Meta-analyses in Economic Psychology: A sustainable approach to cross-cultural differences

Matteo M. Marini / Department of Public Economics and Masaryk University Experimental Economics Laboratory, Masaryk University, Brno, Czech Republic
Giulia Ulivieri / Department of Economics and Statistics, University of Siena, Siena, Italy
Meta-analyses in Economic Psychology: A sustainable approach to cross-cultural differences

Abstract

This manuscript is a methodological work on the state of research using meta-analytic procedures in Economic Psychology, with a focus on the investigation of cross-cultural differences. We review published meta-analyses and introduce a new classification thereof by data source, describing how the different categories relate to the study of cross-cultural differences. We also discuss related opportunities and challenges, proposing a sustainable methodological approach that is then implemented in three case studies where we re-analyze data from published meta-analyses. In doing so, the relevance of culture as a determinant is explored by relating country-level cultural indicators to experimental measures of risk aversion, tax compliance, and prosocial behavior, respectively. It turns out that, after we control for country-level cultural heterogeneity and economic development, country-level individualism predicts these economic outcomes. We discuss possible interpretations of our findings.
Meta-analyses in Economic Psychology:
A sustainable approach to cross-cultural differences

Matteo M. Marini\textsuperscript{a} and Giulia Ulivieri\textsuperscript{b}

\textsuperscript{a}Department of Public Economics & Masaryk University Experimental Economics Laboratory, Masaryk University, Brno, Czech Republic
\textsuperscript{b}Department of Economics and Statistics, University of Siena, Siena, Italy

Abstract

This manuscript is a methodological work on the state of research using meta-analytic procedures in Economic Psychology, with a focus on the investigation of cross-cultural differences. We review published meta-analyses and introduce a new classification thereof by data source, describing how the different categories relate to the study of cross-cultural differences. We also discuss related opportunities and challenges, proposing a sustainable methodological approach that is then implemented in three case studies where we re-analyze data from published meta-analyses. In doing so, the relevance of culture as a determinant is explored by relating country-level cultural indicators to experimental measures of risk aversion, tax compliance, and prosocial behavior, respectively. It turns out that, after we control for country-level cultural heterogeneity and economic development, country-level individualism predicts these economic outcomes. We discuss possible interpretations of our findings.

Keywords: meta-analysis; individualism; fractionalization; Multiple Price List; Tax Evasion Game; Dictator Game.

Correspondence to: Department of Public Economics, Faculty of Economics and Administration, Masaryk University, Lipova 507/41a, 602 00 Brno (Czech Republic).
E-mail addresses: matteom.marini@gmail.com (M.M. Marini); g.ulivieri1@student.unisi.it (G. Ulivieri).

Declarations of interest: none.
**JEL classification codes:** C90; Z10.

**PsycINFO classification codes:** 2260; 2930.

1 Introduction

Meta-analysis as a statistical tool to combine and evaluate empirical results from multiple studies on a specific topic has proliferated in Economics in recent times (Poot, 2012), even though it is still relatively underused if one looks at disciplines such as Psychology and Medicine. Thanks to their ability to synthesize the available evidence and inform the design of new studies, two different types of meta-studies from the viewpoint of data structure have gained popularity over the years: the aggregate data (AD) meta-analysis and the individual participant data (IPD) meta-analysis. While the former involves the synthesis of study-level data and is the most commonly used, the latter relies on individual-level data and is widely regarded as the gold standard for meta-analysis (Riley et al., 2010). From the viewpoint of the methodological approaches adopted in Economics, instead, Geyer-Klingeberg et al. (2020) introduce the following categorization: (i) traditional meta-analysis, which focuses on the estimation of the overall effect size as a weighted average of the effect sizes reported in a set of prior studies;\(^1\) (ii) meta-regression analysis, which goes one step further and explains the simultaneous effect of different heterogeneity drivers through regression analysis; (iii) meta-analytic structural equation modeling, which simultaneously tests relationships between several variables and compares the explanatory power of competing theory frameworks.

