Meta-Analysis of Empirical Estimates of Loss-Aversion

Alexander L. Brown Taisuke Imai Ferdinand M. Vieider
Colin F. Camerer *

January 19, 2021

Abstract

Loss aversion is one of the most widely used concepts in behavioral economics. We conduct a large-scale interdisciplinary meta-analysis, to systematically accumulate knowledge from numerous empirical estimates of the loss aversion coefficient reported during the past couple of decades. We examine 607 empirical estimates of loss aversion from 150 articles in economics, psychology, neuroscience, and several other disciplines. Our analysis indicates that the mean loss aversion coefficient is between 1.8 and 2.1. We also document how reported estimates vary depending on the observable characteristics of the study design.

JEL Classification codes: D81, D90, C90, C11
Keywords: loss aversion, prospect theory, meta-analysis

*Brown: Department of Economics, Texas A&M University, alexbrown@tamu.edu. Imai: Department of Economics, LMU Munich, taisuke.imai@econ.lmu.de. Vieider: Department of Economics, Ghent University, ferdinand.vieider@ugent.be. Camerer (corresponding author): Division of the Humanities and Social Sciences, California Institute of Technology, camerer@hss.caltech.edu. We thank Hyundam Je and Hye Joon Lee for research assistance. We are grateful for the feedback provided by Chishio Furukawa, Klaus Schmidt, Tom Stanley, and the audience at ESA 2019 World Meeting, ESA North American Meeting 2019, MAER-Net Colloquium 2019, D-TEA 2020, and 2020 CESifo Area Conference on Behavioral Economics. Financial support from the National Science Foundation #1757282 (Brown and Camerer) and Deutsche Forschungsgemeinschaft through CRC TRR 190 (Imai) are gratefully acknowledged.

1 Introduction

Loss-aversion is the empirical observation that decisions often reflect a disproportionate distaste for potential losses, compared to equal-sized gains, relative to a point of reference. Loss aversion is a core feature of prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1991, 1992; Wakker, 2010), an explicitly descriptive model of choice under risk and uncertainty which has been widely applied and cited.[1] The strength of aversion to loss compared to attraction to gain is typically captured by a single parameter, λ.

In his popular-science book, Kahneman (2011) says "the concept of loss aversion is certainly the most significant contribution of psychology to behavioral economics" (p. 300). Indeed, loss-aversion has been widely applied to many types of economic decisions and analyses. The most common application is analysis of experimental decisions over monetary risks (as in the original Kahneman and Tversky, 1979). However, loss-aversion and dependence on reference points have also been used well beyond its initial simple application to three-outcome risks. Applications include financial asset prices (Barberis, 2013), the equity premium puzzle (Benartzi and Thaler, 1995), the disposition effect in asset trading (Odean, 1998; Weber and Camerer, 1998), labor supply decisions (Camerer et al., 1997), political power of entitlements change (Romer, 1996), majority voting and politics (Alesina and Passarelli, 2019), sectoral trade policy behavior (Tovar, 2009), and selling-buying price endowment effects in contingent valuation of nontraded goods (Ericson and Fuster, 2014; Tunçel and Hammitt, 2014). It also features prominently in behavioral industrial organization (Heidhues and Kőszegi, 2018).

Several different methods have been used to measure loss-aversion. These include laboratory experiments, representative panel surveys, analyses of natural data, and randomized trials trying to change behavior. Many studies also come from far outside economics, including neuroscience, psychiatry, business and management, and transportation.

Given how widely the concept of loss-aversion has been applied in economics and many other social sciences, it is useful to have the best possible empirical answer about how large loss-aversion is, and how it varies. One of the first empirical estimates of λ is reported in Tversky and Kahneman (1992). The authors elicit the preferences of 25 graduate students from elite west-coast American universities using three sessions of unincentivized lottery-choice experiments. The median λ—no mean nor statistic of dispersion was reported—was $\lambda = 2.25$ (Figure 1). Many analyses, to this day, cite this number as the typical degree of loss aversion. For example, the value $\lambda = 2.25$ is used in numerical simulations of prospect theory in behavioral finance (e.g,

[1] The 1979 paper is the most widely cited empirical economics paper published from 1970-2005 (see Table 2 in Kim et al., 2006). Note also that Fishburn and Kochenberger (1979) also documented loss-aversion in a different sample of preferences elicited for decision analysis.

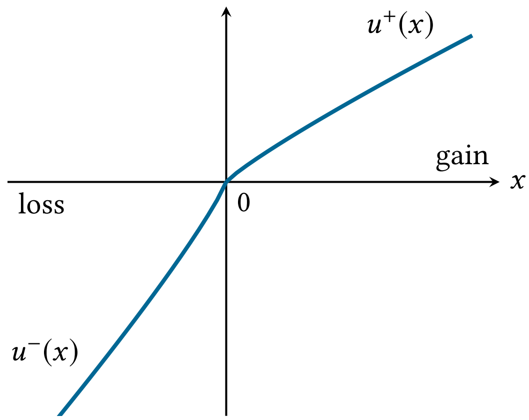

Figure 1: An example prospect theory value function. *Notes*: This is the specification (3) presented in Section 2 with median parameters in Tversky and Kahneman (1992), $\lambda = 2.25$ and $u^{+}(x) = u^{-}(x) = x^{0.88}$.

Barberis et al., 2001; Barberis and Huang, 2001, 2008; Barberis and Xiong, 2009; Barberis et al., 2016, 2020). This peculiar degree of path-dependence is not lost on these authors. As noted in the latter study, "[...] these estimates are almost 30 years old and are based on a small number of participants. Given that the values we assign to these parameters play a significant role in our results, it seems prudent to base these values on a wide range of studies, not just one." (Barberis et al., 2020, p. 25).

What is the best way to cumulate knowledge about λ after thirty years of research? Our view is that *meta-analysis* is an indispensable tool for scientific cumulation.

Meta-analysis is a principled, reproducible, open-science method for accumulating scientific knowledge (and also for detecting nonrandom selective reporting of evidence: Stanley, 2001; Stanley and Doucouliagos, 2012). A meta-analysis uses a clearly specified method of sampling available studies, coding evidence in a way that is compared and compiled across studies, and summarizing both regularity and variation across studies. The idea of synthesizing evidence from multiple studies dates back to the early 1900s (Pearson, 1904; Yates and Cochran, 1938), but the history of modern meta-analysis has its origin in the 1976 AERA presidential address by Gene V. Glass. He introduced the term "meta-analysis" to refer to "the statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings" (Glass, 1976, p. 3). It has been widely-used in evidence-based practices in medicine and policy for at least a decade or two (Gurevitch et al., 2018). However, meta-analysis has been mostly absent from highly-selective journals in empirical economics.[2]

[2]Prominent meta-analyses in economics include value of life (Doucouliagos et al., 2012, 2014), intertemporal elasticity (Havránek, 2015; Havránek et al., 2015), habit formation (Havránek et al., 2017), foreign direct investment (Iršová and Havránek, 2013), minimum wage effects (Doucouliagos and Stanley, 2009), behavior in dictator and ultimatum games (Engel, 2011; Oosterbeek et al., 2004), truth-telling (Abeler et al., 2019), experimentally-measured

An advantage of meta-analysis is that as new evidence arrives it can be easily added to the previous corpus of studies and results can be quickly updated. No serious science can thrive without meta-analysis, and many results of meta-analyses should be broadcast widely.

This paper reports results of a meta-analysis of empirical estimates of loss aversion. The dataset comprises 607 estimates reported in 150 papers in economics, psychology, neuroscience, and several other disciplines.

The toolkit of meta-analysis can give the best available answers to three questions:

1. What is the central tendency in the distribution of λ estimates; and how much do they vary?

2. Does measured λ vary systematically across different methods, definition of λ, utility specifications, domains of choice, and types of participants?

3. Is there evidence of selective reporting, or publication bias, which distorts reported estimates of λ compared to the corpus of ideal evidence without such biases?

The numerical answers to these questions should be intrinsically interesting for scientists interested in loss-aversion. Beyond intrinsic interest, however, the results can help scientists do their work better in several ways.

Imagine an early career researcher (ECR) who is interested in loss-aversion but not quite sure what steps to take to measure it or to apply it. First, the ECR might ask: What method should I use to measure λ? What are the most popular methods? Does it make much difference which one is used?

Results on how estimated λ's vary with characteristics of the measurement method, such as the type of the data (experimental or field), reward (monetary or non-monetary), specification of the utility function, and the definition of loss aversion, can guide the ECR.[3]

Second, the ECR may be doing a behavior change RCT. Then she needs a specific estimate of λ, or a plausible range of values, to use to make a power calculation. Perhaps she is planning to prepay teacher bonuses, which they can later lose, to motivate them to increase student test outcomes (Fryer et al., 2012). Is $\lambda = 2.25$ a good guess or is there a better guess? Is there a more refined estimate of λ for the subset of studies in the meta-analysis which are most like the one she is planning? Meta-analysis can help here too.

Third, suppose the ECR has just read recent review articles about prospect theory and reference-dependent preferences (e.g., Barberis, 2013; DellaVigna, 2018; O'Donoghue and Sprenger, 2018).

discount rates (Matoušek et al., 2020), and present-bias in Convex Time Budget experiments (Imai et al., forthcoming).

[3]These methodological variations can also help us understand the mechanisms behind loss aversion, along with process measures such as response times, psychophysiology, and neuroscientific data.

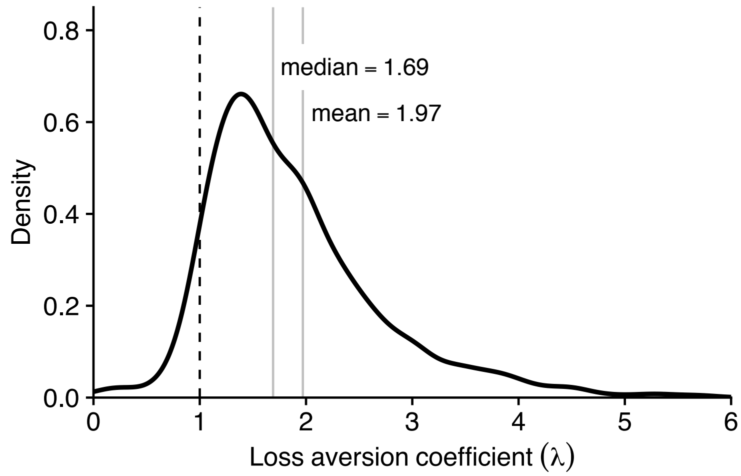

FIGURE 2: Distribution of reported estimates of loss aversion coefficient.

Those reviews have a "narrative" programmatic structure in which results of early and key studies raise fundamental questions that later studies are designed to answer. The reader is usually left with an understanding of the historical intellectual trajectory, and what the next wave of studies should try to understand better. The ECR wonders, is anything important left out of the narrative? The meta-analysis helps answer this question too. (However, the comparison and complementarities of meta-analysis and narrative review are subtle and important, so we will return to them in the conclusion.)

Figure 2 shows the distribution of the loss aversion coefficient λ in our dataset with median value of 1.69 and the mean of 1.97. The distribution is right-skewed and has a substantial mass (93.9%) on the range $\lambda > 1$, corresponding to loss aversion (as opposed to loss tolerance, $\lambda < 1$). Applying a Bayesian hierarchical approach taking into account the uncertainty surrounding the measurements, we find that the average λ in the literature lies between 1.7 and 1.9. Taking into account the fact that many papers reported more than one estimate (thus producing correlation among estimates), the average is between 1.8 and 2.1. We also examine whether observed heterogeneity in reported λ can be attributed to some of the observable characteristics of the study design. The results do not show many strong reliable effects.

Even to economists unfamiliar with meta-analysis, the method should be appealing. It is essentially an application of some special econometrics to literature review (see Stanley and Doucouliagos, 2012). Like any empirical study, the greatest concern should be the inclusion criteria of the dataset (i.e., selection). While our broad inclusion criteria are independent of estimates of λ, we have little control over publication decisions (which makes papers more prominent and easier to find) and whether a study is written at all (i.e., "the file drawer problem"; Rosenthal, 1979), which could be dependent on the values of estimated λ. We use funnel plots to consider these issues and how they might affect our analysis. As noted elsewhere, no technique, not even the alternative narrative review, can address this issues perfectly (Borenstein et al., 2009). Meta-analysis at least has the tools to examine these possible issues analytically.

Related papers. There are two previous meta-analyses of loss aversion, to which we contribute a newer and broader scope. Neumann and Böckenholt (2014) conduct a meta-analysis of 109 estimates of loss aversion from 33 studies about consumer brand choice. As we do later in this paper, they use a multi-level, random-effects technique to account for variability of estimates, both within and between studies, of the logged-λ parameter. They report a base model $\lambda = 1.49$ and an "enhanced model" $\lambda = 1.73$ accounting for sources of estimate variability. Perhaps because of their narrow focus on consumer choice, their meta-regression controls explain nearly all of the variability within their data. Notably, the use of external vs. internal reference points, estimates derived from models that account for both heterogeneity in taste and process, and unpublished vs. published studies all lead to lower loss aversion.

A different approach was used by Walasek et al. (2018). Their analysis used only published experimental studies of mixed lotteries of gains and losses where original raw data were available for reanalysis. Their corpus is 19 estimates from 17 articles.[4] Rather than meta-analyzing estimates from the original papers, they re-estimated parameters for a single model of cumulative prospect theory (i.e., power utility function with symmetric curvature, $\alpha = \beta$, see equation (4) in Section 2) using the original data. Their random-effects meta-analysis on the 19 estimates has an average $\lambda = 1.31$, which is substantially smaller than the one found in the current paper. Despite their rather strict restrictions, the authors note that there are high levels of variability between studies (the data is not conducive to looking at this question within studies) in both estimates and procedures.

The rest of the paper is organized as follows. Section 2 introduces the concept of loss aversion. Section 3 describes how we assembled the dataset of empirical estimates of loss aversion. Section 4 provides results and Section 5 discusses their implications.

[4]Though not specifically excluded by the aforementioned criteria, the authors also excluded studies that relied on adaptive questions because of concerns about how such techniques would affect their maximum likelihood estimation procedures (i.e., Abdellaoui et al., 2008; Wakker and Deneffe, 1996).

2 Conceptual Framework

We consider a situation where an agent makes a choice under risk between prospects with at most two distinct outcomes. This simplified structure still captures a wide range of empirical studies examined here. Let $L = (x, p; y)$ denote a *simple lottery* which gives outcome x with probability p and outcome y with probability $1 - p$. A key assumption of prospect theory is that outcomes are evaluated as gains and losses relative to a *reference point*. For simplicity of exposition, in this section, we assume the reference point to be 0, so that the sign of the outcome indicates whether it is a gain or a loss. We call a lottery *non-mixed* if two outcomes have the same sign (i.e., either $x, y > 0$ or $x, y < 0$) and *mixed* if one of the outcomes is positive and the other outcome is negative. Without loss of generality, we assume that $x > 0 > y$ when we deal with a mixed lottery.

In this setup, both original prospect theory by Kahneman and Tversky (1979) (hereafter OPT) and its modern incarnation, cumulative prospect theory of Tversky and Kahneman (1992) (hereafter PT), postulate that the agent evaluates non-mixed prospects $L = (x, p; y)$ with $x, y > 0$ or $x, y < 0$ as

$$V(L) = w^i(p) \times v(x) + \left(1 - w^i(p)\right) \times v(y), \tag{1}$$

and mixed prospects $L = (x, p; y)$ with $x > 0 > y$ as

$$V(L) = w^+(p) \times v(x) + w^-(1 - p) \times v(y), \tag{2}$$

where $w^i : [0, 1] \longrightarrow [0, 1]$ is a probability weighting function for gains ($i = +$) or for losses ($i = -$), with $w^i(0) = 0$ and $w^i(1) = 1$. Note that $w^+ = w^-$ is assumed under OPT.

In the literature (both empirical and applied-theoretical), the definition of the value function $v : \mathbf{R} \longrightarrow \mathbf{R}$ is given by

$$v(x) = \begin{cases} u^+(x) & \text{if } x \geq 0 \\ -\lambda u^-(-x) & \text{if } x < 0 \end{cases}, \tag{3}$$

where λ is a positive constant, $u^+ : \mathbf{R}_+ \longrightarrow \mathbf{R}$ and $u^- : \mathbf{R}_+ \longrightarrow \mathbf{R}$ are both monotonically increasing and $u^+(0) = u^-(0) = 0$. The parameter λ is the *loss aversion coefficient*, the target variable of interest in this study. Notice that mixed prospects are necessary to identify loss aversion, since λ cancels out in evaluation of pure-loss prospects such as in equation (1).

There are several definitions of loss aversion proposed in the literature. Kahneman and Tversky's (1979) notion of loss aversion is that the agent prefers a lottery $(y, 0.5; -y)$ over $(x, 0.5; -x)$ if $x > y \geq 0$ (i.e., the value function for losses is steeper than the value function for gains). The

OPT value function gives us

$$v(x) + v(-x) < v(y) + v(-y) \quad \Longleftrightarrow \quad v(x) - v(y) < v(-y) - v(-x).$$

Under this interpretation, the agent is loss averse if the coefficient

$$\lambda_{\mathrm{KT}} = \frac{-v(-x)}{v(x)}$$

is greater than one for all $x > 0$.

Under PT, decision weights naturally enter the definition of loss aversion (Schmidt and Zank, 2005). Suppose an agent is indifferent between a mixed lottery $L = (x, p; y)$ with $x > 0 > y$ and a sure outcome of 0. Then, from equations (2) and (3), we have

$$w^{+}(p) \times u^{+}(x) - \lambda w^{-}(1-p) \times u^{-}(-y) = 0 \quad \Longleftrightarrow \quad \lambda = \frac{w^{+}(p)}{w^{-}(1-p)} \times \frac{u^{+}(x)}{u^{-}(-y)}.$$

Tversky and Kahneman (1992) take the ratio of the utility of a loss of one monetary unit and a gain of one monetary unit, $\lambda_{\mathrm{TK}} = -v(1)/v(1)$. This definition is motivated from the power specification they used: value function (3) together with $u^{+}(x) = x^{\alpha}$ and $u^{-}(x) = x^{\beta}$. Köbberling and Wakker (2005) take the ratio of the left and the right derivatives of v at the reference point, $\lambda_{\mathrm{KW}} = \lim_{x \nearrow 0} v'(x) / \lim_{x \searrow 0} v'(x)$ (which was informally introduced in Benartzi and Thaler, 1995). Wakker and Tversky's (1993) definition of loss aversion requires that the slope of the utility function at any loss outcome is larger than the slope at the equal-sized gain: $v'(-x) > v'(x)$ for all $x > 0$. Finally, Neilson (2002) defines (weak) loss aversion by $-v(x)/x > v(y)/y$ for all positive $x, y > 0$, which does not give us a straightforward definition of the loss aversion coefficient.

A particularly popular functional apparatus is the one using different power utility parameters for gains and losses, following the approach of Tversky and Kahneman (1992):

$$v(x) = \begin{cases} x^{\alpha} & \text{if } x \geq 0 \\ -\lambda(-x)^{\beta} & \text{if } x < 0 \end{cases}. \tag{4}$$

In this formulation, the loss aversion parameter is dependent on the scale of the data, and thus not uniquely defined due to scaling issues (see Wakker, 2010, Section 9.6, for a theoretical discussion). If, on the other hand, the two power parameters are assumed to be identical, i.e. $\alpha = \beta$, this issue does not occur. It also does not occur for different utility parameters using alternative functional forms, such as exponential utility (Köbberling and Wakker, 2005).

3 Data

3.1 Identification and Selection of Relevant Studies

In order to deliver an unbiased meta-analysis, we first identified and selected relevant papers following unambiguously specified inclusion criteria. The main criterion is to include "all empirical papers that estimate a coefficient of loss aversion." Note that, under this criterion, we include papers that use choice data from laboratory or field experiments and also non-experimental, naturally occurring data including stock prices, TV game shows, and surveys on transportation.

We searched for relevant papers on the scientific citation indexing database Web of Science. The initial search, made in the summer of 2017, returned a total hits of 1,547 papers. As a first step of paper identification, we went through titles and abstracts and threw out 910 papers that were clearly irrelevant for our study. We then read the remaining papers, applied our inclusion criteria based on the content, and then coded information (described in Section 3.2 below). We also used IDEAS/RePEc and Google Scholar to search for unpublished working papers. Finally, we posted a message on the email list of the Economic Science Association to ask for relevant papers (in February 2018).

In the initial search phase, we cast a net also to identify papers that estimate the degree of loss aversion in riskless choices, through measuring the discrepancy between willingness-to-pay (WTP) and willingness-to-accept (WTA). However, in reading many of these papers, we discovered that while the authors had measured WTP and WTA for a given item, they had not intended to do so as a way to estimate the loss aversion coefficient. While it is straightforward to impose a certain linear utility structure on such papers, we viewed it as problematic to impose our own assumptions on the work of others, and thus, faced with no better option, did not included these papers.

The search and selection procedure is summarized in Online Appendix A.1. We identified 150 papers at the end of this process. Twenty papers are unpublished at the time of the initial data collection (summer 2017).

3.2 Data Construction

We assembled the dataset for our meta-analysis by coding relevant information—estimates of the loss aversion coefficient, characteristics of the data, and measurement methods. The primary variable of interests are estimates of the loss aversion coefficient λ. These estimates come in two different forms: (i) *aggregate-level*, where a single λ for the "representative" agent/subject is estimated by pooling data from all subjects in a study; (ii) *individual-level*, where λ is estimated for each subject in a study and the summary statistics of empirical distribution, typically mean or median, are reported. We have a dummy variable capturing the type of reported estimates. We also coded standard errors (SEs) of parameter estimates in order to control simply for the quality of the study/uncertainty surrounding any given estimate in the meta-analysis. When SEs are not reported, we reconstructed them from other available information such as standard deviation (SD), p-value (of the null hypothesis of loss neutrality), or the inter-quartile range (IQR).[5]

We also coded variables describing characteristics of the data and measurement methods. These variables include: type of the data (e.g., experimental, non-experimental, TV game show); location of the experiment (e.g., laboratory, field, online); types of reward (e.g., real or hypothetical, money, health, time); subject population (e.g., children, college students, general population, farmers); definition of loss aversion coefficient; utility specification; and several others. Table A.1 in the Online Appendix lists all variables coded in the study.

The set of papers we included span a wide range of disciplines (see Table A.3 in the Online Appendix). Since fields/journals have different reporting cultures and standards, we could not always retrieve all the necessary information from reading papers. We thus emailed the authors of the papers when some essential summary statistics of the loss aversion coefficient or sample size information were missing.[6]

3.3 Descriptive Statistics

We identified 150 articles that report an estimate of the loss aversion coefficient λ (see Online Appendix E for the full list of articles). Among these, 130 articles were published in 78 journal outlets (including eight articles published in the "Top 5" journals in economics). The dataset includes papers from a variety of disciplines: economics, management, psychology, neuroscience, medicine, psychiatry, agriculture, environment, transportation, and operations research (see the list of journals and their classifications in Table A.2 in the Online Appendix).

We also identified where the data (either experimental or survey) were collected for 147 articles in the dataset. Most of these articles report estimates from data collected in a single country. Ten of them collected data from two to three countries/regions, and three of them (l'Haridon and Vieider, 2019; Rieger et al., 2017; Wang et al., 2017) conducted large-scale cross-country studies,

[5]We calculated/approximated 68 SEs from other available information: 64 from IQR and sample size and four from p-values. Use of the IQR to infer the SE of the mean (from an approximation of the standard deviation, SD $\approx 1.35 \times$ IQR) is in principle only legitimate if the parameters are normally distributed in the population. That assumption is clearly a stretch. Nevertheless, obtaining even an "approximated" SE seemed preferable to dropping the observation entirely, or to making other, even stronger, assumptions allowing us to keep the observation.

[6]We contacted the authors of 51 papers asking for additional clarification on 175 estimates (i.e., mean, SE, or number of obs). If the authors did not respond to our initial request, we sent an additional email. Overall, we received 39 responses. Of those, 28 responses were ultimately useful and we recovered additional information on 78 estimates. The remainder had to be imputed.

Table 1: Study characteristics.

