysis. For instance, effect sizes are not consistently reported in a fashion that allows comparisons across studies. Consequently, any calculations of overall mean effect sizes should be regarded as tentative and likely somewhat imprecise. Future meta-analytic work would greatly benefit from a move towards some standardisation when it comes to reporting results. Given the apparent scientific interest in the topic, this relatively banal goal would be one of the most efficient ways to promote scientific progress in the field.

Ironically, some of the most ambitious and most prominently published studies such as Malmendier & Nagel, 2011, or Callen et al., 2014, are among the ones that are least well suited for meta-analytic review. While the reasons for that – supreme rigour and nuance – are difficult to criticise, it is nevertheless unfortunate that the evidence produced by some of the most brilliant minds working on the topic contributes little to the overall answers to some of the most basic questions. This inverse relationship between undeniable quality and rigour on the one hand and eligibility for meta-analysis on the other implies that the studies that can easily be included in meta-analyses may be more likely to suffer from problematic specifications as well as inaccurate date on disaster exposure causing imprecise causal estimations (Kuroishi & Sawada, 2019).

Furthermore, I need to mention that any meta-analysis conducted by a single master's student who is generally literate in the field but still far from a subject matter expert will be incomplete and lacking in many regards, starting with general expertise and sensitivity to important issues and ending with insurmountable limitations such as the inability to conduct proper reliability checks for the coding of study characteristics. Further research could also try to add more articles to the sample by scanning other sources such as EconLit, dissertation and thesis databases as well as lists provided by funding agencies. Additionally, future research could include more detailed information on study quality such as more sophisticated assessments of internal validity (e.g. use of random assignment, plausibility of causal identification, condition concealment, attrition), external validity (e.g. use of random sampling procedures, samples based on specific subpopulations), construct validity (e.g. reliability of measures (for correction rather than exclusion or moderator analyses) and other relevant measurement characteristics).

ii. Scientific implications

On the basis of the results and my experiences over the course of the research process, I would like to highlight two conclusions that may be relevant for the scientific community to take into account when conducting future research in this field.

The first one is that the substantial between-study-heterogeneity which cannot be explained by differences in study characteristics implies that the question about the *general* relationship between shock *in general* and risk preferences *in general* cannot be answered confidently at this point in time. For now, we may have to be content with the preliminary and possibly somewhat unsatisfactory conclusion that there is no general true effect of any kind of shock on risk preferences. This would not imply that no predictions about behaviour in the field can be made as there are by now many studies on behaviour in a wide range of populations, locations, and circumstances. It does, however, mean that the predictions should for now be based more on the studies that share the most characteristics with the respective situation instead of the combined evidence on shocks and risk preferences.

Future work may try to conduct more nuanced analyses with a richer and more granular set of variables in order to make more confident claims about the direction and magnitude of a general effect. This, however, will only yield convincing results if the primary research and associated research and publication practices evolve as well. This leads me to my second point.

Given the increasing amount of primary literature on the determinants of risk preferences, a major task for the future will likely be to continuously integrate new findings into the existing body of knowledge. While individual studies will certainly remain important and relevant on for their own sake, the potential for knowledge generation through different kinds of meta-scientific research on the topic will become even more promising.

In order to improve the ease of and prospects for objective, reliable and valid meta-analytic work, I would like to briefly sketch out a few recommendations based on my experiences while searching for, summarising, and coding the publications on the topic.

Most importantly, I would like to encourage authors and editors to not compromise on clear and comprehensive reporting practices. This includes reporting of basic sample characteristics and summary statistics (at least location, size, average/median age, gender, year of collection, time since shock etc.). The same applies to the key results. As much as possible, effects should be reported as standardised and easily interpretable coefficients with standard errors, standard deviations and if possible exact p-values included. Sometimes that may be difficult, for example when non-parametric tests or odds-ratios are better suited (or even the only really suitable) statistical approaches. However, keeping potential meta-analytic review of their studies in mind, researchers could reduce the number of studies without easily interpretable effect sizes.

To researchers who want to portray a nuanced and accurate picture of their subject, it may seem unsatisfactory to highlight *one single coefficient* as the main result of their work. I fully understand that impulse and I would even agree that focus on one single coefficient paint a misleadingly simple picture of the underlying research and it may lead to the article's reception being reduced to it. However, given the increasing importance of meta-analytic review in the field, I would like to urge authors and editors to invest some time into finding an approach that balances accessibility of central information for meta-analytic review with sufficiently nuanced portrayals of methodologies, data, and results.

The issue becomes more complicated and nuanced when it comes to research design and measurement methodologies. I recognise that different circumstances will call for different methodologies. This issue is exacerbated by the fact that the kinds of shocks that are investigated in this line of research are qua their nature unforeseeable events. This means that especially for the longitudinal studies which compare measurements before and after the shock there is no possibility of planning the study in advance. Usually, there will be a first measurement with the goal of studying something entirely different, followed by the unforeseen shock and a second measurement This second measurement then has to follow not necessarily what would be best in terms of adding a datapoint to the existing literature on the effects of shocks on risk preferences but is somewhat constrained by the theory and motivations for conducting the first one.

A potential obstacle for standardisation of methodologies is the incentive to conduct original research. While this is certainly important even in a mature field due to the pressure to improve methodologies, researchers and editors should keep in mind that, maybe apart from an initial period of exploration, replication and aggregation of similar studies may be equally important for the advancement of the field.

As of today, one might even argue that effect sizes should not be a focus of meta-analytic work due to the variety of independent variables and lack of consistency in research designs (see chapter *II.iii*). While this argument certainly should play a role when interpreting the effect sizes of *currently existing* research, it might become less relevant in the future if the scientific community manages to develop more standardised approaches. Furthermore, standardised effect sizes will also be important for detecting publication bias.

Given the increasing amount of primary research in the field, future meta-analytic work may also further investigate the sources of between-study heterogeneity. Further work could also take more seriously the heterogeneity of impacts that may influence risk preferences. Some possible ways forward can be found in (Rockmore & Barrett, 2022).

iii. Policy implications and Conclusion

As hinted at before, anyone working on topics for which the relationship between shocks and risk preferences is important, should be very cautious when it comes to drawing conclusions from the literature about a general relationship between the two variables. Given the relatively large variety of independent variables, contexts and research methodologies, meta-analyses about complex psychological, social, and economic relationships should not necessarily be assumed to be some kind of "gold standard" but rather only one of many potentially informative pieces of evidence. In this respect, meta-analyses in the social sciences differ markedly from their cousins in the medical and some natural sciences. The specific absolute and relative limitations and strengths of a meta-analysis will likely differ from case to case.

At this moment in time, I would likely advise policymakers and other professionals for whom risk preferences are an important factor to instead carefully consider the results of primarily those studies which provide information on a group of people or circumstances which closely resembles the issues at hand.

While this may sound somewhat disappointing, I would caution against understanding it as a reason for despair. Social science is a slow, iterative, and messy process which takes time and effort to develop true and actionable answers to meaningful questions. Given the increasing volume of research on the determinants of risk preferences, I am confident that the next decade will be one of great progress in the field. I hope that meta-analysis on the topic will continue to be pursued and that this will raise the overall standard when it comes to research methodology and reporting.

References

Abatayo, A. L., & Lynham, J. (2020). Risk preferences after a typhoon: An artefactual field experiment with fishers in the Philippines. *Journal of Economic Psychology*, *79*, 102195. https://doi.org/10.1016/j.joep.2019.102195

Adema, J., Nikolka, T., Poutvaara, P., & Sunde, U. (2022). On the stability of risk preferences: Measurement matters. *Economics Letters*, *210*, 110172. https://doi.org/10.1016/j.econlet.2021.110172

Adjei-Mantey, K., & Horioka, C. Y. (2022). Determinants of health insurance enrollment and health expenditure in Ghana: An empirical analysis. *Review of Economics of the Household*, 1–20. https://doi.org/10.1007/s11150-022-09621-x

Ahsan, D. A. (2014). Does natural disaster influence people's risk preference and trust? An experiment from cyclone prone coast of Bangladesh. *International Journal of Disaster Risk Reduction*, *9*, 48–57. https://doi.org/10.1016/j.ijdrr.2014.02.005

Akesaka, M., Eibich, P., Hanaoka, C., & Shigeoka, H. (2021). Temporal Instability of Risk Preference among the Poor: Evidence from Payday Cycles. *National Bureau of Economic Research.* Advance online publication. https://doi.org/10.3386/w28784

Alan, S., Baydar, N., Boneva, T., Crossley, T. F., & Ertac, S. (2017). Transmission of risk preferences from mothers to daughters. *Journal of Economic Behavior & Organization*, *134*, 60–77. https://doi.org/10.1016/j.jebo.2016.12.014

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. https://doi.org/10.1111/1467-8489.12434

Andersen, S., Harrison, G. W., Lau, M. I., & Elisabet Rutström, E. (2008). Lost IN STATE SPACE: Are PREFERENCES STABLE? *International Economic Review*, *49*(3), 1091–1112. https://doi.org/10.1111/j.1468-2354.2008.00507.x

Anderson, L. R., & Mellor, J. M. (2008). Predicting health behaviors with an experimental measure of risk preference. *Journal of Health Economics*, *27*(5), 1260–1274. https://doi.org/10.1016/j.jhealeco.2008.05.011

Anderson, L. R., & Mellor, J. M. (2009). Are risk preferences stable? Comparing an experimental measure with a validated survey-based measure. *Journal of Risk and Uncertainty*, *39*(2), 137–160. https://doi.org/10.1007/s11166-009-9075-z

Angerer, S., Dutcher, E. G., Glätzle-Rützle, D., Lergetporer, P., & Sutter, M. (2021). The Formation of Risk Preferences Through Small-Scale Events. *SSRN Electronic Journal.* Advance online publication. https://doi.org/10.2139/ssrn.3912380

Angrisani, M., Cipriani, M., Guarino, A., Kendall, R., & Ortiz de Zarate, J. (2020). *Risk Preferences at the Time of COVID-19: An Experiment with Professional Traders and Students.* https://doi.org/10.2139/ssrn.3609586

Arslan, R. C., Brümmer, M., Dohmen, T., Drewelies, J., Hertwig, R., & Wagner, G. G. (2020). How people know their risk preference. *Scientific Reports*, *10*(1), 15365. https://doi.org/10.1038/s41598-020-72077-5

Ayton, P., Bernile, G., Bucciol, A., & Zarri, L. (2020). The impact of life experiences on risk taking. *Journal of Economic Psychology*. https://www.ncbi.nlm.nih.gov/pmc/articles/pmc7255124/

Baillon, A., Koellinger, P. D., & Treffers, T. (2016). Sadder but wiser: The effects of emotional states on ambiguity attitudes. *Journal of Economic Psychology*, *53*, 67–82. https://doi.org/10.1016/j.joep.2015.12.008

Banks, J., Bassoli, E., & Mammi, I. (2020). Changing attitudes to risk at older ages: The role of health and other life events. *Journal of Economic Psychology*, *79*, 102208. https://doi.org/10.1016/j.joep.2019.102208

Barseghyan, L., Prince, J., & Teitelbaum, J. C. (2011). Are Risk Preferences Stable across Contexts? Evidence from Insurance Data. *American Economic Review*, *101*(2), 591–631. https://doi.org/10.1257/aer.101.2.591

Barseghyan, L., Teitelbaum, J. C., & Xu, L. (2018). Different Contexts, Different Risk Preferences? *SSRN Electronic Journal*. Advance online publication. https://doi.org/10.2139/ssrn.3200657

Barsky, R. B., Juster, F. T., Kimball, M. S [M. S.], & Shapiro, M. D [M. D.] (1997a). Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study. *The Quarterly Journal of Economics*, *112*(2), 537–579. https://doi.org/10.1162/003355397555280

Barsky, R. B., Juster, F. T., Kimball, M. S [M. S.], & Shapiro, M. D [M. D.] (1997b). Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study. *The Quarterly Journal of Economics*, *112*(2), 537–579. https://doi.org/10.1162/003355397555280

Batista, C., & Umblijs, J. (2014). Migration, risk attitudes, and entrepreneurship: Evidence from a representative immigrant survey. *IZA Journal of Migration*, *3*(1), 1–25. https://doi.org/10.1186/s40176-014-0017-4

Bauernschuster, S., Falck, O., Heblich, S., Suedekum, J., & Lameli, A. (2014). Why are educated and risk-loving persons more mobile across regions? *Journal of Economic Behavior & Organization*, *98*, 56–69. https://doi.org/10.1016/j.jebo.2013.12.011

Bchir, M. A., & Willinger, M. (2013). Does the exposure to natural hazards affect risk and time preferences? Some insights from a field experiment in Perú. *Unpublished Manuscript*. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=4e88b1d6c17ee9021aa3516d073b8380b0172447

Beck, U. (2020). *Risikogesellschaft: Auf dem Weg in eine andere Moderne* (24. Auflage). *Edition Suhrkamp: 1365 = Neue Folge Band 365*. Suhrkamp.

Beine, M. A. R., Charness, G., Dupuy, A., & Joxhe, M. (2020). Shaking Things Up: On the Stability of Risk and Time Preferences. *SSRN Electronic Journal*. Advance online publication. https://doi.org/10.2139/ssrn.3570289

Bellemare, C., & Shearer, B. (2010). Sorting, incentives and risk preferences: Evidence from a field experiment. *Economics Letters*, *108*(3), 345–348. https://doi.org/10.1016/j.econlet.2010.06.010

Bernile, G., Bhagwat, V., & Rau, R. P. (2017). What Doesn't Kill You Will Only Make You More Risk-Loving: Early-life Disasters and CEO Behavior. *The Journal of Finance*, *72*(1), 167–206. https://doi.org/10.1111/jofi.12432

Blasio, G. de, Paola, M. de, Poy, S., & Scoppa, V. (2021). Massive earthquakes, risk aversion, and entrepreneurship. *Small Business Economics*, *57*(1), 295–322. https://doi.org/10.1007/s11187-020-00327-x

Bourdeau-Brien, M., & Kryzanowski, L. (2020). Natural disasters and risk aversion. *Journal of Economic Behavior & Organization*, *177*, 818–835. https://doi.org/10.1016/j.jebo.2020.07.007

Brown, H., & van der Pol, M. (2015). Intergenerational transfer of time and risk preferences. *Journal of Economic Psychology*, *49*, 187–204. https://doi.org/10.1016/j.joep.2015.06.003

Brown, P., Daigneault, A., Tjernström, E., & Zou, W. (2018). Natural Disasters, Social Protection, and Risk Perceptions. *World Development*, *104*, 310–325. https://doi.org/10.1016/j.worlddev.2017.12.002

Brown, R., Montalva, V., Thomas, D., & Velásquez, A. (2019). Impact of Violent Crime on Risk Aversion: Evidence from the Mexican Drug War. *The Review of Economics and Statistics*, *101*(5), 892–904. https://doi.org/10.1162/rest_a_00788

Brunnermeier, M. K., & Nagel, S. (2008). Do Wealth Fluctuations Generate Time-Varying Risk Aversion? Micro-Evidence on Individuals' Asset Allocation. *American Economic Review*, *98*(3), 713–736. https://doi.org/10.1257/aer.98.3.713

Bryan, G., Chowdhury, S., & Mobarak, A. M. (2014). Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh. *Econometrica*, *82*(5), 1671–1748. https://doi.org/10.3982/ECTA10489

Bucciol, A., & Miniaci, R. (2018). Financial Risk Propensity, Business Cycles and Perceived Risk Exposure. *Oxford Bulletin of Economics and Statistics*, *80*(1), 160–183. https://doi.org/10.1111/obes.12193

Cahlíková, J., & Cingl, L. (2017). Risk preferences under acute stress. *Experimental Economics*, *20*(1), 209–236. https://doi.org/10.1007/s10683-016-9482-3

Callen, M. (2015). Catastrophes and time preference: Evidence from the Indian Ocean Earthquake. *Journal of Economic Behavior & Organization*, *118*, 199–214. https://doi.org/10.1016/j.jebo.2015.02.019

Callen, M., Isaqzadeh, M., Long, J. D., & Sprenger, C. (2014). Violence and Risk Preference: Experimental Evidence from Afghanistan. *American Economic Review*, *104*(1), 123–148. https://doi.org/10.1257/aer.104.1.123

Cameron, L., & Shah, M. (2015). Risk-Taking Behavior in the Wake of Natural Disasters. *Journal of Human Resources*, *50*(2), 484–515. https://doi.org/10.3368/jhr.50.2.484

Campos-Vazquez, R. M., & Cuilty, E. (2014). The role of emotions on risk aversion: A Prospect Theory experiment. *Journal of Behavioral and Experimental Economics*, *50*, 1–9. https://doi.org/10.1016/j.socec.2014.01.001

Card, N. A. (2016). *Applied meta-analysis for social science research* (Paperback edition). *Methodology in the social sciences*. The Guilford Press.

Cassar, A., Healy, A., & Kessler, C. von (2017). Trust, Risk, and Time Preferences After a Natural Disaster: Experimental Evidence from Thailand. *World Development*, *94*, 90–105. https://doi.org/10.1016/j.worlddev.2016.12.042

Castillo, M., Dickinson, D. L., & Petrie, R. (2017). Sleepiness, choice consistency, and risk preferences. *Theory and Decision*, *82*(1), 41–73. https://doi.org/10.1007/s11238-016-9559-7

Chantarat, S., Oum, S., Samphantharak, K., & Sann, V. (2019). Natural Disasters, Preferences, and Behaviors: Evidence from the 2011 Mega Flood in Cambodia. *Journal of Asian Economics*, *63*, 44–74. https://doi.org/10.1016/j.asieco.2019.05.001

Charness, G., & Gneezy, U. (2012). Strong Evidence for Gender Differences in Risk Taking. *Journal of Economic Behavior & Organization*, *83*(1), 50–58. https://doi.org/10.1016/j.jebo.2011.06.007

Chetty, R., & Szeidl, A. (2007). Consumption Commitments and Risk Preferences. *The Quarterly Journal of Economics*, *122*(2), 831–877. https://doi.org/10.1162/qjec.122.2.831

Choi, J., & Kim, K.-H. (2022). The Differential Consequences of Fear, Anger, and Depression in Response to COVID-19 in South Korea. *International Journal of Environmental Research and Public Health*, *19*(11), 6723. https://doi.org/10.3390/ijerph19116723

Chuang, Y., & Schechter, L. (2015). Stability of Experimental and Survey Measures of Risk, Time, and Social Preferences: A Review and Some New Results. *Journal of Development Economics*, *117*, 151–170. https://doi.org/10.1016/j.jdeveco.2015.07.008

Cicerale, A., Blanzieri, E., & Sacco, K. (2022). How does decision-making change during challenging times? *PLOS ONE*, *17*(7), e0270117. https://doi.org/10.1371/journal.pone.0270117

Cobb-Clark, D. A., Dahmann, S. C., & Kettlewell, N. (2022). Depression, Risk Preferences, and Risk-Taking Behavior. *Journal of Human Resources*, *57*(5), 1566–1604. https://doi.org/10.3368/jhr.58.1.0419-10183R1

Cohn, A., Engelmann, J., Fehr, E., & Maréchal, M. A. (2015). Evidence for Countercyclical Risk Aversion: An Experiment with Financial Professionals. *American Economic Review*, *105*(2), 860–885. https://doi.org/10.1257/aer.20131314

Conte, A., Levati, M. V., & Nardi, C. (2018). Risk Preferences and the Role of Emotions. *Economica*, *85*(338), 305–328. https://doi.org/10.1111/ecca.12209

Cramer, J., Hartog, J., Jonker, N., & van Praag, C. (2002). Low risk aversion encourages the choice for entrepreneurship: an empirical test of a truism. *Journal of Economic Behavior & Organization*, *48*(1), 29–36. https://doi.org/10.1016/S0167-2681(01)00222-0

Davis, J., Mengersen, K., Bennett, S., & Mazerolle, L. (2014). Viewing systematic reviews and meta-analysis in social research through different lenses. *SpringerPlus*, *3*(1), 511. https://doi.org/10.1186/2193-1801-3-511

Decker, S., & Schmitz, H. (2016). Health shocks and risk aversion. *Journal of Health Economics*, *50*, 156–170. https://doi.org/10.1016/j.jhealeco.2016.09.006

Deole, S. S., & Rieger, M. O. (2023). The immigrant-native gap in risk and time preferences in Germany: Levels, socio-economic determinants, and recent changes. *Journal of Population Economics*, *36*(2), 743–778. https://doi.org/10.1007/s00148-022-00925-x

Di Falco, S., & Vieider, F. M. (2022). Environmental Adaptation of Risk Preferences. *The Economic Journal*, *132*(648), 2737–2766. https://doi.org/10.1093/ej/ueac030

Dohmen, T., Lehmann, H., & Pignatti, N. (2016). Time-varying individual risk attitudes over the Great Recession: A comparison of Germany and Ukraine. *Journal of Comparative Economics*, *44*(1), 182–200. https://doi.org/10.1016/j.jce.2015.10.002

Dupas, P. (2014). Short-run subsidies and long-run adoption of new health products: Evidence from a field experiment. *Econometrica*, *82*(1), 197–228. https://doi.org/10.3982/ECTA9508

Eckel, C. C., El-Gamal, M. A., & Wilson, R. K. (2009). Risk loving after the storm: A Bayesian-Network study of Hurricane Katrina evacuees. *Journal of Economic Behavior & Organization*, *69*(2), 110–124. https://doi.org/10.1016/j.jebo.2007.08.012