The meta-analytic method is so powerful and multifaceted that it lends itself to the pursuit of multiple goals. For example, it allows to synthesize research through increased statistical power and address new research questions not posed by individual studies, not to mention the chance to settle controversies arising from conflicting claims. However, meta-analysis also has the potential to foster another behavior infrequently performed by scholars, namely, the collection of large, cross-cultural samples to assess generalizability of findings. Clearly, the extent to which this goal can be accomplished depends on the data source, but this opportunity appears to be particularly remarkable when it comes to the meta-analysis of experimental evidence, which among other things is able to mitigate the problem of reverse causality typical of the observational studies investigating the economic effects of culture. Indeed, the latter have been object of insufficient academic efforts on the part of economists, who have started to devote attention to the topic only at the beginning of the 21\textsuperscript{st} century (Alesina and Giuliano, 2015; Castellani, 2019).

In sum, it is worth taking stock of the state of research using meta-analytic procedures in

\(^1\)The effect size is the summary statistic measuring between-group differences in the economic outcome under examination. For an enlightening, up-to-date guide to the meta-analytic method, we refer the reader to (Borenstein et al., 2021).
Economic Psychology, which usually hosts meta-analyses of experimental evidence. In doing so, we focus on the investigation of cross-cultural differences and make three distinct contributions to the literature. First, we locate meta-analyses in Economic Psychology through a systematic search and categorize them by data source, describing how the different categories relate to our goal, namely, the study of cross-cultural differences. Second, we discuss opportunities and challenges posed by the investigation of cross-cultural differences in the context of meta-analysis, proposing a sustainable methodological approach. Third, we implement this approach in three case studies where, on the basis of related literature, cultural values and heterogeneity are linked to three specific economic outcomes. More precisely, we explore the potential role of country-level individualism and fractionalization in explaining experimental measures of risk aversion, tax compliance, and prosocial behavior, respectively.

The remainder of the paper is organized as follows. Section 2 performs a literature review and deals with methodological issues, thereby shaping the first two abovementioned contributions. Section 3 addresses the third point and presents the three case studies. Section 4 discusses the results and concludes.

2 Literature review and methodology

To provide a fine-grained picture of the state of research using meta-analytic procedures in Economic Psychology, we construct a dataset that includes all the meta-analyses published in the following seven journals: Journal of Economic Psychology (JoEP), Experimental Economics (Exp Econ), Journal of Economic Behavior & Organization (JEBO), Journal of Behavioral Decision Making (JBDM), Judgment and Decision Making (JDM), Journal of the Economic Science Association (JESA), and Journal of Behavioral and Experimental Economics (JBEE). We choose this set of journals as a reasonably representative sample of outlets for behavioral decision researchers, and use the search engines on the respective websites to locate articles that contain at least one of the following terms: meta-analysis, meta-analytic, or meta-study. This search is conducted in September 2023 and produces lists of 271, 89, 456, 188, 594, 21, and 178 manuscripts in the seven archives, respectively. Importantly, we aim to select all the papers applying any type of meta-analytic technique.

As shown in Table 1, after screening abstract and text of the papers we are able to include 103 articles that, depending on the data source, can be sorted into four categories:

- traditional meta-analyses (e.g., Polman and Wu (2020); Umer (2023)), whose included
Table 1: Meta-analyses by journal

<table>
<thead>
<tr>
<th>Journal</th>
<th>Traditional</th>
<th>Crowd-sourced</th>
<th>Internal</th>
<th>Organizational</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>JoEP</td>
<td>10</td>
<td>0</td>
<td>8</td>
<td>0</td>
<td>18</td>
</tr>
<tr>
<td>Exp Econ</td>
<td>13</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>JEBO</td>
<td>7</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>JBDM</td>
<td>4</td>
<td>0</td>
<td>23</td>
<td>0</td>
<td>27</td>
</tr>
<tr>
<td>JDM</td>
<td>9</td>
<td>1</td>
<td>7</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>JESA</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>JBEE</td>
<td>8</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>56</strong></td>
<td><strong>1</strong></td>
<td><strong>44</strong></td>
<td><strong>2</strong></td>
<td><strong>103</strong></td>
</tr>
</tbody>
</table>

For each journal the columns report number of traditional, crowd-sourced, internal, and organizational meta-analyses.

studies stem from a systematic literature search. They aim to be as inclusive as possible and accordingly represent an all-purpose tool to pursue not only a wide array of goals, which encompass research synthesis through increased statistical power (i.e., oftentimes the central goal), support to the design of evidence-based policies and studies, extension of prior results by addressing research questions not posed by the included studies, settling of controversies arising from conflicting claims in the literature;