	Freq.	%		Freq.	%
Total number of studies	185	100.0			
Data type			*Reward type*		
Lab experiment	98	53.0	Money	154	83.2
Field experiment	29	15.7	Other	14	7.6
Other field data	20	10.8	Consumption good	8	4.3
Classroom experiment	18	9.7	Mixed	5	2.7
Online experiment	17	9.2	Health	2	1.1
Game show	3	1.6	Food	1	0.5
Subject population			Environment	1	0.5
University population	91	49.2	*Continent*		
General	63	34.1	Europe	78	42.2
Farmer	13	7.0	North America	56	30.3
Mixed	5	2.7	Asia	25	13.5
Children	4	2.2	Africa	8	4.3
Elderly	3	1.6	Oceania	7	3.8
Monkey	1	0.5	South America	6	3.2
Unknown	5	2.7	Multiple	3	1.6

Notes: In two studies, the geographic location of the data is unknown. These studies were run online through Amazon's Mechanical Turk or a mobile app, and the authors did not specify what geographic controls were used.

collecting data from more than 30 countries/regions. In total, the estimates of the loss aversion coefficient comprised in our dataset come from 71 countries/regions (Figure C.1 in the Online Appendix).

Next we look at the basic design characteristics of the studies. We have 185 "studies" reported in 150 papers, where a study is defined by a combination of several variables: type of the data, location of data collection, subject pool, type of reward, and continent of data collection. The frequency of each design characteristic is shown in Table 1.

The majority of our data comes from laboratory experiments, but we also have studies using non-experimental data such as surveys, stock market data, and game shows. Subjects were mostly recruited from the pool of university students or the general population. There is also a small set of studies which recruited special populations such as financial professionals, entrepreneurs, managers, and patients with psychiatric disorders or gambling problems. The type of reward used in the studies is mostly monetary. About three-quarters of the studies were conducted in Europe or North America.

Table 2: Types of estimates.

	All estimates		With SE	
	Freq.	Prop.	Freq.	Prop.
Aggregate-level	286	0.471	225	0.542
Individual-level mean	159	0.262	125	0.301
Individual-level median	162	0.267	65	0.157
Total	607	1.000	415	1.000

Notes: There are 85 cases where both mean and median of the distribution of individual-level estimates are reported. "With SE" indicates the observations where SEs are available. In addition, there are four aggregate-level estimates for which SEs are approximated with reported p-values, and 64 individual-level medians for which SEs are approximated with IQR. SEs are imputed for the rest of 124 observations (see Section 4.2).

Next, we look at the main variable of interest, the estimated coefficient of loss aversion. We have a total of 607 estimates in the dataset (Table 2). About half of these estimate the degree of loss aversion of a "representative" subject by pooling data from all subjects together (we call these aggregate-level, or simply aggregate, estimates). The other half estimated the coefficient for each individual subject in the study and reported summary statistics of the distribution, either mean or median. There are 85 cases where we have both the mean and the median of the distribution of the loss aversion coefficients estimated at the individual level.

Finally, we look at the specification of the functional form and the definition of loss aversion λ (Table 3). There are 302 observations which assume the CRRA form for the basic utility functions u^+ and u^- as in equation (4), following Tversky and Kahneman (1992), but 190 of them assume and estimate the common curvature for gains and losses ($\alpha = \beta$). We observe less variation in the specification of reference points and the definition of loss aversion coefficients. Three-quarters of the observations set the reference point at zero, but our dataset also includes studies where reference points are assumed to be subjects' status quo or expectations. More than 80% of the observations estimate the loss aversion coefficient λ as Kahneman and Tversky (1979) define (equation (3)).

4 Results

We structure the results into three distinct parts. We start from a non-parametric analysis of the reported loss aversion coefficients and their SEs. We subsequently fit random-effects meta-analytic distributions to the data, and document the estimated mean and median loss aversion.

Table 3: Utility specification.

	Freq.	%		Freq.	%
Total number of estimates	522	100.0			
Loss aversion λ			*Functional form of u*		
Kahneman-Tversky	445	85.2	CRRA	302	57.9
Köbberling-Wakker	37	7.1	CARA	53	10.2
Kőszegi-Rabin	9	1.7	Linear	73	14.0
Other	3	0.6	Other parametric	32	6.1
Not reported	28	5.4	Nonparametric	16	3.1
Reference point			Not reported	46	8.8
Zero	394	75.5			
Status quo	60	11.5			
Expectation	18	3.4			
Other / Not reported	50	9.6			

Notes: There are 85 cases where both mean and median of the distribution of individual-level estimates are reported. We keep only one measure from each of these observations.

Finally, we conduct a series of meta-regressions to see to what extent we can explain the estimated between-study variance.

4.1 Nonparametric Analysis

We start our presentation of the results by showing some non-parametric patterns in the reported loss aversion coefficient λ. Since we do not need SEs for this analysis, we can make use of all the estimates of loss aversion we coded in the dataset. Figure 3A shows the distribution of all the encoded loss aversion parameters. The mean of the parameters is 1.97. Given the right skew in the distribution, the median is considerably lower than the mean, at 1.69.

Figure 3B shows the same estimates, but now plots separate density functions for the estimates obtained from aggregate-level means ($N = 286$), individual-level means ($N = 159$), and medians ($N = 162$). Aggregate-level estimates can be seen to have the lowest mode (around 1.36), with individual-level medians having a slightly higher mode around 1.89. Means of individual-level estimates show a fat right tail, indicating a higher frequency of larger values. Table 4 shows summary statistics of reported λ for each type of measurement. The means of aggregate-level estimates and individual-level medians are close together, at 1.96 and 1.83, respectively. The somewhat lower mean of the latter results from fewer very large observations amongst medians than amongst aggregate estimates. The individual-level means have the largest variation (SD = 2.61 versus 1.67 for aggregate means and 0.76 for individual-level medians), including some of

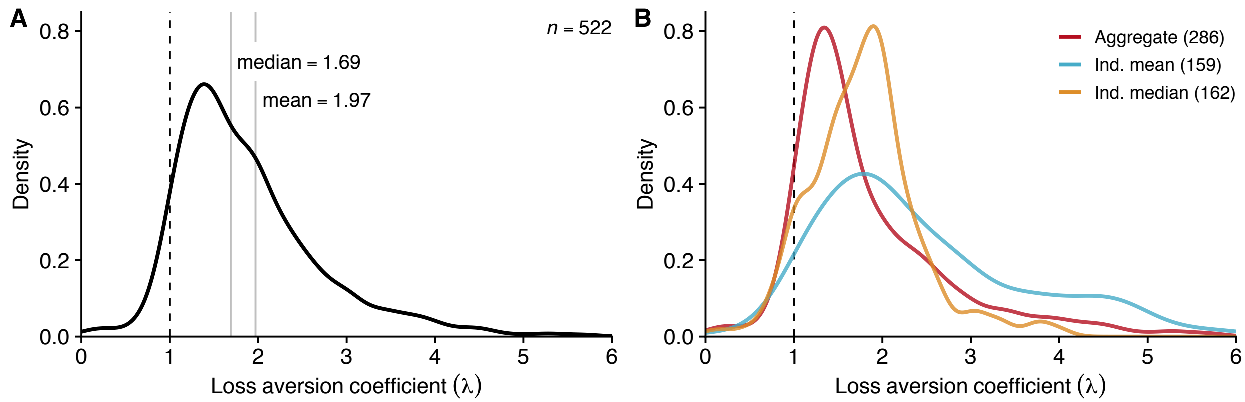

FIGURE 3: Distribution of loss aversion parameters. (A) All reported estimates combined. (B) Separated by the type of estimates. *Notes*: There are 85 cases that report both individual-level mean and median. We keep individual-level medians from these cases in panel A, while panel B includes all 607 estimates in the data. The x-axis is cut off at 6 for better visual rendering. In the Online Appendix, Figure C.3 shows density plots of $\log(\lambda)$ and Figure C.4 shows the empirical CDF of λ for each type of estimates.

TABLE 4: Summary statistics of reported λ.

Type	N	Mean	SD	Q25	Median	Q75	Min	Max
Aggregate-level	286	1.955	1.667	1.280	1.545	2.214	0.040	23.460
Individual-level mean	159	2.945	2.610	1.625	2.180	3.420	0.110	19.861
Individual-level median	162	1.829	0.759	1.404	1.800	2.000	0.110	7.500
All	607	2.181	1.857	1.340	1.790	2.375	0.040	23.460

Notes: There are 85 cases that report both individual-level mean and median.

the smallest estimates as well as some of the largest. The truncated nature of the distribution then results in the highest mean by far, 2.95.

The differences between measurement types above cannot be interpreted causally. That is, the different measurements generally derive from different studies and are based on different data, so that the observed differences cannot be directly ascribed to the type of measurement used. To gain insight into the effect of measurement type, we can conduct an analysis based on the 85 studies for which both means and medians are reported. With a mean of the means of 3.47 (median of means 2.08) and a mean of the medians of 1.71 (median of medians 1.69), the results confirm the ones for the overall sample (see Figure C.5 in the Online Appendix). That is, the individual-level estimates tend to be rightward skewed, and this strongly affects the aggregate estimate reported in a paper when means of the individual-level estimates are used instead of medians. This issue, however, will be at least partly remedied by the observation that means of individual-level estimates also tend to come with increased SEs, which will in turn lead to increased pooling of the larger estimates in our meta-analytic estimations.

The earliest evidence recorded in our dataset is Tversky and Kahneman's (1992) famous 2.25 (see Figure C.2 in the Online Appendix for the time trend of reported estimates). Half of the estimates in the data are found in papers published after 2015. Individual-level estimates appear in the dataset after 2006, in part due to the rise of common experimental elicitation procedures. The raw data do not reveal a clear time trend, suggesting that estimates have remained the same on average for the last 30 years.

Just looking at the raw reported λ, Figure 3 and Table 4 suggest that the "average" loss aversion coefficient λ would locate somewhere between 1.8 and 2.9. At the same time, there is high dispersion in the reported λ. These rough initial estimates, however, do not take the quality of the estimate into account. The latter can be assessed objectively by means of the standard error associated to each estimate, which we can use to calculate a proper "meta-analytic average" by weighing each estimate by the inverse of its standard error. Estimates falling far from the mean will then be given less weight to the extent that they have large associated standard errors.

4.2 Precision of Estimates

Figure 4 illustrates the relationship between reported estimates and their associated SEs using data points where we have both a value for λ and an associated SE (either reported in the paper or re-constructed from other available information by us).

Figure 4 tells us several things about empirical estimates of loss aversion. Loss aversion may be imprecisely estimated, for instance, because of a small sample size, or because of other characteristics of the design. In our data, estimates that are far from 1 have higher SEs. They are not precisely estimated coefficients due to, for example, small sample size or other issues with the design or behavior. More precisely estimated λ's tend to cluster between 1 and 2. Second, 398 out of 419 estimates (95%) are (weakly) larger than 1, producing a massive asymmetry in the funnel plot. Third, about 76.6% (305 out of 398) of $\lambda \geq 1$ estimates report results that are statistically significantly different from 1 (two-sided Wald test with a significance level of 0.05).

Table 2 above shows that 192 out of our 607 recorded estimates are missing SEs. We approximate SEs from IQRs and p-values for 68 of these observations (footnote 5), which leave 124 SEs missing. Since standard errors are a fundamental ingredient for meta-analysis because they provide weights for the observations, we thus risk losing many observations, including the iconic measure of 2.25 reported by Tversky and Kahneman (1992). If studies not reporting any SEs are different from studies reporting them, we may furthermore distort our estimates systematically. To overcome this issue, we impute the missing SEs using the subset of the data for which we have

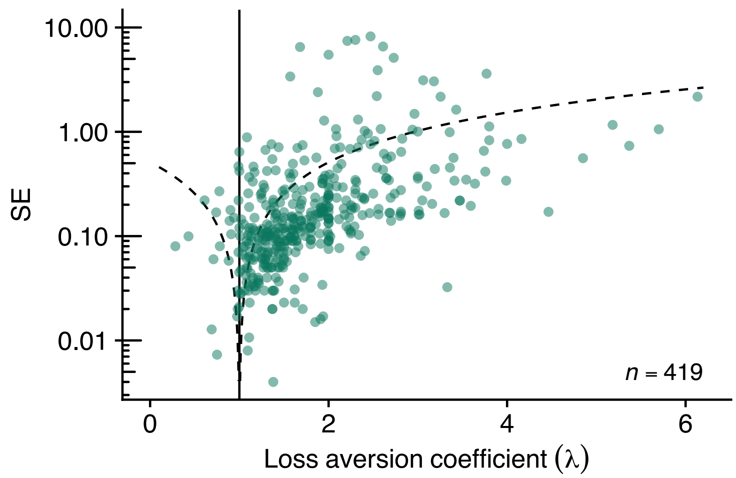

FIGURE 4: Relationship between reported λ and associated SE. Estimates with corresponding SEs reported are included ($n = 419$). *Notes*: For observations which have both individual-level mean and median, we keep the median in this figure. The vertical solid line corresponds to no loss aversion $\lambda = 1$. Two dashed curves represent the boundaries for statistically-significant loss aversion/tolerance ($\lambda \neq 1$). The x-axis is cut off at 6 and the y-axis is displayed in the log-scale for better visualization.

both λ and its associated SE. See Online Appendix A.2 for the detail of SE imputation. We will, from now on, make use of the full set of observations.

4.3 Average Loss Aversion in the Literature

The main goal of our meta-analysis is first to obtain the "best available" estimate of the loss aversion coefficient λ combining the available information in the literature and then to understand the heterogeneity of reported estimates across studies. To this end, we analyze the data using a *Bayesian hierarchical modeling* approach.

Setup. Consider the dataset $(\lambda_i, se_i^2)_{i=1}^m$, where λ_i is the ith *measurement* (or *observation*) of the loss aversion coefficient in the dataset and se_i is the associated standard error that captures the uncertainty surrounding the estimate. In the benchmark model, we assume that the ith reported estimate λ_i is normally distributed around the parameter $\overline{\lambda}_i$:

$$\lambda_i \mid \overline{\lambda}_i, se_i \sim \mathcal{N}(\overline{\lambda}_i, se_i^2), \tag{5}$$

where the variability is due to the sampling variation captured by the known standard error se_i.[7]

[7]The parameter $\overline{\lambda}_i$ is often referred to as the "true effect size" in the random-effects meta-analysis.

Sampling variation is part of the observed variation in the reported estimates $(\lambda_i)_{i=1}^m$, but it may not be all, since there is a possibility of "genuine" heterogeneity across measurements (due to different settings, for example). We model this by assuming that each $\overline{\lambda}_i$ is normally distributed, adding another level to the hierarchy:

$$\overline{\lambda}_i \mid \lambda_0, \tau \sim \mathcal{N}(\lambda_0, \tau^2), \tag{6}$$

where λ_0 is the *overall mean* of the estimated loss aversion parameters $\overline{\lambda}_i$, and τ is its standard deviation, capturing the variation between observations in the data. The overall variance in the data, therefore, consists of two parts, the between-observation variance, τ^2, and the individual sampling variation coming from measurement uncertainty, se^2. This can be clearly seen by combining expressions (5) and (6) into one:

$$\lambda_i \mid \lambda_0, \tau, se_i \sim \mathcal{N}(\lambda_0, \tau^2 + se_i^2).$$

Model estimation. We start from fitting the model expressed as equations (5) and (6), re-stated as model M1 here, to the data:

$$\begin{aligned} \lambda_i \mid \overline{\lambda}_i, se_i &\sim \mathcal{N}(\overline{\lambda}_i, se_i^2), \\ \overline{\lambda}_i \mid \lambda_0, \tau &\sim \mathcal{N}(\lambda_0, \tau^2), \\ \lambda_0 &\sim \text{half } \mathcal{N}(1, 5), \\ \tau &\sim \text{half } \mathcal{N}(0, 5), \end{aligned} \tag{M1}$$

where "half $\mathcal{N}$" indicates a normal distribution folded at 0 to exclude non-permissible negative values.[8] This model incorporates the assumption that every observation is statistically independent, and that the observations are normally distributed. We later relax these rather strong assumptions.

The estimated overall mean λ_0 is 1.810 with a 95% credible interval (CrI) of [1.742, 1.880].[9,10] The mean is clearly and significantly lower than the non-parametric result that we saw above,

[8] We estimate the model in `Stan` (Carpenter et al., 2017) using the Hamiltonian Monte Carlo simulations, an algorithm for Markov Chain Monte Carlo, and launch it from `R` (`https://www.r-project.org/`) using `RStan` (Stan Development Team, 2020). Priors for the population-level parameters are chosen in such a way as to be mildly regularizing, i.e., they are informative, but typically encompass ranges that are one order of magnitude larger than the estimated values we expect based on the range of the data (McElreath, 2016).

[9] A Bayesian credible interval (CrI) of size $1 - \alpha$ given data D is an interval $[L(D), U(D)]$ such that $\Pr[L(D) \leq \theta \leq U(D)] = 1 - \alpha$, where θ is the parameter of interest. Unlike the (frequentist) confidence interval, CrI has a literal probabilistic interpretation: given the data, there is a $100 \times (1 - \alpha)\%$ probability that the true parameter value is in the interval.

[10] Results from the "classical" (frequentist) random-effects meta-analysis are presented in Online Appendix D. We obtain virtually identical estimates.

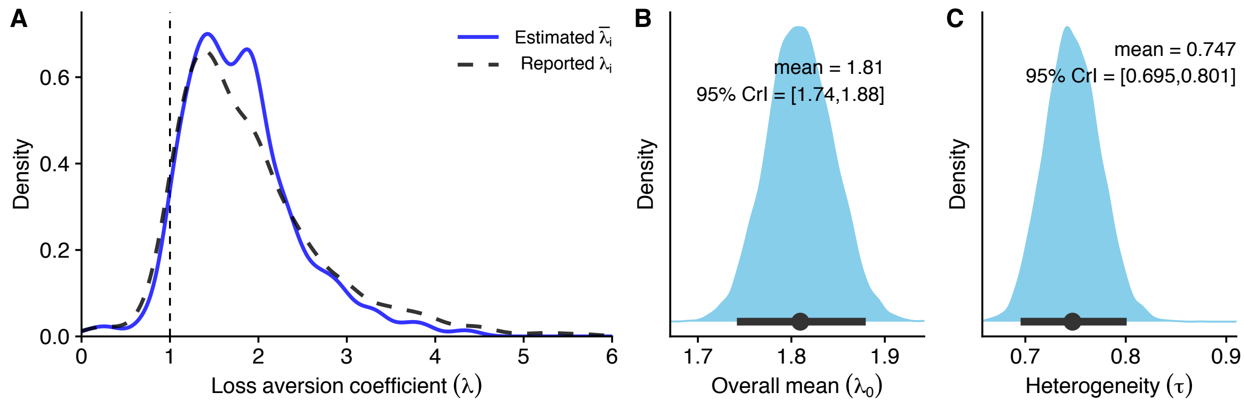

FIGURE 5: (A) Density plot of loss aversion coefficient, estimated $\bar{\lambda}_i$ versus reported λ_i. (B) Posterior draws of the overall mean λ_0. (C) Posterior draws of the heterogeneity parameter τ. *Notes:* The dashed curve in panel A indicates the density of the observed loss aversion parameters, λ_i; the solid curve in panel A indicates the density of the estimated parameters, $\bar{\lambda}_i$. Observations above six are not shown in panel A for better visualization. The black dots and lines in panels B and C represent the posterior means and the 95% credible intervals of λ_0 and τ, respectively.

which was 1.965. This is shown in Figure 5AB. The density of estimated $\bar{\lambda}_i$ is lower than the one of observed λ_i for values above 2.5. The same occurs for values below one. This is meta-analytic pooling at work—estimates that fall far from the mean are shrunk towards more plausible values, with the amount of shrinkage proportional to the standard error. See discussion in Online Appendix B.2.

The estimates produced are of course only valid conditional on our assumptions. We already know that the normality assumption seems a stretch, given the skewed distribution of the reported λ. To see this, we can take a look at the *posterior predictive distribution*—the distribution of loss aversion coefficients we would expect new observations λ_{new} to display, provided that the characteristics of the studies from which these observations are obtained are similar on average to those of past studies—and compare it to the distribution of actual observations.[11] This is shown in Figure 6A. Relatively to the actual observations—either as reported (λ_i), or as estimated ($\bar{\lambda}_i$)—the posterior predictive distribution overestimates the likelihood of values smaller than 1, while it underestimates the likelihood of intermediate values between 1 and 2. It does not attribute any

[11]Formally, the posterior predictive distribution is written:

$$\pi\left(\lambda_{\text{new}} \mid (\lambda_i)_{i=1}^m\right) = \int \pi\left(\lambda_{\text{new}} \mid \theta\right) \pi\left(\theta \mid (\lambda_i)_{i=1}^m\right) d\theta,$$

where $\theta = ((\bar{\lambda}_i)_{i=1}^m, \lambda_0, \tau)$ is a vector of model parameters (Gelman et al., 2014). Evaluating this integral is difficult, but we approximate it by drawing $\lambda_{\text{new}}^{(s)} \sim \mathcal{N}(\lambda_0^{(s)}, \tau_{(s)}^2)$ using posterior simulations $(\lambda_0^{(s)}, \tau_{(s)})$, $s = 1, \ldots, N$. We have 8,000 draws (2,000 iterations $\times$ 4 chains) of λ_{new}.

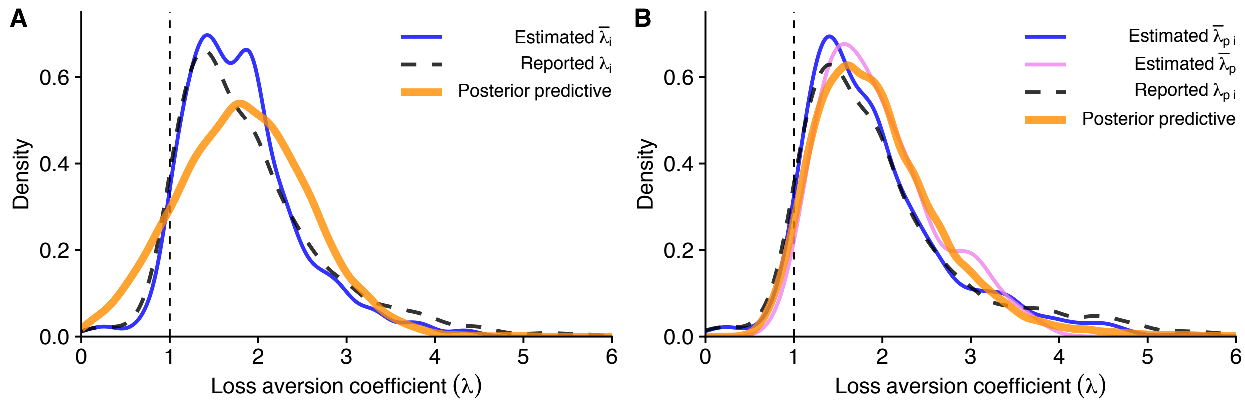

FIGURE 6: Distributions of reported and estimated λ, and posterior predictive distribution of λ. (A) Assuming a normal distribution for the population level (model M1). (B) Assuming a log-normal distribution for the population level (model M2).

probability to values beyond 4, which are not uncommon in the data. It thus seems desirable to look for a model that may provide a better fit to the data.

We now extend the baseline model M1 in two ways. First, we use a log-normal distribution for the population-level distribution. Second, we explicitly model the nesting of observations in papers, in order to overcome any potential distortions deriving from non-independence of observations. Remember that our 607 observations have been obtained from 150 distinct papers, the largest number of observations in a single paper being 53 (Rieger et al., 2017; Wang et al., 2017). The independence assumption seems rather heroic in this case. To do this, we introduce paper-level estimates as an additional hierarchical level. Let λ_{pi} be the ith estimate reported in paper p. We formulate a model as follows:

$$\begin{aligned}
\lambda_{pi} \mid \overline{\lambda}_{pi}, se_{pi} &\sim \mathcal{N}(\overline{\lambda}_{pi}, se_{pi}^2), \\
\overline{\lambda}_{pi} \mid df, \overline{\lambda}_p, \sigma_p &\sim t(df, \overline{\lambda}_p, \sigma_p^2), \\
\overline{\lambda}_p \mid \lambda_0^\ell, \tau_\ell &\sim \log \mathcal{N}(\lambda_0^\ell, \tau_\ell^2), \\
\lambda_0^\ell &\sim \mathcal{N}(1, 5), \\
\tau_\ell &\sim \text{half } \mathcal{N}(0, 5), \\
df &\sim \text{half } \mathcal{N}(0, 5), \\
\sigma_p &\sim \text{half } \mathcal{N}(0, 5).
\end{aligned} \tag{M2}$$

The system M2 now explicitly models the nesting of the estimated observation-level parameters, $\overline{\lambda}_{pi}$, in paper-level estimates, $\overline{\lambda}_p$. The former are modeled as following a robust student-t

Table 5: Summary of estimation results.