Einav, L., Finkelstein, A., Pascu, I., & Cullen, M. R. (2012). How General are Risk Preferences? Choices under Uncertainty in Different Domains. *American Economic Review*, *102*(6), 2606-2038. https://doi.org/10.1257/aer.102.6.2606

Enrique Fatas, Nathaly Jiménez, Lina Restrepo-Plaza, & Gustavo Rincón (2021). The Behavioral Consequences of Conflict Exposure on Risk Preferences. Advance online publication. https://doi.org/10.1093/acrefore/9780190625979.013.637

Ert, E., & Haruvy, E. (2017). Revisiting risk aversion: Can risk preferences change with experience? *Economics Letters*, *151*, 91–95. https://doi.org/10.1016/j.econlet.2016.12.008

Festjens, A., Bruyneel, S., Diecidue, E., & Dewitte, S. (2015). Time-based versus money-based decision making under risk: An experimental investigation. *Journal of Economic Psychology*, *50*, 52–72. https://doi.org/10.1016/j.joep.2015.07.003

Filipski, M., Jin, L., Zhang, X., & Chen, K. Z. (2019). Living like there's no tomorrow: The psychological effects of an earthquake on savings and spending behavior. *European Economic Review*, *116*, 107–128. https://doi.org/10.1016/j.euroecorev.2019.04.004

Finger, R., Wüpper, D., & McCallum, C. (2023). The (in)stability of farmer risk preferences. *Journal of Agricultural Economics*, *74*(1), 155–167. https://doi.org/10.1111/1477-9552.12496

Frondel, M., Osberghaus, D., & Sommer, S. (2021). *Corona and the Stability of Personal Traits and Preferences: Evidence From Germany*. https://doi.org/10.2139/ssrn.3820484

Galizzi, M. M., & Miniaci, R. (2016). Temporal Stability, Cross-Validity, and External Validity of Risk Preferences Measures: Experimental Evidence from a UK Representative Sample. *SSRN Electronic Journal*. Advance online publication. https://doi.org/10.2139/ssrn.2822613

Gao, M., Liu, Y.-J., & Shi, Y. (2020). Do people feel less at risk? Evidence from disaster experience. *Journal of Financial Economics*, *138*(3), 866–888. https://doi.org/10.1016/j.jfineco.2020.06.010

Garagnani, M. (2020). The Predictive Power of Risk Elicitation Tasks. *SSRN Electronic Journal*. Advance online publication. https://doi.org/10.2139/ssrn.3692455

Gardner, M., & Steinberg, L. (2005). Peer influence on risk taking, risk preference, and risky decision making in adolescence and adulthood: An experimental study. *Developmental Psychology*, *41*(4), 625–635. https://doi.org/10.1037/0012-1649.41.4.625

Gassmann, X., Malézieux, A., Spiegelman, E., & Tisserand, J.-C. (2022). Preferences after pan(dem)ics: Time and risk in the shadow of COVID-19. *Judgment and Decision Making*, *17*(4), 745–767. https://doi.org/10.1017/s1930297500008925

Gerrans, P., Faff, R., & Hartnett, N. (2015). Individual financial risk tolerance and the global financial crisis. *Accounting & Finance*, *55*(1), 165–185. https://doi.org/10.1111/acfi.12053

Graeber, D., Schmidt, U., Schroeder, C., & Seebauer, J. (2020). The Effect of a Major Pandemic on Risk Preferences - Evidence from Exposure to COVID-19. *SSRN Electronic Journal.* Advance online publication. https://doi.org/10.2139/ssrn.3724461

Gordon, M., Lumley, T., & Gordon, M. M. (2019). Package 'forestplot'. Advanced forest plot using 'grid'graphics. The Comprehensive R Archive Network, Vienna.

Guiso, L., Sapienza, P., & Zingales, L. (2018). Time varying risk aversion. *Journal of Financial Economics*, *128*(3), 403–421. https://doi.org/10.1016/j.jfineco.2018.02.007

Habib, M., Cassotti, M., Moutier, S., Houdé, O., & Borst, G. (2015). Fear and anger have opposite effects on risk seeking in the gain frame. *Frontiers in Psychology, 6,* 253. https://doi.org/10.3389/fpsyg.2015.00253

Hackethal, A., Kirchler, M., Laudenbach, C., Razen, M., & Weber, A. (2023). On the role of monetary incentives in risk preference elicitation experiments. *Journal of Risk and Uncertainty, 66*(2), 189–213. https://doi.org/10.1007/s11166-022-09377-w

Haile, K. K., Nillesen, E., & Tirivayi, N. (2020). Impact of formal climate risk transfer mechanisms on risk-aversion: Empirical evidence from rural Ethiopia. *World Development, 130,* 104930. https://doi.org/10.1016/j.worlddev.2020.104930

Hanaoka, C., Shigeoka, H., & Watanabe, Y. (2018). Do Risk Preferences Change? Evidence from the Great East Japan Earthquake. *American Economic Journal: Applied Economics, 10*(2), 298–330. https://doi.org/10.1257/app.20170048

Harrer, M., Cuijpers, P., Furukawa, T., & Ebert, D. (2021). *Doing meta-analysis in R: a hands-on guide.* Routledge. bookdown.org/MathiasHarrer/Doing_Meta_Analysis_in_R/rob-plots.html.

Harrison, G. W., Lau, M. I., & Rutström, E. E. (2005). *Dynamic Consistency in Denmark: A Longitudinal Field Experiment.* https://doi.org/10.2139/ssrn.698483

Harrison, G. W., Lau, M. I., & Yoo, H. I. (2020). Risk Attitudes, Sample Selection, and Attrition in a Longitudinal Field Experiment. *The Review of Economics and Statistics, 102*(3), 552–568. https://doi.org/10.1162/rest_a_00845

Havránek, T., Stanley, T. D., Doucouliagos, H., Bom, P., Geyer-Klingeberg, J., Iwasaki, I., Reed, W. R., Rost, K., & Aert, R. C. M. (2020). Reporting Guidelines for Meta-Analysis in Economcis. *Journal of Economic Surveys, 34*(3), 469–475. https://doi.org/10.1111/joes.12363

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2-3), 61-83; discussion 83-135. https://doi.org/10.1017/s0140525x0999152x

Hermann, D. (2017). Determinants of financial loss aversion: The influence of prenatal androgen exposure (2D:4D). *Personality and Individual Differences*, *117*, 273–279. https://doi.org/10.1016/j.paid.2017.06.016

Hertwig, R., Wulff, D. U., & Mata, R. (2019). Three gaps and what they may mean for risk preference. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *374*(1766), 20180140. https://doi.org/10.1098/rstb.2018.0140

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*(11), 1539–1558. https://doi.org/10.1002/sim.1186

Holden, S. T., & Tilahun, M. (2021). *Shocks and stability of risk preferences* (Centre for Land Tenure Studies Working Paper 05/21). Ås: Norwegian University of Life Sciences (NMBU), Centre for Land Tenure Studies (CLTS). https://www.econstor.eu/handle/10419/262021

Huang, L., Zhou, Y., Han, Y., Hammitt, J. K., Bi, J., & Liu, Y. (2013). Effect of the Fukushima nuclear accident on the risk perception of residents near a nuclear power plant in China. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(49), 19742–19747. https://doi.org/10.1073/pnas.1313825110

Ikeda, S., Yamamura, E., & Tsutsui, Y. (2020). COVID-19 Enhanced Diminishing Sensitivity in Prospect-Theory Risk Preferences: A Panel Analysis. *SSRN Electronic Journal.* Advance online publication. https://doi.org/10.2139/ssrn.3715236

Ingwersen, N., Frankenberg, E., & Thomas, D. (2023). Evolution of Risk Aversion over Five Years after a Major Natural Disaster. *National Bureau of Economic Research.* Advance online publication. https://doi.org/10.3386/w31102

Jackson, K. C., & Mackevicius, C. L. (2023). What impacts can we expect from school spending policy? Evidence from evaluations in the US. *American Economic Journal: Applied Economics*. https://www.aeaweb.org/articles?id=10.1257/app.20220279

Jaeger, D. A., Dohmen, T., Falk, A., Huffman, D., Sunde, U., & Bonin, H. (2010). Direct Evidence on Risk Attitudes and Migration. *The Review of Economics and Statistics*, *92*(3), 684–689. https://doi.org/10.1162/REST_a_00020

Jakiela, P., & Ozier, O. (2019). The Impact of Violence on Individual Risk Preferences: Evidence from a Natural Experiment. *The Review of Economics and Statistics*, *101*(3), 547–559. https://doi.org/10.1162/rest_a_00763

Jetter, M., Magnusson, L. M., & Roth, S. (2020). Becoming sensitive: Males' risk and time preferences after the 2008 financial crisis. *European Economic Review*, *128*, 103512. https://doi.org/10.1016/j.euroecorev.2020.103512

Josef, A. K., Richter, D., Samanez-Larkin, G. R., Wagner, G. G., Hertwig, R., & Mata, R. (2016). Stability and change in risk-taking propensity across the adult life span. *Journal of Personality and Social Psychology, 111*(3), 430–450. https://doi.org/10.1037/pspp0000090

Kadlec, D., Sainani, K. L., & Nimphius, S. (2023). With Great Power Comes Great Responsibility: Common Errors in Meta-Analyses and Meta-Regressions in Strength & Conditioning Research. *Sports Medicine*, *53*(2), 313–325. https://doi.org/10.1007/s40279-022-01766-0

Kahneman, D., & Tversky, A. (2013). Prospect Theory: An Analysis of Decision Under Risk. In *Handbook of the Fundamentals of Financial Decision Making* (pp. 99–127). WORLD SCIENTIFIC. https://doi.org/10.1142/9789814417358_0006

Kahsay, G. A., & Osberghaus, D. (2018). Storm Damage and Risk Preferences: Panel Evidence from Germany. *Environmental and Resource Economics*, *71*(1), 301–318. https://doi.org/10.1007/s10640-017-0152-5

Kettlewell, N. (2019). Risk preference dynamics around life events. *Journal of Economic Behavior & Organization*, *162*, 66–84. https://doi.org/10.1016/j.jebo.2019.04.018

Kettlewell, N., Rijsdijk, F., Siribaddana, S., Sumathipala, A., Tymula, A., Zavos, H., & Glozier, N. (2023). Natural disaster and risk preferences: Evidence from Sri Lankan twins. *Applied Economics*, 1–24. https://doi.org/10.1080/00036846.2023.2168616

Khan, S. U., Ramella, S., Rahman, H. U., & Hyder, Z. (2022). Household Portfolio Allocations: Evidence on Risk Preferences from the Household, Income, and Labour Dynamics in Australia (HILDA) Survey Using Tobit Models. *Journal of Risk and Financial Management*, *15*(4), 161. https://doi.org/10.3390/jrfm15040161

Khwaja, A., Sloan, F., & Salm, M. (2006). Evidence on preferences and subjective beliefs of risk takers: The case of smokers. *International Journal of Industrial Organization*, *24*(4), 667–682. https://doi.org/10.1016/j.ijindorg.2005.10.001

Kim, Y.-I., & Lee, J. (2014). The long-run impact of a traumatic experience on risk aversion. *Journal of Economic Behavior & Organization*, *108*, 174–186. https://doi.org/10.1016/j.jebo.2014.09.009

Kimball, M. S [Miles S.], Sahm, C. R., & Shapiro, M. D [Matthew D.] (2008). Imputing Risk Tolerance From Survey Responses. *Journal of the American Statistical Association*, *103*(483), 1028–1038. https://doi.org/10.1198/016214508000000139

Knight, F. H. (1921). *Risk, uncertainty and profit*. University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship. https://ssrn.com/abstract=1496192

Krčál, O., Staněk, R., & Slanicay, M. (2019). Made for the job or by the job? A lab-in-the-field experiment with firefighters. *Research in Economics*, *73*(4), 271–276. https://doi.org/10.1016/j.rie.2019.08.001

Kuroishi, Y., & Sawada, Y. (2019). On the stability of preferences: Experimental evidence from two disasters. http://www.cirje.e.u-tokyo.ac.jp/research/dp/2019/2019cf1130.pdf

l'Haridon, O., & Vieider, F. M. (2019). All over the map: A worldwide comparison of risk preferences. *Quantitative Economics*, *10*(1), 185–215. https://doi.org/10.3982/QE898

Lerner, J. S [J. S.], & Keltner, D. (2001). Fear, anger, and risk. *Journal of Personality and Social Psychology*, *81*(1), 146–159. https://doi.org/10.1037/0022-3514.81.1.146

Lerner, J. S [Jennifer S.], Gonzalez, R. M., Small, D. A., & Fischhoff, B. (2003). Effects of fear and anger on perceived risks of terrorism: A national field experiment. Advance online publication. https://doi.org/10.1037/e633872013-121

Li, J.-Z., Li, S., Wang, W.-Z., Rao, L.-L., & Liu, H. (2011). Are people always more risk averse after disasters? Surveys after a heavy snow-hit and a major earthquake in China in 2008. *Applied Cognitive Psychology*, *25*(1), 104–111. https://doi.org/10.1002/acp.1648

Li, K. K., Huang, B., Tam, T., & Hong, Y. (2020). Does the COVID-19 Pandemic Affect People's Social and Economic Preferences? Evidence from China. *SSRN Electronic Journal.* Advance online publication. https://doi.org/10.2139/ssrn.3690072

Lohmann, P., Gsottbauer, E., You, J., & Kontoleon, A. (2020). Social Preferences and Economic Decision-Making in the Wake of COVID-19: Experimental Evidence from China. *SSRN Electronic Journal.* Advance online publication. https://doi.org/10.2139/ssrn.3705264

Lönnqvist, J.-E., Verkasalo, M., Walkowitz, G., & Wichardt, P. C. (2015). Measuring individual risk attitudes in the lab: Task or ask? An empirical comparison. *Journal of Economic Behavior & Organization*, *119*, 254–266. https://doi.org/10.1016/j.jebo.2015.08.003

Love, R. O., & Robison, L. J. (1983). An Empirical Analysis of the Intertemporal Stability of Risk Preference. *Journal of Agricultural and Applied Economics*, *16*(1), 159–166. https://doi.org/10.1017/s0081305200016654

Malmendier, U., & Nagel, S. (2011). Depression Babies: Do Macroeconomic Experiences Affect Risk Taking? *The Quarterly Journal of Economics*, *126*(1), 373–416. https://doi.org/10.1093/qje/qjq004

Mata, R., Frey, R., Richter, D., Schupp, J., & Hertwig, R. (2018). Risk Preference: A View from Psychology. *Journal of Economic Perspectives*, *32*(2), 155–172. https://doi.org/10.1257/jep.32.2.155

Meister, L., & Schulze, K. (2022). *How shocks affect stock market participation* (DIW Roundup: Politik im Fokus No. 142). Berlin: Deutsches Institut für Wirtschaftsforschung (DIW). https://www.econstor.eu/handle/10419/267487

Meunier, L., & Ohadi, S. (2021). The Impact of the COVID-19 Crisis on Individuals' Risk and Time Preferences. http://www.accessecon.com/pubs/eb/2021/volume41/eb-21-v41-i3-p90.pdf

Moya, A. (2018). Violence, psychological trauma, and risk attitudes: Evidence from victims of violence in Colombia. *Journal of Development Economics*, *131*, 15–27. https://doi.org/10.1016/j.jdeveco.2017.11.001

Mudzingiri, C., Guvuriro, S., & Gomo, C. (2021). Exploring Association between Self-Reported Financial Status and Economic Preferences Using Experimental Data. *Journal of Risk and Financial Management*, *14*(6), 243. https://doi.org/10.3390/jrfm14060243

Mudzingiri, C., & Koumba, U. (2021). Eliciting Risk Preferences Experimentally versus Using a General Risk Question. Does Financial Literacy Bridge the Gap? *Risks*, *9*(8), 140. https://doi.org/10.3390/risks9080140

Necker, S., & Ziegelmeyer, M. (2016). Household risk taking after the financial crisis. *The Quarterly Review of Economics and Finance*, *59*, 141–160. https://doi.org/10.1016/j.qref.2015.03.006

Page, L., Savage, D. A., & Torgler, B. (2012). *Variation in Risk Seeking Behavior in a Natural Experiment on Large Losses Induced by a Natural Disaster.* https://doi.org/10.2139/ssrn.2111922

Qiu, Y., Colson, G., & Grebitus, C. (2014). Risk preferences and purchase of energy-efficient technologies in the residential sector. *Ecological Economics*, *107*, 216–229. https://doi.org/10.1016/j.ecolecon.2014.09.002

Raghunathan, R., & Pham, M. T. (1999). All Negative Moods Are Not Equal: Motivational Influences of Anxiety and Sadness on Decision Making. *Organizational Behavior and Human Decision Processes*, *79*(1), 56–77. https://doi.org/10.1006/obhd.1999.2838

Revelle, W., & Revelle, M. W. (2015). Package 'psych'. The comprehensive R archive network, 337, 338.

Reynaud, A., & Aubert, C. (2020). Does flood experience modify risk preferences? Evidence from an artefactual field experiment in Vietnam. *The Geneva Risk and Insurance Review*, *45*(1), 36–74. https://doi.org/10.1057/s10713-019-00044-w

Reynaud, A., & Couture, S. (2012). Stability of risk preference measures: Results from a field experiment on French farmers. *Theory and Decision*, *73*(2), 203–221. https://doi.org/10.1007/s11238-012-9296-5

Rockmore, M., & Barrett, C. B. (2022). The implications of aggregate measures of exposure to violence for the estimated impacts on individual risk preferences. *World Development*, *157*, 105925. https://doi.org/10.1016/j.worlddev.2022.105925

Said, F., Afzal, U., & Turner, G. (2015). Risk taking and risk learning after a rare event: Evidence from a field experiment in Pakistan. *Journal of Economic Behavior & Organization*, *118*, 167–183. https://doi.org/10.1016/j.jebo.2015.03.001

Schaller, J., & Stevens, A. H. (2015). Short-run effects of job loss on health conditions, health insurance, and health care utilization. *Journal of Health Economics*, *43*, 190–203. https://doi.org/10.1016/j.jhealeco.2015.07.003

Schildberg-Hörisch, H. (2018). Are Risk Preferences Stable? *Journal of Economic Perspectives*, *32*(2), 135–154. https://doi.org/10.1257/jep.32.2.135

Schmidt, U., Neyse, L., & Aleknonyte, M. (2019). Income inequality and risk taking: The impact of social comparison information. *Theory and Decision*, *87*(3), 283–297. https://doi.org/10.1007/s11238-019-09713-8

Schwarzer, G. (2007). meta: An R package for meta-analysis. R news, 7(3), 40-45.

Shachat, J., Walker, M. J., & Wei, L. (2021). How the onset of the Covid-19 pandemic impacted pro-social behaviour and individual preferences: Experimental evidence from China. *Journal of Economic Behavior & Organization*, *190*, 480–494. https://doi.org/10.1016/j.jebo.2021.08.001

She, S., Eimontaite, I., Zhang, D., & Sun, Y. (2017). Fear, Anger, and Risk Preference Reversals: An Experimental Study on a Chinese Sample. *Frontiers in Psychology*, *8*, 1371. https://doi.org/10.3389/fpsyg.2017.01371

Shigeoka, H. (2019). Long-Term Consequences of Growing up in a Recession on Risk Preferences. *National Bureau of Economic Research.* Advance online publication. https://doi.org/10.3386/w26352

Shupp, R., Loveridge, S., Skidmore, M., Lim, J., & Rogers, C. (2017). Risk, Loss, and Ambiguity Aversion after a Natural Disaster. *Economics of Disasters and Climate Change*, *1*(2), 121–142. https://doi.org/10.1007/s41885-017-0013-2

Sterne, J. A. C., Sutton, A. J., Ioannidis, J. P. A., Terrin, N., Jones, D. R., Lau, J., Carpenter, J., Rücker, G., Harbord, R. M., Schmid, C. H., Tetzlaff, J., Deeks, J. J., Peters, J., Macaskill, P., Schwarzer, G., Duval, S., Altman, D. G., Moher, D., & Higgins, J. P. T. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ*, *343*, d4002. https://doi.org/10.1136/bmj.d4002

Thamarapani, D., & Rockmore, M. (2022). The stability and evolution of risk attitudes and time preferences after a disaster. *International Journal of Disaster Risk Reduction*, *70*, 102791. https://doi.org/10.1016/j.ijdrr.2022.102791

Treffers, T., Koellinger, P. D., & Picot, A. (2016). Do Affective States Influence Risk Preferences? *Schmalenbach Business Review*, *17*(3-4), 309–335. https://doi.org/10.1007/s41464-016-0018-3

Tsutsui, Y., & Tsutsui-Kimura, I. (2022). How does risk preference change under the stress of COVID-19? Evidence from Japan. *Journal of Risk and Uncertainty*, *64*(2), 191–212. https://doi.org/10.1007/s11166-022-09374-z

van den Berg, M., Fort, R., & Burger, K. (2009). Natural hazards and risk aversion: Experimental evidence from Latin America. Advance online publication. https://doi.org/10.22004/ag.econ.51394

Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the metafor Package. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Volland, B. (2017). The role of risk and trust attitudes in explaining residential energy demand: Evidence from the United Kingdom. *Ecological Economics*, *132*, 14–30. https://doi.org/10.1016/j.ecolecon.2016.10.002

Voors, M. J., Nillesen, E. E. M., Verwimp, P., Bulte, E. H., Lensink, R., & van Soest, D. P. (2012). Violent Conflict and Behavior: A Field Experiment in Burundi. *American Economic Review*, *102*(2), 941–964. https://doi.org/10.1257/aer.102.2.941

Wehrung, D. A., Maccrimmon, K. R., & Brothers, K. M. (1984). Utility Assessment: Domains, Stability, And Equivalence Procedures. *INFOR: Information Systems and Operational Research*, *22*(2), 98–115. https://doi.org/10.1080/03155986.1984.11731917

Wickham, H. (2011). ggplot2. Wiley interdisciplinary reviews: computational statistics, 3(2), 180-185.

Willinger, M., Bchir, M. A., & Heitz, C. (2013). Risk and time preferences under the threat of background risk: a case-study of lahars risk in central Java. *Unpublished Manuscript*. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=643b43543d373547f498bb2150a61a75e2bf021b

Wilson, D. B., Ph.D. (n.d.). Practical Meta-Analysis Effect Size Calculator [Online calculator]. Retrieved June 9, 2023, from https://campbellcollaboration.org/research-resources/effect-size-calculator.html

Witte, B.-C. (2012). Fund Managers—Why the Best Might Be the Worst: On the Evolutionary Vigor of Risk-Seeking Behavior. *Economics*(1). https://www.degruyter.com/document/doi/10.5018/economics-ejournal.ja.2012-24/html

Yao, S., Wang, Z., Sun, M., Liao, J., & Cheng, F. (2020). Top executives' early-life experience and financial disclosure quality: Impact from the Great Chinese Famine. *Accounting & Finance*, *60*(5), 4757–4793. https://doi.org/10.1111/acfi.12659

Yarberry, W., & Yarberry, W. (2021). Dplyr. CRAN Recipes: DPLYR, Stringr, Lubridate, and RegEx in R, 1-58.

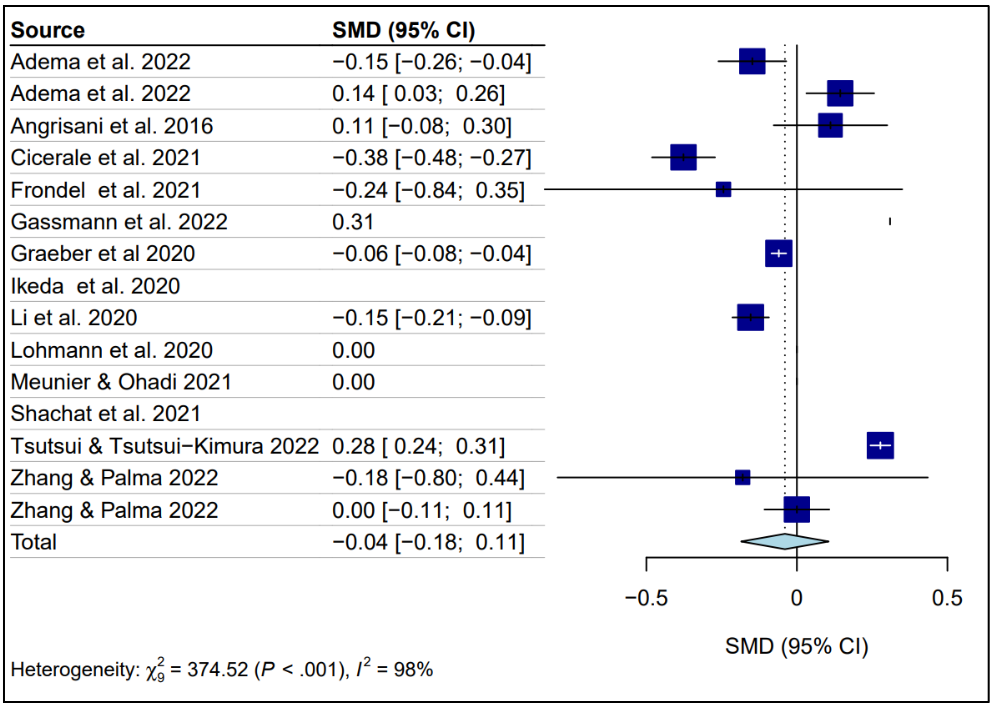
Zhang, P., & Palma, M. A. (2021). Stability of Risk Preferences During COVID-19: Evidence From Four Measurements. *Frontiers in Psychology*, *12*, 702028. https://doi.org/10.3389/fpsyg.2021.702028

Appendix

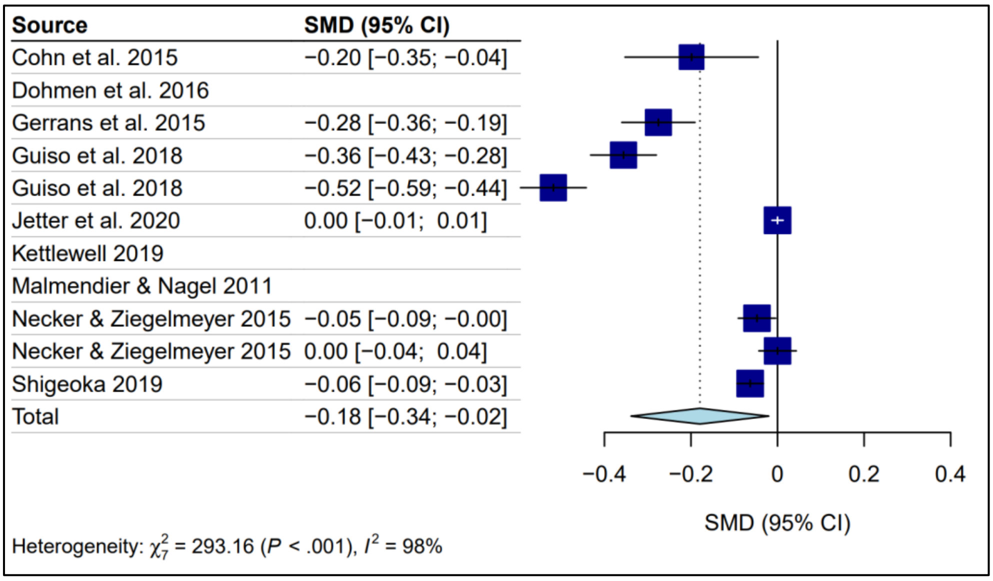
i. Figures

1. Forest Plots

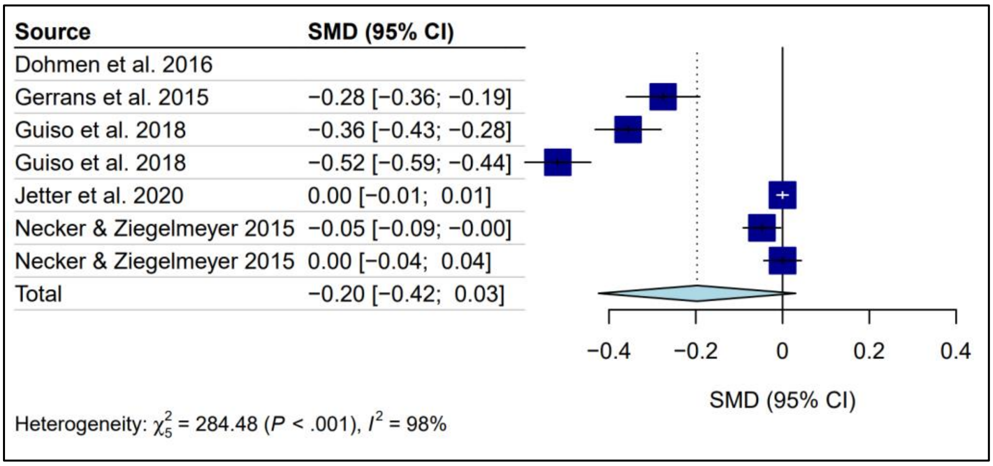
a. Pandemic / COVID-19



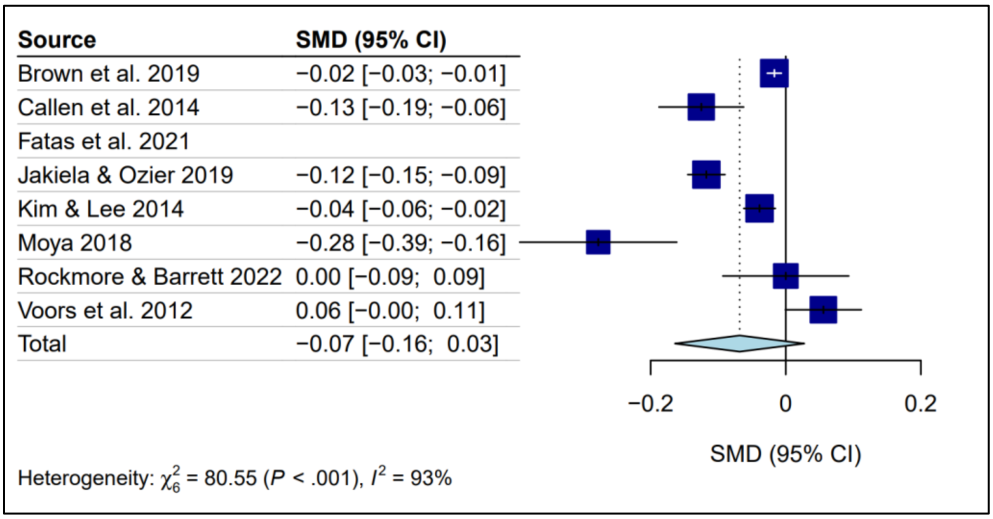
b. Economic shocks



c. Global Financial Crisis

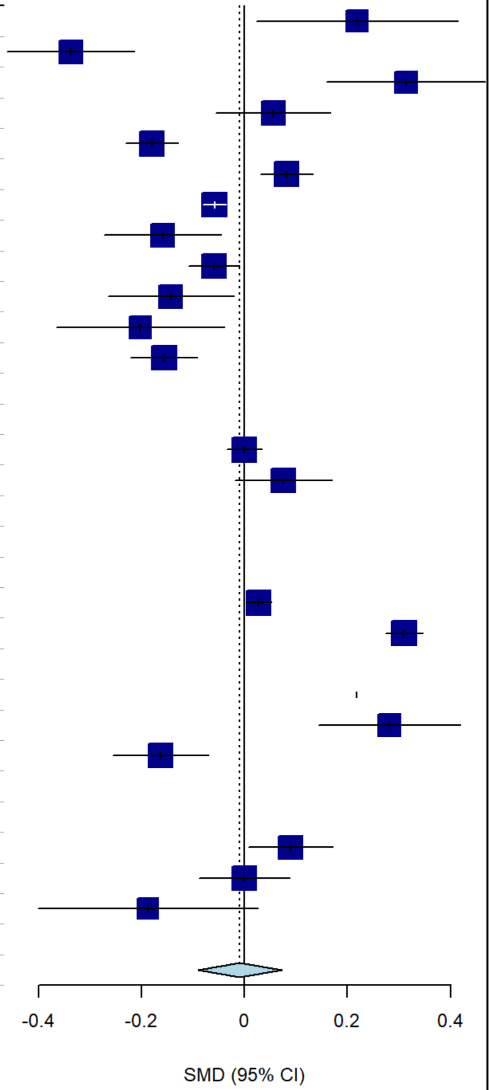


d. Conflict & Violence



a. Natural Disasters

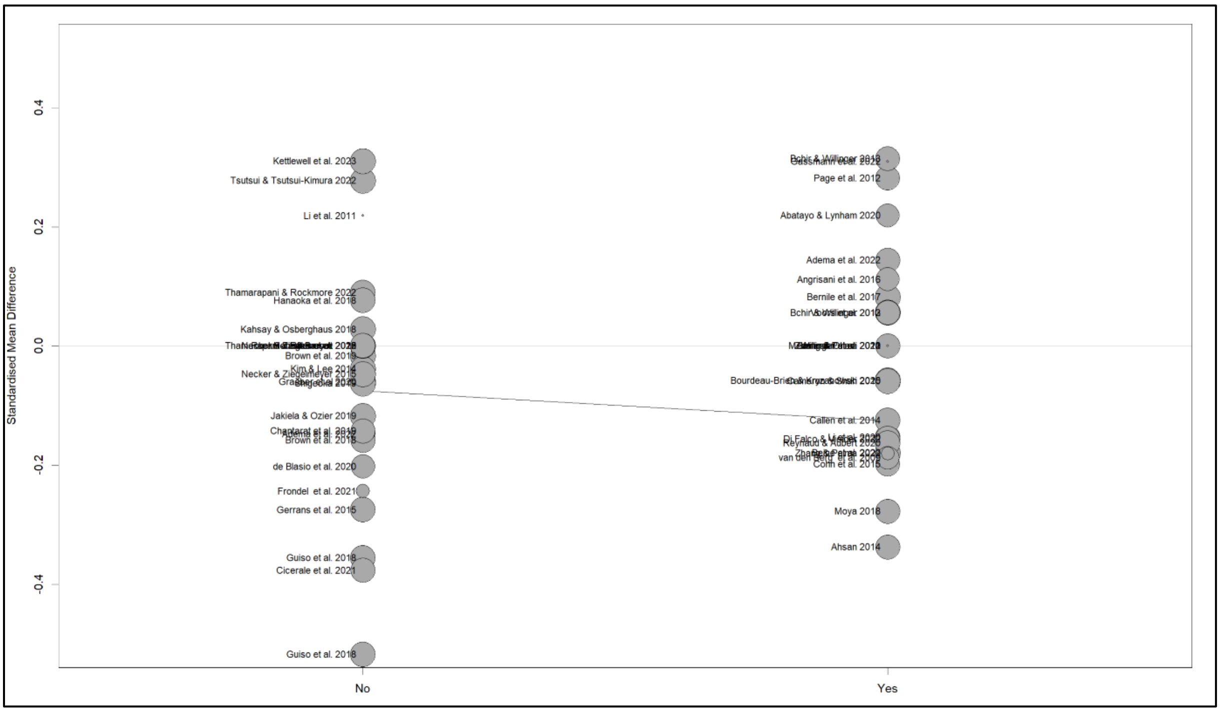
Source	SMD (95% CI)
Abatayo & Lynham 2020	0.22 [0.02; 0.41]
Ahsan 2014	-0.34 [-0.46; -0.21]
Bchir & Willinger 2013	0.31 [0.16; 0.47]
Bchir & Willinger 2013	0.06 [-0.06; 0.17]
Beine et al. 2020	-0.18 [-0.23; -0.13]
Bernile et al. 2017	0.08 [0.03; 0.13]
Bourdeau-Brien & Kryzanowski 2020	-0.06 [-0.08; -0.04]
Brown et al. 2018	-0.16 [-0.27; -0.04]
Cameron & Shah 2015	-0.06 [-0.11; -0.01]
Chantarat et al. 2019	-0.14 [-0.27; -0.02]
de Blasio et al. 2020	-0.20 [-0.37; -0.04]
Di Falco & Vieider 2022	-0.16 [-0.22; -0.09]
Eckel et al. 2009	
Finger et al. 2022	
Hanaoka et al. 2018	0.00 [-0.03; 0.03]
Hanaoka et al. 2018	0.08 [-0.02; 0.17]
Holden & Tilahun 2021	
Ingwersen et al. 2023	
Ingwersen et al. 2023	
Kahsay & Osberghaus 2018	0.03 [0.00; 0.05]
Kettlewell et al. 2023	0.31 [0.27; 0.35]
Kuroishi & Sawada 2020	
Li et al. 2011	0.22
Page et al. 2012	0.28 [0.14; 0.42]
Reynaud & Aubert 2020	-0.16 [-0.26; -0.07]
Said et al. 2015	
Shupp et al. 2017	
Thamarapani & Rockmore 2022	0.09 [0.01; 0.17]
Thamarapani & Rockmore 2022	0.00 [-0.09; 0.09]
van den Berg et al. 2009	-0.19 [-0.40; 0.03]
Willinger et al. 2013	0.00
Total	-0.01 [-0.09; 0.07]



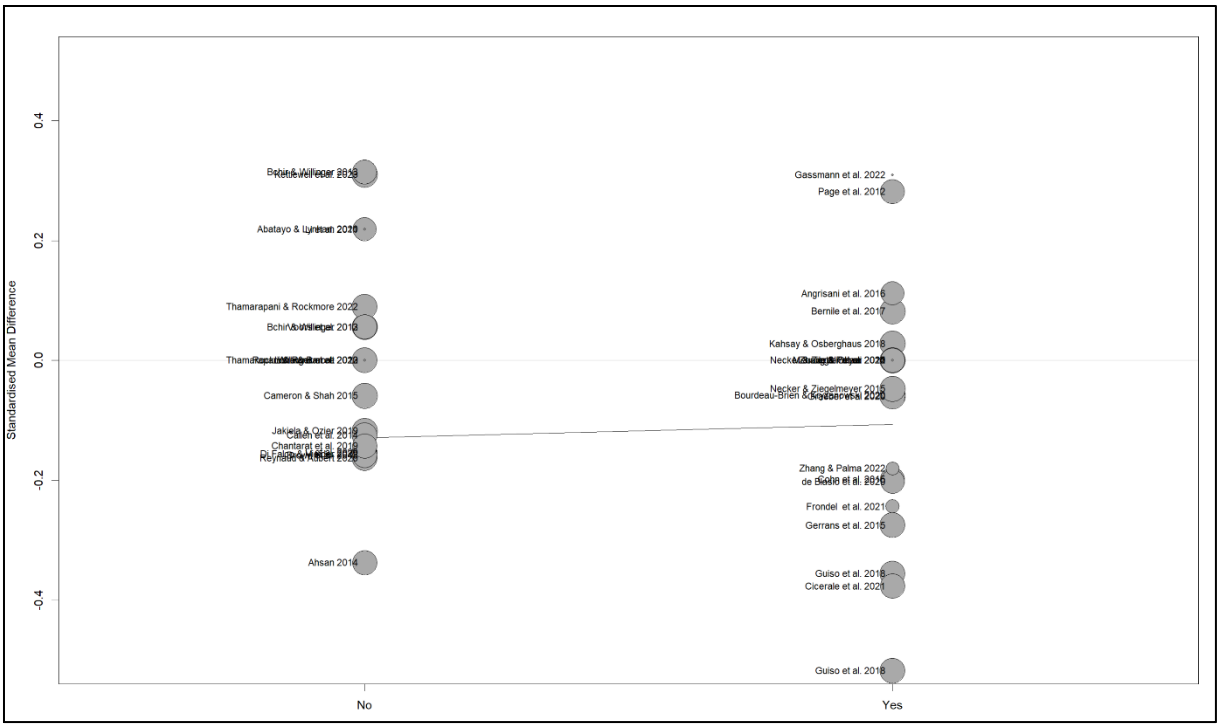
Heterogeneity: $\chi^2_{20} = 505.69$ ($P < .001$), $I^2 = 96\%$