	Distributional assumption			Posterior of λ_0				Posterior of τ			
Model	Obs. level	Paper level	Pop. level	Mean	SD	2.5%	97.5%	Mean	SD	2.5%	97.5%
M1	Normal		Normal	1.810	0.036	1.741	1.880	0.747	0.027	0.695	0.801
M2	Normal	Student-t	Log-normal	1.955	0.072	1.824	2.104	0.743	0.342	0.605	0.907

Notes: In Model M2, (λ_0, τ) are calculated from the log-normal parameters $(\lambda_0^\ell, \tau_\ell)$ by $\lambda_0 = \exp(\lambda_0^\ell + \tau_\ell^2/2)$ and $\tau^2 = [\exp(\tau_\ell^2) - 1] \exp(2\lambda_0^\ell + \tau_\ell^2)$.

distribution instead of a normal distribution to account for observed outliers.[12] The latter are modeled as following a log-normal distribution. Note the super-/sub-scripts ℓ in the location and scale parameters $(\lambda_0^\ell, \tau_\ell^2)$ of the log-normal distribution. We can calculate the mean and the median of the distribution by $\exp(\lambda_0^\ell + \tau_\ell^2/2)$ and $\exp(\lambda_0^\ell)$, respectively. The mean of the log-normal distribution is given by $\lambda_0 \equiv \exp(\lambda_0^\ell + \tau_\ell^2/2)$.

We again start by examining the fit of the model to the data, and by summarizing the population-level parameters. The model fit is shown in Figure 6B. The log-normal distribution can now be seen to fit the estimated paper-level data well. The distribution of the paper-level observations has more probability mass between about 1 and 3, but less beyond that point, compared to the actual study-level observations. The degrees of freedom of the student-t distribution are estimated at 1.32, thus vindicating the use of the robust distribution. The mean loss aversion parameter obtained from this estimation is 1.955, with a 95% CrI of [1.824, 2.104]. Notice, however, that even though this estimate is virtually identical to the one obtained under the standard model at the outset, that occurs by coincidence rather than being a feature of the model. One can further see that there is now increased uncertainty surrounding the prediction interval. This is indeed natural, since the paper-level estimates are surrounded themselves by larger amounts of uncertainty, which is then passed up the hierarchy to the aggregate parameters.

Robustness checks. Online Appendix B.2 presents estimation results for two additional models, but our preferred model M2 fits better than these "intermediate" models. We also estimated the models under different priors or using the "complete" data including only observations where associated SEs are available, and obtained similar conclusions. These robustness checks are presented in Online Appendix B.3.

[12] We estimate the degrees of freedom of the distribution, df, endogenously from the data. This allows us to determine whether the student-t distribution provides a good fit, which is the case if the degrees of freedom are small, or whether it converges to a normal distribution, which is the case for large degrees of freedom (Kruschke, 2010).

4.4 Explaining Heterogeneity

We observe a non-negligible amount of between-paper heterogeneity (expressed in estimated τ in Table 5, model M2) among reported estimates of λ. In this section, we seek to understand the source of this variability in order to provide a tentative answer to our second key question: "Do reported estimates of λ systematically vary by underlying design characteristics for measurement of loss aversion?"

Remember that we coded several features about the characteristics of study design (Table A.1 in Online Appendix). Figure C.8 provides a first look into how these features are related to reported estimates of λ. Each panel presents how the reported λ varies by underlying design characteristics. We do observe some patterns in the figure, but the effects appear rather weak and it is not clear if these relations are systematic and robust.

We approach this question with a *random-effects meta-regression*, which extends our previous random-effects model by incorporating coded features of the observation or the paper into the model. More precisely, we set up a new model, which expands model M2 by allowing for the location of the observations to be systematically shifted depending on observed characteristics of the observation or the paper. The model looks as follows:

$$
\begin{aligned}
\lambda_{pi} \mid \overline{\lambda}_{pi}, se_{pi} &\sim \mathcal{N}(\overline{\lambda}_{pi}, se_{pi}^2), \\
\overline{\lambda}_{pi} \mid df, \overline{\lambda}_p, \sigma_p, \beta &\sim t(df, \overline{\lambda}_p + X_{pi}\beta, \sigma_p^2), \\
\overline{\lambda}_p \mid \lambda_0^\ell, \tau_\ell &\sim \log\mathcal{N}(\lambda_0^\ell, \tau_\ell^2), \\
\lambda_0^\ell &\sim \mathcal{N}(1, 5), \\
\tau_\ell &\sim \text{half } \mathcal{N}(0, 5), \\
df &\sim \text{half } \mathcal{N}(0, 5), \\
\sigma_p &\sim \text{half } \mathcal{N}(0, 5),
\end{aligned}
$$

where X_{pi} is a vector of study characteristics associated with ith observation reported in paper p. These characteristics consist mostly of dummy variables taking the value of 0 or 1, with β a vector of coefficients. To facilitate the interpretation of the constant, non-dummy independent variables included in vector X_{pi} are mean-centered—each coefficient in the vector β then captures the additive effect on the paper-level mean $\overline{\lambda}_p$, relative to the "baseline study" (characterized by the omitted categories in dummy variables and the means of non-dummy variables) which will become clear later.

Estimation results are presented in Figure 7. First, the posterior mean of the estimated $\overline{\lambda}_{p^*}$ for the benchmark study p^* is 1.991 (SD = 0.211, 95% CrI = [1.577, 2.420]). Each estimated coefficient in β captures the effect of the study characteristic from this benchmark value.

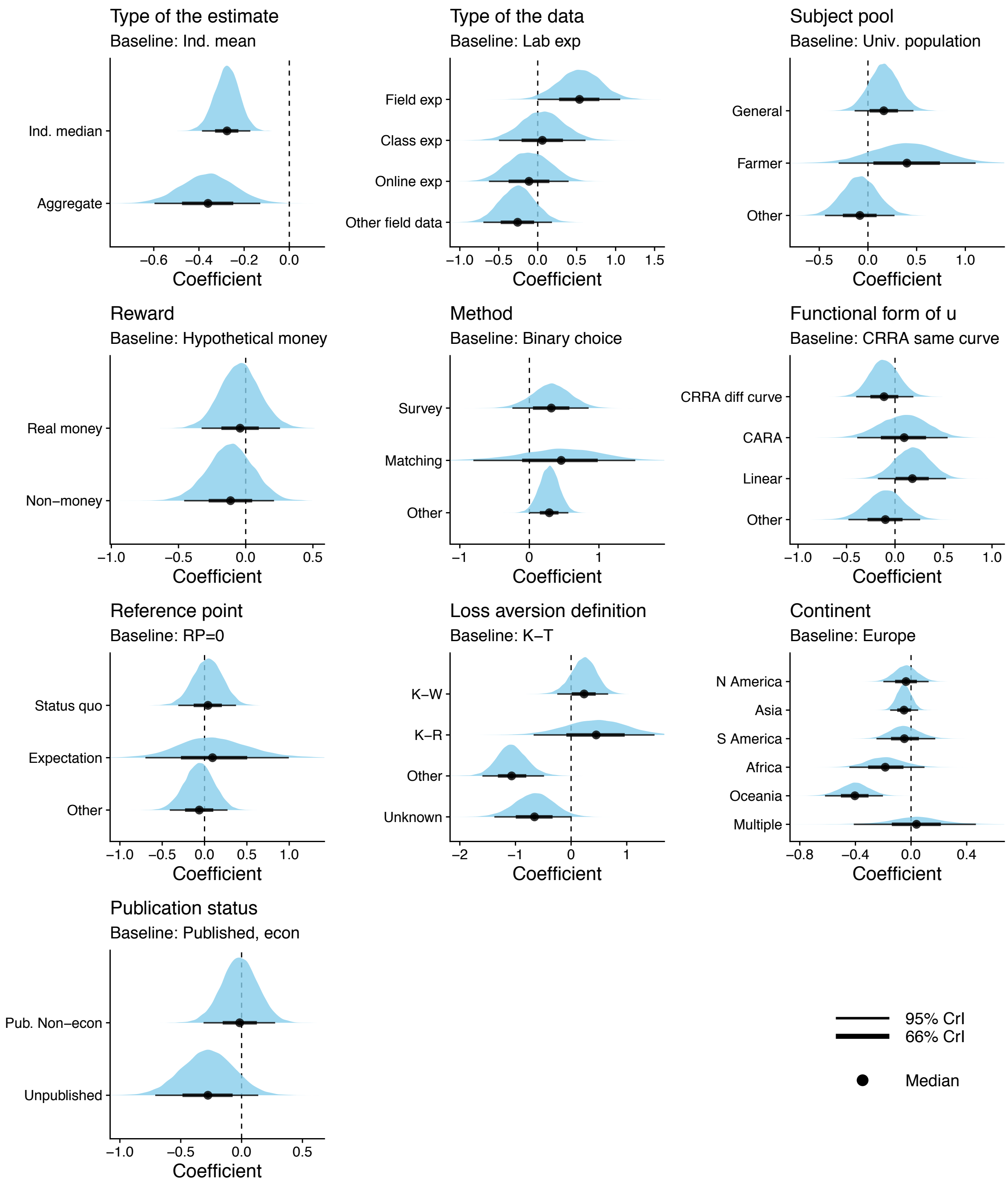

FIGURE 7: Bayesian random-effects meta-regression. Posterior distributions of coefficients β, together with posterior medians (black dot), 66% (thick solid line) and 95% (thin solid line) credible intervals, are shown. Table C.1 in the Online Appendix presents the result in a table format.

As we see above in Section 4.1, the type of estimates reliably captures the variation in reported λ— individual-level means tend to be higher than the other two types of estimates, due to skewed distributions of individually-estimated λ. We also find that field experiments are associated with higher λ compared to laboratory experiments and studies recruiting general population samples are also associated with higher values of λ compared to the studies with a population of university students. We do not observe differences between studies using monetary rewards and non-monetary rewards, but survey studies tend to produce lower estimates of λ than the binary lottery choice tasks which are common in laboratory experiments. In terms of the specification of the value function, it does not seem to matter much which functional form (CRRA, CARA, etc.) one assumes for the basic utility functions u^+ and u^-, or whether reference points are assumed to be zero, status quo, or expectations. Studies estimating λ following the definition by Köbberling and Wakker (2005) produce higher λ compared to the standard Tversky and Kahneman's (1992) definition, but the effect is modest.

Taken together, our Bayesian meta-regression analysis uncovers some factors that are associated with the size of reported loss aversion coefficients, but it is still a difficult task to draw a complete picture of the observed heterogeneity. We note that 14.4% of the between observation variance is explained by covariates.

4.5 Publication Bias

The cumulation of scientific knowledge is threatened by selective reporting or publication of findings. When a theory makes a strong prediction about the sign or magnitude of a certain effect and the literature takes such effects to be the norm, then new studies that find "unusual" results may not be written up at all, may not be reported in papers, and academic journals may be reluctant to publish papers into which such estimates are included. We will refer to such selective reporting of scientific findings collectively as "publication bias".

In the context of loss aversion, one might suspect that researchers preferentially report evidence for loss aversion ($\lambda \geq 1$) and put evidence for loss tolerance ($\lambda < 1$) "in the file drawer" because such results contradict the initial hypothesis. Other hypotheses are, however, possible, for instance because researchers in some disciplines may be motivated to attack the "prevailing paradigm" of loss aversion, or because some journals may be interested in publishing results that conform to the economic orthodoxy of no loss aversion. Even coming up with a null hypothesis is thus more complex in our case than it would be when trying to simply ascertain the effect of a treatment or its absence.

A tool often used to detect publication bias are funnel plots (Stanley and Doucouliagos, 2012). A *funnel plot* is a scatter plot of the estimates of effect sizes against their precision. Figure 8

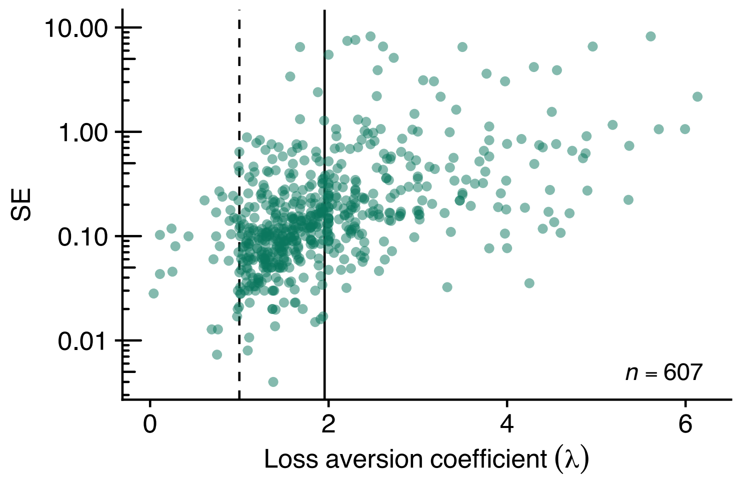

FIGURE 8: Funnel plot. *Notes*: The vertical dotted line corresponds to no loss aversion $\lambda = 1$. The vertical solid line corresponds to the estimated mean λ_0 from model M2, 1.955. The x-axis is cut off at 6 and the y-axis is displayed in the log-scale for better visualization.

shows such a plot for all data points for which we have both an estimate of loss aversion and an associated standard error (thus excluding estimates with imputed SEs). In the absence of bias, we would expect the observations at the bottom of the graph to be concentrated around the mean estimate, indicated by the solid vertical line. As we move up in the graph and the precision of the studies decreases, we would then expect an increase in the degree of dispersion. In the absence of publication bias, this dispersion ought to be symmetric around the mean. A larger number of observations in the upper right side of the graph compared to the upper left side would then be an indicator of classical publication bias, whereby estimates of loss aversion that fall closer to 1 and are not significant are less likely to be reported.

At first sight, there would indeed appear to be such an asymmetry in the graph. We clearly observe some large estimates in the upper right part of the graph, and hardly any corresponding estimates in the upper left part. This may, however, in part be due to differences in the reported measures. Figure 9 shows separate funnel plots for the different types of measures—aggregate estimates, individual-level means, and medians. The asymmetry just discussed seems indeed most pronounced for individual-level means, as well as aggregate estimates. Median estimates tend to show less of a bias, providing at least some indication that the observed asymmetry stems partially from the measurement type (plotting the log of loss aversion against the precision instead does not affect any of these insights). Since the largest estimates also have the largest SEs on average, the asymmetry shown in this graph will be at least partially taken care of by the pooling inherent in meta-analytic estimation.

Even more pronounced than these (weak) indications of traditional publication bias, how-

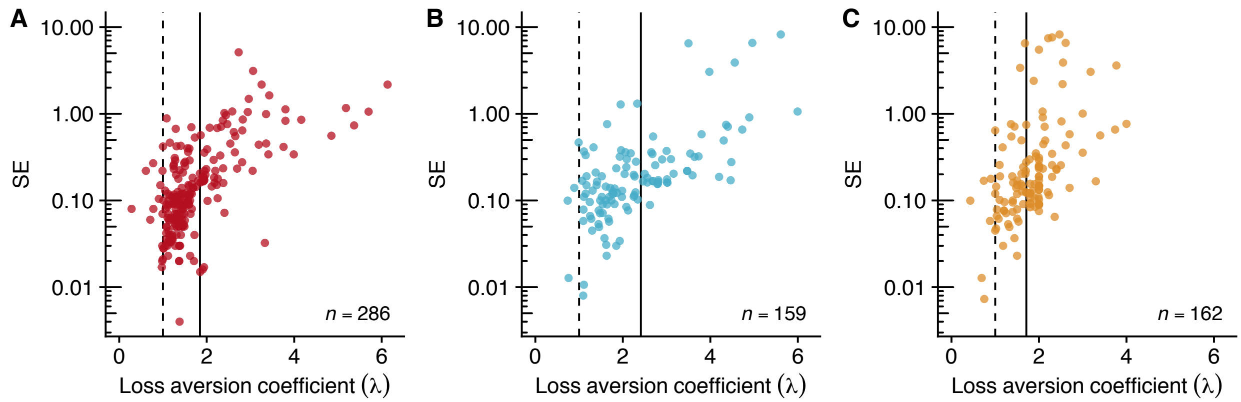

Figure 9: Funnel plot by the type of estimates. Vertical lines correspond to mean estimates M2 applied to each subset of data separately. (A) Aggregate-level estimates. (B) Individual-level means. (C) Individual-level medians.

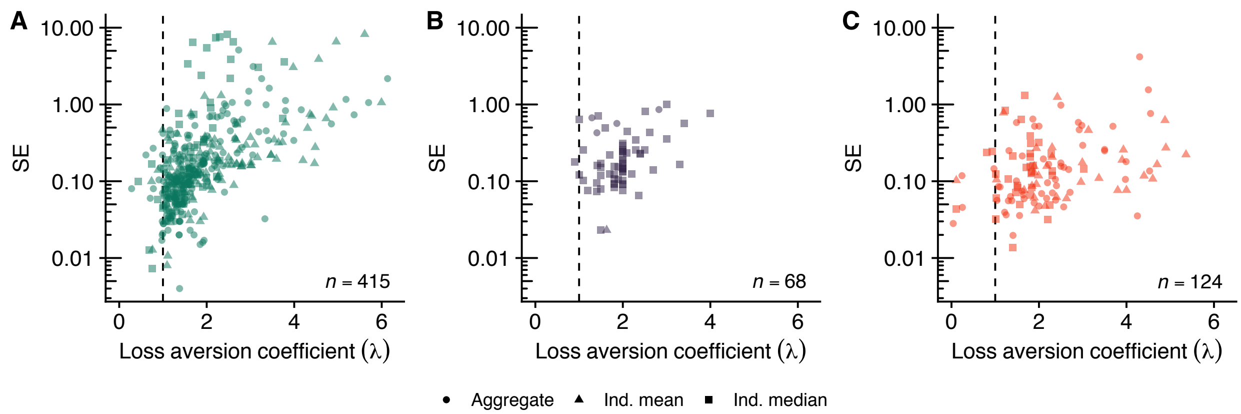

Figure 10: Funnel plot by the type of SEs. (A) SEs reported in the paper. (B) SEs approximated by available information (IQR or p-value). (C) SEs imputed.

ever, is the large cluster of studies at the bottom left corner of the graph. One might be concerned they indicate a model of publication bias where the true λ is around 0.8–1 but editorial bias favors publication of studies with higher estimates at lower precision. (In such a model, low-precision, low-estimate studies would not be published because of the judgment of journal referees or editors, and high-precision, high-estimate studies would not be published because they are statistically improbable.) While this finding would be disturbing, there are, however, better alternative explanations. Figure 10 separates the funnel plot by how the SE of the estimate were obtained: either directly from the paper, estimated by percentiles (e.g., IQR, median, etc.), or imputed entirely (via number of observations and other factors, see Section 3.2). Notably, the asymmetry is only visible in the graph where SEs are reported (panel A). Since it is not consistent with any story of publication bias for journal editors to only be biased toward studies that report standard errors, it is difficult to reconcile this figure with any story of publication bias. A more plausible explanation is simply that the non-normality of λ and the truncation of estimates at 0 mechanically induces smaller SE calculations for lower values. When researchers instead report non-parametric measures like IQR, these high-precision estimates disappear.

Our sample of working papers is small, but it might be suggestive of a form of publication bias, just at a much smaller magnitude. Publication bias would suggest that the estimates found in published papers would vary from working papers. Our meta-regressions (presented in Section 4.4) predict observed loss aversion coefficients to be 0.27 points lower in working papers than in published studies, roughly about 1.7 (though the two numbers are not statistically different at the 5% level). Even if this number were taken at face value, we must remember the "true" value of lambda would need to be a weighted average of the working paper estimate and published estimate, so 1.7 can still be thought of as a lower bound of the "true" estimate. Notably, it is within the confidence interval of estimated lambda and much closer to our estimates than the 0.8–1 range which we dismissed above.

5 Discussion

Loss aversion is an important concept in behavioral economics and has been applied widely. This paper reports a meta-analysis of empirical estimates of the loss aversion coefficient λ. Our preferred specification indicates a mean $\lambda = 1.955$ and a 95% credible interval of $[1.824, 2.104]$. Many other specifications are within 0.1–0.3 of this finding and produce confidence intervals that do not include 1 or 2.25. The former number is consistent with no degree of loss aversion; the latter is an early estimate from Tversky and Kahneman (1992) which seems a bit too high. While there is a wide degree of heterogeneity across estimates, in general, no single factor emerges from meta-regression that greatly changes estimated loss aversion. Estimates derived from non-university populations, field experiments, and means of individual elicitations (compared to aggregates) are correlated with a modest increase in the loss aversion parameter.

With any empirical analysis, a key concern is differential selection of reported data. The outlined criteria for inclusion in our dataset are explicit and objective; we have no reason to believe, ex-ante, it should be correlated with our parameter of interest. However, our dataset is dominated by published studies. To the extent that publication may vary with a reported λ parameter, our analysis may suffer from bias.[13] Because the parameter of interest here is bounded by zero and positively skewed, it is difficult to use standard meta-analysis technique like funnel

[13] It is important to note that any research synthesis technique would suffer from this issue, not just meta-analysis (Borenstein et al., 2009).

plot asymmetry (see Figure 8) and regression to measure publication bias. Nonetheless, those techniques do not indicate a degree of bias that invalidate our general findings. An important test of publication bias is whether unpublished working papers report reliably different results than published ones. There appears to be a modest difference; unpublished papers are about 0.28 lower ($\lambda \approx 1.7$) which, notably, is still within the confidence interval of our current estimate.

In the introduction we promised to return to subtle, important comparisons between meta-analysis and narrative review. We will do that now.

Our stance is that many readers will not know much about meta-analysis and might be skeptical about it, especially in comparison to the familiar style we call "narrative review." The discussion anticipates such skepticism and offers counter-arguments. The unabashed goal is to advocate for more appreciation of the underdog method of meta-analysis.

We will start with fears about narrative review. Such reviews could be influenced by biases in remembering recent (and perhaps socially-connected) salient, recently-encountered data. Meta-analysis is a partial antidote.

There are two natural fears about meta-analysis. One is that it is somehow a mistake to reduce all research in a field to a single value or a range of numbers. The second fear is that simple apples-to-apples comparisons across studies neglects heterogeneity of methods (see Borenstein et al., 2009, for a history of such critiques). These fears are natural, but meta-analysis techniques have evolved to allay such fears.

The "single number" fear is misguided because humans—including scientists—value simplicity. If meta-analysis did not provide a carefully and transparently derived value (or range), researchers will imagine another simplified value, one way or another. It is also often necessary to choose *some* value as an input into structural models within economics, and to make power calculations during pre-registration. Furthermore, the purpose of this or any meta-analysis is not to simply provide one number, but to demonstrate how such numbers vary given other factors that categorize studies, using meta-regression techniques (e.g., Figure 7). Stanley and Doucouliagos (2012) reminds us that meta-regressions simply take a well-known and established technique to understand variation in many kinds of economic data, and just apply it to the data our own profession generates.

Next we turn to a sharper comparison between meta-analysis and narrative review. In a narrative review, there is no explicit attempt to canvas all studies based on stated criteria. Instead, an expert reviewer chooses studies that seem to be of especially high-quality or pivot the scientific trajectory in a useful new direction. It is similar to an historical analysis of progress in a field. To relate the two methods, a narrative review is simply a meta-analyis with an altered subjective weighting system. The subjective weights are judgments by the reviewer of what findings readers should care the most about, and the reasons why.

One analogy is to sports commentary. Meta-analysis is like a "play-by-play" analyst who describes every action on the field with an even tone. Narrative review is like a "color analyst" who picks out certain plays which are unusually important and explains why they are special. The color analyst *adds* the dramatic emphasis that the play-by-play analyst suppresses. The analogy should make clear why both kinds of commentary are useful; the two together are better than each one alone.[14] A similar analogy is to the proud divide in newspapers between the "news" side (meta-analysis) and "opinion" (narrative review).

A tricky, interesting question is what meta-analysis and narrative review can say about influential "breakthrough" papers? Narrative reviews often remark on how a particular study represents a breakthrough in using new methods or presents a surprising finding that should be prioritized to be studied further. Because meta-analysis is backward-looking, it is not ideally-equipped to identify useful breakthroughs. Because narrative review is subjective, it can look forward and might do better.

An example is De Martino et al. (2010). They found that two patients with damage to the amygdala area of the brain were not averse to losses at all. This is a tiny finding with a large standard error; it's a teaspoon of water added to a swimming pool. Meta-regression is worthless because there is no power to detect an "amygdala damage" vs. "no amygdala damage" difference if the study were added to our corpus. But despite the tiny $N = 2$, it is a clue that could shed light on the fundamental mechanism underlying loss-aversion, and hence deserves further study. Narrative reviews may miss some opportunities to amplify such interesting studies too, but meta-analysis will always neglect them.

We will conclude with one idea about how meta-analysis can guide future research. Meta-regression can actually pinpoint where studies are plentiful and where an additional study would have the greatest new effect on collective knowledge. A low standard error on a meta-regression coefficient means we do not need to learn more. A high standard error means that we do need to learn more. From our dataset, studies (i) other than lab and field experiments, (ii) focusing on specific, non-University student populations, (iii) on continents of South America, Africa and Oceania, (iv) involving rewards not expressed in monetary terms, (v) obtaining preferences in methods other than sequential binary methods, (vi) using utility other than CRRA with equal curvature and (vii) with loss-aversion specifications other than Kahneman and Tversky are the areas where we have uncovered the least data. Interesting new findings about the loss aversion

[14] Another useful analogy comes from performance evaluation in personnel economics. It is well-known Baker et al. (1994) that objective and subjective measure of performance can be complements. Objective measures rein in over-the-top subjective evaluation, and subjective evaluations can add in idiosyncratic information that objective measures are blind to.

parameter are *ex ante* more likely to be in those areas. We encourage future research there.

References

Abdellaoui, M., H. Bleichrodt, and O. l'Haridon (2008): "A Tractable Method to Measure Utility and Loss Aversion under Prospect Theory," *Journal of Risk and Uncertainty*, 36, 245–266.

Abeler, J., D. Nosenzo, and C. Raymond (2019): "Preferences for Truth-Telling," *Econometrica*, 87, 1115–1153.

Alesina, A. and F. Passarelli (2019): "Loss Aversion in Politics," *American Journal of Political Science*, 63, 936–947.

Baker, G., R. Gibbons, and K. J. Murphy (1994): "Subjective Performance Measures in Optimal Incentive Contracts," *The Quarterly Journal of Economics*, 109, 1125–1156.

Barberis, N. and M. Huang (2001): "Mental Accounting, Loss Aversion, and Individual Stock Returns," *The Journal of Finance*, 56, 1247–1292.

——— (2008): "Stocks as Lotteries: The Implications of Probability Weighting for Security Prices," *American Economic Review*, 98, 2066–2100.

Barberis, N., M. Huang, and T. Santos (2001): "Prospect Theory and Asset Prices," *The Quarterly Journal of Economics*, 116, 1–53.

Barberis, N., L. Jin, and B. Wang (2020): "Prospect Theory and Stock Market Anomalies," NBER Working Paper No. 27155.

Barberis, N., A. Mukherjee, and B. Wang (2016): "Prospect Theory and Stock Returns: An Empirical Test," *The Review of Financial Studies*, 29, 3068–3107.

Barberis, N. and W. Xiong (2009): "What Drives the Disposition Effect? An Analysis of a Long-Standing Preference-Based Explanation," *The Journal of Finance*, 64, 751–784.

Barberis, N. C. (2013): "Thirty Years of Prospect Theory in Economics: A Review and Assessment," *Journal of Economic Perspectives*, 27, 173–96.

Benartzi, S. and R. H. Thaler (1995): "Myopic Loss Aversion and the Equity Premium Puzzle," *The Quarterly Journal of Economics*, 110, 73–92.

Borenstein, M., L. V. Hedges, J. P. Higgins, and H. R. Rothstein (2009): *Introduction to Meta-Analysis*, John Wiley & Sons.

Camerer, C., L. Babcock, G. Loewenstein, and R. Thaler (1997): "Labor Supply of New York City Cabdrivers: One Day at a Time," *The Quarterly Journal of Economics*, 112, 407–441.

Carpenter, B., A. Gelman, M. D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. Brubaker, J. Guo, P. Li, and A. Riddell (2017): "Stan: A Probabilistic Programming Language," *Journal of Statistical Software*, 76.

De Martino, B., C. F. Camerer, and R. Adolphs (2010): "Amygdala Damage Eliminates Monetary Loss Aversion," *Proceedings of the National Academy of Sciences*, 107, 3788–3792.

DellaVigna, S. (2018): "Structural Behavioral Economics," in *Handbook of Behavioral Economics - Foundations and Applications*, ed. by B. D. Bernheim, S. DellaVigna, and D. Laibson, Amsterdam: Elsevier/North-Holland, vol. 1, 613–725.

Doucouliagos, C., T. D. Stanley, and M. Giles (2012): "Are Estimates of the Value of a Statistical Life Exaggerated?" *Journal of Health Economics*, 31, 197–206.

Doucouliagos, H. and T. D. Stanley (2009): "Publication Selection Bias in Minimum-Wage Research? A Meta-Regression Analysis," *British Journal of Industrial Relations*, 47, 406–428.

Doucouliagos, H., T. D. Stanley, and W. K. Viscusi (2014): "Publication Selection and the Income Elasticity of the Value of a Statistical Life," *Journal of Health Economics*, 33, 67–75.

Engel, C. (2011): "Dictator Games: A Meta Study," *Experimental Economics*, 14, 583–610.

Ericson, K. M. M. and A. Fuster (2014): "The Endowment Effect," *Annual Review of Economics*, 6, 555–579.

Fishburn, P. C. and G. A. Kochenberger (1979): "Two-Piece von Neumann-Morgenstern Utility Functions," *Decision Sciences*, 10, 503–518.

Fryer, R. G., S. D. Levitt, J. List, and S. Sadoff (2012): "Enhancing the Efficacy of Teacher Incentives through Loss Aversion: A Field Experiment," NBER Working Paper No. 18237.

Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin (2014): *Bayesian Data Analysis*, Boca Raton: CRC Press.

Glass, G. V. (1976): "Primary, Secondary, and Meta-Analysis of Research," *Educational Researcher*, 5, 3–8.

Gurevitch, J., J. Koricheva, S. Nakagawa, and G. Stewart (2018): "Meta-Analysis and the Science of Research Synthesis," *Nature*, 555, 175–182.

Havránek, T. (2015): "Measuring Intertemporal Substitution: The Importance of Method Choices and Selective Reporting," *Journal of the European Economic Association*, 13, 1180–1204.

Havránek, T., R. Horváth, Z. Iršová, and M. Rusnak (2015): "Cross-Country Heterogeneity in Intertemporal Substitution," *Journal of International Economics*, 96, 100–118.

Havránek, T., M. Rusnak, and A. Sokolova (2017): "Habit Formation in Consumption: A Meta-Analysis," *European Economic Review*, 95, 142–167.

Heidhues, P. and B. Kőszegi (2018): "Behavioral Industrial Organization," in *Handbook of Behavioral Economics - Foundations and Applications*, ed. by B. D. Bernheim, S. DellaVigna, and D. Laibson, Amsterdam: Elsevier/North-Holland, vol. 1, 517–612.

Imai, T., T. A. Rutter, and C. F. Camerer (forthcoming): "Meta-Analysis of Present-Bias Estimation Using Convex Time Budgets," *The Economic Journal*.

Iršová, Z. and T. Havránek (2013): "Determinants of Horizontal Spillovers from FDI: Evidence from a Large Meta-Analysis," *World Development*, 42, 1–15.

Kahneman, D. (2011): *Thinking, Fast and Slow*, New York: Farrar, Straus and Giroux.

Kahneman, D. and A. Tversky (1979): "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, 47, 263–292.

Kim, E. H., A. Morse, and L. Zingales (2006): "What Has Mattered to Economics Since 1970," *Journal of Economic Perspectives*, 20, 189–202.

Köbberling, V. and P. P. Wakker (2005): "An Index of Loss Aversion," *Journal of Economic Theory*, 122, 119–131.

Kruschke, J. K. (2010): "Bayesian Data Analysis," *Wiley Interdisciplinary Reviews: Cognitive Science*, 1, 658–676.

l'Haridon, O. and F. M. Vieider (2019): "All Over the Map: A Worldwide Comparison of Risk Preferences," *Quantitative Economics*, 10, 185–215.

Matoušek, J., T. Havránek, and Z. Iršová (2020): "Individual Discount Rates: A Meta-Analysis of Experimental Evidence," Unpublished manuscript.

McElreath, R. (2016): *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*, CRC Press.

Neilson, W. S. (2002): "Comparative Risk Sensitivity with Reference-Dependent Preferences," *Journal of Risk and Uncertainty*, 24, 131–142.

Neumann, N. and U. Böckenholt (2014): "A Meta-Analysis of Loss Aversion in Product Choice," *Journal of Retailing*, 90, 182–197.

Odean, T. (1998): "Are Investors Reluctant to Realize their Losses?" *The Journal of Finance*, 53, 1775–1798.

O'Donoghue, T. and C. Sprenger (2018): "Reference-Dependent Preferences," in *Handbook of Behavioral Economics - Foundations and Applications*, ed. by B. D. Bernheim, S. DellaVigna, and D. Laibson, Amsterdam: Elsevier/North-Holland, vol. 1.

Oosterbeek, H., R. Sloof, and G. Van De Kuilen (2004): "Cultural Differences in Ultimatum Game Experiments: Evidence from a Meta-Analysis," *Experimental Economics*, 7, 171–188.

Pearson, K. (1904): "Report on Certain Enteric Fever Inoculation Statistics," *BMJ*, 2, 1243–1246.

Rieger, M. O., M. Wang, and T. Hens (2017): "Estimating Cumulative Prospect Theory Parameters from an International Survey," *Theory and Decision*, 82, 567–596.

Romer, P. M. (1996): "Preferences, Promises, and the Politics of Entitlement," in *Individual and Social Responsibility: Child Care, Education, Medical Care, and Long-Term Care in America*, University of Chicago Press, 195–228.

Rosenthal, R. (1979): "The File Drawer Problem and Tolerance for Null Results," *Psychological Bulletin*, 86, 638–641.

Schmidt, U. and H. Zank (2005): "What Is Loss Aversion?" *Journal of Risk and Uncertainty*, 30, 157–167.

Stan Development Team (2020): "RStan: the R interface to Stan," R package version 2.21.2, http://mc-stan.org/.

Stanley, T. D. (2001): "Wheat from Chaff: Meta-analysis as Quantitative Literature Review," *Journal of Economic Perspectives*, 15, 131–150.

Stanley, T. D. and H. Doucouliagos (2012): *Meta-Regression Analysis in Economics and Business*, London: Routledge.

Tovar, P. (2009): "The Effects of Loss Aversion on Trade Policy: Theory and Evidence," *Journal of International Economics*, 78, 154–167.

Tunçel, T. and J. K. Hammitt (2014): "A New Meta-Analysis on the WTP/WTA Disparity," *Journal of Environmental Economics and Management*, 68, 175–187.

Tversky, A. and D. Kahneman (1991): "Loss Aversion in Riskless Choice: A Reference-Dependent Model," *The Quarterly Journal of Economics*, 106, 1039–1061.

——— (1992): "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 5, 297–323.

Wakker, P. and D. Deneffe (1996): "Eliciting von Neumann-Morgenstern utilities when probabilities are distorted or unknown," *Management Science*, 42, 1131–1150.

Wakker, P. and A. Tversky (1993): "An Axiomatization of Cumulative Prospect Theory," *Journal of Risk and Uncertainty*, 7, 147–175.

Wakker, P. P. (2010): *Prospect Theory: For Risk and Ambiguity*, Cambridge: Cambridge University Press.

Walasek, L., T. L. Mullett, and N. Stewart (2018): "A Meta-Analysis of Loss Aversion in Risky Contexts," Unpublished manuscript.

Wang, M., M. O. Rieger, and T. Hens (2017): "The Impact of Culture on Loss Aversion," *Journal* *of Behavioral Decision Making*, 30, 270–281.

Weber, M. and C. F. Camerer (1998): "The Disposition Effect in Securities Trading: An Experimental Analysis," *Journal of Economic Behavior & Organization*, 33, 167–184.

Yates, F. and W. G. Cochran (1938): "The Analysis of Groups of Experiments," *The Journal of Agricultural Science*, 28, 556–580.

Supplementary Online Material

Meta-Analysis of Empirical Estimates of Loss-Aversion

Alexander L. Brown   Taisuke Imai   Ferdinand M. Vieider
Colin F. Camerer

Contents

A Data

A.1 Paper Search and Inclusion

We searched for relevant papers on the scientific citation indexing database Web of Science. We used, after several trial-and-error to fine-tune, the following combination of query terms.

```
(   (loss AND avers*)
 OR "loss aversion coefficient"
 OR "loss aversion index"
 OR ("loss avers*" AND ("willingness to pay" OR "willingness to accept")) )
                              AND
       (estimat* OR measur* OR experiment* OR survey)
```

Figure A.1: Keywords used in the search.

The initial search, made in the summer of 2017, returned a total hits of 1,547 papers. As a first step of paper identification, we went through titles and abstracts and threw out 910 papers that were clearly irrelevant for our study. We then read the remaining papers, applied our inclusion criteria based on the content, and then coded information. We also used IDEAS/RePEc and Google Scholar to search for unpublished working papers. Finally, we posted a message on the email list of the Economic Science Association to ask for relevant papers (in February 2018).

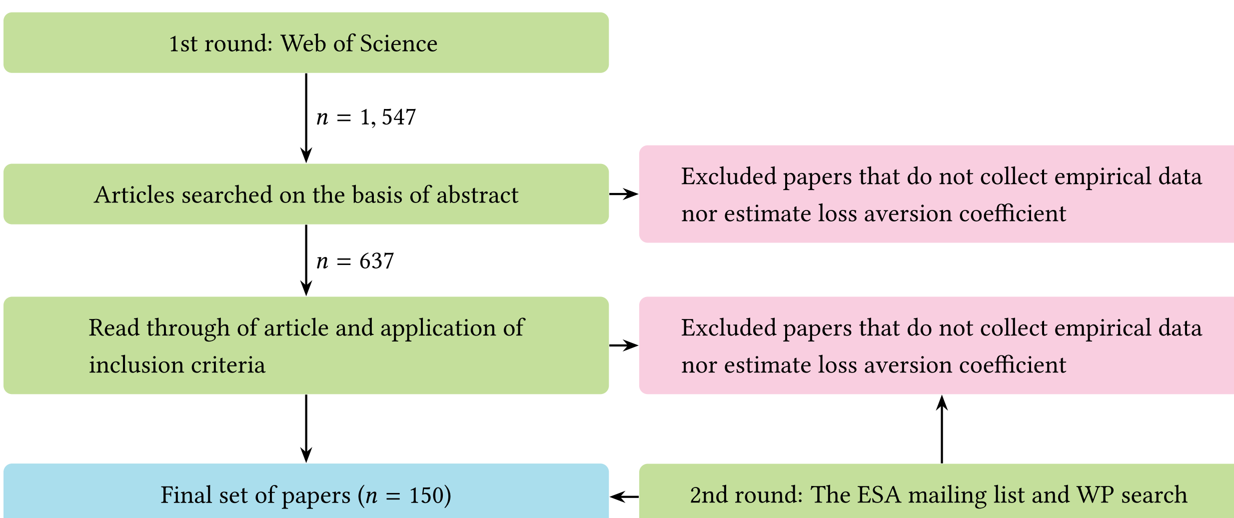

Figure A.2: Paper search and data construction.

A.2 Approximation and Imputation of Missing Standard Errors

The dataset includes 192 estimates (out of 607) of loss aversion coefficient without corresponding standard errors (SEs). In order to keep these observations in our meta-analysis, we approximated and imputed missing SEs using other available information.

First, we calculated SEs of four observations from p-values of the two-sided test for the null hypothesis $H_0 : \lambda = 1$, from

$$se = \frac{|\lambda - 1|}{\Phi^{-1}(1-p)},$$

where Φ^{-1} is the quantile function of the standard normal distribution.

Second, we approximated 64 SEs from inter-quartile range (IQR) and sample size, using

$$se \approx \frac{1.35 \times \mathrm{IQR}}{\sqrt{n}}.$$

Note that the use of this approximation formula is legitimate if the parameters are normally distributed in the population, which is a strong assumption in our dataset. Nevertheless, obtaining even an "approximated" SE seemed preferable to dropping the observation entirely, or to making other, even stronger, assumptions allowing us to keep the observation.

Finally, we imputed remaining 124 missing SEs. The basic idea is to estimate the parameters characterizing their distribution in the data, $\log(se_o) \sim \mathcal{N}(\mu_{se}, \sigma^2_{se})$. Using these distributional parameters, we can then estimate the missing values in SE by letting $\log(se_m) \sim \mathcal{N}(\widehat{\mu}_{se}, \widehat{\sigma}^2_{se})$, where the subscripts o and m stand for *observed* and *missing*, respectively, and $(\widehat{\mu}_{se}, \widehat{\sigma}_{se})$ are estimated quantities.

Implementing this estimation, we will thus obtain values for the missing observations in SE that have the same mean and variance. We can, however, do much better than that if we can find other variables in our dataset that are significantly associated with SEs (McElreath, 2016). As it turns out, the single best predictor of the SE is the loss aversion estimate itself. Once it is controlled for, no other predictor—including the measurement type and the square root of the number of observations—is significant. The loss aversion coefficient explains 51% of the variance in SEs. By letting $\mu_{se} = \alpha_{se} + \beta_{se}\lambda$, we can thus get much better imputation results than by only using the distributional characteristics.

Figure A.4 shows the imputed standard errors juxtaposed with the observed standard errors, and plotted against the loss aversion coefficient. The solid line indicates the regression line of the SE on loss aversion in the subset of data for which we observe the SE. The estimates of loss aversion with and without SEs exhibit systematic difference ($p = 0.002$, Wilcoxon rank sum test; Figure A.3 and Figure A.4B) but, as we would expect, the imputed SEs are no different than the observed SEs on average ($p = 0.458$, Wilcoxon rank sum test; Figure A.4C).

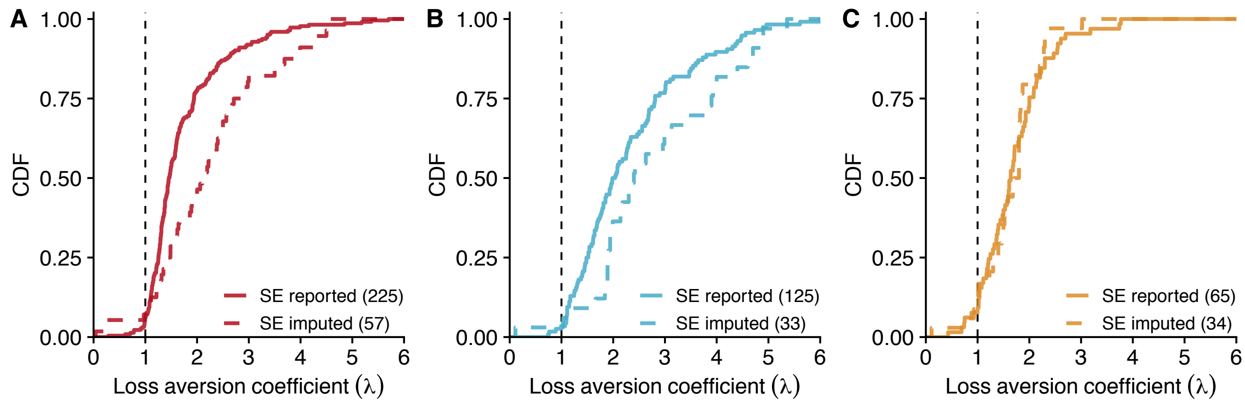

Figure A.3: Empirical CDF of reported loss aversion coefficient λ by the type of estimates and by the type of SE (solid line: imputed; dashed line: reported).

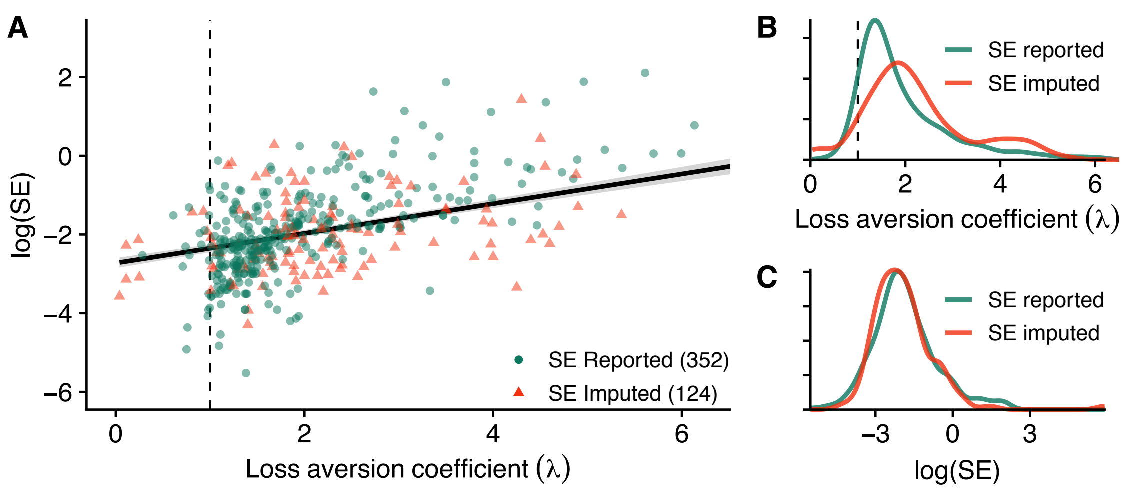

Figure A.4: Imputation of standard errors. *Notes*: The solid line is the regression line of the standard errors on loss aversion in the data with observed standard errors. The x-axis is cut off at 6 for better visualization.

A.3 Coded Variables

Table A.1: List of coded variables.