2. Bubble Plots

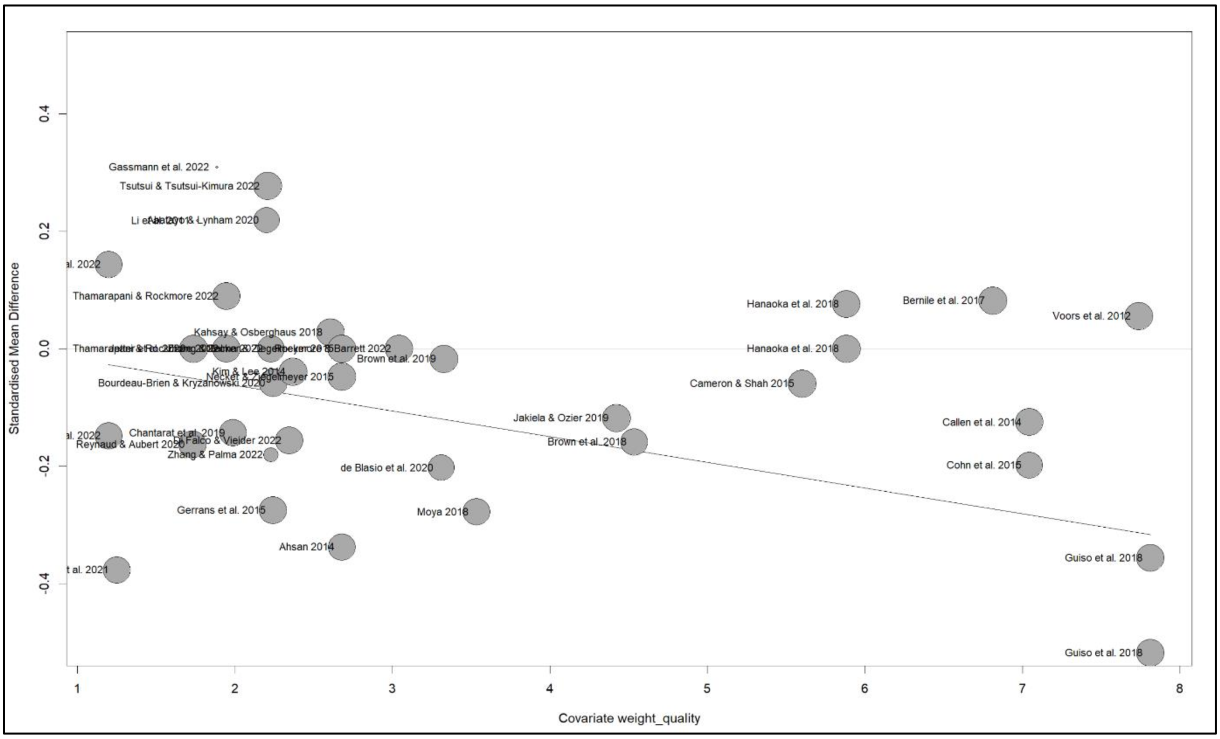
a. Measurement incentivised



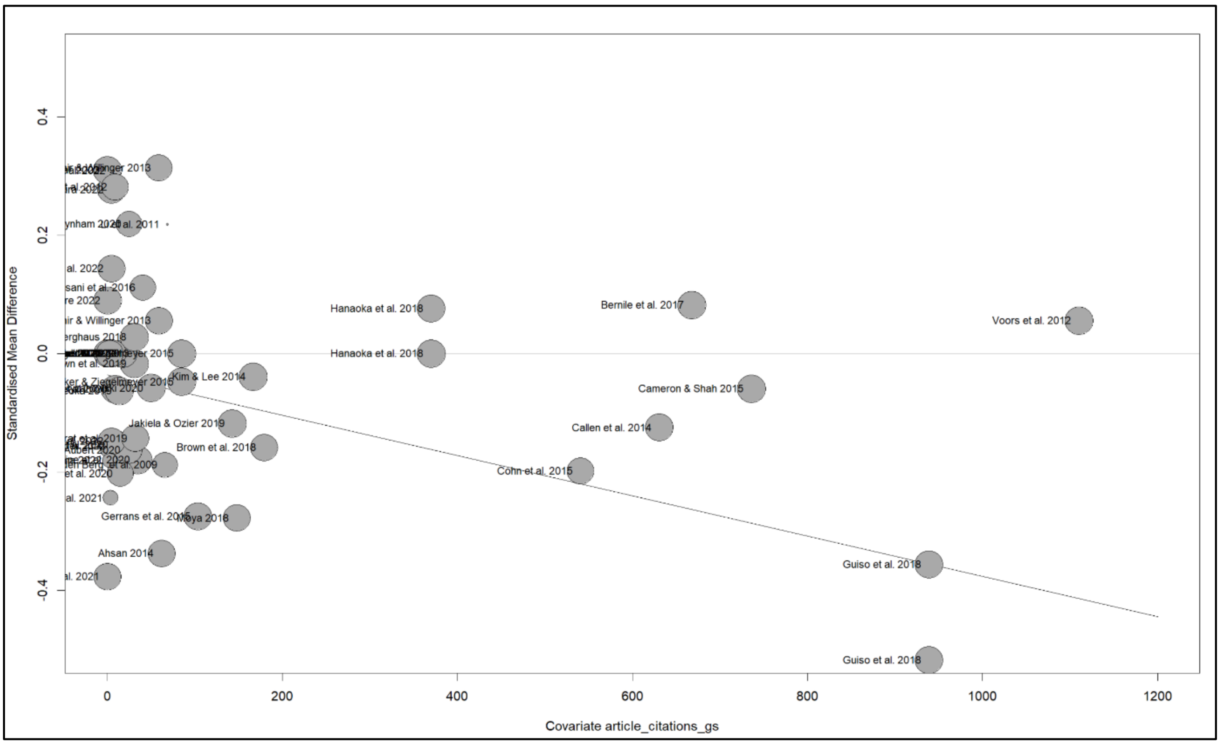
b. WEIRD sample



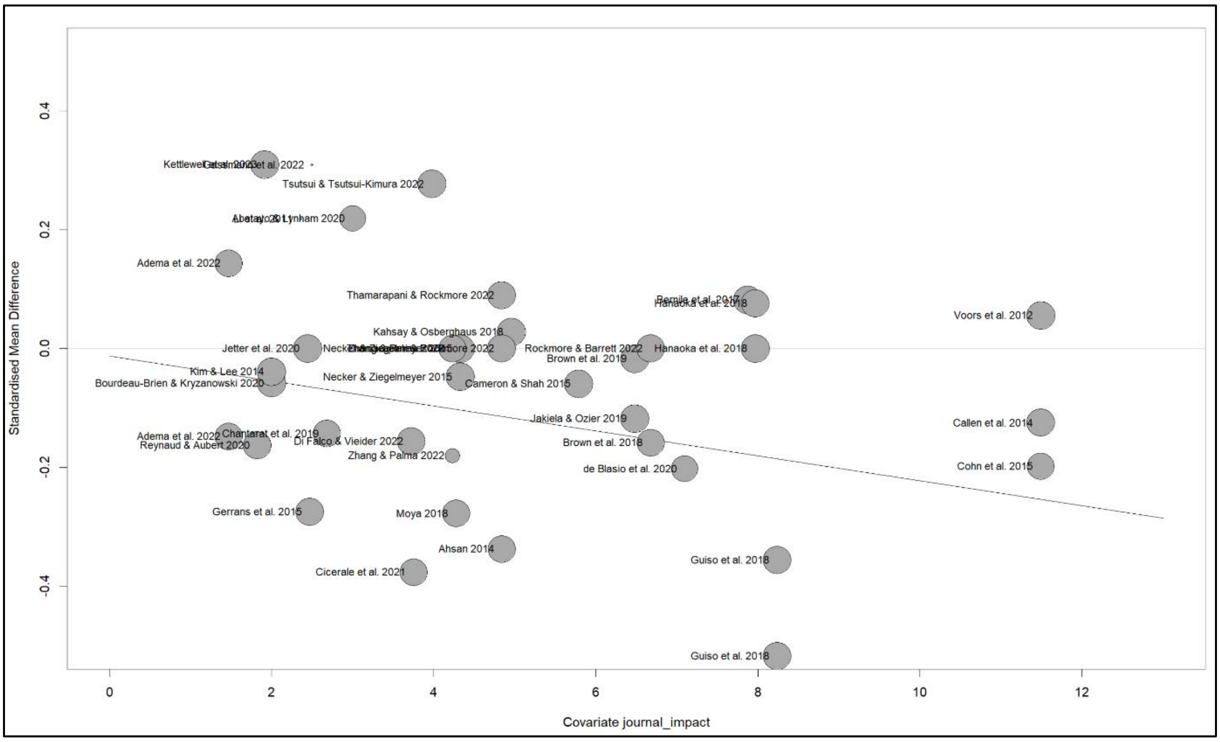
c. Quality index



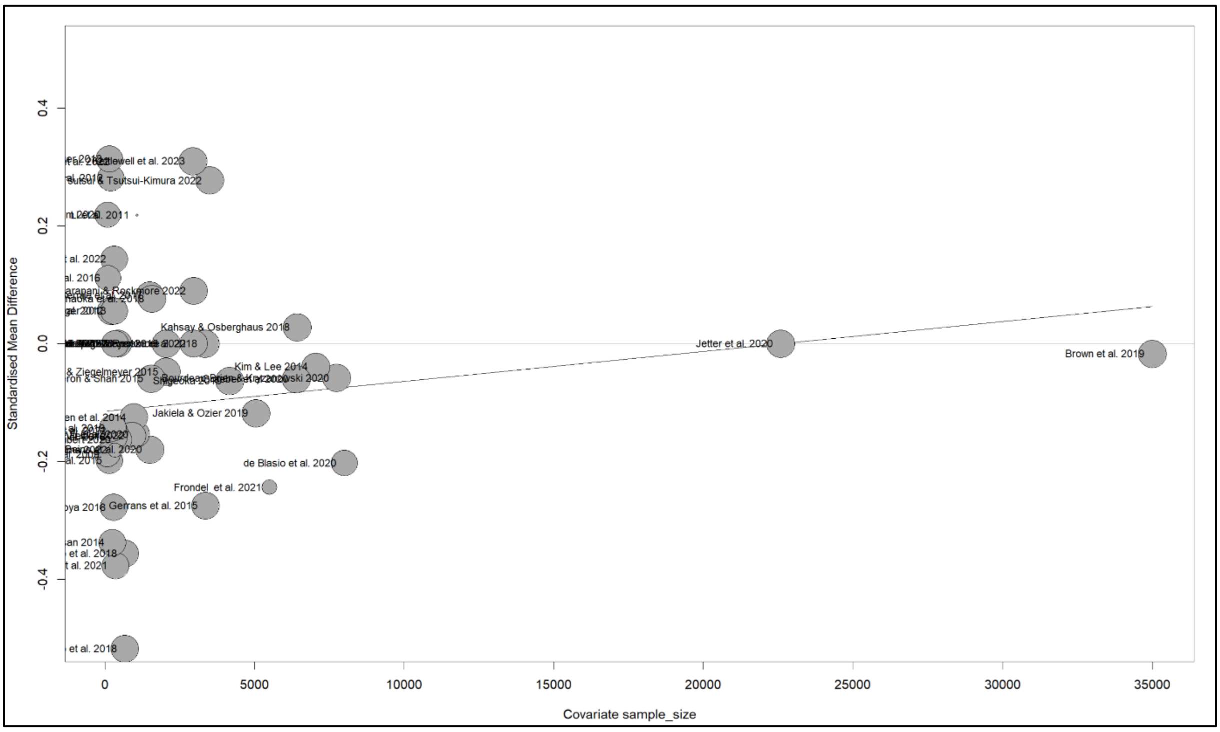
d. Article citations



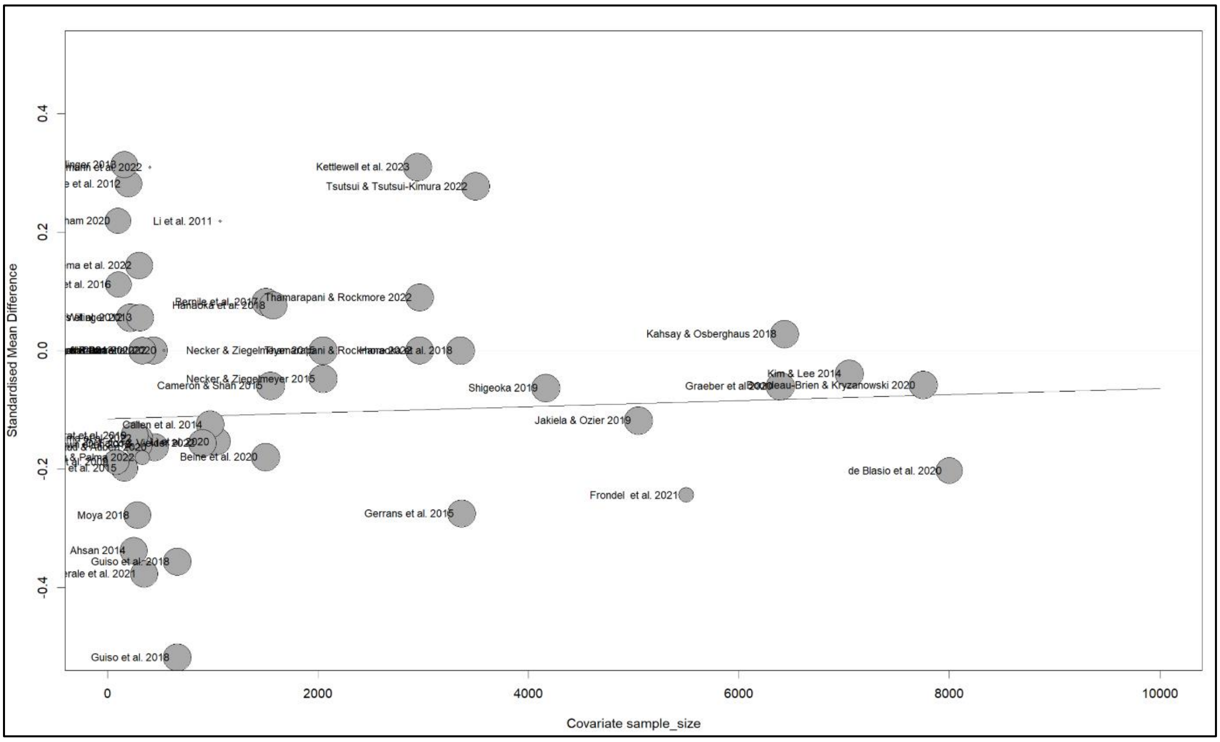
e. Journal impact



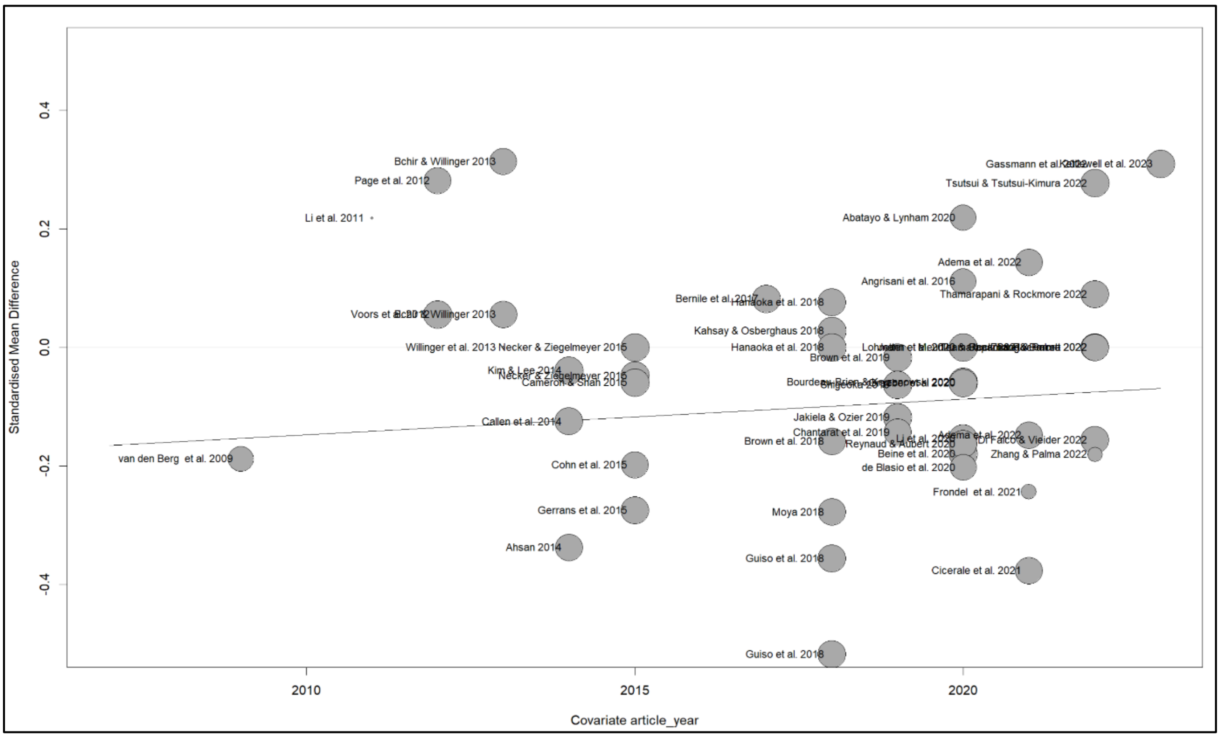
f. Sample size



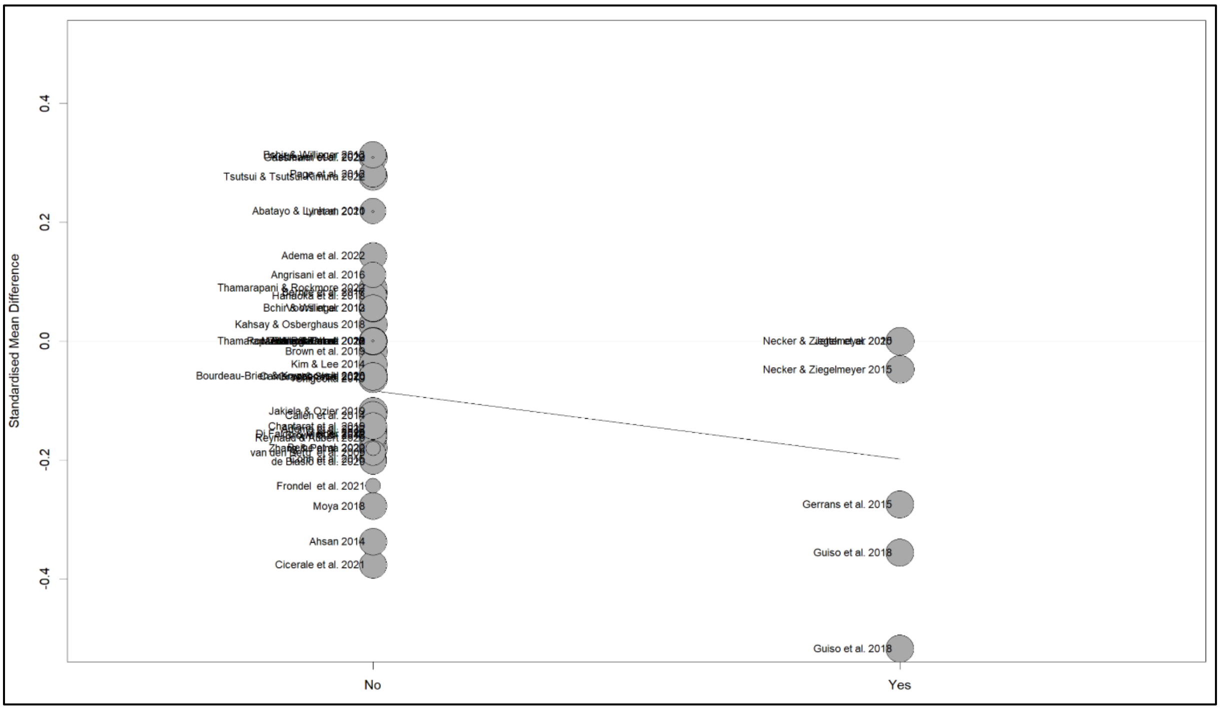
g. Sample size – zoomed in



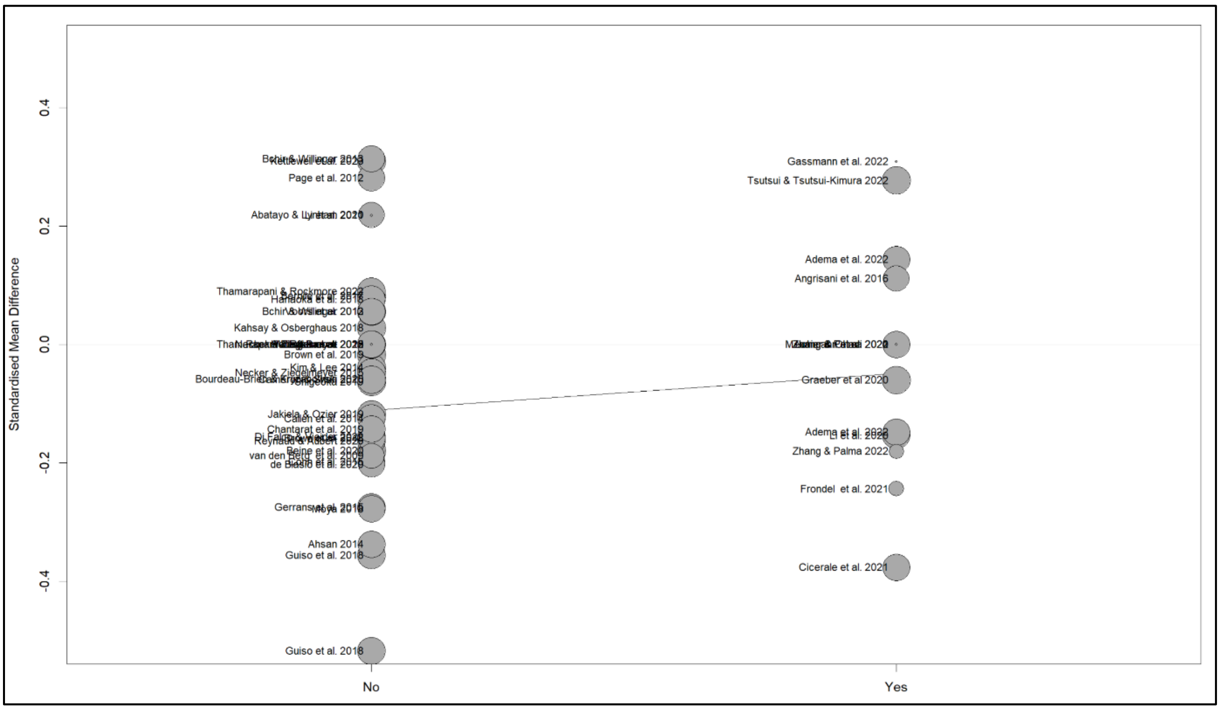
h. Year of publication



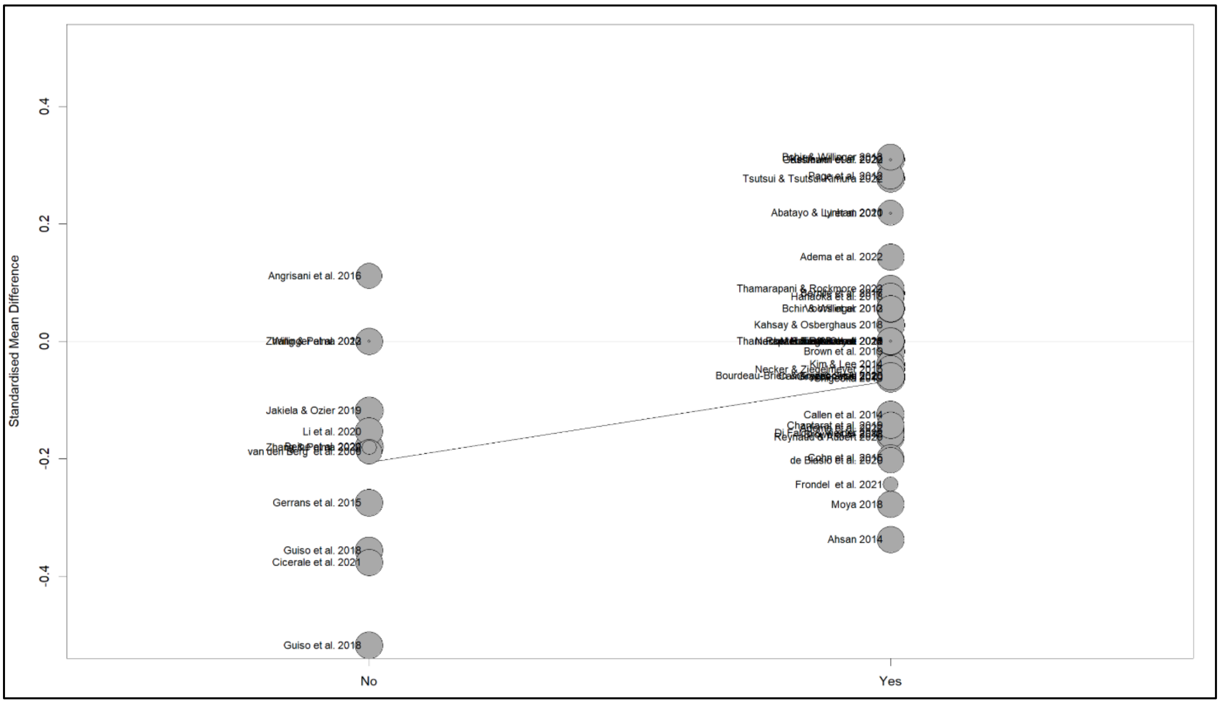
i. Global financial crisis



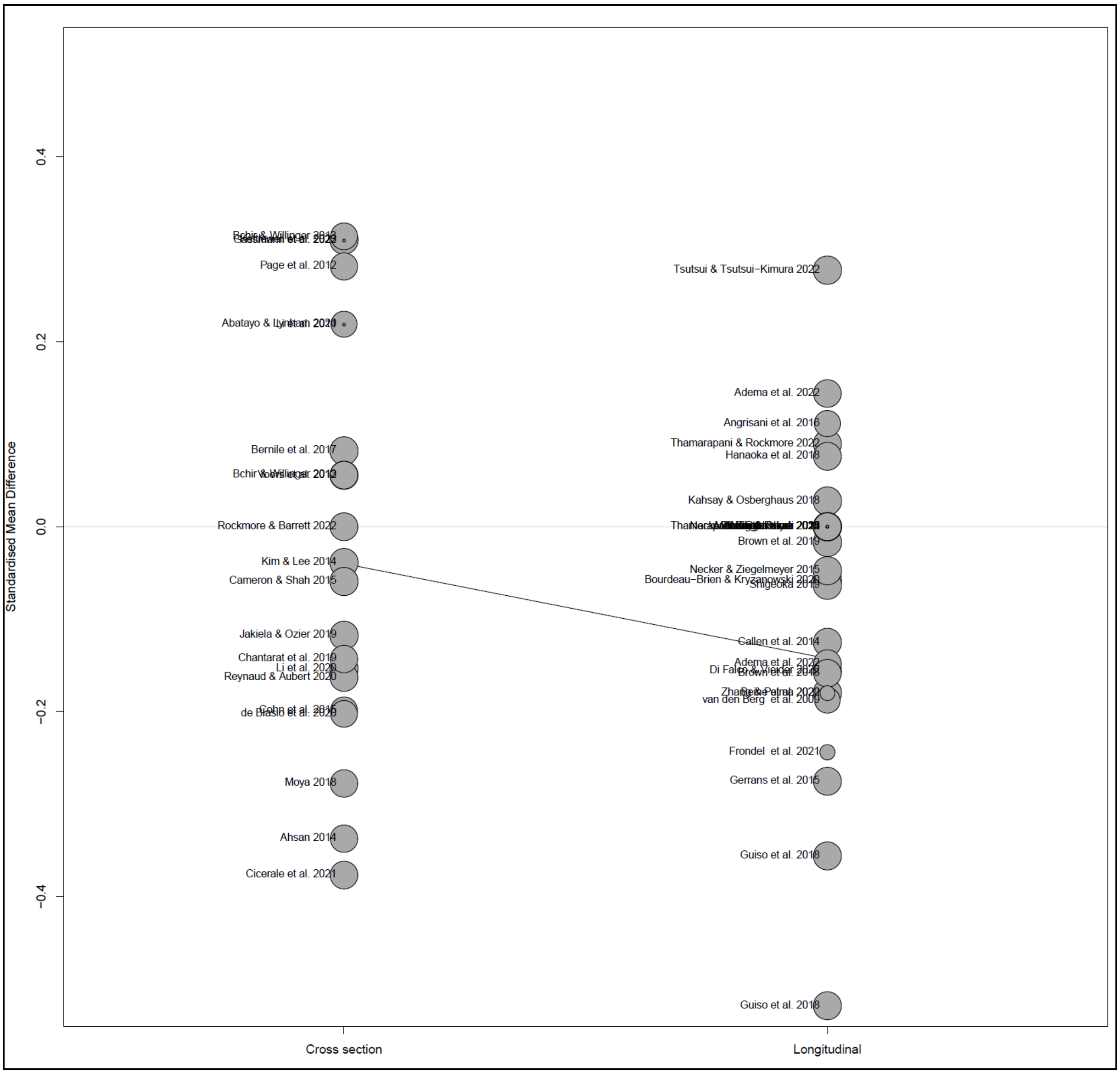
j. COVID-19



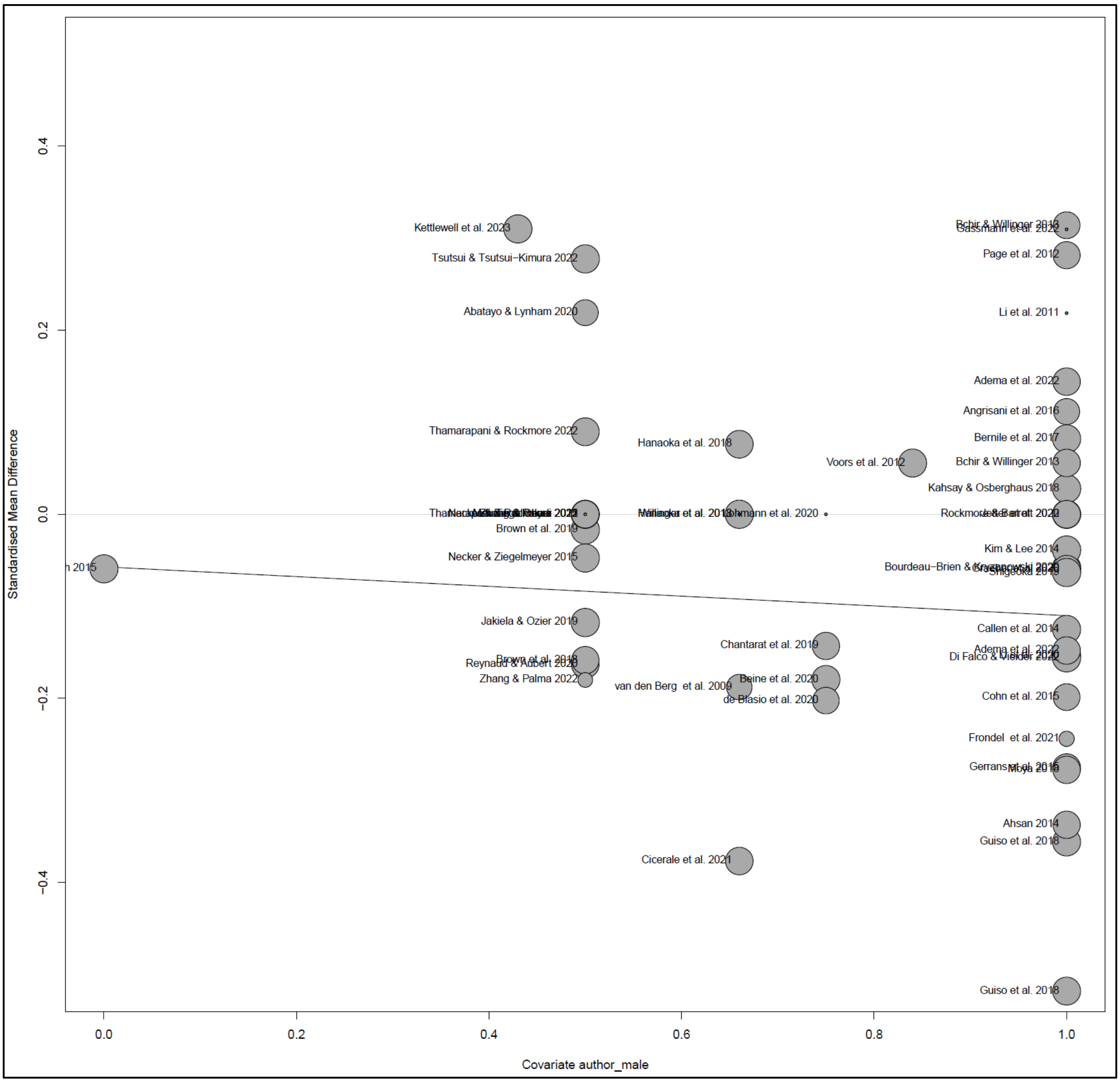
k. Causal identification


Standardised Mean Difference
No
Yes
Angrisani et al. 2016
Jakiela & Ozier 2019
Li et al. 2020
Gerrans et al. 2015
Guiso et al. 2018
Cicerale et al. 2021
Guiso et al. 2018
Tsutsui & Tsutsui-Kimura 2022