Variable	Description
Atricle meta data	
main_lastnames	Last names of the authors
main_firstnames	First names of the authors
main_title	Title of the paper
main_published	= 1 if published
main_yearpub	Year of publication
main_journal	Journal
main_length	Number of pages (excluding online appendices)
main_affliations	Affiliations of the authors
main_fund	Funding sources
Type of data	
type_lab_exp	= 1 if laboratory experiment
type_field_exp	= 1 if field experiment
type_class_exp	= 1 if classroom experiment
type_online_exp	= 1 if online experiment
type_gameshow	= 1 if TV game show
type_field_other	= 1 if other field data
Location of the experiment	
loc_lab	= 1 if laboratory study
loc_field	= 1 if field study
loc_online	= 1 if online study
loc_class	= 1 if classroom study
loc_country	Country
loc_city	City
loc_state	State

Variable	Description
Subject pool	
subject_children	= 1 if subjects are children
subject_uni	= 1 if subjects are university students/staffs
subject_elderly	= 1 if elderly population
subject_gen	= 1 if general population
subject_farm	= 1 if subjects are farmers
subject_mixed	= 1 if mixed subject population
subject_monkey	= 1 if subjects are monkeys
subject_unknown	= 1 if unknown population
Reward	
reward_real	= 1 if real reward
reward_deception	= 1 if deception is used
reward_money	= 1 if monetary reward
reward_food	= 1 if food reward
reward_cons_good	= 1 if consumption goods
reward_env_good	= 1 if environmental goods
reward_health	= 1 if health
reward_mixed	= 1 if mixed type
reward_other	= 1 if other type of reward
Method	
method_question	= 1 if questionnaire
method_seqbin	= 1 if sequential binary choice
method_mpl	= 1 if multiple price list
method_bdm	= 1 if BDM
method_matching	= 1 if matching
method_gp	= 1 if Gneezy-Potters
method_multi	= 1 if combination of multiple methods
method_other	= 1 if other method
method_other_type	Description of the method (if method_other = 1)

Variable	Description
Utility specifications	
spec_u_est	= 1 if utility function is parametrically estimated
spec_nonparametric	= 1 if utility function is nonparametrically recovered
spec_u_crra	= 1 if CRRA is assumed
spec_u_cara	= 1 if CARA is assumed
spec_u_linear	= 1 if linear utility is assumed
spec_u_other	= 1 if other parametric form is assumed
Loss aversion	
loss_kahneman_tversky	= 1 if λ is defined à la Kahneman and Tversky (1979)
loss_neilson	= 1 if λ is defined à la Neilson (2002)
loss_wakker_tversky	= 1 if λ is defined à la Wakker and Tversky (1993)
loss_bowman	= 1 if λ is defined à la Bowman et al. (1999)
loss_kobberling_wakker	= 1 if λ is defined à la Köbberling and Wakker (2005)
loss_koszegi_rabin	= 1 if λ is defined à la Kőszegi and Rabin (2006)
Estimates	
la	Loss aversion coefficient λ
se	SE of λ (reported/calculated)
se_imp	SE of λ (reported/calculated and imputed)
se_type	Type of SE (reported/calculated or imputed)
sd	Standard deviation of λ
iqr	Inter-quartile range of λ
pv	p-value of the null hypothesis $H_0 : \lambda = 1$ (if reported)
N	Sample size
la_type	Type of reported λ

Variable	Description
Aggregate-level estimation results	
ares_present	= 1 if aggregate-level estimates is reported
ares_N	Number of observations
ares_loss	Estimated loss aversion coefficient λ
ares_loss_error	Standard error of estimated λ
ares_ucurv_loss	Estimated u^-
ares_ucurv_loss_error	Standard error of estimated u^-
ares_ucurv_gain	Estimated u^+
ares_ucurv_gain_error	Standard error of estimated u^+
ares_ucurv_eq	= 1 if $u^+ = u^-$ is assumed
Individual-level estimation results	
ires_present	= 1 if individual-level estimates is reported
ires_N	Number of subjects
ires_loss_mean	Mean of the distribution of λ
ires_loss_median	Median of the distribution of λ
ires_loss_var	Variance of the distribution of λ
ires_loss_sd	Standard deviation of the distribution of λ
ires_loss_error	Standard error of mean λ
ires_loss_25_percentile	Q1 of the distribution of λ
ires_loss_75_percentile	Q3 of the distribution of λ
ires_loss_iqr	IQR of the distribution of λ
ires_ucurv_loss_mean	Mean of the distribution of u^-
ires_ucurv_loss_median	Median of the distribution of u^-
ires_ucurv_loss_var	Variance of the distribution of u^-
ires_ucurv_loss_error	Standard error of the distribution of u^-
ires_ucurv_gain_mean	Mean of the distribution of u^+
ires_ucurv_gain_median	Median of the distribution of u^+
ires_ucurv_gain_var	Variance of the distribution of u^+
ires_ucurv_gain_error	Standard error of the distribution of u^+
ires_ucurv_eq	= 1 if $u^+ = u^-$ is assumed

A.4 Journals

Table A.2: List of journals and disciplines.

	Journal	Category
1	Addiction	Substance Abuse
2	Addictive Behaviors and Decision Making	Psychology, Applied
3	American Economic Journal: Economic Policy	Economics
4	American Economic Journal: Microeconomics	Economics
5	American Economic Review	Economics
6	American Journal of Agricultural Economics	Agriculture/Agronomy
7	Behavioral Neuroscience	Neurosciences
8	Brain	Neurosciences
9	Cognition & Emotion	Psychology
10	Consciousness and Cognition	Psychology, Experimental
11	Current Biology	Cell Biology
12	Developmental Cognitive Neuroscience	Psychology, Development
13	Ecological Economics	Ecology
14	Economic Inquiry	Economics
15	Economics Letters	Economics & Business
16	Ekonomický časopis	Economics
17	Emotion	Psychology, Experimental
18	Environment and Development Economics	Economics
19	European Economic Review	Economics
20	European Journal of Operational Research	Operations Research & Management Science
21	European Journal of Transport and Infrastructure Research	Social Sciences, General
22	European Review of Agricultural Economics	Economics & Business
23	Experimental Economics	Economics
24	Frontiers in Human Neuroscience	Psychology
25	Frontiers in Psychology	Psychology, Multidisciplinary
26	Games and Economic Behavior	Economics
27	International Economic Review	Economics
28	International Journal of Applied Behavioral Economics	Economics & Business
29	International Journal of Research in Marketing	Economics & Business
30	Journal of African Economics	Agricultural Sciences
31	Journal of Banking & Finance	Business, Finance
32	Journal of Behavioral and Experimental Economics	Economics
33	Journal of Behavioral Decision Making	Psychology, Applied
34	Journal of Behavioral Finance	Business, Finance
35	Journal of Business and Economic Statistics	Business & Economics
36	Journal of Consumer Research	Economics
37	Journal of Development Economics	Economics
38	Journal of Developmental Studies	Social Sciences, General
39	Journal of Economic Behavior & Organization	Economics
40	Journal of Economic Dynamics and Control	Economics

	Journal	Category
41	Journal of Economic Psychology	Economics
42	Journal of Empirical Finance	Economics
43	Journal of Experimental Psychology: General	Psychology
44	Journal of Gambling Studies	Substance Abuse
45	Journal of Health Economics	Economics & Business
46	Journal of International Economics	Economics
47	Journal of Marketing Research	Economics
48	Journal of Mathematical Psychology	Psychology, Mathematical
49	Journal of Political Economy	Economics
50	Journal of Risk and Uncertainty	Business & Economics
51	Judgment and Decision Making	Psychiatry/Psychology
52	Management Science	Management
53	Marketing Science	Economics
54	Nature	Multidisciplinary Sciences
55	NeuroImage	Neurosciences
56	Neuron	Neurosciences
57	Neuropsychiatric Disease and Treatment	Psychiatry
58	Organizational Behavior and Human Decision Processes	Management
59	PLoS Computational Biology	Biochemical Research Methods
60	PloS One	Multidisciplinary Sciences
61	PNAS	Multidisciplinary Sciences
62	Proceedings of the Royal Society B: Biological Sciences	Evolutionary Biology
63	PsicolÃ³gica	
64	Psychiatry Research	Psychiatry/Psychology
65	Psychological Science	Psychology
66	Psychology and Aging	Gerontology
67	Quantitative Finance	Economics
68	Quarterly Journal of Economics	Economics
69	Rationality and Society	Social Sciences, General
70	Review of Managerial Science	Management
71	Revista Espanola de Financiacion y Contabilidad	Business, Finance
72	Science	Multidisciplinary Sciences
73	The Review of Economics and Statistics	Economics
74	Theory and Decision	Economics
75	Tourism Management	Hospitality, Leisure, Sport & Tourism
76	Transportation Research Part B	Transportation Science & Technology
77	Transportation Research Record	Transportation Science & Technology
78	World Development	Economics

Notes: Journal categories are based on classification provided by The Master Journal List (https://mjl.clarivate.com/home).

Table A.3: Disciplines.

	#	%
Economics	62	47.7
Business/Management	21	16.2
Psychology	16	12.3
Multi-disciplinary	10	7.7
Psychiatry/Medicine	6	4.6
Neuroscience	4	3.1
Agriculture	2	1.5
Transportation/Tourism	3	2.3
Other	6	4.6
Total	130	100.0

B Bayesian Hierarchical Model

B.1 Modeling Framework

The main goal of our meta-analysis is first to obtain the "best available" estimate of the loss aversion coefficient λ combining the available information in the literature and then to understand the heterogeneity of reported estimates across studies. To this end, we analyze the data using a *Bayesian hierarchical modeling* approach.

Meta-analysis is naturally hierarchical. The effect sizes reported in different studies are summary measures of individual-level behavior. We summarize these measures by estimating their mean and variation based on a given model. Additional hierarchical levels can be introduced e.g., to deal with statistical dependence in estimates, such as when one and the same paper or study reports multiple estimates.

Hierarchical models, in turn, are naturally Bayesian (Gelman and Hill, 2006; McElreath, 2016). To see this, one can picture the estimated aggregate mean as an endogenous prior, that will then influence the estimates of the "true" study-level effect, depending on the uncertainty surrounding the estimate itself—a statistical procedure known as "shrinkage" or "pooling". One of the great advantages of the Bayesian approach is further that the estimate emerging from the meta-analysis—the posterior mean of our analysis—can serve as a prior for future empirical studies, and is easy to update with additional evidence. This is conductive to the systematic quantitative accumulation of knowledge—the prime objective of meta-analysis.

Consider the dataset $(\lambda_i, se_i^2)_{i=1}^m$, where λ_i is the ith *measurement* (or *observation*) of the loss aversion coefficient in the dataset and se_i is the associated standard error that captures the uncertainty surrounding the estimate. We assume that the ith reported estimate λ_i is normally distributed around the parameter $\overline{\lambda}_i$:

$$\lambda_i \mid \overline{\lambda}_i, se_i \sim \mathcal{N}(\overline{\lambda}_i, se_i^2), \tag{B.1}$$

where the variability is due to the sampling variation captured by the known standard error se_i.[1]

Sampling variation is part of the observed variation in the reported estimates $(\lambda_i)_{i=1}^m$, but it may not be all, since there is a possibility of "genuine" heterogeneity across measurements (due to different settings, for example). We model this by assuming that each $\overline{\lambda}_i$ is normally distributed, adding another level to the hierarchy:

$$\overline{\lambda}_i \mid \lambda_0, \tau \sim \mathcal{N}(\lambda_0, \tau^2), \tag{B.2}$$

[1]The parameter $\overline{\lambda}_i$ is often referred to as the "true effect size" in the random-effects meta-analysis.

where λ_0 is the *overall mean* of the estimated loss aversion parameters $\overline{\lambda}_i$, and τ is its standard deviation, capturing the variation between observations in the data. The overall variance in the data, therefore, consists of two parts, the between-observation variance, τ^2, and the individual sampling variation coming from measurement uncertainty, se^2. This can be clearly seen by combining expressions (B.1) and (B.2) into one:

$$\lambda_i \mid \lambda_0, \tau, se_i \sim \mathcal{N}(\lambda_0, \tau^2 + se_i^2).$$

Note that this formulation is mathematically equivalent to the classical formulation of random-effects meta-analysis (DerSimonian and Laird, 1986), which is typically expressed as

$$\lambda_i = \overline{\lambda}_i + \xi_i = \lambda_0 + \varepsilon_i + \xi_i,$$

where $\xi_i \sim \mathcal{N}(0, se_i^2)$ is a sampling error of λ_i as an estimate of $\overline{\lambda}_i$, and each observation-specific "true" effect $\overline{\lambda}_i$ is decomposed into λ_0 (the overall mean) and the sampling error ξ_j. It is further assumed that $\varepsilon_i \sim \mathcal{N}(0, \tau^2)$, where τ^2 is the between-observation heterogeneity, beyond the mere sampling variance. When $\tau = 0$, this model reduces to a fixed-effect meta-analysis. This highlights the central assumption underlying fixed-effect meta-analysis—that different estimates differ only based on random sampling variation—which clearly does not seem adequate for the diverse set of estimates included in our meta-analysis. We thus conduct a random-effects analysis, allowing for both random sampling variation and systematic differences between studies and estimates.

In this model, each observation λ_i in the data will be "pooled" towards the overall mean λ_0, with the extent of the pooling depending on two factors: (i) the standard error associated with the estimate; and (ii) how far the estimate lies from the estimated mean, λ_0. As we see above, the variance across observations is decomposed into two parts—variance due to error in estimation, and the genuine between-observation heterogeneity. The pooling equation for a specific observation i takes the following form

$$\overline{\lambda}_i = (1 - \omega_i)\lambda_i + \omega_i\lambda_0, \tag{B.3}$$

where ω_i is the "pooling factor" (Gelman and Pardoe, 2006), defined as

$$\omega_i = \frac{se_i^2}{\tau^2 + se_i^2}. \tag{B.4}$$

The equation makes it clear that, the larger the SE *ceteris paribus*, the larger the pooling factor, and thus the closer the estimate will be drawn to the overall mean estimate of the population, indicated by λ_0. At the same time, the smaller the between study variation captured by τ^2, the more pooling towards the population mean. This makes intuitive sense—estimates are pooled more to the extent that all estimates in the population are similar to each other, and to the extent that they are characterized by a low degree of precision.

It is now straightforward to account for variation across estimates driven by observable characteristics—commonly referred to as meta-regression—by letting

$$\overline{\lambda}_i = \kappa_i + X_i\beta + \varepsilon_i, \tag{B.5}$$

where κ_i is the intercept of the regression, X_i a matrix of observable study characteristics for observation i, and β is a vector of regression coefficients. Notice that the residual is distributed as $\varepsilon_i \sim \mathcal{N}(0, \tau^2)$. By comparing the variance in this model to the variance estimated in a model empty of covariates, i.e. where X_i contains no entries, we will be able to assess what extent of the overall variance between observations is explained by the observation-level characteristics encoded in X_i. In particular, the variance explained is given by $1 - (\tau_1^2/\tau_0^2)$, where τ_0^2 is the estimated variance between observation in a model empty of covariates, and τ_1^2 is the equivalent variance term estimated in the meta-regression model.

While this normal-normal structure expressed in equations (B.1, B.2) is the benchmark setup we use, it will quickly become interesting to relax the modelling assumptions described here, e.g., by replacing the normal distribution with a robust student-t distribution or an asymmetric log-normal distribution, and by allowing for additional hierarchical levels to account for the lack of independence in the observations in our data.

We estimate the model in `Stan` (Carpenter et al., 2017) using the Hamiltonian Monte Carlo simulations, an algorithm for Markov Chain Monte Carlo, and launch it from `R` (`https://www.r-project.org/`) using `RStan` (Stan Development Team, 2020). Priors for the population-level parameters are chosen in such a way as to be mildly regularizing, i.e., they are informative, but typically encompass ranges that are one order of magnitude larger than the estimated values we expect based on the range of the data (McElreath, 2016). Priors for lower-level parameters are then constituted by the endogenously estimated population-level parameters. The estimates we report are not sensitive to the choice of the particular priors we use (Section B.3.3 below).

B.2 Estimation

In Section 4.3, we started from fitting the benchmark model expressed as equations (B.1) and (B.2):

$$\begin{aligned} \lambda_i \mid \overline{\lambda}_i, se_i &\sim \mathcal{N}(\overline{\lambda}_i, se_i^2), \\ \overline{\lambda}_i \mid \lambda_0, \tau &\sim \mathcal{N}(\lambda_0, \tau^2), \\ \lambda_0 &\sim \text{half } \mathcal{N}(1, 5), \\ \tau &\sim \text{half } \mathcal{N}(0, 5). \end{aligned} \tag{M1a}$$

(This model was called M1 in Section 4.3.) The estimated overall mean λ_0 is 1.810 with a 95% credible interval (CrI) of [1.742, 1.880].

Pooling. Equations (B.3) and (B.4) describe the mechanism underlying the pooling. The amount of pooling applied to an observation—i.e. the extent to which an estimated parameter $\overline{\lambda}_i$ is drawn towards the overall mean λ_0 from its observed value λ_i—will depend on the SE associated with the observation, and its distance from the mean. This is illustrated in Figure B.1, which shows a scatter plot of the estimated loss aversion parameter, $\overline{\lambda}_i$, against the observed parameter, λ_i. For standard errors up to 0.4, almost no pooling is observed even for values that fall relatively far from the mean. Pooling increases for larger SEs between 0.4 and 1, and becomes very strong for even larger SEs. The farther an observation falls from the mean, the more it is pooled, ceteris paribus. We further observe very strong pooling for large observations because the standard errors themselves tend to increase with loss aversion, as detailed above.

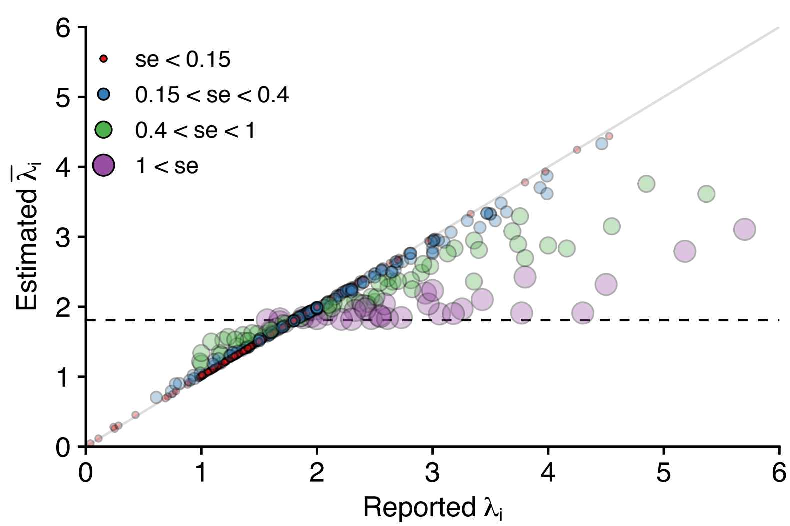

Figure B.1: Pooling of estimates by SE. *Notes*: The horizontal dotted line corresponds to the estimated overall mean $\lambda_0 = 1.810$. The axes are cut off at six for better visualization.

Additional models. In addition to models M1a and M2 discussed in Section 4.3, we estimated two additional "intermediate" models.

The first alternative model is a straightforward extension of model M1a, replacing the normal distribution with a log-normal distribution:

$$\begin{aligned} \lambda_i \mid \overline{\lambda}_i, se_i &\sim \mathcal{N}(\overline{\lambda}_i, se_i^2), \\ \overline{\lambda}_i \mid \lambda_0^\ell, \tau_\ell &\sim \log \mathcal{N}(\lambda_0^\ell, \tau_\ell^2), \\ \lambda_0^\ell &\sim \mathcal{N}(1, 5), \\ \tau_\ell &\sim \text{half } \mathcal{N}(0, 5). \end{aligned} \tag{M1b}$$

Note the super-/sub-scripts ℓ in the location and scale parameters $(\lambda_0^\ell, \tau_\ell^2)$ of the log-normal distribution. We can calculate the mean and the median of the distribution by $\exp(\lambda_0^\ell + \tau_\ell^2/2)$ and $\exp(\lambda_0^\ell)$, respectively. The mean of the log-normal distribution is given by $\lambda_0 \equiv \exp(\lambda_0^\ell + \tau_\ell^2/2)$.

The second alternative model tries to address the non-independence of reported estimates by explicitly modeling the nesting of observations in papers. To do this, we introduce paper-level estimates as an additional hierarchical level. Let λ_{pi} be the ith estimate reported in paper p. We formulate a model as follows:

$$\begin{aligned} \lambda_{pi} \mid \overline{\lambda}_{pi}, se_{pi} &\sim \mathcal{N}(\overline{\lambda}_{pi}, se_{pi}^2), \\ \overline{\lambda}_{pi} \mid \overline{\lambda}_p, \sigma_p &\sim \mathcal{N}(\overline{\lambda}_p, \sigma_p^2), \\ \overline{\lambda}_p \mid \lambda_0^\ell, \tau_\ell &\sim \log \mathcal{N}(\lambda_0^\ell, \tau_\ell^2), \\ \lambda_0^\ell &\sim \mathcal{N}(1, 5), \\ \tau_\ell &\sim \text{half } \mathcal{N}(0, 5), \\ \sigma_p &\sim \text{half } \mathcal{N}(0, 5). \end{aligned} \tag{M1c}$$

The system now explicitly models the nesting of the estimated observation-level parameters, $\overline{\lambda}_{pi}$, in paper-level estimates, $\overline{\lambda}_p$. The latter are then modeled as following a log-normal distribution, just as previously. Figure B.2 illustrates the idea behind this formulation.

Figure B.3 below summarizes all four models we estimated.

Estimating model M1b, we obtain a mean λ_0 of 1.826, with a 95% CrI of [1.741, 1.880]. Figure B.4 (top right panel) shows the posterior predictive distribution from the estimation of this model. The fit can be seen to be much better than that of the baseline normal-normal model shown above, and to fit the actual observations closely. We thus conclude that a log-normal distribution provides a good fit to the data. The mean loss aversion λ_0 is 2.052 (95% CrI [1.911, 2.209]), under model M1c. The fit to the data, however, appears to be a little off, allowing room for improvement (Figure B.4, bottom left panel). The posterior predictive distribution puts much larger likelihood on values between 1.8 and 3 while it puts smaller likelihood on values below 1.5.

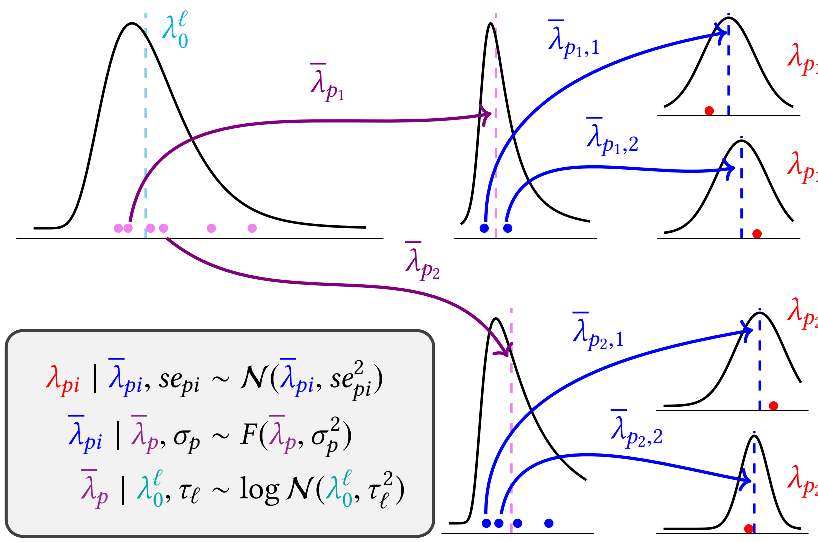

FIGURE B.2: Illustration of the nesting structure in models M1c and M2. For the paper-level distribution F in the middle layer, model M1c assumes a normal distribution and model M2 assumes a student-t distribution with additional parameter df.

Model M1a

$$\begin{aligned}
\lambda_i \mid \overline{\lambda}_i, se_i &\sim \mathcal{N}(\overline{\lambda}_i, se_i^2),\\
\overline{\lambda}_i \mid \lambda_0, \tau &\sim \mathcal{N}(\lambda_0, \tau^2),\\
\lambda_0 &\sim \text{half } \mathcal{N}(1, \nu),\\
\tau &\sim \text{half } \mathcal{N}(0, \nu).
\end{aligned}$$

Model M1b

$$\begin{aligned}
\lambda_i \mid \overline{\lambda}_i, se_i &\sim \mathcal{N}(\overline{\lambda}_i, se_i^2),\\
\overline{\lambda}_i \mid \lambda_0^\ell, \tau_\ell &\sim \log \mathcal{N}(\lambda_0^\ell, \tau_\ell^2),\\
\lambda_0^\ell &\sim \mathcal{N}(1, \nu),\\
\tau_\ell &\sim \text{half } \mathcal{N}(0, \nu).
\end{aligned}$$

Model M1c

$$\begin{aligned}
\lambda_{pi} \mid \overline{\lambda}_{pi}, se_{pi} &\sim \mathcal{N}(\overline{\lambda}_{pi}, se_{pi}^2),\\
\overline{\lambda}_{pi} \mid \overline{\lambda}_p, \sigma_p &\sim \mathcal{N}(\overline{\lambda}_p, \sigma_p^2),\\
\overline{\lambda}_p \mid \lambda_0^\ell, \tau_\ell &\sim \log \mathcal{N}(\lambda_0^\ell, \tau_\ell^2),\\
\lambda_0^\ell &\sim \mathcal{N}(1, \nu),\\
\tau_\ell &\sim \text{half } \mathcal{N}(0, \nu),\\
\sigma_p &\sim \text{half } \mathcal{N}(0, \nu).
\end{aligned}$$

Model M2

$$\begin{aligned}
\lambda_{pi} \mid \overline{\lambda}_{pi}, se_{pi} &\sim \mathcal{N}(\overline{\lambda}_{pi}, se_{pi}^2),\\
\overline{\lambda}_{pi} \mid df, \overline{\lambda}_p, \sigma_p &\sim t(df, \overline{\lambda}_p, \sigma_p^2),\\
\overline{\lambda}_p \mid \lambda_0^\ell, \tau_\ell &\sim \log \mathcal{N}(\lambda_0^\ell, \tau_\ell^2),\\
\lambda_0^\ell &\sim \mathcal{N}(1, \nu),\\
\tau_\ell &\sim \text{half } \mathcal{N}(0, \nu),\\
df &\sim \text{half } \mathcal{N}(0, \nu),\\
\sigma_p &\sim \text{half } \mathcal{N}(0, \nu).
\end{aligned}$$

FIGURE B.3: Summary of models.

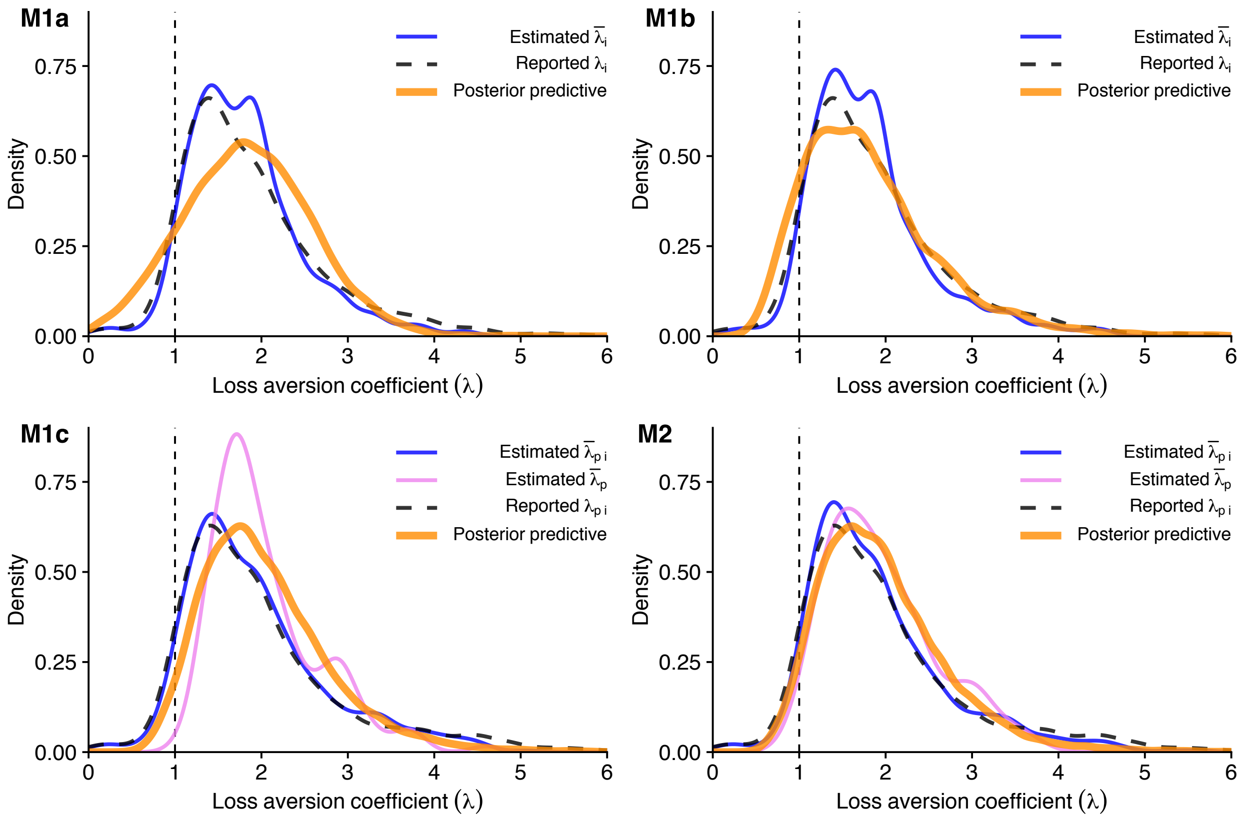

FIGURE B.4: Distributions of reported and estimated λ, and posterior predictive distribution of λ.

TABLE B.1: Summary of estimation results.

	Distributional assumption			Posterior of λ_0				Posterior of τ			
Model	Obs. level	Paper level	Pop. level	Mean	SD	2.5%	97.5%	Mean	SD	2.5%	97.5%
M1a	Normal		Normal	1.810	0.036	1.741	1.880	0.747	0.027	0.695	0.801
M1b	Normal		Log-normal	1.826	0.039	1.754	1.904	0.816	0.260	0.740	0.899
M1c	Normal	Normal	Log-normal	2.052	0.076	1.911	2.209	0.752	0.358	0.602	0.929
M2	Normal	Student-t	Log-normal	1.955	0.072	1.824	2.104	0.743	0.342	0.605	0.907

Notes: In Models M1c and M2, (λ_0, τ) are calculated from the log-normal parameters $(\lambda_0^\ell, \tau_\ell)$ by $\lambda_0 = \exp(\lambda_0^\ell + \tau_\ell^2/2)$ and $\tau^2 = [\exp(\tau_\ell^2) - 1]\exp(2\lambda_0^\ell + \tau_\ell^2)$.

B.3 Robustness Checks

B.3.1 Estimation Using Subsets of the Dataset

TABLE B.2: Estimation result for each type of reported λ.

		Distributional assumption			Posterior of λ_0			
Model	Type	Obs. level	Paper level	Pop. level	Mean	SD	2.5%	97.5%
M1a	Aggregate	Normal		Normal	1.704	0.044	1.619	1.792
	Individual-mean	Normal		Normal	2.439	0.103	2.242	2.642
	Individual-median	Normal		Normal	1.704	0.046	1.612	1.796
M2	Aggregate	Normal	Student-t	Log-normal	1.846	0.111	1.650	2.083
	Individual-mean	Normal	Student-t	Log-normal	2.412	0.148	2.148	2.731
	Individual-median	Normal	Student-t	Log-normal	1.711	0.085	1.557	1.890

Notes: In Model M2, λ_0 is calculated from the log-normal location parameter λ_0^ℓ by $\lambda_0 = \exp(\lambda_0^\ell + \tau_\ell^2/2)$.

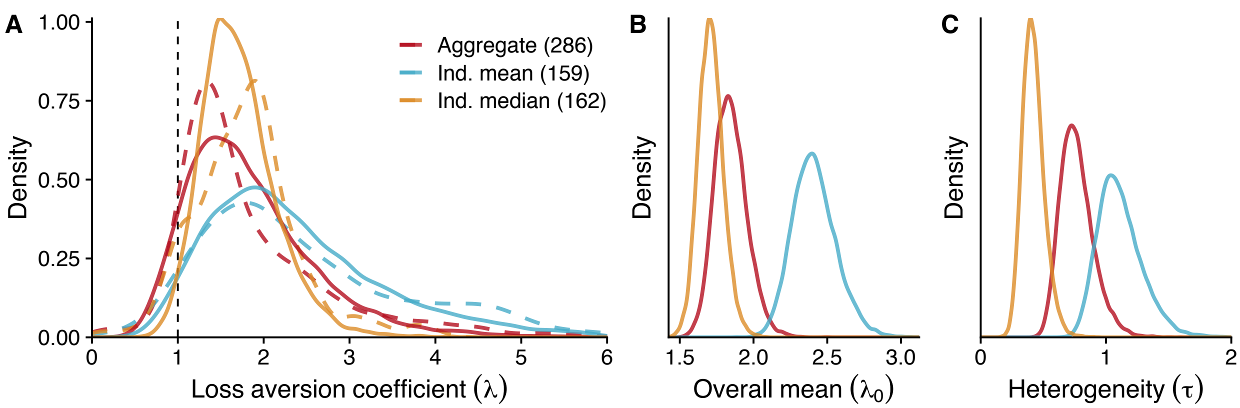

FIGURE B.5: Estimation of model M2 for each type of reported loss aversion coefficient separately. (A Distributions of reported λ_{pi} (dashed lines) and posterior predictive distributions (solid lines). (BC) Posterior distributions of λ_0 and τ. *Notes*: (λ_0, τ) are calculated from the log-normal parameters $(\lambda_0^\ell, \tau_\ell)$ are calculated by $\lambda_0 = \exp(\lambda_0^\ell + \tau_\ell^2/2)$ and $\tau^2 = [\exp(\tau_\ell^2) - 1]\exp(2\lambda_0^\ell + \tau_\ell^2)$.

B.3.2 Estimation Using the "Complete" Dataset

TABLE B.3: Sensitivity to SE imputation. (A) All data, including observations with imputed SEs (identical to Table 5). (C) Complete data, including only observations where associated SEs are available.

		Distributional assumption			Posterior of λ_0				Posterior of τ			
Model		Obs. level	Paper level	Pop. level	Mean	SD	2.5%	97.5%	Mean	SD	2.5%	97.5%
M1a	A	Normal		Normal	1.810	0.036	1.741	1.880	0.747	0.027	0.695	0.801
	C	Normal		Normal	1.713	0.041	1.632	1.793	0.714	0.033	0.653	0.780
M1b	A	Normal		Log-normal	1.826	0.039	1.754	1.904	0.816	0.260	0.740	0.899
	C	Normal		Log-normal	1.709	0.039	1.636	1.789	0.670	0.229	0.598	0.750
M1c	A	Normal	Normal	Log-normal	2.052	0.076	1.911	2.209	0.752	0.358	0.602	0.929
	C	Normal	Normal	Log-normal	2.041	0.097	1.866	2.242	0.877	0.449	0.684	1.122
M2	A	Normal	Student-t	Log-normal	1.955	0.072	1.824	2.104	0.743	0.342	0.605	0.907
	C	Normal	Student-t	Log-normal	1.959	0.092	1.791	2.152	0.822	0.418	0.642	1.046

Notes: In Models M1c and M2, (λ_0, τ) are calculated from log-normal parameters $(\lambda_0^\ell, \tau_\ell)$ by $\lambda_0 = \exp(\lambda_0^\ell + \tau_\ell^2/2)$ and $\tau^2 = [\exp(\tau_\ell^2) - 1]\exp(2\lambda_0^\ell + \tau_\ell^2)$.

B.3.3 Prior-Sensitivity Analysis

TABLE B.4: Sensitivity to prior specifications. The standard deviation for the half-normal distribution is set at $\nu \in \{5, 10\}$.

		Distributional assumption			Posterior of λ_0				Posterior of τ			
Model	ν	Obs. level	Paper level	Pop. level	Mean	SD	2.5%	97.5%	Mean	SD	2.5%	97.5%
M1a	5	Normal		Normal	1.810	0.036	1.741	1.880	0.747	0.027	0.695	0.801
	10	Normal		Normal	1.809	0.035	1.742	1.878	0.747	0.028	0.695	0.804
M1b	5	Normal		Log-normal	1.826	0.039	1.754	1.904	0.816	0.260	0.740	0.899
	10	Normal		Log-normal	1.826	0.038	1.754	1.901	0.815	0.255	0.741	0.895
M1c	5	Normal	Normal	Log-normal	2.052	0.076	1.911	2.209	0.752	0.358	0.602	0.929
	10	Normal	Normal	Log-normal	2.051	0.076	1.906	2.209	0.753	0.362	0.601	0.936
M2	5	Normal	Student-t	Log-normal	1.955	0.072	1.824	2.104	0.743	0.342	0.605	0.907
	10	Normal	Student-t	Log-normal	1.954	0.071	1.822	2.102	0.742	0.341	0.605	0.903

Notes: In Models M1c and M2, (λ_0, τ) are calculated from log-normal parameters $(\lambda_0^\ell, \tau_\ell)$ by $\lambda_0 = \exp(\lambda_0^\ell + \tau_\ell^2/2)$ and $\tau^2 = [\exp(\tau_\ell^2) - 1]\exp(2\lambda_0^\ell + \tau_\ell^2)$.

C Additional Figures and Tables

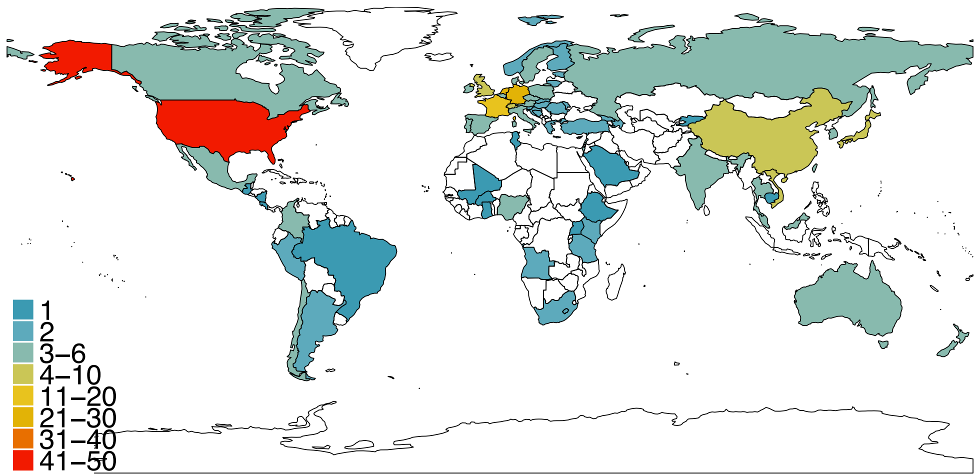

FIGURE C.1: Study location. *Notes*: It is possible that the same dataset was used in two or more papers (e.g., a cross-country dataset from Rieger et al. (2017) and Wang et al. (2017) mentioned above) to estimate model parameters. In such a case, countries are counted multiple times.

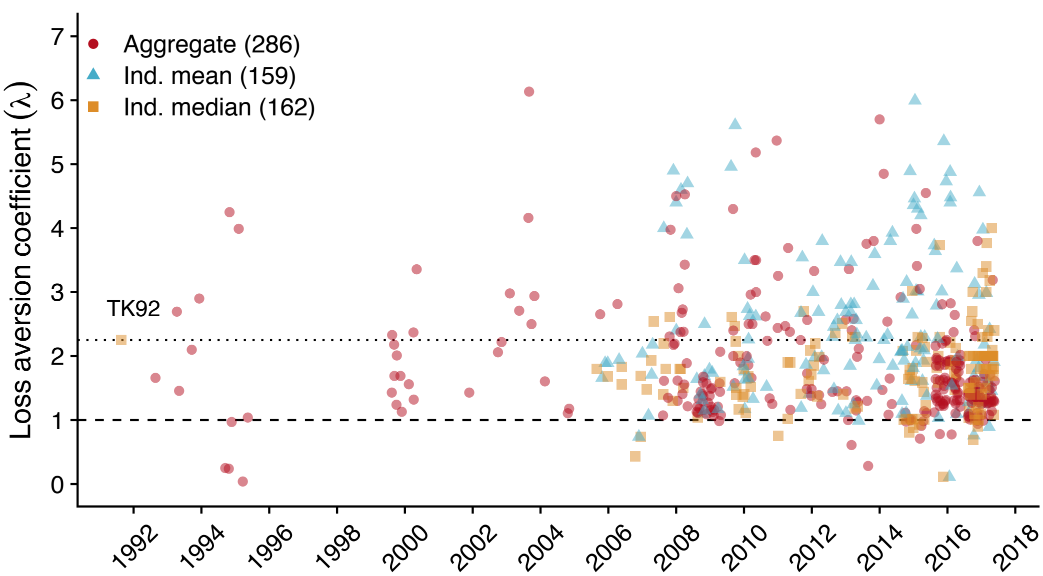

FIGURE C.2: Reported loss aversion coefficients (λ) over time. *Notes*: The y-axis is cut off at 7 for better visualization. The first observation, labeled "TK92", corresponds to the estimate 2.25 from Tversky and Kahneman (1992).

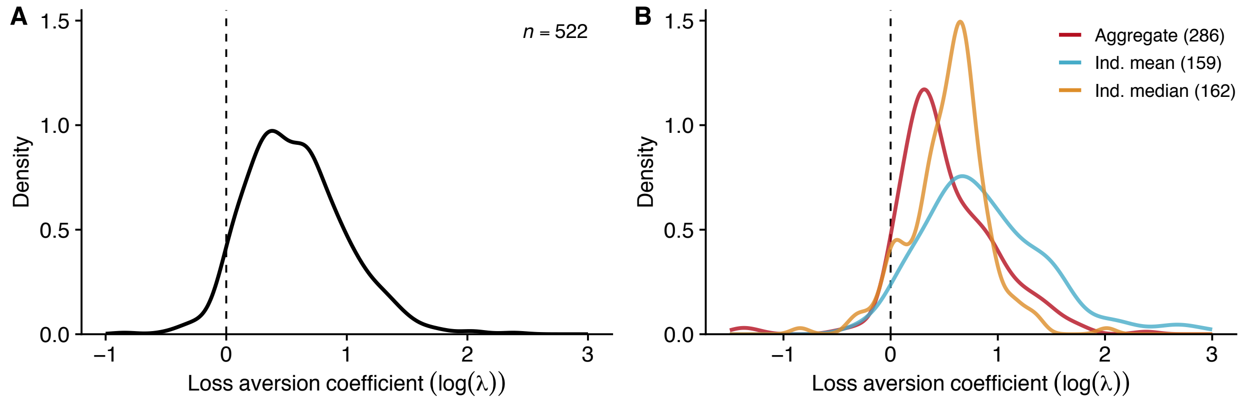

FIGURE C.3: Distribution of logged loss aversion coefficient $\log(\lambda)$. C.f. Figure 3.

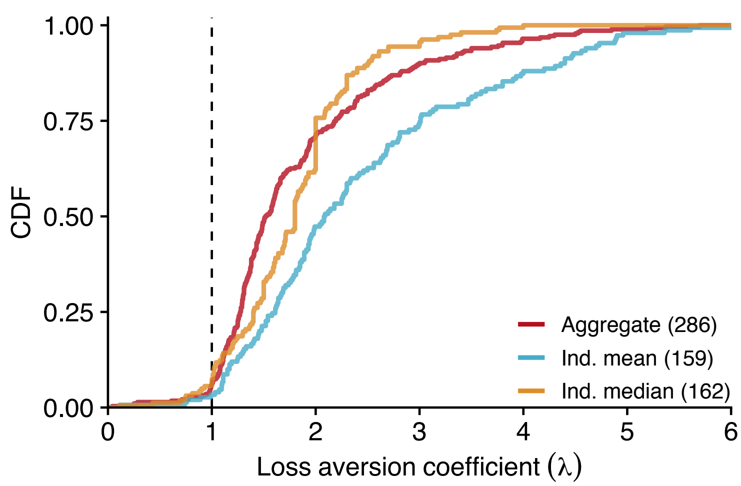

FIGURE C.4: Empirical CDF of reported loss aversion coefficient λ by the type of estimates.

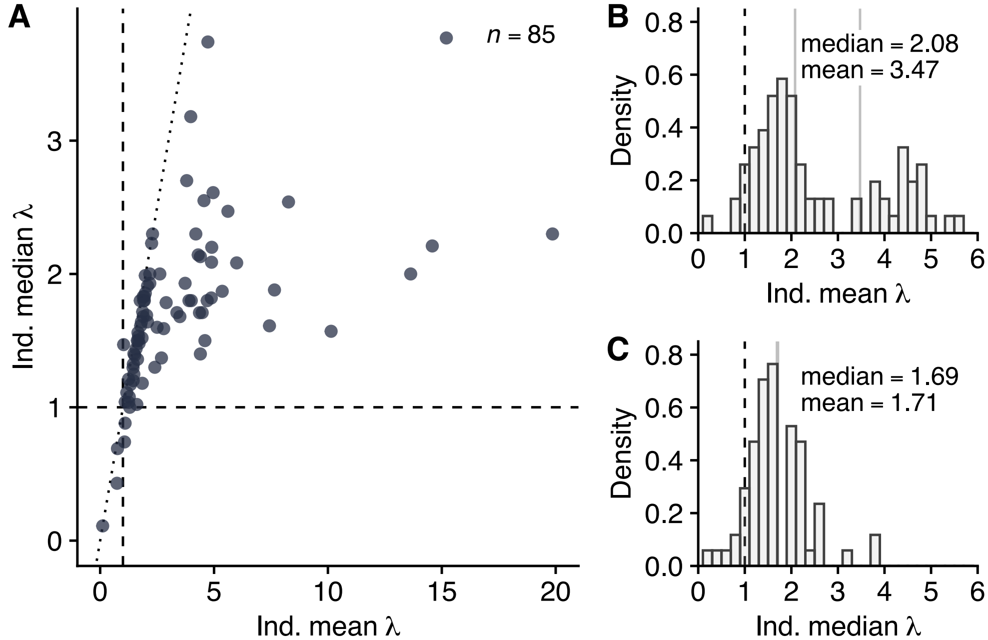

FIGURE C.5: Comparing 85 pairs of individual-level means and medians in 34 papers that report both. *Notes*: The mean is larger than the median in 94% (80 out of 85) of the pairs.

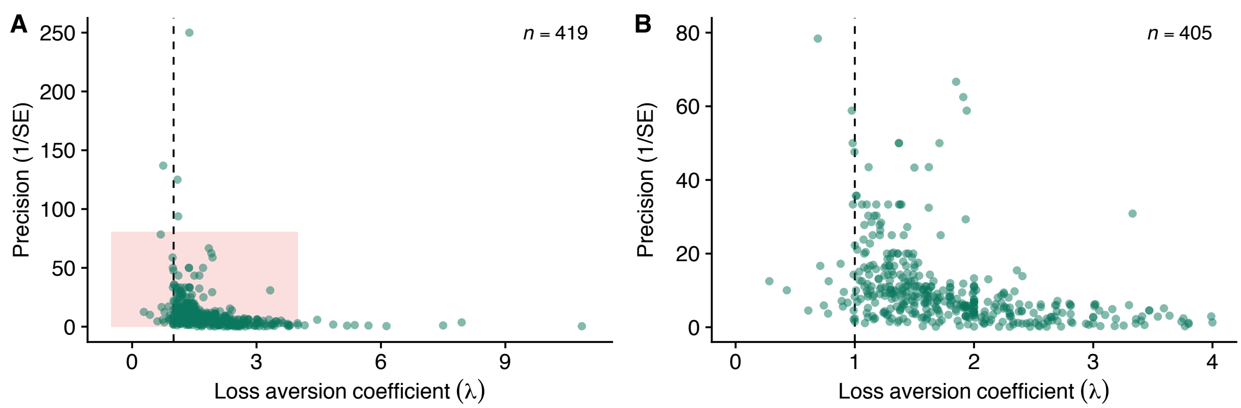

FIGURE C.6: Estimated λ and precision ($1/se$). (A) Complete dataset with reported SE. (B) A subset of the complete dataset (inside the red box in panel A; $\lambda \leq 4$ and $1/se \leq 80$).