Adema et al. 2022
Kahsay & Osberghaus 2018
Brown et al. 2019
Callen et al. 2014
Frondel et al. 2021
Moya 2018
Ahsan 2014

l. Design

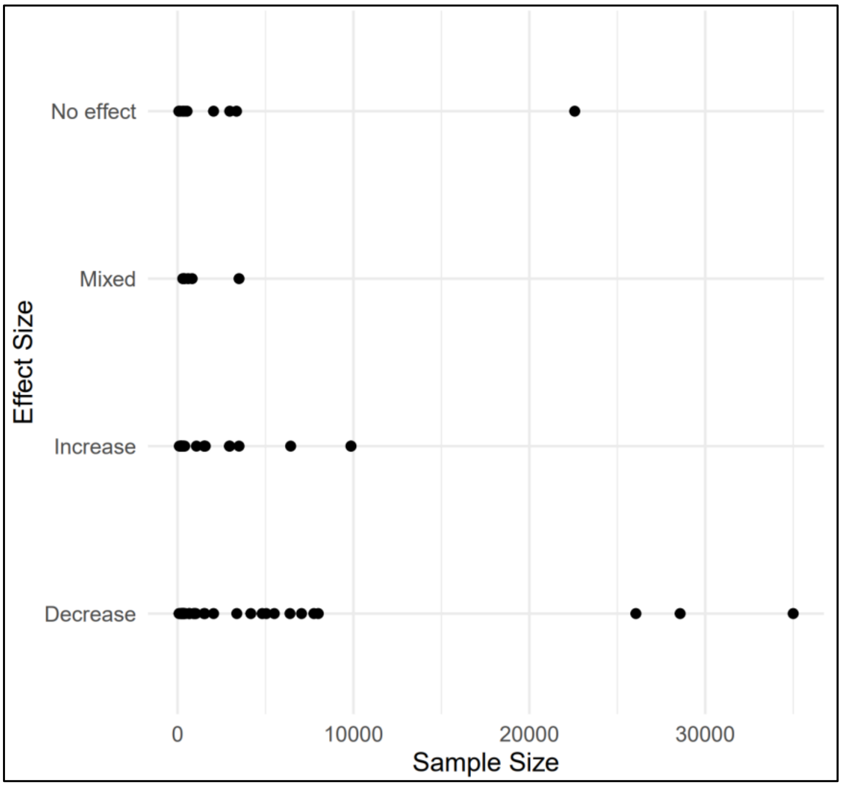


m. Author's gender (ratio of male authors)

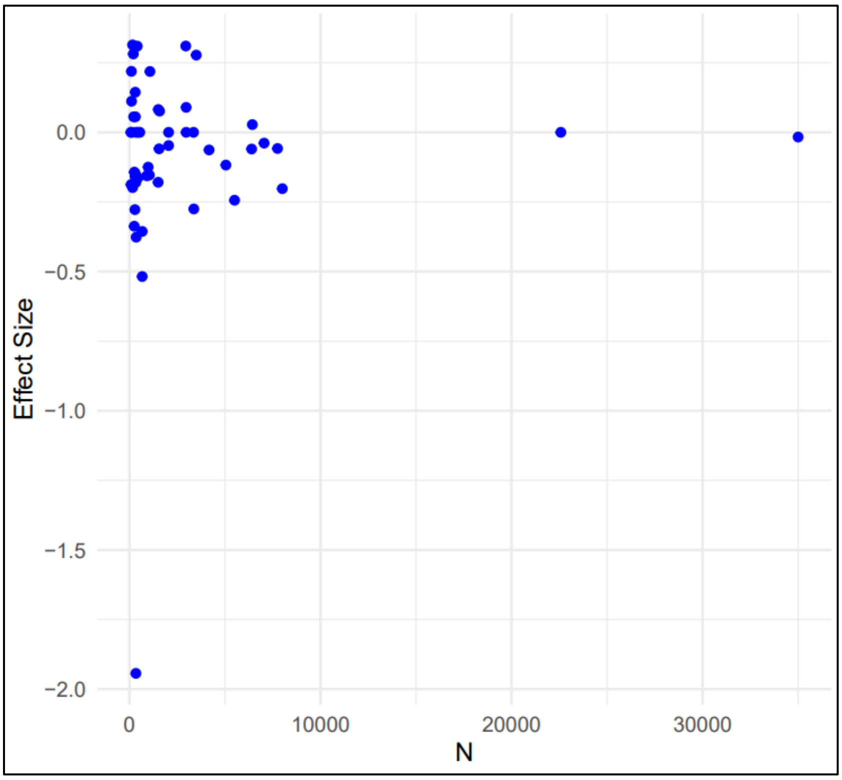

Standardised Mean Difference
Covariate author_male
0.4
0.2
0.0
−0.2
−0.4
0.0
0.2
0.4
0.6
0.8
1.0
Kettlewell et al. 2023
Tsutsui & Tsutsui−Kimura 2022
Abatayo & Lynham 2020
Thamarapani & Rockmore 2022
Hanaoka et al. 2018
Voors et al. 2012
Brown et al. 2019
Necker & Ziegelmeyer 2015
Jakiela & Ozier 2019
Chantarat et al. 2019
Reynaud & Aubert 2020
Zhang & Palma 2022
van den Berg et al. 2009
de Blasio et al. 2020
Cicerale et al. 2021
Page et al. 2012
Li et al. 2011
Adema et al. 2022
Angrisani et al. 2016
Bernile et al. 2017
Bchir & Willinger 2013
Kahsay & Osberghaus 2018
Kim & Lee 2014
Callen et al. 2014
Cohn et al. 2015
Frondel et al. 2021
Ahsan 2014
Guiso et al. 2018

3. Scatter Plots

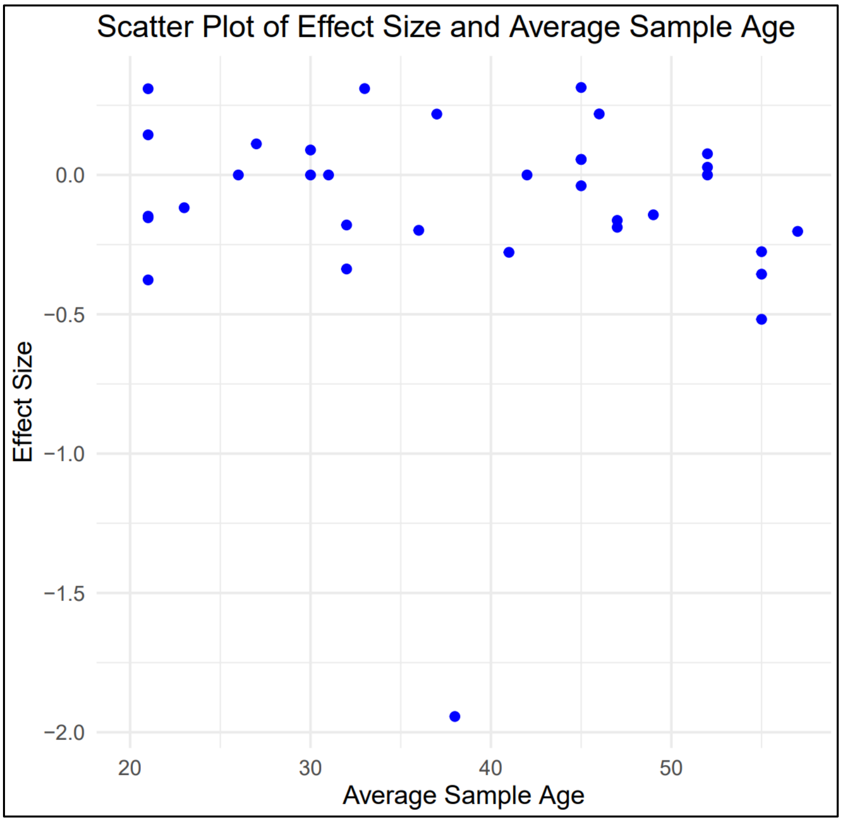
a. Sample size and effect direction



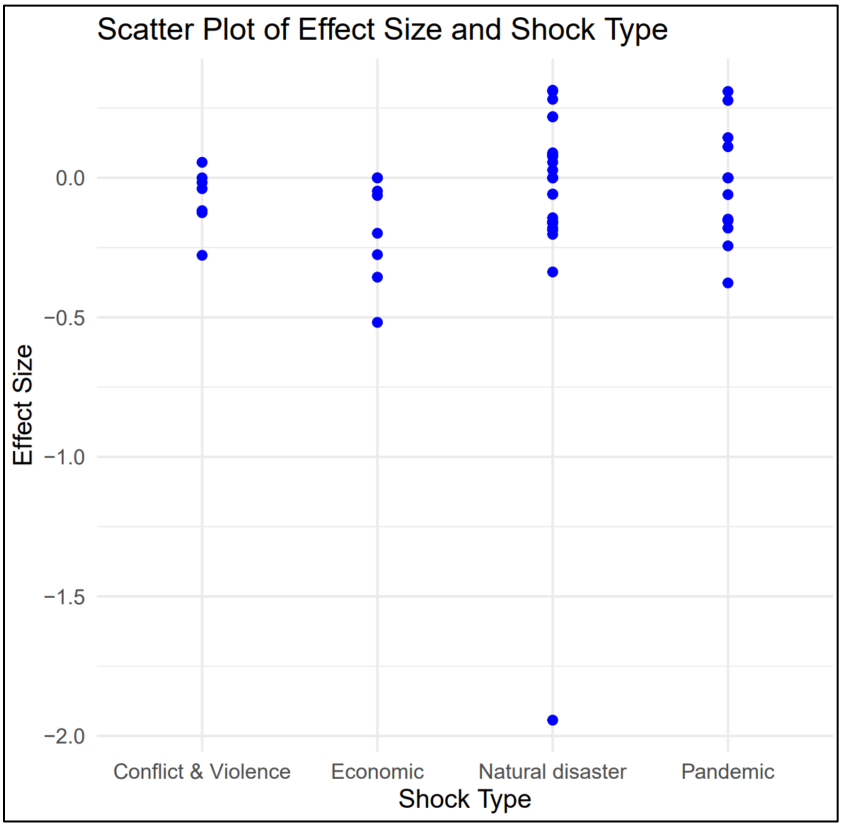
b. Sample size and effect size



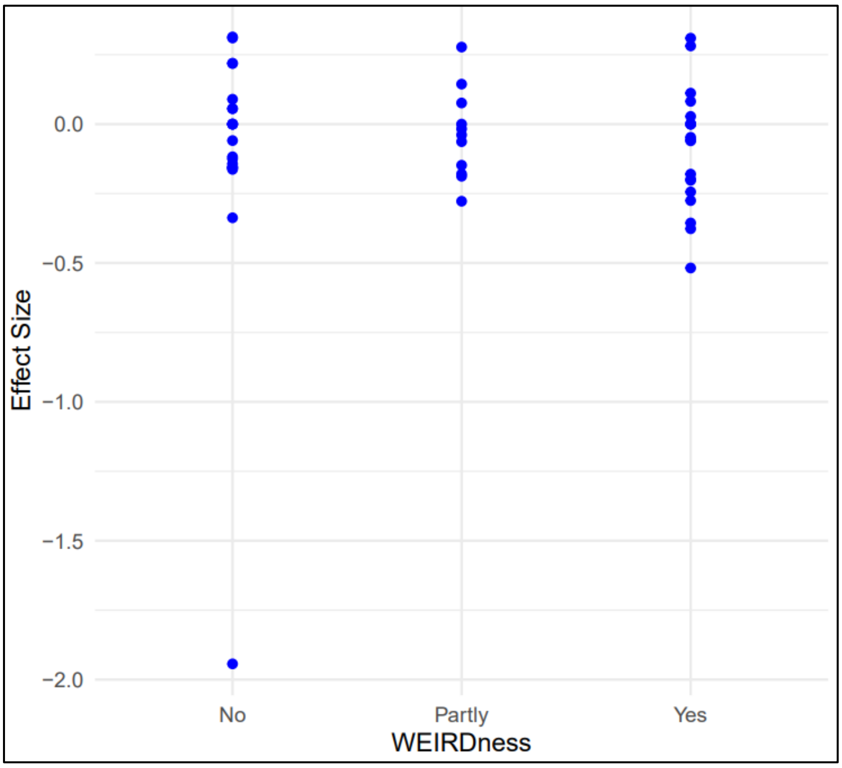
c. Sample age and effect size



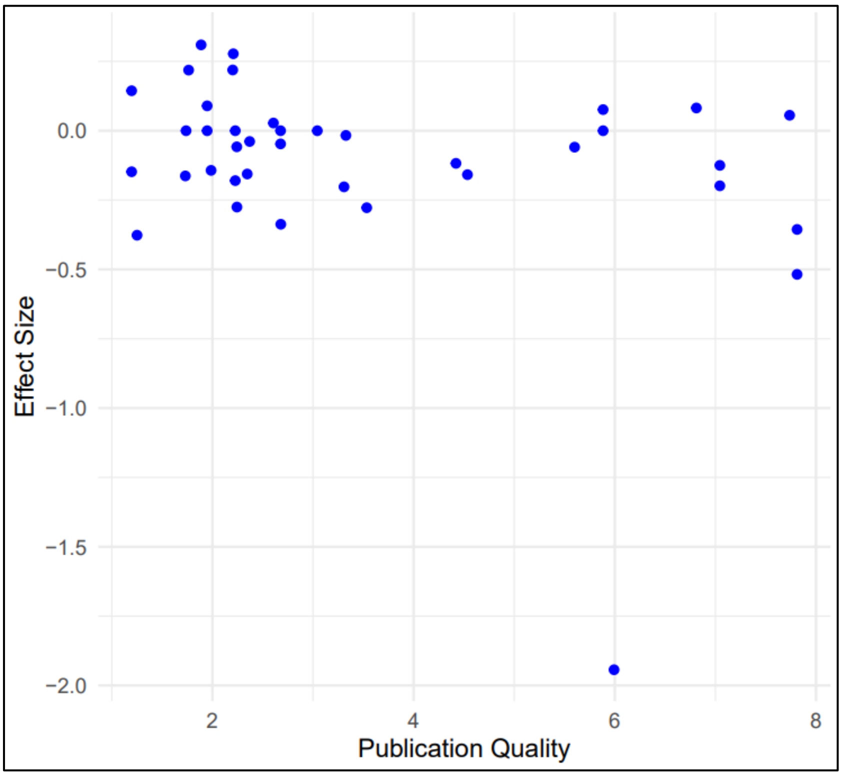
d. **Shock type and effect size**



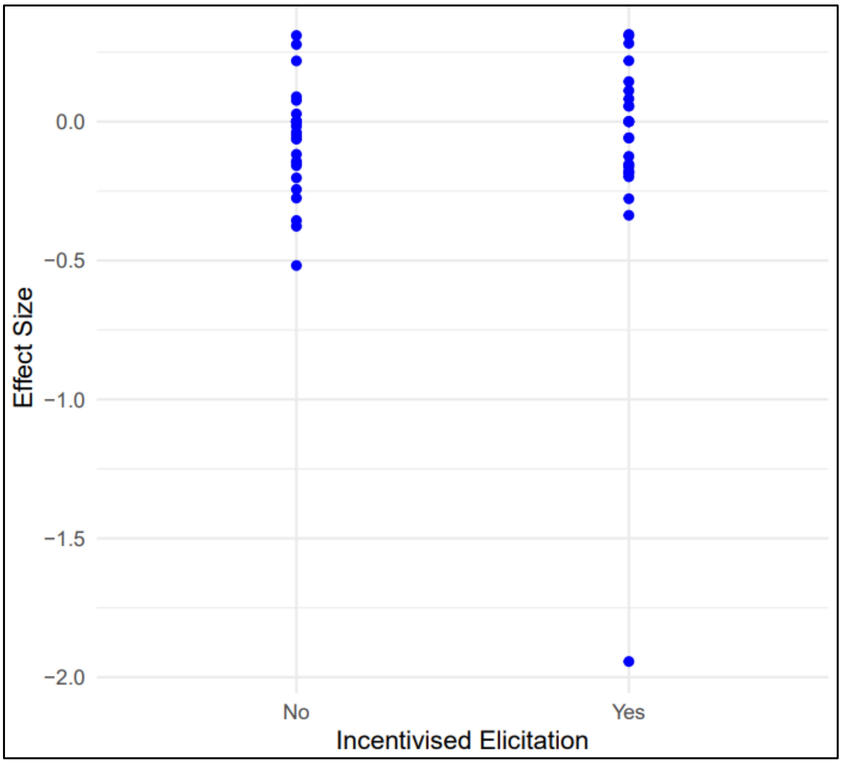
e. WEIRD sample and effect size



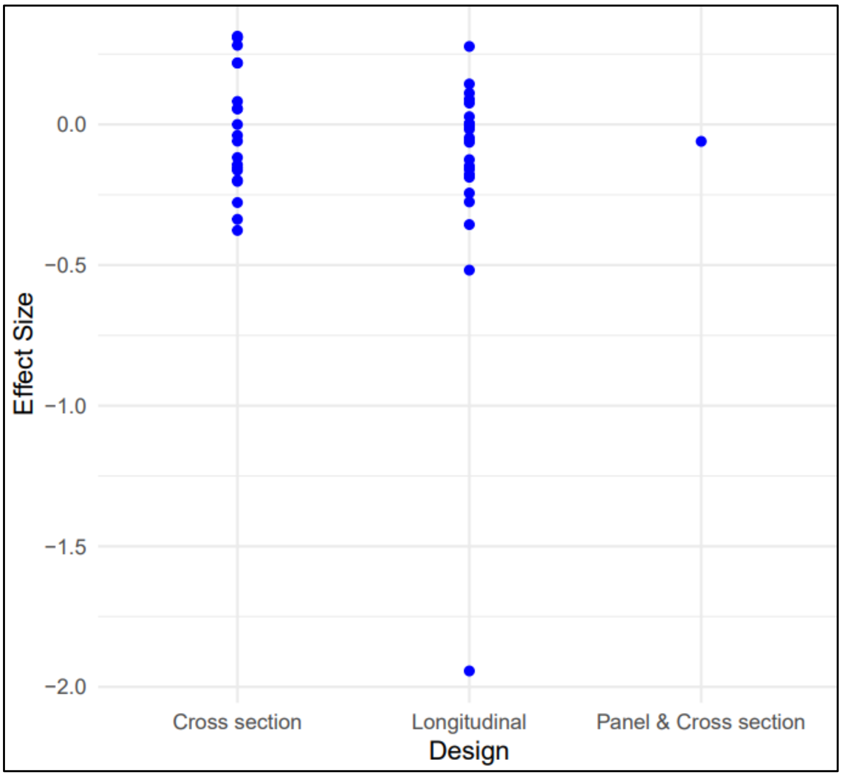
f. Publication quality and effect size



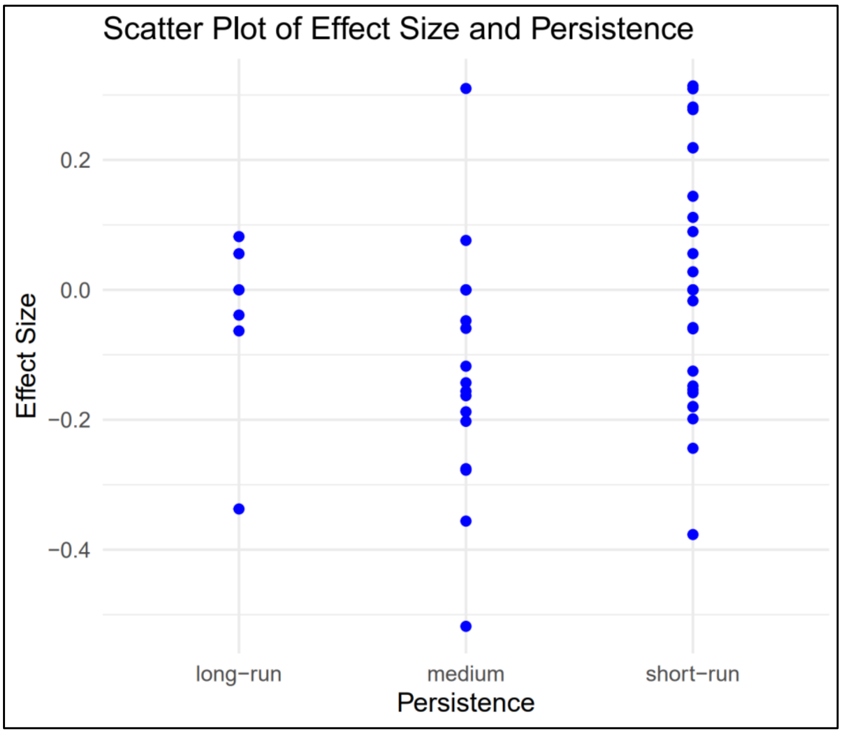
g. Measurement incentivised



h. Design and effect size



i. Persistence



ii. Tables

1. Variables in final dataset

Category	*Variable*
Authors	Names, Share of male authors, Main academic discipline
Source characteristics	Publication status, Peer review, Citations on Google Scholar and Web of Science, Journal, Journal's impact factor
Article results	Title, Year of publication, Focus
	Reaction to positive events (if applicable), Reaction to negative events, (Main) Sources of within-study-heterogeneity, Statistical tools, Reported coefficient, Standardised effect size, Categorical classification of effect size, p-value, Significant effect, Reported standard error, Standard deviation, Standardised standard error, Persistence
Additional weights	Factor for double-counting, Quality index
Sample information	Location, WEIRDness, Size, Year, Sampling lag, Share of male participants, Mean age
Design	Experiment/Survey, Longitudinal/Cross-sectional, Source of variation, Causal identification
Dependent variable	Elicitation method, Incentivised
Independent variable	Shock type, GFC, COVID-19

2. List of articles in final dataset

Article	*Status*	*Effect Direction*	*Sources of within-study-heterogeneity*	*Effect Size*	*Location*	*Sample Size*	*Sampling lag*	*Shock Type*
Abatayo & Lynham 2020	Published	Increase	gender	0.219	Philippines	100	1-3 years	Natural disaster
Adema et al. 2022	Published	Decrease	measurement	-0.148	India, Mexico, Europe	303	1 year	Pandemic
Adema et al. 2022	Published	Increase	measurement	0.14408	India, Mexico, Europe	303	1 year	Pandemic
Ahsan 2014	Published	Decrease		-0.33731	Bangladesh	250	3 years	Natural disaster
Angrisani et al. 2016	Working Paper	No effect		0.11156	USA	108	< 3 months	Pandemic
Bchir & Willinger 2013	Working Paper	Increase	income	0.314	Peru	162	< 1 year	Natural disaster
Bchir & Willinger 2013	Working Paper	No effect	income	0.05576	Peru	309		Natural disaster

Beine et al. 2020	Working Paper	Decrease	exposure intensity	-0.17956	Albania	1502	< 3 months	Natural disaster
Bernile et al. 2017	Published	Increase		0.08194	USA	1508	> 10 years	Natural disaster
Bourdeau-Brien et al. 2020	Published	Decrease		-0.05806	USA	7750	< 3 months	Natural disaster
Brown et al. 2018	Published	Decrease	ethnicity	-0.15849	Fiji	295	< 3 months	Natural disaster
Brown et al. 2019	Published	Decrease		-0.01695	Mexico	35000	< 3 months	Conflict & Violence
Callen et al. 2014	Published	Decrease		-0.12506	Afghanistan	977	< 3 months	Conflict & Violence
Cameron & Shah 2015	Published	Decrease		-0.05927	Indonesia	1550	3 years	Natural disaster
Cassar et al. 2017	Published	Decrease		-1.94342	Thailand	334	3 years	Natural disaster
Chantarat et al. 2019	Published	Decrease		-0.14303	Cambodia	256	3 years	Natural disaster
Cicerale et al. 2021	Published	Decrease		-0.3767	Italy	350	< 3 months	Pandemic
Cohn et al. 2015	Published	Decrease	certainty	-0.19862	Switzerland	162	< 3 months	Economic
de Blasio et al. 2020	Published	Decrease		-0.20241	Italy	8000	2 years	Natural disaster
Di Falco & Vieider 2022	Published	Decrease		-0.15607	Ethiopia	906	< 1 year	Natural disaster
Dohmen et al. 2016	Published	Decrease			Ukraine & Germany	26056	1-3 years	Economic
Eckel et al. 2009	Published	Mixed	gender		USA	352	1 year	Natural disaster
Fatas et al. 2021	Published	Increase	exposure intensity		Colombia	207	> 10 years	Conflict & Violence
Finger et al. 2022	Published	Increase	exposure type		Switzerland	1530	< 3 months	Natural disaster
Frondel et al. 2021	Working Paper	Decrease	exposure intensity	-0.2439	Germany	5500	< 1 year	Pandemic
Gassmann et al. 2022	Published	Increase		0.3095	France	406	< 3 months	Pandemic
Gerrans et al. 2015	Published	Decrease		-0.2753	A, NZ, NA, UK	3368	1-3 years	Economic
Graeber et al 2020	Working Paper	Decrease	gender, income	-0.06	Germany	6393	< 3 months	Pandemic
Guiso et al. 2018	Published	Decrease	gender, age, inc., educ.	-0.356	Italy	666	1-3 years	Economic
Guiso et al. 2018	Published	Decrease	age, education	-0.518	Italy	666	1-3 years	Economic
Hanaoka et al. 2018	Published	No effect	gender, time	0	Japan	3352	1-5 years	Natural disaster
Hanaoka et al. 2018	Published	Increase	gender, time	0.0761	Japan	1575	1-5 years	Natural disaster
Holden & Tilahun 2021	Working Paper	Mixed	measurement		Ethiopia	830	1-3 years	Natural disaster
Ikeda et al. 2020	Working Paper	Mixed	stakes		Japan	3495	< 1 year	Pandemic
Ingwersen et al. 2023	Working Paper	Increase	exposure type, exposure intensity		Indonesia	9860	< 1 year	Natural disaster
Ingwersen et al. 2023	Working Paper	Increase	exposure type, exposure intensity		Indonesia	9860	5 years	Natural disaster
Jakiela & Ozier 2019	Published	Decrease		-0.11757	Kenya	5047	1-3 years	Conflict & Violence
Jetter et al. 2020	Published	No effect	gender	0	Australia	22579	5 years	Economic
Kahsay & Osberghaus 2018	Published	Increase	exposure intensity	0.02777	Germany	6431	1-3 years	Natural disaster
Kettlewell 2019	Published	Decrease	type of exposure, time		Australia	4810	< 1 year	Economic
Kettlewell et al. 2023	Published	Increase		0,31	Sri Lanka	2946	1-3 years	Natural disaster
Kim & Lee 2014	Published	Decrease	exposure intensity	-0.03883	Korea	7047	> 10 years	Conflict & Violence
Kuroishi & Sawada 2020	Working Paper	Increase			Japan & Philippines	344	6 years	Natural disaster
Li et al. 2011	Published	Increase		0.2185	China	1072	< 3 months	Natural disaster
Li et al. 2020	Working Paper	Decrease	age	-0.1536	China	1040	< 3 months	Pandemic
Lohmann et al. 2020	Working Paper	No effect		0	China	539	< 3 months	Pandemic
Malmendier & Nagel 2011	Published	Decrease			USA	28571	> 10 years	Economic
Meunier & Ohadi 2021	Published	No effect	domain	0	USA, UK, A, EU	72	< 3 months	Pandemic
Moya 2018	Published	Decrease	time, domain, certainty	-0.27764	Colombia	284	1-5 years	Conflict & Violence
Necker & Ziegelmeyer 2015	Published	Decrease	exposure type	-0.04753	Germany	2047	1-3 years	Economic
Necker & Ziegelmeyer 2015	Published	No effect	exposure type	0	Germany	2047	1-3 years	Economic
Page et al. 2012	Published	Increase		0.28144	Australia	202	< 3 months	Natural disaster
Reynaud & Aubert 2020	Published	Decrease	domain	-0.163	Vietnam	448	1-5 years	Natural disaster
Rockmore & Barrett 2022	Published	No effect	exposure type	0	Uganda	442	> 10 years	Conflict & Violence
Said et al. 2015	Published	Mixed	time, exposure intensity		Pakistan	384	1-3 years	Natural disaster
Shachat et al. 2021	Published	Mixed	domain		China	602	< 3 months	Pandemic
Shigeoka 2019	Working Paper	Decrease		-0.06324	Japan	4165	> 10 years	Economic
Shupp et al. 2017	Published	Mixed	exposure type		USA	295	< 3 months	Natural disaster
Thamarapani & Rockmore 2022	Published	Increase		0.08963	Indonesia	2966	< 1 year	Natural disaster
Thamarapani & Rockmore 2022	Published	No effect		0	Indonesia	2966	1-3 years	Natural disaster
Tsutsui & Tsutsui-Kimura 2022	Published	Increase	time	0.27743	Japan	3495	< 1 year	Pandemic
van den Berg et al. 2009	Working Paper	Decrease		-0.18775	Nicaragua & Peru	84	1-3 years	Natural disaster
Voors et al. 2012	Published	Increase	domain	0.05565	Burundi	220	> 10 years	Conflict & Violence
Willinger et al. 2013	Working Paper	No effect		0	Indonesia	131	< 1 year	Natural disaster
Zhang & Palma 2022	Published	Decrease	gender, measurement	-0.18	USA	331	< 1 year	Pandemic
Zhang & Palma 2022	Published	No effect	gender, measurement	0	USA	331	< 1 year	Pandemic

3. List of excluded articles and brief justifications

Article	*Source*	*Justification for exclusion*
Akesaka et al. 2021	NBER Working Paper No. 28784	No focus on stability after shocks
Bucciol & Miniaci 2018	Oxford Bulletin of Economics and Statistics	No focus on stability after shocks
Ert & Haruvy 2017	Economic Letters	No focus on stability after shocks
Love & Robinson 1984	Southern Journal of Agricultural Economics	too old and only focus on stability
Reynaud & Couture 2012	Theory	Focus only on measurement
Angerer et al. 2021	ESifo Working Papers	No focus on stability after shocks
Meier 2022	American Economic Journal: Applied Economics	Not in the field
Krčál et al. 2019	Research in Economics	No focus on stability after shocks
l'Haridon & Vieder 2019	Quantitative Economics	No focus on stability after shocks
Schmidt et al. 2019	Theory and Decision	No focus on stability after shocks
Haile et al. 2020	World Development	No focus on stability after shocks
Harrison et al. 2020	The Review of Economics and Statistics	No focus on stability after shocks
Einav et al. 2012	American Economic Review	No focus on stability after shocks
Castillo et al. 2017	Theory	No focus on stability after shocks
Cahlikova & Cingl 2017	Experimental Economics	Not in the same field, no focus on stability after shocks
Cobb-Clark et al. 2019	The Journal of Human Resources	No focus on stability after shocks
Barseghyan et al. 2011	The American Economic Review	No focus on stability after shocks
Chetty & Szeidl 2007	Quarterly Journal of Economics	No focus on stability after shocks
Levin et al. 2007	Journal of Behavioral Decision Making	No focus on stability after shocks
Wehrung et al. 1984	INFOR	No focus on stability after shocks
Andersen et al. 2008	International Economic Review	No focus on stability after shocks
Josef et al. 2016	Journal of Personality and Social Psychology	Focus on long-term stability over the lifetime
Callen 2014	Journal of Economic Behavior & Organization	Focus on time preferences
Shavit et al. 2014	Journal of Economic Psychology	Focus on time preferences & no exogenous influence
Brunnermeier & Nagel 2008	American Economic Review	No focus on stability after shocks
Dohmen et al. 2011	Journal of the European Economic Association	No focus on stability after shocks
Huang et al. 2013	Proceedings of the National Academy of Sciences	No direct exposure to shock. Risk measures only shock related
Gao et al. 2020	Journal of Financial Economics	No risky decisions. Focus on risk perception.
Goebel et al. 2015	Journal of Population Economics	Focus only on disaster related risk perception/attitudes
Filipski et al. 2019	European Economic Review	No focus on risk preferences but on outcomes
Weber et al. 2013	Review of Finance	No focus on stability after shocks
Hoffmann et al. 2013	Journal of Banking & Finance	Focus on behavioural outcomes

Decker & Schmitz 2016	Journal of Health Economics	No focus on stability after external shocks
Li et al. 2008	Applied Cognitive Psychology	No access to publication