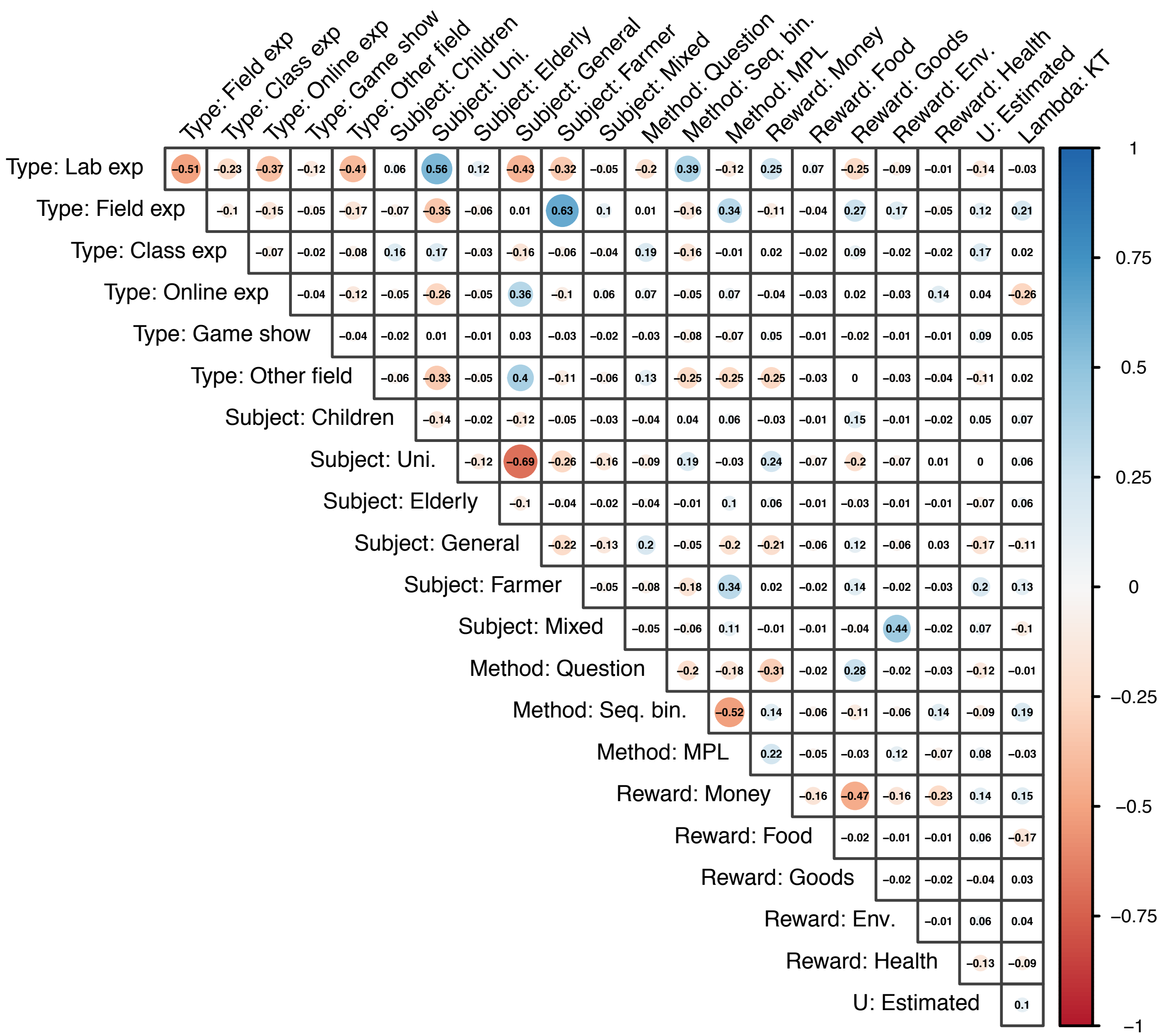

FIGURE C.7: Co-occurences of design characteristics.

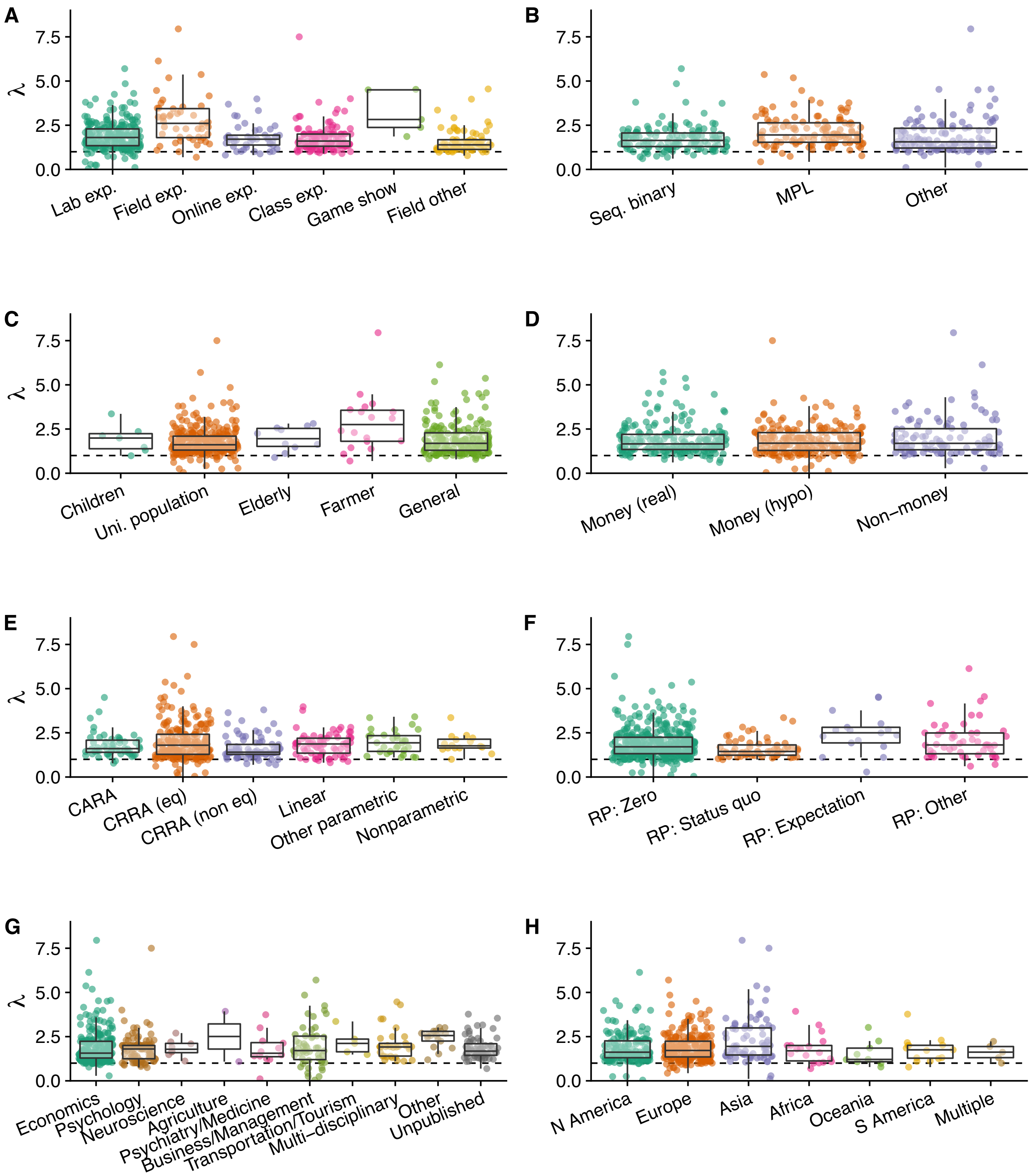

FIGURE C.8: Estimated loss aversion coefficient λ and study characteristics. (A) Type of data. (B) Elicitation method. (C) Subject population. (D) Reward type. (E) Specification of basic utility u. (F) Reference point. (G) Journal discipline. (H) Continent. *Notes*: The horizontal dashed line in each panel corresponds to $\lambda = 1$. The y-axis is cut off at 9 for visual clarity of lower numbers.

Table C.1: Random-effects meta-regression. Posterior distributions of coefficients β.

Category	Variable	Median	2.5%	16.5%	83.5%	97.5%
Type of the estimate	Individual-level mean			baseline		
	Individual-level median	−0.276	−0.386	−0.329	−0.226	−0.173
	Aggregate-level	−0.360	−0.597	−0.475	−0.248	−0.129
Type of the data	Lab experiment			baseline		
	Field experiment	0.536	−0.003	0.274	0.792	1.059
	Class experiment	0.060	−0.497	−0.207	0.324	0.614
	Online experiment	−0.112	−0.626	−0.374	0.150	0.397
	Other field data	−0.258	−0.697	−0.475	−0.046	0.183
Subject pool	Univ. population			baseline		
	General	0.163	−0.137	0.017	0.309	0.467
	Farmer	0.400	−0.300	0.056	0.740	1.106
	Other	−0.083	−0.441	−0.257	0.089	0.273
Reward	Hypothetical money			baseline		
	Real money	−0.042	−0.328	−0.181	0.098	0.255
	Non-money	−0.113	−0.458	−0.276	0.048	0.211
Method	Binary choice			baseline		
	Survey	0.315	−0.244	0.051	0.575	0.852
	Matching	0.458	−0.805	−0.104	0.983	1.522
	Other	0.287	0.003	0.150	0.419	0.561
Functional form of u	CRRA same curvature			baseline		
	CRRA diff curvature	−0.113	−0.400	−0.254	0.034	0.189
	CARA	0.095	−0.390	−0.146	0.320	0.543
	Linear	0.179	−0.177	0.006	0.348	0.526
	Other	−0.099	−0.481	−0.283	0.077	0.259
Reference point	Zero			baseline		
	Status quo	0.041	−0.310	−0.129	0.205	0.372
	Expectation	0.094	−0.698	−0.279	0.504	0.995
	Other	−0.062	−0.410	−0.230	0.103	0.271
Definition of λ	Kahneman and Tversky (1979)			baseline		
	Köbberling and Wakker (2005)	0.234	−0.249	0.011	0.442	0.666
	Kőszegi and Rabin (2006)	0.451	−0.673	−0.086	0.963	1.499
	Other	−1.069	−1.594	−1.315	−0.805	−0.487
	Unknown	−0.658	−1.379	−0.994	−0.333	0.011
Continent	Europe			baseline		
	North America	−0.036	−0.199	−0.115	0.042	0.126
	Asia	−0.050	−0.150	−0.100	−0.001	0.051
	South America	−0.048	−0.249	−0.144	0.056	0.174
	Africa	−0.186	−0.443	−0.309	−0.055	0.097
	Oceania	−0.404	−0.619	−0.504	−0.306	−0.202
	Multiple	0.039	−0.412	−0.138	0.214	0.466
Publication status	Econ journal			baseline		
	Non-econ journal	−0.016	−0.312	−0.155	0.126	0.276
	Unpublished	−0.277	−0.709	−0.486	−0.074	0.137

D "Classical" Meta-Analysis

The random-effects meta-analysis (DerSimonian and Laird, 1986) assumes that

$$\lambda_i = \mu_i + \varepsilon_i = \lambda_0 + \xi_i + \varepsilon_i, \tag{D.1}$$

where $\varepsilon_i \sim \mathcal{N}(0, se_i^2)$ is a sampling variation of λ_i as an estimate of μ_i, and the observation-specific "true" effect μ_i is decomposed into λ_0 (the overall mean) and the sampling variation ξ_i. It is assumed that $\xi_i \sim \mathcal{N}(0, \tau^2)$, where τ^2 is the genuine heterogeneity, beyond the mere sampling variance, that must be estimated. Note that the random-effects model (D.1) reduces to the fixed-effect model when $\tau^2 = 0$. The random-effects estimate of λ_0 is calculated by the weighted average of individual estimates:

$$\lambda_0^{RE} = \frac{\sum_{i=1}^m g_i \lambda_i}{\sum_{i=1}^m g_i},$$

where the weight is given by $g_i = 1/(se_i^2 + \hat{\tau}^2)$ and $\hat{\tau}^2$ is the estimate of τ^2. As we explained in Section B.1 above, the model (D.1) is mathematically equivalent to model M1a. Note also that our dataset includes *statistically dependent* estimates. In order to account for such dependency, we use cluster-robust variance estimation to account for correlation of estimates among each study (Hedges et al., 2010).

We also apply a three-level modeling to handle statistically-dependent estimates. Let λ_{pi} denote the ith estimate of λ from paper p. The first level is $\lambda_{pi} = \mu_{pi} + \varepsilon_{pi}$, where μ_{pi} is the "true" loss aversion coefficient and $\varepsilon_{pi} \sim \mathcal{N}(0, se_{pi}^2)$ for the ith estimate in paper p. The second level is $\mu_{pi} = \overline{\lambda}_p + \xi_{pi}^{(2)}$, where $\overline{\lambda}_p$ is the average loss aversion in paper p and $\xi_{pi}^{(2)} \sim \mathcal{N}(0, \tau_{(2)}^2)$. Finally, the third level is $\overline{\lambda}_p = \lambda_0 + \xi_p^{(3)}$, where λ_0 is the population average of λ and $\xi_p^{(3)} \sim \mathcal{N}(0, \tau_{(3)}^2)$. These equations are combined into a single model:

$$\lambda_{pi} = \lambda_0 + \xi_{pi}^{(2)} + \xi_p^{(3)} + \varepsilon_{pi}. \tag{D.2}$$

We estimate a random-effects model (B.1) and a multi-level model (D.2). Results are presented in Table D.1: columns (1) and (2) use the subset of data where both λ and SE are reported (or reconstructed from other available information), and columns (3) and (4) use the full data where missing SEs are imputed as described above.

Random-effects estimate shows the average loss aversion coefficient between 1.7 and 1.8. The null hypothesis of no loss aversion (i.e., $H_0 : \lambda = 1$) is rejected at the conventional 5% significance level. We also look at the I^2 statistic (Higgins and Thompson, 2002), which measures the amount of heterogeneity relative to the total amount of variance in the observed effects. Formally, the I^2 statistic is computed by

$$I^2 = \frac{\hat{\tau}^2}{\hat{\tau}^2 + s^2} \times 100,$$

Table D.1: "Classical" meta-analytic average of loss aversion coefficient.

	SE reported		All data	
	(1)	(2)	(3)	(4)
	RE	ML	RE	ML
λ_0	1.7124	1.8854	1.8088	1.9373
	(0.0874)	(0.0811)	(0.0761)	(0.0669)
p-value	< 0.0001	< 0.0001	< 0.0001	< 0.0001
$\hat{\tau}^2$	0.5074		0.5562	
$se(\hat{\tau}^2)$	(0.0432)		(0.0386)	
I^2	99.5940		99.5407	
I^2 (within paper)		15.4057		34.0375
I^2 (between paper)		84.2990		65.5952
Observations	352	352	521	521
Clusters	114	114	150	150

Notes: p-values are from the two-sided test of the null hypothesis $H_0 : \lambda_0 = 1$. Standard errors in parentheses are cluster-robust (Hedges et al., 2010). For observations which have both individual-level mean and median, we keep the median. Columns (1)-(2), "SE reported", use the complete dataset where SEs are reported in the paper. Columns (3)-(4), "All data", use the full dataset where missing SEs are approximated from available information or imputed. One observation with $\lambda = 23.46$ (the maximum value in the dataset) is excluded because the very large imputed SE produces an error in the estimation code.

where $\hat{\tau}^2$ is the estimated value of τ^2 and

$$s^2 = \frac{(m-1)\sum g_i}{(\sum g_i)^2 + \sum g_i^2}$$

is the "typical" sampling variance of the observed effect sizes, where $g_i = 1/se_i^2$. We observe that 99% of the total variability in estimates are due to between-observation heterogeneity.

Taking into account the hierarchical structure of our dataset, the multi-level model provides the average loss aversion coefficient about 1.9, that is slightly higher than random-effect estimates discussed above. The heterogeneity measure I^2 adapted to the multi-level specification shows that 84% of of total variance is due to between-paper heterogeneity, 15% is due to within-paper heterogeneity, and the rest (less than 1%) is sampling variation (column (2)). The contribution of between-paper heterogeneity decreases to 66% when we use the full dataset with imputed standard errors (column (4)).

E List of Articles Included in the Meta-Analysis

Abdellaoui, M., H. Bleichrodt, and H. Kammoun (2013a): "Do Financial Professionals Behave According to Prospect Theory? an Experimental Study," *Theory and Decision*, 74, 411–429.

Abdellaoui, M., H. Bleichrodt, and O. l'Haridon (2008): "A Tractable Method to Measure Utility and Loss Aversion under Prospect Theory," *Journal of Risk and Uncertainty*, 36, 245.

Abdellaoui, M., H. Bleichrodt, O. l'Haridon, and C. Paraschiv (2013b): "Is There One Unifying Concept of Utility? an Experimental Comparison of Utility under Risk and Utility over Time," *Management Science*, 59, 2153–2169.

Abdellaoui, M., H. Bleichrodt, O. l'Haridon, and D. Van Dolder (2016): "Measuring Loss Aversion under Ambiguity: A Method to Make Prospect Theory Completely Observable," *Journal of Risk and Uncertainty*, 52, 1–20.

Abdellaoui, M., H. Bleichrodt, and C. Paraschiv (2007): "Loss Aversion under Prospect Theory: A Parameter-Free Measurement," *Management Science*, 53, 1659–1674.

Abdellaoui, M. and E. Kemel (2014): "Eliciting Prospect Theory when Consequences are Measured in Time Units: 'Time is not Money'," *Management Science*, 60, 1844–1859.

Abdellaoui, M., O. L'Haridon, and C. Paraschiv (2011): "Experienced vs. Described Uncertainty: Do We Need Two Prospect Theory Specifications?" *Management Science*, 57, 1879–1895.

Andersson, O., H. J. Holm, J.-R. Tyran, and E. Wengström (2016): "Deciding for Others Reduces Loss Aversion," *Management Science*, 62, 29–36.

Andrikogiannopoulou, A. and F. Papakonstantinou (2016): "Heterogeneity in Risk Preferences: Evidence from a Real-World Betting Market," Unpublished manuscript.

Arbel, Y., D. Ben-Shahar, and S. Gabriel (2016): "Are the Disabled Less Loss Averse? Evidence from a Natural Policy Experiment," *Economic Inquiry*, 54, 1291–1318.

Arkes, H. R., D. Hirshleifer, D. Jiang, and S. S. Lim (2010): "A Cross-Cultural Study of Reference Point Adaptation: Evidence from China, Korea, and the US," *Organizational Behavior and Human Decision Processes*, 112, 99–111.

Armantier, O. and A. Boly (2015): "Framing of Incentives and Effort Provision," *International Economic Review*, 56, 917–938.

Attema, A. E., W. B. Brouwer, and O. l'Haridon (2013): "Prospect Theory in the Health Domain: A Quantitative Assessment," *Journal of Health Economics*, 32, 1057–1065.

Attema, A. E., W. B. Brouwer, O. l'Haridon, and J. L. Pinto (2015): "Estimating Sign-Dependent Societal Preferences for Quality of Life," *Journal of Health Economics*, 43, 229–243.

Baláž, V., V. Bačová, E. Drobná, K. Dudeková, and K. Adamík (2013): "Testing Prospect Theory Parameters," *Ekonomicky časopis*, 61, 655–671.

Baltussen, G., M. J. van den Assem, and D. van Dolder (2016): "Risky Choice in the Limelight," *Review of Economics and Statistics*, 98, 318–332.

Banerji, A. and N. Gupta (2014): "Detection, Identification, and Estimation of Loss Aversion: Evidence from an Auction Experiment," *American Economic Journal: Microeconomics*, 6, 91–133.

Barkley-Levenson, E. E., L. Van Leijenhorst, and A. Galván (2013): "Behavioral and Neural Correlates of Loss Aversion and Risk Avoidance in Adolescents and Adults," *Developmental Cognitive Neuroscience*, 3, 72–83.

Bartczak, A., S. Chilton, and J. Meyerhoff (2015): "Wildfires in Poland: The Impact of Risk Preferences and Loss Aversion on Environmental Choices," *Ecological Economics*, 116, 300–309.

Bateman, H., R. Stevens, and A. Lai (2015): "Risk Information and Retirement Investment Choice Mistakes under Prospect Theory," *Journal of Behavioral Finance*, 16, 279–296.

Baucells, M., M. Weber, and F. Welfens (2011): "Reference-Point Formation and Updating," *Management Science*, 57, 506–519.

Beauchamp, J. P., D. J. Benjamin, C. F. Chabris, and D. I. Laibson (2012): "How Malleable are Risk Preferences and Loss Aversion," Unpublished manuscript.

Bediou, B., I. Comeig, A. Jaramillo-Gutiérrez, and D. Sander (2013): "The Role of «perceived Loss» Aversion on Credit Screening: An Experiment," *Spanish Journal of Finance and Accounting/Revista Española De Financiación y Contabilidad*, 42, 83–97.

Bell, D. R. and J. M. Lattin (2000): "Looking for Loss Aversion in Scanner Panel Data: The Confounding Effect of Price Response Heterogeneity," *Marketing Science*, 19, 185–200.

Berkelaar, A. B., R. Kouwenberg, and T. Post (2004): "Optimal Portfolio Choice under Loss Aversion," *Review of Economics and Statistics*, 86, 973–987.

Bocquého, G., F. Jacquet, and A. Reynaud (2014): "Expected Utility or Prospect Theory Maximisers? Assessing Farmers' Risk Behaviour from Field-Experiment Data," *European Review of Agricultural Economics*, 41, 135–172.

Booij, A. S. and G. Van de Kuilen (2009): "A Parameter-Free Analysis of the Utility of Money for the General Population under Prospect Theory," *Journal of Economic Psychology*, 30, 651–666.

Booij, A. S., B. M. Van Praag, and G. Van de Kuilen (2010): "A Parametric Analysis of Prospect Theory's Functionals for the General Population," *Theory and Decision*, 68, 115–148.

Bougherara, D., X. Gassmann, L. Piet, and A. Reynaud (2017): "Structural Estimation of Farmers' Risk and Ambiguity Preferences: a Field Experiment," *European Review of Agricultural Economics*, 44, 782–808.

Brooks, A. M., C. M. Capra, and G. S. Berns (2012): "Neural Insensitivity to Upticks in Value is Associated with the Disposition Effect," *NeuroImage*, 59, 4086–4093.

Brown, J. K., J. A. Waltz, G. P. Strauss, R. P. McMahon, M. J. Frank, and J. M. Gold (2013): "Hypothetical Decision Making in Schizophrenia: The Role of Expected Value Computation and 'Irrational' Biases," *Psychiatry Research*, 209, 142–149.

Buchanan, J. and D. Houser (2017): "What If Wages Fell During a Recession?" Unpublished manuscript.

Campos-Vazquez, R. M. and E. Cuilty (2014): "The Role of Emotions on Risk Aversion: a Prospect Theory Experiment," *Journal of Behavioral and Experimental Economics*, 50, 1–9.

Casey, J. T. (1994): "Buyers' Pricing Behavior for Risky Alternatives: Encoding Processes and Preference Reversals," *Management Science*, 40, 730–749.

——— (1995): "Predicting Buyer-Seller Pricing Disparities," *Management Science*, 41, 979–999.

Cecchi, F., R. Lensink, and E. Slingerland (2017): "Loss and Ambiguity Aversion and the Willingness to Pay for Index Insurance: Experimental Evidence from Rural Kenya," Unpublished manuscript.

Charpentier, C. J., C. Hindocha, J. P. Roiser, and O. J. Robinson (2016): "Anxiety Promotes Memory for Mood-Congruent Faces but Does not Alter Loss Aversion," *Scientific Reports*, 6, 24746.

Chen, M. K., V. Lakshminarayanan, and L. R. Santos (2006): "How Basic are Behavioral Biases? Evidence from Capuchin Monkey Trading Behavior," *Journal of Political Economy*, 114, 517–537.

Chib, V. S., B. De Martino, S. Shimojo, and J. P. O'Doherty (2012): "Neural Mechanisms Underlying Paradoxical Performance for Monetary Incentives are Driven by Loss Aversion," *Neuron*, 74, 582–594.

Chiong, W., K. A. Wood, A. J. Beagle, M. Hsu, A. S. Kayser, B. L. Miller, and J. H. Kramer (2015): "Neuroeconomic Dissociation of Semantic Dementia and Behavioural Variant Frontotemporal Dementia," *Brain*, 139, 578–587.

d'Acremont, M., E. Fornari, and P. Bossaerts (2013): "Activity in Inferior Parietal and Medial Prefrontal Cortex Signals the Accumulation of Evidence in a Probability Learning Task," *PLoS Computational Biology*, 9, e1002895.

Dadzie, S., K. Balcombe, N. Bardsley, and I. Fraser (2017): "Loss Aversion and Risk Preferences Under Prospect Theory: A Study of Ghanaian Farmers," Unpublished manuscript.