4. Effect direction

item	count	percent	cum_count	cum_percent
Decrease	31	0.47	31	0.47
Increase	18	0.27	49	0.74
No effect	11	0.17	60	0.91
Mixed	6	0.09	66	1.00

5. Effect size

var	type	label	n	NA.prc	mean	sd	se	md	trimmed	range	iqr	skew
dd	numeric	dd	52	21.21	-0.08	0.32	0.04	-0.04	-0.05	2.26 (-1.94-0.31)	0.22	-3.84

6. Effect size categorical

item	count	percent	cum_count	cum_percent
small	27	0.41	27	0.41
no effect	11	0.17	38	0.58
very small	11	0.17	49	0.74
mixed	5	0.08	54	0.82
small-to-medium	5	0.08	59	0.89
NA	4	0.06	63	0.95
large	1	0.02	64	0.97
medium	1	0.02	65	0.98
very large	1	0.02	66	1.00

7. Significant effects

item	count	percent	cum_count	cum_percent
yes	56	0.85	56	0.85
no	10	0.15	66	1.00

8. Shock type

item	count	percent	cum_count	cum_percent
Natural disaster	32	0.48	32	0.48
Pandemic	15	0.23	47	0.71
Economic	11	0.17	58	0.88
Conflict & Violence	8	0.12	66	1.00

9. Persistence

item	count	percent	cum_count	cum_percent
short-run	32	0.48	32	0.48
medium	23	0.35	55	0.83
long-run	11	0.17	66	1.00

10. Sampling lag

item	count	percent	cum_count	cum_percent
< 3 months	18	0.28	18	0.28
1-3 years	14	0.22	32	0.49
< 1 year	11	0.17	43	0.66
> 10 years	7	0.11	50	0.77
1-5 years	4	0.06	54	0.83
3 years	4	0.06	58	0.89
1 year	3	0.05	61	0.94
5 years	2	0.03	63	0.97
2 years	1	0.02	64	0.98
6 years	1	0.02	65	1.00

11. Design

item	count	percent	cum_count	cum_percent
Longitudinal	40	0.61	40	0.61
Cross section	25	0.38	65	0.98
Panel & Cross section	1	0.02	66	1.00

12. Sample locations

item	count	percent	cum_count	cum_percent
USA	8	0.12	8	0.12
Indonesia	6	0.09	14	0.21
Germany	5	0.08	19	0.29
Japan	5	0.08	24	0.36
China	4	0.06	28	0.42
Italy	4	0.06	32	0.48
Australia	3	0.05	35	0.53
Colombia	2	0.03	37	0.56
Czechia, India, Mexico, Spain	2	0.03	39	0.59
Ethiopia	2	0.03	41	0.62
Peru	2	0.03	43	0.65
Switzerland	2	0.03	45	0.68
Afghanistan	1	0.02	46	0.70
Albania	1	0.02	47	0.71
Australia &NZ (50.2%), North America (33.2%), UK (15.8%), Others (0.8%)	1	0.02	48	0.73
Bangladesh	1	0.02	49	0.74
Burundi	1	0.02	50	0.76
Cambodia	1	0.02	51	0.77
Fiji	1	0.02	52	0.79
France	1	0.02	53	0.80
Japan & Philippines	1	0.02	54	0.82
Kenya	1	0.02	55	0.83
Korea	1	0.02	56	0.85
Mexico	1	0.02	57	0.86
Nicaragua & Peru	1	0.02	58	0.88
Pakistan	1	0.02	59	0.89
Philippines	1	0.02	60	0.91
Sri Lanka	1	0.02	61	0.92
Thailand	1	0.02	62	0.94
Uganda	1	0.02	63	0.95
Ukraine & Germany	1	0.02	64	0.97
USA, UK, Australia, Europe	1	0.02	65	0.98
Vietnam	1	0.02	66	1.00

13. Sample size descriptive statistics

Sample Size Descriptive Statistics

var	type	label	n	NA.prc	mean	sd	se	md	trimmed	range	iqr	skew
dd	numeric	dd	66	0	3578.47	6831.4	840.89	868	1918.44	34928 (72-35000)	3158.75	3.2

iv. R-Code

```
### PACKAGES

library(tidyverse)
library(readr)
library(dplyr)
library(tidyr)
library(ggplot2)
library(meta)
library(metafor)
library(readxl)
library(kableExtra)
library(stargazer)
library(psych)
library(effects)
library(AER)
library(stargazer)
library(knitr)
library(kableExtra)
library(gtsummary)
library(forestplot)
library(freqtables)
library(gridExtra)
library(clean)

# IMPORT DATASET

setwd("~/2_Studium/1_Tokyo/Academics/Thesis/Data")

library(readxl)
shocksandrisk <- read_excel("shocksandrisk.xlsx")
View(shocksandrisk)

# modify variables
shocksandrisk$author_male <- as.numeric(shocksandrisk$author_male)

shocksandrisk$publication_status <- as.factor(shocksandrisk$publication_status)
shocksandrisk$article_peer_reviewed <- as.factor(shocksandrisk$article_peer_reviewed)
shocksandrisk$positive_event_effect <- as.factor(shocksandrisk$positive_event_effect)
shocksandrisk$negative_event_effect <- as.factor(shocksandrisk$negative_event_effect)
shocksandrisk$main_stat_method <- as.factor(shocksandrisk$main_stat_method)
shocksandrisk$effect_size_categorical <- as.factor(shocksandrisk$effect_size_categorical)
shocksandrisk$effect_significant <- as.factor(shocksandrisk$effect_significant)
shocksandrisk$effect_persistence <- as.factor(shocksandrisk$effect_persistence)
shocksandrisk$effect_size_categorical <- as.factor(shocksandrisk$effect_size_categorical)
shocksandrisk$inclusion_direction <- as.factor(shocksandrisk$inclusion_direction)
shocksandrisk$inclusion_effect_size_categorical <-
as.factor(shocksandrisk$inclusion_effect_size_categorical)
shocksandrisk$inclusion_effect_size_quantitative <-
as.factor(shocksandrisk$inclusion_effect_size_quantitative)
shocksandrisk$sample_lag <- as.factor(shocksandrisk$sample_lag)
shocksandrisk$design_general <- as.factor(shocksandrisk$design_general)
shocksandrisk$design_variation_source <- as.factor(shocksandrisk$design_variation_source)
```

```
shocksandrisk$design_long_cross <- as.factor(shocksandrisk$design_long_cross)
shocksandrisk$design_identification_causal <- as.factor(shocksandrisk$design_identification_causal)
shocksandrisk$depvar_incentivised <- as.factor(shocksandrisk$depvar_incentivised)
shocksandrisk$indepvar_shock_type <- as.factor(shocksandrisk$indepvar_shock_type)
shocksandrisk$indepvar_shock_type_covid19 <-
as.factor(shocksandrisk$indepvar_shock_type_covid19)
shocksandrisk$indepvar_shock_type_gfc <- as.factor(shocksandrisk$indepvar_shock_type_gfc)
shocksandrisk$WEIRD <- as.factor(shocksandrisk$WEIRD)


levels(shocksandrisk$publication_status)
levels(shocksandrisk$article_peer_reviewed)
levels(shocksandrisk$positive_event_effect)
levels(shocksandrisk$negative_event_effect)
levels(shocksandrisk$effect_size_categorical)
levels(shocksandrisk$effect_significant)
levels(shocksandrisk$effect_persistence)
levels(shocksandrisk$effect_size_categorical)
levels(shocksandrisk$inclusion_direction)
levels(shocksandrisk$inclusion_effect_size_categorical)
levels(shocksandrisk$inclusion_effect_size_quantitative)
levels(shocksandrisk$sample_lag)
levels(shocksandrisk$design_general)
levels(shocksandrisk$design_variation_source)
levels(shocksandrisk$design_identification_causal)
levels(shocksandrisk$depvar_incentivised)
levels(shocksandrisk$indepvar_shock_type)
levels(shocksandrisk$indepvar_shock_type_covid19)
levels(shocksandrisk$indepvar_shock_type_gfc)
```

DESCRIPTIVE STATISTICS

```
shocksandrisk_subset <- shocksandrisk %>%
  select(article_author_year, journal_name, publication_status, negative_event_effect, heterogeneity,
effect_size, sample_size, sample_location, sample_lag, design_long_cross, depvar_incentivised,
indepvar_shock_type) %>%
  rename("Study" = article_author_year,
        "Journal" = journal_name,
        "Status" = publication_status,
        "Effect Direction" = negative_event_effect,
        "Heterogeneity" = heterogeneity,
        "Effect Size" = effect_size,
        "N" = sample_size,
        "Location" = sample_location,
        "Lag" = sample_lag,
        "Design" = design_long_cross,
        "Incentivised" = depvar_incentivised,
        "Shock Type" = indepvar_shock_type)


shocksandrisk_table <- kable(shocksandrisk_subset, format = "html", digits = 3) %>%
  kable_styling(latex_options = c("striped", "hold_position"), font_size = 11) %>%
  column_spec(1, bold = TRUE) %>%
  column_spec(2:11, width = "3cm")
```

```
shocksandrisk_table


shocksandrisk_results <- shocksandrisk %>%
  select(article_author_year, negative_event_effect, effect_size, sample_size, sample_location,
design_long_cross, depvar_incentivised, indepvar_shock_type) %>%
  rename("Study" = article_author_year,
         "Effect Direction" = negative_event_effect,
         "Effect Size" = effect_size,
         "N" = sample_size,
         "Location" = sample_location,
         "Design" = design_long_cross,
         "Incentivised" = depvar_incentivised,
         "Shock Type" = indepvar_shock_type)


shocksandrisk_table_results <- kable(shocksandrisk_results, format = "html", digits = 3) %>%
  kable_styling(latex_options = c("striped", "hold_position"), font_size = 11) %>%
  column_spec(1, bold = TRUE) %>%
  column_spec(2:8, width = "4cm")

shocksandrisk_table_results
```

#results descriptives

```
decriptives_effect_size <- descr(shocksandrisk$effect_size)
kable(decriptives_effect_size, caption = "Sample Size Descriptive Statistics", align = "c", digits = 2)
%>%
  kable_styling(bootstrap_options = "striped", full_width = FALSE) %>%
  column_spec(1, bold = TRUE) %>%
  column_spec(2, width = "2cm") %>%
  row_spec(0, bold = TRUE) %>%
  row_spec(1:nrow(decriptives_effect_size))


frequencies_effect <- freq(shocksandrisk$negative_event_effect)
kable(frequencies_effect, caption = "Frequencies of effect directions", align = "c", digits = 2) %>%
  kable_styling(bootstrap_options = "striped", full_width = FALSE) %>%
  column_spec(1, bold = TRUE) %>%
  column_spec(2:3, width = "2cm") %>%
  row_spec(0, bold = TRUE) %>%
  row_spec(1:nrow(frequencies_effect))
ggplot(shocksandrisk, aes(x=negative_event_effect)) +
  geom_bar() +
  labs(title="Frequencies of effect directions", x="Effect Direction", y="Frequencies") +
  theme_minimal()


frequencies_effect_categorical <- freq(shocksandrisk$effect_size_categorical)
kable(frequencies_effect_categorical, caption = "Frequencies of effect sizes (categorical)", align = "c",
digits = 2) %>%
  kable_styling(bootstrap_options = "striped", full_width = FALSE) %>%
  column_spec(1, bold = TRUE) %>%
  column_spec(2:3, width = "2cm") %>%
  row_spec(0, bold = TRUE) %>%
```

```
  row_spec(1:nrow(frequencies_effect_categorical))
ggplot(shocksandrisk, aes(x=effect_size_categorical)) +
  geom_bar() +
  labs(title="Frequencies of effect size (categorical)", x="Categorical Effect Size", y="Frequencies") +
  theme_minimal()

frequencies_significant_results <- freq(shocksandrisk$effect_significant)
kable(frequencies_significant_results, caption = "Frequencies of significant results", align = "c", digits
= 2) %>%
  kable_styling(bootstrap_options = "striped", full_width = FALSE) %>%
  column_spec(1, bold = TRUE) %>%
  column_spec(2:3, width = "2cm") %>%
  row_spec(0, bold = TRUE) %>%
  row_spec(1:nrow(frequencies_significant_results))

frequencies_effect_persistence <- freq(shocksandrisk$effect_persistence)
kable(frequencies_effect_persistence, caption = "Frequencies of effect persistence", align = "c", digits
= 2) %>%
  kable_styling(bootstrap_options = "striped", full_width = FALSE) %>%
  column_spec(1, bold = TRUE) %>%
  column_spec(2:3, width = "2cm") %>%
  row_spec(0, bold = TRUE) %>%
  row_spec(1:nrow(frequencies_effect_persistence))
ggplot(shocksandrisk, aes(x=effect_persistence)) +
  geom_bar() +
  labs(title="Frequencies of effect persistence", x="Persistence", y="Frequencies") +
  theme_minimal()
```

#publication descriptives

```
shocksandrisk_publication <- shocksandrisk %>%
  select(article_author_year, journal_name, publication_status, article_peer_reviewed, author_male,
article_citations_gs, article_citations_wos, journal_name, journal_impact) %>%
  rename("Study" = article_author_year,
         "Journal" = journal_name,
         "Status" = publication_status,
         "Peer Reviewed" = article_peer_reviewed,
         "Male share of authors" = author_male,
         "Google Scholar Citations" = article_citations_gs,
         "Web of Science Citations" = article_citations_wos,
         "Journal impact factor" = journal_impact)

shocksandrisk_publication_table <- kable(shocksandrisk_publication, format = "html", digits = 3)
%>%
  kable_styling(latex_options = c("striped", "hold_position"), font_size = 11) %>%
  column_spec(1, bold = TRUE)%>%
  column_spec(3:8, width = "3cm")

shocksandrisk_publication_table
```

#sample descriptives

```
shocksandrisk_sample <- shocksandrisk %>%
```

```
  select(article_author_year, sample_year, sample_size, sample_location, WEIRD, sample_lag,
sample_male, sample_age) %>%
  rename("Study" = article_author_year,
         "Sample Size (N)" = sample_size,
         "Location" = sample_location,
         "WEIRD" = WEIRD,
         "Sampling Year" = sample_year,
         "Lag" = sample_lag,
         "Male Share" = sample_male,
         "Average Age" = sample_age)

shocksandrisk_sample_table <- kable(shocksandrisk_sample, format = "html", digits = 3) %>%
  kable_styling(latex_options = c("striped", "hold_position"), font_size = 11) %>%
  column_spec(1, bold = TRUE)%>%
  column_spec(2:8, width = "3cm")

shocksandrisk_sample_table


decriptives_sample_size <- descr(shocksandrisk$sample_size)
kable(decriptives_sample_size, caption = "Sample Size Descriptive Statistics", align = "c", digits = 2)
%>%
  kable_styling(bootstrap_options = "striped", full_width = FALSE) %>%
  column_spec(1, bold = TRUE) %>%
  column_spec(2, width = "2cm") %>%
  row_spec(0, bold = TRUE) %>%
  row_spec(1:nrow(decriptives_sample_size)) %>%
  footnote(general = "")
ggplot(shocksandrisk, aes(x = sample_size)) +
  geom_histogram(binwidth = 100)
ggplot(shocksandrisk, aes(x = sample_size)) +
  geom_histogram(binwidth = 1000)

ggplot(shocksandrisk, aes(x=sample_size, y=negative_event_effect)) +
  geom_point() +
  labs(title="Sample size distribution", x="Sample Size", y="Effect Size") +
  theme_minimal()

frequencies_location <- freq(shocksandrisk$sample_location)
kable(frequencies_location, caption = "Frequencies of sample locations", align = "c", digits = 2) %>%
  kable_styling(bootstrap_options = "striped", full_width = FALSE) %>%
  column_spec(1, bold = TRUE) %>%
  column_spec(2:3, width = "2cm") %>%
  row_spec(0, bold = TRUE) %>%
  row_spec(1:nrow(frequencies_location))

frequencies_sample_lag <- freq(shocksandrisk$sample_lag)
kable(frequencies_sample_lag, caption = "Frequencies of sampling lag", align = "c", digits = 2) %>%
  kable_styling(bootstrap_options = "striped", full_width = FALSE) %>%
  column_spec(1, bold = TRUE) %>%
  column_spec(2:3, width = "2cm") %>%
  row_spec(0, bold = TRUE) %>%
  row_spec(1:nrow(frequencies_sample_lag))
```

#design and variables descriptives

```
frequencies_shock_type <- freq(shocksandrisk$indepvar_shock_type)
kable(frequencies_shock_type, caption = "Frequencies of shock type", align = "c", digits = 2) %>%
  kable_styling(bootstrap_options = "striped", full_width = FALSE) %>%
  column_spec(1, bold = TRUE) %>%
  column_spec(2:3, width = "2cm") %>%
  row_spec(0, bold = TRUE) %>%
  row_spec(1:nrow(frequencies_shock_type))

frequencies_design_longcross <- freq(shocksandrisk$design_long_cross)
kable(frequencies_design_longcross, caption = "Frequencies of design type", align = "c", digits = 2)
%>%
  kable_styling(bootstrap_options = "striped", full_width = FALSE) %>%
  column_spec(1, bold = TRUE) %>%
  column_spec(2:3, width = "2cm") %>%
  row_spec(0, bold = TRUE) %>%
  row_spec(1:nrow(frequencies_design_longcross))


### BASIC META-ANALYSIS

# Basic random effects meta-analysis
m.gen <- metagen(TE = effect_size_unweighted,
                 seTE = depvar_se_standardised,
                 studlab = article_author_year,
                 data = shocksandrisk,
                 sm = "SMD",
                 fixed = FALSE,
                 random = TRUE,
                 method.tau = "REML",
                 hakn = TRUE,
                 title = "Shocks and Risk Preferences")
summary(m.gen)

m.gen_for_tables <- metagen(TE = effect_size,
                 seTE = depvar_se_standardised,
                 studlab = article_author_year,
                 data = shocksandrisk,
                 sm = "SMD",
                 fixed = FALSE,
                 random = TRUE,
                 method.tau = "REML",
                 hakn = TRUE,
                 title = "Shocks and Risk Preferences")

m.gen_table <- kable(m.gen_for_tables, format = "html", digits = 3) %>%
  kable_styling(latex_options = c("striped", "hold_position"), font_size = 12) %>%
  column_spec(1, bold = TRUE) %>%
  column_spec(2:11, width = "2.5cm")
m.gen_table

forest.meta(m.gen_for_tables, layout = "JAMA", fontsize = 10)
pdf("forest_plot.pdf", width = 10, height = 20)
print(forest.meta(m.gen_for_tables, layout = "JAMA", fontsize = 10))
dev.off()
```

```
drapery(m.gen_for_tables,
        labels = "studlab",
        type = "pval",
        legend = FALSE)
```

 # Excluding Cassar 2017 due to doubts about effect size calculation

```
shocksandrisk_nocassar <- shocksandrisk[-c(15), ]

m.gen_nocassar <- metagen(TE = effect_size_unweighted,
                 seTE = depvar_se_standardised,
                 studlab = article_author_year,
                 data = shocksandrisk_nocassar,
                 sm = "SMD",
                 fixed = FALSE,
                 random = TRUE,
                 method.tau = "REML",
                 hakn = TRUE,
                 title = "Shocks and Risk Preferences")

summary(m.gen_nocassar)

m.gen_nocassar_for_tables <- metagen(TE = effect_size,
                          seTE = depvar_se_standardised,
                          studlab = article_author_year,
                          data = shocksandrisk_nocassar,
                          sm = "SMD",
                          fixed = FALSE,
                          random = TRUE,
                          method.tau = "REML",
                          hakn = TRUE,
                          title = "Shocks and Risk Preferences")
forest.meta(m.gen_nocassar_for_tables, layout = "JAMA", fontsize = 10)
pdf("forest_plot_nocassar.pdf", width = 10, height = 20)
print(forest.meta(m.gen_nocassar_for_tables, layout = "JAMA", fontsize = 10))
dev.off()

drapery(m.gen_nocassar_for_tables,
        labels = "studlab",
        type = "pval",
        legend = FALSE)
```

#excluding NAs

```
shocksandrisk_effects <- shocksandrisk[-c(15,21,22,23,24,26,46,47,33,34,35,36,40,43,48,55,56,58), ]

m.gen_effects_for_tables <- metagen(TE = effect_size,
                          seTE = depvar_se_standardised,
                          studlab = article_author_year,
                          data = shocksandrisk_effects,
                          sm = "SMD",
                          fixed = FALSE,
                          random = TRUE,
                          method.tau = "REML",
                          hakn = TRUE,
                          title = "Shocks and Risk Preferences")
```

```
forest.meta(m.gen_effects_for_tables, layout = "JAMA", fontsize = 9)
forest.meta(m.gen, layout = "JAMA", fontsize = 10)
pdf("forest_plot_effects.pdf", width = 10, height = 20)
print(forest.meta(m.gen_effects_for_tables, layout = "JAMA", fontsize = 10))
dev.off()
```

###BETWEEN STUDY HETEROGENEITY

```
m.gen <- update.meta(m.gen, prediction = TRUE)
summary(m.gen)
```