De Langhe, B. and S. Puntoni (2015): "Bang for the Buck: Gain-Loss Ratio as a Driver of Judgment and Choice," *Management Science*, 61, 1137–1163.

De Martino, B., C. F. Camerer, and R. Adolphs (2010): "Amygdala Damage Eliminates Monetary Loss Aversion," *Proceedings of the National Academy of Sciences*, 107, 3788–3792.

Delle Site, P. and F. Filippi (2011): "Stochastic User Equilibrium and Value-of-Time Analysis with Reference-Dependent Route Choice," *European Journal of Transport and Infrastructure Research*, 11, 194–218.

Dimmock, S. G. and R. Kouwenberg (2010): "Loss-Aversion and Household Portfolio Choice," *Journal of Empirical Finance*, 17, 441–459.

Engström, P., K. Nordblom, H. Ohlsson, and A. Persson (2015): "Tax Compliance and Loss Aversion," *American Economic Journal: Economic Policy*, 7, 132–64.

Erev, I., E. Ert, and E. Yechiam (2008): "Loss Aversion, Diminishing Sensitivity, and the Effect of Experience on Repeated Decisions," *Journal of Behavioral Decision Making*, 21, 575–597.

Erner, C., A. Klos, and T. Langer (2013): "Can Prospect Theory be Used to Predict an Investor's Willingness to Pay?" *Journal of Banking & Finance*, 37, 1960–1973.

Ernst, M., R. C. Plate, C. O. Carlisi, E. Gorodetsky, D. Goldman, and D. S. Pine (2014): "Loss Aversion and 5htt Gene Variants in Adolescent Anxiety," *Developmental Cognitive Neuroscience*, 8, 77–85.

Fahle, S., S. Sautua, et al. (2017): "How Do Risk Attitudes Affect Pro-Social Behavior? Theory and Experiment," Unpublished manuscript.

Farber, H. S. (2015): "Why you Can't Find a Taxi in the Rain and Other Labor Supply Lessons from Cab Drivers," *The Quarterly Journal of Economics*, 130, 1975–2026.

Feess, E., H. Müller, and C. Schumacher (2016): "Estimating Risk Preferences of Bettors with Different Bet Sizes," *European Journal of Operational Research*, 249, 1102–1112.

Festjens, A., S. Bruyneel, E. Diecidue, and S. Dewitte (2015): "Time-Based Versus Money-Based Decision Making under Risk: An Experimental Investigation," *Journal of Economic Psychology*, 50, 52–72.

Fleming, S. M. and R. J. Dolan (2010): "Effects of Loss Aversion on Post-Decision Wagering: Implications for Measures of Awareness," *Consciousness and Cognition*, 19, 352–363.

Freudenreich, H., O. Musshoff, and B. Wiercinski (2017): "The Relationship between Farmers' Shock Experiences and Their Uncertainty Preferences – Experimental Evidence from Mexico," Unpublished manuscript.

Frydman, C., C. Camerer, P. Bossaerts, and A. Rangel (2010): "MAOA-L Carriers are Better at

Making Optimal Financial Decisions under Risk," *Proceedings of the Royal Society B: Biological Sciences*, 278, 2053–2059.

Füllbrunn, S. C. and W. J. Luhan (2017): "Decision Making for Others: The Case of Loss Aversion," *Economics Letters*, 161, 154–156.

Gächter, S., E. J. Johnson, and A. Herrmann (2010): "Individual-Level Loss Aversion in Riskless and Risky Choices," Unpublished manuscript.

Gill, D. and V. Prowse (2012): "A Structural Analysis of Disappointment Aversion in a Real Effort Competition," *American Economic Review*, 102, 469–503.

Giorgetta, C., A. Grecucci, A. Rattin, C. Guerreschi, A. G. Sanfey, and N. Bonini (2014): "To Play or not to Play: a Personal Dilemma in Pathological Gambling," *Psychiatry Research*, 219, 562–569.

Glöckner, A. and T. Pachur (2012): "Cognitive Models of Risky Choice: Parameter Stability and Predictive Accuracy of Prospect Theory," *Cognition*, 123, 21–32.

Goeree, J. K. and T. Offerman (2003): "Winner's Curse without Overbidding," *European Economic Review*, 47, 625–644.

Goldstein, D. G., E. J. Johnson, and W. F. Sharpe (2008): "Choosing Outcomes Versus Choosing Products: Consumer-Focused Retirement Investment Advice," *Journal of Consumer Research*, 35, 440–456.

Gurevich, G., D. Kliger, and O. Levy (2009): "Decision-Making under Uncertainty–A Field Study of Cumulative Prospect Theory," *Journal of Banking & Finance*, 33, 1221–1229.

Habib, M. A. and E. J. Miller (2009): "Reference-Dependent Residential Location Choice Model Within a Relocation Context," *Transportation Research Record*, 2133, 92–99.

Hardie, B. G., E. J. Johnson, and P. S. Fader (1993): "Modeling Loss Aversion and Reference Dependence Effects on Brand Choice," *Marketing Science*, 12, 378–394.

Harrison, G., S. Mortiz, and R. Pibernik (2010): "Characterizing Risk Attitudes of Industrial Managers," Unpublished manuscript.

Harrison, G. W. and E. E. Rutström (2009): "Expected Utility Theory and Prospect Theory: One Wedding and a Decent Funeral," *Experimental Economics*, 12, 133.

Harrison, G. W. and J. T. Swarthout (2016): "Cumulative Prospect Theory in the Laboratory: A Reconsideration," Unpublished manuscript.

Hastings, J. S. and J. M. Shapiro (2013): "Fungibility and Consumer Choice: Evidence from Commodity Price Shocks," *The Quarterly Journal of Economics*, 128, 1449–1498.

Imas, A., S. Sadoff, and A. Samek (2016): "Do People Anticipate Loss Aversion?" *Management*

Science, 63, 1271–1284.

Jakusch, S. T., S. Meyer, and A. Hackethal (2016): "Taming models of prospect theory in the wild? Estimation of Vlcek and Hens (2011)," Unpublished manuscript.

Jindal, P. (2015): "Risk Preferences and Demand Drivers of Extended Warranties," *Marketing Science*, 34, 39–58.

Johnson, E. J., S. Gächter, and A. Herrmann (2006): "Exploring the Nature of Loss Aversion," Unpublished manuscript.

Kapoor, S. (2017): "Are Firms Loss Averse? Evidence from the Restaurant Industry," Unpublished manuscript.

Kemel, E. and C. Paraschiv (2013): "Prospect Theory for Joint Time and Money Consequences in Risk and Ambiguity," *Transportation Research Part B: Methodological*, 56, 81–95.

Kim, M.-s., B.-N. Kang, and J. Y. Lim (2016): "Decision-Making Deficits in Patients with Chronic Schizophrenia: Iowa Gambling Task and Prospect Valence Learning Model," *Neuropsychiatric Disease and Treatment*, 12, 1019.

Kivetz, R., O. Netzer, and V. Srinivasan (2004): "Alternative Models for Capturing the Compromise Effect," *Journal of Marketing Research*, 41, 237–257.

Klapper, D., C. Ebling, and J. Temme (2005): "Another Look at Loss Aversion in Brand Choice Data: Can We Characterize the Loss Averse Consumer?" *International Journal of Research in Marketing*, 22, 239–254.

Kliger, D. and O. Levy (2009): "Theories of Choice under Risk: Insights from Financial Markets," *Journal of Economic Behavior & Organization*, 71, 330–346.

Koudstaal, M., R. Sloof, and M. Van Praag (2016): "Risk, Uncertainty, and Entrepreneurship: Evidence from a Lab-in-the-Field Experiment," *Management Science*, 62, 2897–2915.

Krajbich, I., C. Camerer, and A. Rangel (2017): "Exploring the Scope of Neurometrically Informed Mechanism Design," *Games and Economic Behavior*, 101, 49–62.

Krčál, O., M. Kvasnička, and R. Staněk (2016): "External Validity of Prospect Theory: The Evidence from Soccer Betting," *Journal of Behavioral and Experimental Economics*, 65, 121–127.

Kudryavtsev, A. and J. Pavlodsky (2012): "Description-Based and Experience-Based Decisions: Individual Analysis." *Judgment & Decision Making*, 7.

Lanchava, L., K. Carlson, B. Šebánková, J. Flegr, and G. Nave (2015): "No Evidence of Association Between Toxoplasma Gondii Infection and Financial Risk Taking in Females," *PLoS One*, 10, e0136716.

Lee, B. and Y. Veld-Merkoulova (2016): "Myopic Loss Aversion and Stock Investments: An

Empirical Study of Private Investors," *Journal of Banking & Finance*, 70, 235–246.

Lee, S., M. J. Lee, B. W. Kim, J. M. Gilman, J. K. Kuster, A. J. Blood, C. M. Kuhnen, and H. C. Breiter (2015): "The Commonality of Loss Aversion Across Procedures and Stimuli," *PloS One*, 10, e0135216.

Levy, M. (2010): "Loss Aversion and the Price of Risk," *Quantitative Finance*, 10, 1009–1022.

l'Haridon, O. and F. M. Vieider (2016): "All Over the Map: A Worldwide Comparison of Risk Preferences," Unpublished manuscript.

Li, Q., W. Nan, J. Taxer, W. Dai, Y. Zheng, and X. Liu (2016): "Problematic Internet Users Show Impaired Inhibitory Control and Risk Taking with Losses: Evidence from Stop Signal and Mixed Gambles Tasks," *Frontiers in Psychology*, 7, 370.

Li, Y., M. Baldassi, E. J. Johnson, and E. U. Weber (2013): "Complementary Cognitive Capabilities, Economic Decision Making, and Aging." *Psychology and Aging*, 28, 595.

Liebenehm, S. and H. Waibel (2014): "Simultaneous Estimation of Risk and Time Preferences among Small-Scale Cattle Farmers in West Africa," *American Journal of Agricultural Economics*, 96, 1420–1438.

Lim, S.-L. and A. S. Bruce (2015): "Prospect Theory and Body Mass: Characterizing Psychological Parameters for Weight-Related Risk Attitudes and Weight-Gain Aversion," *Frontiers in Psychology*, 6, 330.

Liu, E. M. (2013): "Time to Change What to Sow: Risk Preferences and Technology Adoption Decisions of Cotton Farmers in China," *Review of Economics and Statistics*, 95, 1386–1403.

Liu, E. M. and J. Huang (2013): "Risk Preferences and Pesticide Use by Cotton Farmers in China," *Journal of Development Economics*, 103, 202–215.

Liu, E. M., J. Meng, and J. T.-y. Wang (2014): "Confucianism and Preferences: Evidence from Lab Experiments in Taiwan and China," *Journal of Economic Behavior & Organization*, 104, 106–122.

Lorains, F. K., N. A. Dowling, P. G. Enticott, J. L. Bradshaw, J. S. Trueblood, and J. C. Stout (2014): "Strategic and Non-Strategic Problem Gamblers Differ on Decision-Making under Risk and Ambiguity," *Addiction*, 109, 1128–1137.

Nakamoto, Y. and M. Sato (2016): "Gender Differences, Social Loss Aversion and Sports Performance in Japanese Schoolchildren," *International Journal of Applied Behavioral Economics*, 5, 14–30.

Nguyen, Q. (2010): "How Nurture can Shape Preferences: An Experimental Study on Risk Preferences of Vietnamese Fishers," *Environment and Development Economics*, 15, 609–631.

——— (2011): "Does Nurture Matter: Theory and Experimental Investigation on the Effect of

Working Environment on Risk and Time Preferences," *Journal of Risk and Uncertainty*, 43, 245–270.

——— (2016): "Linking Loss Aversion and Present Bias with Overspending Behavior of Tourists: Insights from a Lab-in-the-Field Experiment," *Tourism Management*, 54, 152–159.

Nguyen, Q., M. C. Villeval, and H. Xu (2012): "Trust and Trustworthiness Under the Prospect Theory: A Field Experiment in Vietnam," Unpublished manuscript.

Nilsson, H., J. Rieskamp, and E.-J. Wagenmakers (2011): "Hierarchical Bayesian Parameter Estimation for Cumulative Prospect Theory," *Journal of Mathematical Psychology*, 55, 84–93.

Pachur, T., R. Mata, and R. Hertwig (2017): "Who Dares, Who Errs? Disentangling Cognitive and Motivational Roots of Age Differences in Decisions under Risk," *Psychological Science*, 28, 504–518.

Peón, D., M. Antelo, and A. Calvo (2016): "Overconfidence and Risk Seeking in Credit Markets: An Experimental Game," *Review of Managerial Science*, 10, 511–552.

Petraud, J., S. Boucher, and M. Carter (2015): "Competing Theories of Risk Preferences and the Demand for Crop Insurance: Experimental Evidence from Peru," Unpublished manuscript.

Post, T., M. J. Van den Assem, G. Baltussen, and R. H. Thaler (2008): "Deal or No Deal? Decision Making under Risk in a Large-Payoff Game Show," *American Economic Review*, 98, 38–71.

Rau, H. A. (2014): "The Disposition Effect and Loss Aversion: Do Gender Differences Matter?" *Economics Letters*, 123, 33–36.

——— (2015): "The Disposition Effect in Team Investment Decisions: Experimental Evidence," *Journal of Banking & Finance*, 61, 272–282.

Rieger, M. O., M. Wang, and T. Hens (2017): "Estimating Cumulative Prospect Theory Parameters from an International Survey," *Theory and Decision*, 82, 567–596.

Rosenblatt-Wisch, R. (2008): "Loss Aversion in Aggregate Macroeconomic Time Series," *European Economic Review*, 52, 1140–1159.

Rutledge, R. B., P. Smittenaar, P. Zeidman, H. R. Brown, R. A. Adams, U. Lindenberger, P. Dayan, and R. J. Dolan (2016): "Risk Taking for Potential Reward Decreases Across the Lifespan," *Current Biology*, 26, 1634–1639.

Santos-Pinto, L., A. Bruhin, J. Mata, and T. Åstebro (2015): "Detecting Heterogeneous Risk Attitudes with Mixed Gambles," *Theory and Decision*, 79, 573–600.

Schmidt, U. and S. Traub (2002): "An Experimental Test of Loss Aversion," *Journal of Risk and Uncertainty*, 25, 233–249.

Schulreich, S., H. Gerhardt, and H. R. Heekeren (2016): "Incidental Fear Cues Increase Monetary Loss Aversion." *Emotion*, 16, 402.

Schunk, D. (2009): "Behavioral Heterogeneity in Dynamic Search Situations: Theory and Experimental Evidence," *Journal of Economic Dynamics and Control*, 33, 1719–1738.

Sheremenko, G. and N. Magnan (2015): "Gender-Specific Risk Preferences and Fertilizer Use in Kenyan Farming Households," Unpublished manuscript.

Shimokawa, S. (2013): "Two Asymmetric and Conflicting Learning Effects of Calorie Posting on Overeating: Laboratory Snack Choice Experiment," Unpublished manuscript.

Sokol-Hessner, P., C. A. Hartley, J. R. Hamilton, and E. A. Phelps (2015): "Interoceptive Ability Predicts Aversion to Losses," *Cognition and Emotion*, 29, 695–701.

Sokol-Hessner, P., M. Hsu, N. G. Curley, M. R. Delgado, C. F. Camerer, and E. A. Phelps (2009): "Thinking Like a Trader Selectively Reduces Individuals' Loss Aversion," *Proceedings of the National Academy of Sciences*, 106, 5035–5040.

Song, C. (2016): "An Experiment on Reference Points and Expectations," Unpublished manuscript.

Sprenger, C. (2015): "An Endowment Effect for Risk: Experimental Tests of Stochastic Reference Points," *Journal of Political Economy*, 123, 1456–1499.

Stanton, S. J., O. A. Mullette-Gillman, R. E. McLaurin, C. M. Kuhn, K. S. LaBar, M. L. Platt, and S. A. Huettel (2011): "Low- and High-Testosterone Individuals Exhibit Decreased Aversion to Economic Risk," *Psychological Science*, 22, 447–453.

Takeuchi, H., R. Kawada, K. Tsurumi, N. Yokoyama, A. Takemura, T. Murao, T. Murai, and H. Takahashi (2016): "Heterogeneity of Loss Aversion in Pathological Gambling," *Journal of Gambling Studies*, 32, 1143–1154.

Takeuchi, H., K. Tsurumi, T. Murao, A. Takemura, R. Kawada, S.-i. Urayama, T. Aso, G.-i. Sugihara, J. Miyata, T. Murai, and H. Takahashi (2017): "Common and Differential Brain Abnormalities in Gambling Disorder Subtypes Based on Risk Attitude," *Addictive Behaviors*, 69, 48–54.

Tanaka, T., C. F. Camerer, and Q. Nguyen (2010): "Risk and Time Preferences: Linking Experimental and Household Survey Data from Vietnam," *American Economic Review*, 100, 557–71.

Tanaka, Y. and A. Munro (2013): "Regional Variation in Risk and Time Preferences: Evidence from a Large-Scale Field Experiment in Rural Uganda," *Journal of African Economies*, 23, 151–187.

Tom, S. M., C. R. Fox, C. Trepel, and R. A. Poldrack (2007): "The Neural Basis of Loss Aversion in Decision-Making under Risk," *Science*, 315, 515–518.

Tovar, P. (2009): "The Effects of Loss Aversion on Trade Policy: Theory and Evidence," *Journal of International Economics*, 78, 154–167.

Tversky, A. and D. Kahneman (1992): "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 5, 297–323.

Van Assen, M. A. L. M. and C. C. P. Snijders (2010): "The Effect of Nonlinear Utility on Behaviour in Repeated Prisoner's Dilemmas," *Rationality and Society*, 22, 301–332.

van de Kaa, E. J. (2010): "Sign-Dependent Value of Time in Stated Preference: Judgment Bias or Exposure of Genuine Preference?" *European Journal of Transport and Infrastructure Research*, 10.

Viegas, R. G., A. M. Oliveira, A. Garriga-Trillo, and A. Grieco (2012): "A Functional Model for the Integration of Gains and Losses under Risk: Implications for the Measurement of Subjective Value," *Psicológica*, 33, 711–733.

Vieider, F. M. (2012): "Moderate Stake Variations for Risk and Uncertainty, Gains and Losses: Methodological Implications for Comparative Studies," *Economics Letters*, 117, 718–721.

Vieider, F. M., C. Villegas-Palacio, P. Martinsson, and M. Mejía (2016): "Risk Taking for Oneself and Others: A Structural Model Approach," *Economic Inquiry*, 54, 879–894.

Viswanathan, V., S. Lee, J. M. Gilman, B. W. Kim, N. Lee, L. Chamberlain, S. L. Livengood, K. Raman, M. J. Lee, J. Kuster, D. B. Stern, B. Calder, F. J. Mulhern, A. J. Blood, and H. C. Breiter (2015): "Age-Related Striatal BOLD Changes Without Changes in Behavioral Loss Aversion," *Frontiers in Human Neuroscience*, 9, 176.

Voigt, G., C. Montag, S. Markett, and M. Reuter (2015): "On the Genetics of Loss Aversion: An Interaction Effect of BDNF Val66Met and DRD2/ANKK1 Taq1a." *Behavioral Neuroscience*, 129, 801.

von Gaudecker, H.-M., A. van Soest, and E. Wengstrom (2011): "Heterogeneity in Risky Choice Behavior in a Broad Population," *American Economic Review*, 101, 664–94.

von Gaudecker, H.-M., A. van Soest, and E. Wengström (2012): "Experts in Experiments: How Selection Matters for Estimated Distributions of Risk Preferences," *Journal of Risk and Uncertainty*, 45, 159–190.

Walasek, L. and N. Stewart (2015): "How to Make Loss Aversion Disappear and Reverse: Tests of the Decision by Sampling Origin of Loss Aversion." *Journal of Experimental Psychology: General*, 144, 7–11.

Wang, M., M. O. Rieger, and T. Hens (2017): "The Impact of Culture on Loss Aversion," *Journal of Behavioral Decision Making*, 30, 270–281.

WANG, S. W., M. FILIBA, AND C. F. CAMERER (2010): "Dynamically Optimized Sequential Experimentation (DOSE) for Estimating Economic Preference Parameters," Unpublished manuscript.

WARD, P. S., D. L. ORTEGA, D. J. SPIELMAN, AND V. SINGH (2014): "Heterogeneous Demand for Drought-Tolerant Rice: Evidence from Bihar, India," *World Development*, 64, 125–139.

WARD, P. S. AND V. SINGH (2015): "Using Field Experiments to Elicit Risk and Ambiguity Preferences: Behavioural Factors and the Adoption of New Agricultural Technologies in Rural India," *The Journal of Development Studies*, 51, 707–724.

WORTHY, D. A., B. PANG, AND K. A. BYRNE (2013): "Decomposing the Roles of Perseveration and Expected Value Representation in Models of the Iowa Gambling Task," *Frontiers in Psychology*, 4, 640.

YOGO, M. (2008): "Asset Prices under Habit Formation and Reference-Dependent Preferences," *Journal of Business & Economic Statistics*, 26, 131–143.

ZEISBERGER, S., T. LANGER, AND M. WEBER (2012a): "Why Does Myopia Decrease the Willingness to Invest? is it Myopic Loss Aversion or Myopic Loss Probability Aversion?" *Theory and Decision*, 72, 35–50.

ZEISBERGER, S., D. VRECKO, AND T. LANGER (2012b): "Measuring the Time Stability of Prospect Theory Preferences," *Theory and Decision*, 72, 359–386.

ZHANG, Y., K. DONOHUE, AND T. H. CUI (2016): "Contract Preferences and Performance for the Loss-Averse Supplier: Buyback vs. Revenue Sharing," *Management Science*, 62, 1734–1754.

References

Bowman, D., D. Minehart, and M. Rabin (1999): "Loss Aversion in a Consumption-Savings Model," *Journal of Economic Behavior & Organization*, 38, 155–178.

Carpenter, B., A. Gelman, M. D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. Brubaker, J. Guo, P. Li, and A. Riddell (2017): "Stan: A Probabilistic Programming Language," *Journal of Statistical Software*, 76.

DerSimonian, R. and N. Laird (1986): "Meta-Analysis in Clinical Trials," *Controlled Clinical Trials*, 7, 177–188.

Gelman, A. and J. Hill (2006): *Data Analysis Using Regression and Multilevel/Hierarchical Models*, Cambridge: Cambridge University Press.

Gelman, A. and I. Pardoe (2006): "Bayesian Measures of Explained Variance and Pooling in Multilevel (Hierarchical) Models," *Technometrics*, 48, 241–251.

Hedges, L. V., E. Tipton, and M. C. Johnson (2010): "Robust Variance Estimation in Meta-Regression with Dependent Effect Size Estimates," *Research Synthesis Methods*, 1, 39–65.

Higgins, J. P. T. and S. G. Thompson (2002): "Quantifying Heterogeneity in a Meta-Analysis," *Statistics in Medicine*, 21, 1539–1558.

Kahneman, D. and A. Tversky (1979): "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, 47, 263–292.

Köbberling, V. and P. P. Wakker (2005): "An Index of Loss Aversion," *Journal of Economic Theory*, 122, 119–131.

Kőszegi, B. and M. Rabin (2006): "A model of Reference-Dependent Preferences," *The Quarterly Journal of Economics*, 121, 1133–1165.

McElreath, R. (2016): *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*, CRC Press.

Neilson, W. S. (2002): "Comparative Risk Sensitivity with Reference-Dependent Preferences," *Journal of Risk and Uncertainty*, 24, 131–142.

Rieger, M. O., M. Wang, and T. Hens (2017): "Estimating Cumulative Prospect Theory Parameters from an International Survey," *Theory and Decision*, 82, 567–596.

Stan Development Team (2020): "RStan: the R interface to Stan," R package version 2.21.2, http://mc-stan.org/.

Tversky, A. and D. Kahneman (1992): "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 5, 297–323.

Wakker, P. and A. Tversky (1993): "An Axiomatization of Cumulative Prospect Theory," *Journal of Risk and Uncertainty*, 7, 147–175.

Wang, M., M. O. Rieger, and T. Hens (2017): "The Impact of Culture on Loss Aversion," *Journal of Behavioral Decision Making*, 30, 270–281.