SUBGROUP ANALYSIS

#filter for different shock types

```
shocksandrisk_natural_disaster <- filter(shocksandrisk_nocassar, indepvar_shock_type %in% c("Natural disaster"))
shocksandrisk_pandemic <- filter(shocksandrisk_nocassar, indepvar_shock_type %in% c("Pandemic"))
shocksandrisk_conflict <- filter(shocksandrisk_nocassar, indepvar_shock_type %in% c("Conflict & Violence"))
shocksandrisk_economic <- filter(shocksandrisk_nocassar, indepvar_shock_type %in% c("Economic"))
shocksandrisk_GFC <- filter(shocksandrisk_nocassar, !indepvar_shock_type_gfc %in% c("No"))
shocksandrisk_Covid19 <- filter(shocksandrisk_nocassar, !indepvar_shock_type_covid19 %in% c("No"))

        #forest plots for different shock types

m.gen_natural_disaster <- metagen(TE = effect_size,
                 seTE = depvar_se_standardised,
                 studlab = article_author_year,
                 data = shocksandrisk_natural_disaster,
                 sm = "SMD",
                 fixed = FALSE,
                 random = TRUE,
                 method.tau = "REML",
                 hakn = TRUE,
                 title = "Shocks and Risk Preferences - Natural Disaster")
summary(m.gen_natural_disaster)
forest.meta(m.gen_natural_disaster, layout = "JAMA", fontsize = 9)
pdf("forest_plot_natural_disaster.pdf", width = 10, height = 20)
print(forest.meta(m.gen_natural_disaster, layout = "JAMA", fontsize = 10))
dev.off()

m.gen_pandemic <- metagen(TE = effect_size,
                      seTE = depvar_se_standardised,
                      studlab = article_author_year,
                      data = shocksandrisk_pandemic,
                      sm = "SMD",
                      fixed = FALSE,
                      random = TRUE,
                      method.tau = "REML",
                      hakn = TRUE,
```

```
                    title = "Shocks and Risk Preferences - Pandemics")
summary(m.gen_pandemic)
forest.meta(m.gen_pandemic, layout = "JAMA", fontsize = 9)
pdf("forest_plot_pandemic.pdf", width = 10, height = 20)
print(forest.meta(m.gen_pandemic, layout = "JAMA", fontsize = 10))
dev.off()

m.gen_conflict <- metagen(TE = effect_size,
                    seTE = depvar_se_standardised,
                    studlab = article_author_year,
                    data = shocksandrisk_conflict,
                    sm = "SMD",
                    fixed = FALSE,
                    random = TRUE,
                    method.tau = "REML",
                    hakn = TRUE,
                    title = "Shocks and Risk Preferences - Conflict & Violence")
summary(m.gen_conflict)
forest.meta(m.gen_conflict, layout = "JAMA", fontsize = 9)
pdf("forest_plot_conflict.pdf", width = 10, height = 20)
print(forest.meta(m.gen_conflict, layout = "JAMA", fontsize = 10))
dev.off()


m.gen_economic <- metagen(TE = effect_size,
                    seTE = depvar_se_standardised,
                    studlab = article_author_year,
                    data = shocksandrisk_economic,
                    sm = "SMD",
                    fixed = FALSE,
                    random = TRUE,
                    method.tau = "REML",
                    hakn = TRUE,
                    title = "Shocks and Risk Preferences - Economic Influences")
summary(m.gen_economic)
forest.meta(m.gen_economic, layout = "JAMA", fontsize = 9)
pdf("forest_plot_economic.pdf", width = 10, height = 20)
print(forest.meta(m.gen_economic, layout = "JAMA", fontsize = 10))
dev.off()


m.gen_GFC <- metagen(TE = effect_size,
                    seTE = depvar_se_standardised,
                    studlab = article_author_year,
                    data = shocksandrisk_GFC,
                    sm = "SMD",
                    fixed = FALSE,
                    random = TRUE,
                    method.tau = "REML",
                    hakn = TRUE,
                    title = "Shocks and Risk Preferences - Global Financial Crisis")
summary(m.gen_GFC)
forest.meta(m.gen_GFC, layout = "JAMA", fontsize = 9)
pdf("forest_plot_GFC.pdf", width = 10, height = 20)
print(forest.meta(m.gen_GFC, layout = "JAMA", fontsize = 10))
dev.off()
```

```
m.gen_Covid19 <- metagen(TE = effect_size,
                         seTE = depvar_se_standardised,
                         studlab = article_author_year,
                         data = shocksandrisk_Covid19,
                         sm = "SMD",
                         fixed = FALSE,
                         random = TRUE,
                         method.tau = "REML",
                         hakn = TRUE,
                         title = "Shocks and Risk Preferences - COVID-19")
summary(m.gen_Covid19)
forest.meta(m.gen_Covid19, layout = "JAMA", fontsize = 9)
pdf("forest_plot_pandemic.pdf", width = 10, height = 20)
print(forest.meta(m.gen_Covid19, layout = "JAMA", fontsize = 10))
dev.off()

#filter for persistence (long, medium, short)
shocksandrisk_shortrun <- filter(shocksandrisk_nocassar, effect_persistence %in% c("short-run"))
shocksandrisk_mediumrun <- filter(shocksandrisk_nocassar, effect_persistence %in% c("medium"))
shocksandrisk_longrun <- filter(shocksandrisk_nocassar, effect_persistence %in% c("long-run"))

    #forest plots for persistence
m.gen_shortrun <- metagen(TE = effect_size,
                    seTE = depvar_se_standardised,
                    studlab = article_author_year,
                    data = shocksandrisk_shortrun,
                    sm = "SMD",
                    fixed = FALSE,
                    random = TRUE,
                    method.tau = "REML",
                    hakn = TRUE,
                    title = "Shocks and Risk Preferences - short-run")
summary(m.gen_shortrun)
forest.meta(m.gen_shortrun, layout = "JAMA", fontsize = 9)
pdf("forest_plot_shortrun.pdf", width = 10, height = 20)
print(forest.meta(m.gen_shortrun, layout = "JAMA", fontsize = 10))
dev.off()

m.gen_mediumrun <- metagen(TE = effect_size,
                     seTE = depvar_se_standardised,
                     studlab = article_author_year,
                     data = shocksandrisk_mediumrun,
                     sm = "SMD",
                     fixed = FALSE,
                     random = TRUE,
                     method.tau = "REML",
                     hakn = TRUE,
                     title = "Shocks and Risk Preferences - medium-run")
summary(m.gen_mediumrun)
forest.meta(m.gen_mediumrun, layout = "JAMA", fontsize = 9)
pdf("forest_plot_shortrun.pdf", width = 10, height = 20)
print(forest.meta(m.gen_mediumrun, layout = "JAMA", fontsize = 10))
dev.off()
```

```
m.gen_longrun <- metagen(TE = effect_size,
                  seTE = depvar_se_standardised,
                  studlab = article_author_year,
                  data = shocksandrisk_longrun,
                  sm = "SMD",
                  fixed = FALSE,
                  random = TRUE,
                  method.tau = "REML",
                  hakn = TRUE,
                  title = "Shocks and Risk Preferences - long-run")
summary(m.gen_longrun)
forest.meta(m.gen_longrun, layout = "JAMA", fontsize = 9)
pdf("forest_plot_shortrun.pdf", width = 10, height = 20)
print(forest.meta(m.gen_longrun, layout = "JAMA", fontsize = 10))
dev.off()


# basic subgroup meta analysis

update.meta(m.gen,
        subgroup = publication_status,
        tau.common = FALSE)

update.meta(m.gen,
        subgroup = author_discipline,
        tau.common = FALSE)

update.meta(m.gen,
        subgroup = effect_persistence,
        tau.common = FALSE)

update.meta(m.gen,
        subgroup = design_general,
        tau.common = FALSE)

update.meta(m.gen,
        subgroup = design_variation_source,
        tau.common = FALSE)

update.meta(m.gen,
        subgroup = design_long_cross,
        tau.common = FALSE)

update.meta(m.gen,
        subgroup = design_identification_causal,
        tau.common = FALSE)

update.meta(m.gen,
        subgroup = depvar_incentivised,
        tau.common = FALSE)

update.meta(m.gen,
        subgroup = indepvar_shock_type,
        tau.common = FALSE)

update.meta(m.gen,
```

```
        subgroup = indepvar_shock_type_gfc,
        tau.common = FALSE)

update.meta(m.gen,
        subgroup = indepvar_shock_type_covid19,
        tau.common = FALSE)

update.meta(m.gen,
        subgroup = WEIRD,
        tau.common = FALSE)

update.meta(m.gen,
        subgroup = author_male,
        tau.common = FALSE)
```

###META-REGRESSIONS

 # incentivised

```
m.gen.reg_incentive <- metareg (m.gen, ~depvar_incentivised)
m.gen.reg_incentive
m.gen.reg_incentive <- metareg (m.gen_for_tables, ~depvar_incentivised)
bubble(m.gen.reg_incentive, studlab = TRUE, ylim = c(-0.5, 0.5))

pdf("bubble_plot_incentive", width = 15, height = 15)
print(bubble(m.gen.reg_incentive, studlab = TRUE, ylim = c(-0.5, 0.5)))
dev.off()

ggplot(shocksandrisk, aes(x=depvar_incentivised, y=effect_size)) +
  geom_point(color="blue") +
  labs(title="Scatter Plot of Effect Size and Incentivisation", x="Incentivised Elicitation", y="Effect
Size") +
  theme_minimal()
```

#sample

```
m.gen.reg_WEIRD <- metareg (m.gen, ~WEIRD)
m.gen.reg_WEIRD
m.gen.reg_WEIRD <- metareg (m.gen_for_tables, ~WEIRD)
bubble(m.gen.reg_WEIRD, studlab = TRUE, ylim = c(-0.5, 0.5))

pdf("bubble_plot_WEIRD", width = 15, height = 15)
print(bubble(m.gen.reg_WEIRD, studlab = TRUE, ylim = c(-0.5, 0.5)))
dev.off()

ggplot(shocksandrisk, aes(x=WEIRD, y=effect_size)) +
  geom_point(color="blue") +
  labs(title="Scatter Plot of Effect Size and Sample WEIRDness", x="WEIRDness", y="Effect Size")
+
  theme_minimal()
```

#quality (citations, journal, sample size)

```
m.gen.reg_citations <- metareg (m.gen, ~article_citations_gs)
m.gen.reg_citations
m.gen.reg_citations <- metareg (m.gen_for_tables, ~article_citations_gs)
bubble(m.gen.reg_citations, studlab = TRUE, ylim = c(-0.5, 0.5), xlim = c(0, 1200))
```

```
pdf("bubble_plot_Citations", width = 15, height = 15)
print(bubble(m.gen.reg_citations, studlab = TRUE, ylim = c(-0.5, 0.5)))
dev.off()

m.gen.reg_size <- metareg (m.gen, ~sample_size)
m.gen.reg_size
m.gen.reg_size <- metareg (m.gen_for_tables, ~sample_size)
bubble(m.gen.reg_size, studlab = TRUE, ylim = c(-0.5, 0.5))
pdf("bubble_plot_sample_size", width = 15, height = 15)
print(bubble(m.gen.reg_size, studlab = TRUE, ylim = c(-0.5, 0.5)))
dev.off()
bubble(m.gen.reg_size, studlab = TRUE, ylim = c(-0.5, 0.5), xlim = c(0, 10000))
pdf("bubble_plot_sample_size_zoom", width = 15, height = 15)
print(bubble(m.gen.reg_size, studlab = TRUE, ylim = c(-0.5, 0.5), xlim = c(0, 10000)))
dev.off()

ggplot(shocksandrisk, aes(x=sample_size, y=effect_size)) +
  geom_point(color="blue") +
  labs(title="Scatter Plot of Effect Size and Sample Size", x="N", y="Effect Size") +
  theme_minimal()

m.gen.reg_impact <- metareg (m.gen, ~journal_impact)
m.gen.reg_impact
m.gen.reg_impact <- metareg (m.gen_for_tables, ~journal_impact)
bubble(m.gen.reg_impact, studlab = TRUE, ylim = c(-0.5, 0.5), xlim = c(0, 13))
pdf("bubble_plot_journal_impact", width = 15, height = 15)
print(bubble(m.gen.reg_impact, studlab = TRUE, ylim = c(-0.5, 0.5)))
dev.off()

m.gen.reg_quality <- metareg (m.gen, ~weight_quality)
m.gen.reg_quality
m.gen.reg_quality <- metareg (m.gen_for_tables, ~weight_quality)
bubble(m.gen.reg_quality, studlab = TRUE, ylim = c(-0.5, 0.5))
pdf("bubble_plot_quality", width = 15, height = 15)
print(bubble(m.gen.reg_quality, studlab = TRUE, ylim = c(-0.5, 0.5)))
dev.off()

ggplot(shocksandrisk, aes(x=weight_quality, y=effect_size)) +
  geom_point(color="blue") +
  labs(title="Scatter Plot of Effect Size and Publication Quality", x="Publication Quality", y="Effect
Size") +
  theme_minimal()

m.gen.author_male <- metareg (m.gen, ~author_male)
m.gen.author_male
m.gen.author_male <- metareg (m.gen_for_tables, ~author_male)
bubble(m.gen.author_male, studlab = TRUE, ylim = c(-0.5, 0.5))
pdf("bubble_plot_auhor_gender", width = 15, height = 15)
print(bubble(m.gen.author_male, studlab = TRUE, ylim = c(-0.5, 0.5)))
dev.off()

#year
m.gen.reg_year <- metareg (m.gen, ~article_year)
m.gen.reg_year
m.gen.reg_year <- metareg (m.gen_for_tables, ~article_year)
bubble(m.gen.reg_year, studlab = TRUE, ylim = c(-0.5, 0.5), xlim = c(2007, 2023))
```

```
pdf("bubble_plot_year", width = 15, height = 15)
print(bubble(m.gen.reg_year, studlab = TRUE, ylim = c(-0.5, 0.5)))
dev.off()

ggplot(shocksandrisk, aes(x=article_year, y=effect_size)) +
  geom_point(color="blue") +
  labs(title="Scatter Plot of Effect Size and Year of Publication", x="Year of Publication", y="Effect
Size") +
  theme_minimal()
```

#independent variable

```
m.gen.reg_shocktype <- metareg (m.gen, ~indepvar_shock_type)
m.gen.reg_shocktype
m.gen.reg_shocktype <- metareg (m.gen_for_tables, ~indepvar_shock_type)
bubble(m.gen.reg_shocktype, studlab = TRUE, ylim = c(-0.5, 0.5))
pdf("bubble_plot_shock_type", width = 15, height = 15)
print(bubble(m.gen.reg_shocktype, studlab = TRUE, ylim = c(-0.5, 0.5)))
dev.off()

ggplot(shocksandrisk, aes(x=indepvar_shock_type, y=effect_size)) +
  geom_point(color="blue") +
  labs(title="Scatter Plot of Effect Size and Shock Type", x="Shock Type", y="Effect Size") +
  theme_minimal()

m.gen.reg_shocktype_gfc <- metareg (m.gen, ~indepvar_shock_type_gfc)
m.gen.reg_shocktype_gfc
m.gen.reg_shocktype_gfc <- metareg (m.gen_for_tables, ~indepvar_shock_type_gfc)
bubble(m.gen.reg_shocktype_gfc, studlab = TRUE, ylim = c(-0.5, 0.5))
pdf("bubble_plot_gfc", width = 15, height = 15)
print(bubble(m.gen.reg_shocktype_gfc, studlab = TRUE, ylim = c(-0.5, 0.5)))
dev.off()

m.gen.reg_shocktype_covid19 <- metareg (m.gen, ~indepvar_shock_type_covid19)
m.gen.reg_shocktype_covid19
m.gen.reg_shocktype_covid19 <- metareg (m.gen_for_tables, ~indepvar_shock_type_covid19)
bubble(m.gen.reg_shocktype_covid19, studlab = TRUE, ylim = c(-0.5, 0.5))
pdf("bubble_plot_covid19", width = 15, height = 15)
print(bubble(m.gen.reg_shocktype_covid19, studlab = TRUE, ylim = c(-0.5, 0.5)))
dev.off()
```

 # design

```
m.gen.reg_causal <- metareg (m.gen, ~design_identification_causal)
m.gen.reg_causal
m.gen.reg_causal <- metareg (m.gen_for_tables, ~design_identification_causal)
bubble(m.gen.reg_causal, studlab = TRUE, ylim = c(-0.5, 0.5))
pdf("bubble_plot_causal", width = 15, height = 15)
print(bubble(m.gen.reg_causal, studlab = TRUE, ylim = c(-0.5, 0.5)))
dev.off()

ggplot(shocksandrisk, aes(x=design_identification_causal, y=effect_size)) +
  geom_point(color="blue") +
  labs(title="Scatter Plot of Effect Size and Causal Identification", x="Causal Identification", y="Effect
Size") +
  theme_minimal()

m.gen.reg_longcross <- metareg (m.gen, ~design_long_cross)
```

```
m.gen.reg_longcross
m.gen.reg_longcross <- metareg (m.gen_for_tables, ~design_long_cross)
bubble(m.gen.reg_longcross, studlab = TRUE, ylim = c(-0.5, 0.5))
pdf("bubble_plot_design", width = 15, height = 15)
print(bubble(m.gen.reg_longcross, studlab = TRUE, ylim = c(-0.5, 0.5)))
dev.off()

ggplot(shocksandrisk, aes(x=design_long_cross, y=effect_size)) +
  geom_point(color="blue") +
  labs(title="Scatter Plot of Effect Size and Design", x="Design", y="Effect Size") +
  theme_minimal()

# further meta-regressions (caution for ecological bias!)

#age
m.gen.reg_age <- metareg (m.gen, ~sample_age)
m.gen.reg_age
m.gen.reg_age <- metareg (m.gen_for_tables, ~sample_age)
bubble(m.gen.reg_age, studlab = TRUE, ylim = c(-0.5, 0.5), xlim = c(10, 60))
pdf("bubble_plot_age", width = 15, height = 15)
print(bubble(m.gen.reg_age, studlab = TRUE, ylim = c(-0.5, 0.5)))
dev.off()

ggplot(shocksandrisk, aes(x=sample_age, y=effect_size)) +
  geom_point(color="blue") +
  labs(title="Scatter Plot of Effect Size and Average Sample Age", x="Average Sample Age",
y="Effect Size") +
  theme_minimal()


### PUBLICATION BIAS

#including unpublished papers!!!
funnel.meta(m.gen,
            xlim = c(-0.7, 0.7),
            studlab = TRUE)
title("Funnel Plot (Shocks and Risk Preferences - including unpublished papers)")

funnel.meta(m.gen,
            xlim = c(-0.7, 0.7),
            ylim = c(0.11,0),
            studlab = TRUE)
title("Funnel Plot (Shocks and Risk Preferences - including unpublished papers) - zoomed in")

#only published papers --> Publication Bias
shocksandrisk_published <- filter(shocksandrisk_nocassar, !publication_status %in% c("Working
Paper"))

m.gen_published <- metagen(TE = effect_size_unweighted,
                           seTE = depvar_se_standardised,
                           studlab = article_author_year,
                           data = shocksandrisk_published,
                           sm = "SMD",
                           fixed = FALSE,
                           random = TRUE,
                           method.tau = "REML",
```

```
                    hakn = TRUE,
                    title = "Shocks and Risk Preferences")

summary(m.gen_published)

forest.meta(m.gen_published)
pdf("forest_plot_published.pdf", width = 10, height = 20)
print(forest.meta(m.gen_published, layout = "JAMA", fontsize = 10))
dev.off()

funnel.meta(m.gen_published,
         xlim = c(-0.7, 0.7),
         ylim = c(0.11,0),
         studlab = TRUE)
title("Funnel Plot (Shocks and Risk Preferences - only published papers)")
```

#only unpublished papers

```
shocksandrisk_unpublished <- filter(shocksandrisk, !publication_status %in% c("Published"))

m.gen_unpublished <- metagen(TE = effect_size_unweighted,
                    seTE = depvar_se_standardised,
                    studlab = article_author_year,
                    data = shocksandrisk_unpublished,
                    sm = "SMD",
                    fixed = FALSE,
                    random = TRUE,
                    method.tau = "REML",
                    hakn = TRUE,
                    title = "Shocks and Risk Preferences")
summary(m.gen_unpublished)

forest.meta(m.gen_unpublished)
pdf("forest_plot_unpublished.pdf", width = 10, height = 20)
print(forest.meta(m.gen_unpublished, layout = "JAMA", fontsize = 10))
dev.off()

funnel.meta(m.gen_unpublished,
         xlim = c(-0.7, 0.7),
         ylim = c(0.11,0),
         studlab = TRUE)
title("Funnel Plot (Shocks and Risk Preferences - only working papers)")
```

 # exclude Frondel 2021 & Zhang&Palma 2022 due to unusually large standard errors

```
shocksandrisk_nofrondelzhang <- shocksandrisk[-c(25,65,66), ]

m.gen_nofrondelzhang <- metagen(TE = effect_size_unweighted,
                        seTE = depvar_se_standardised,
                        studlab = article_author_year,
                        data = shocksandrisk_nofrondelzhang,
                        sm = "SMD",
                        fixed = FALSE,
                        random = TRUE,
                        method.tau = "REML",
                        hakn = TRUE,
                        title = "Shocks and Risk Preferences")
```

```
summary(m.gen_nofrondelzhang)

funnel.meta(m.gen_nofrondelzhang,
            xlim = c(-0.5, 0.5),
            studlab = TRUE)
title("Funnel Plot (Shocks and Risk Preferences - no Frondel & Zhang)")
```