- meta-analyses of *crowd-sourced* data (e.g., Yilmaz and Alper (2019); Huber et al. (2023)), whereby a core coordination team of researchers invite multiple independent contributors to collaborate on selected research questions. This approach relies on the horizontal distribution of ownership, resources, and expertise, “enabling the conduct of large-scale research projects, democratizing who contributes to science, and assessing the robustness of findings” (Uhlmann et al., 2019);

- *internal* meta-analyses (e.g., Mulder et al. (2020); Arechar and Rand (2022)), by which one’s own studies within a manuscript or larger project are combined. Their peculiar goals are to summarize the project results through increased statistical power, as well as give credibility to null findings and foster the inclusion thereof when a paper consists of multiple studies. In addition, they also allow to extend prior results by addressing questions not posed by the included studies, and settle controversies arising from conflicting claims;\(^2\)

\(^2\)These meta-analyses are called mini (Goh et al., 2016) or single-paper meta-analyses (McShane and
• meta-analyses typical of the *organizational* research literature (e.g., Sinha (1994); Gooding et al. (1996)), which draw upon databases of financial information on global companies and aggregate results across multiple industries treated as studies. These meta-analyses aim to test theories relevant to the private sector and evaluate the generalizability of findings (Hunter and Schmidt, 2004).

Examining absolute frequencies, our sample of meta-analyses is rather homogeneous when it comes to data collection methods of the original studies, since 89 of the 103 papers meta-analyze experimental evidence. Also, 100 of the 103 papers implement either traditional or internal meta-analyses. Nevertheless, by their very nature the latter do not lend themselves to the investigation of how outcomes vary across geographic regions, since studies from the same research groups are likely to be conducted in the same countries. On the contrary, traditional meta-analyses aim to be as comprehensive as possible, and accordingly represent the natural tool to account for heterogeneity in terms of geographic differences. Meta-analyses of crowd-sourced data are also suitable for this purpose, but may not possess the strength of traditional meta-analyses when it comes to summarizing the literature. Similarly, the use of organizational meta-analyses to study cross-cultural differences, albeit possible, is restricted to the specific research questions characterizing organizational settings.

Of particular interest is Figure 1, which illustrates an overall upward trend in meta-analyses in Economic Psychology. As shown in histogram (1), the number of studies applying meta-analytic procedures has at least doubled since 2015, with this surge being only partially driven by the impactful primers of Goh et al. (2016) and McShane and Böckenholt (2017) on internal meta-analyses. Indeed, the green bars show a visible increase in traditional meta-analyses as well. On the contrary, just 25 of the 103 meta-analyses do not ignore *geographic differences* in the location of data collection, and are highlighted in red in histogram (2). What is more, in 19 of these 25 manuscripts the authors only control for geographic differences through dummy variables in meta-regressions and subgroup analyses, meaning that any related findings can hardly be interpreted and lots of research opportunities lie in this area. It is unsurprising that the remaining 6 papers out of 25 (Oosterbeek et al., 2004; Yilmaz and Alper, 2019; Cochard et al., 2021; Markowsky and Beblo, 2022; Im and Chen, 2022; Marini, 2023) mostly implement traditional meta-analyses, given the suitability of the latter for explaining heterogeneity in Böckenholt, 2017) when conclusively summarizing multiple studies in the same manuscript. They are instead called single-project meta-analyses (Corazzini and Marini, 2022) when self-contained and the included studies are published as separate papers.
In terms of geographic differences.\textsuperscript{3} This is definitely a desirable direction for future research, in light of the evidence that subject pools from Western, educated, industrialized, rich, and democratic (WEIRD) societies may not duly represent the global population with regard to biases, patterns, and preferences in economic decisions (Henrich et al., 2010). Whether these six papers apply AD or IPD meta-analyses, they share a common approach to the investigation of cross-country differences in economic outcomes, since they all use country-level continuous indicators of culture or economic development as covariates of meta-regressions.

Along with democracy, legal system, and other regulatory institutions, culture is one of the essential components characterizing countries, as well as an under-studied predictor of a wide array of economic outcomes (Castellani, 2019). As pointed out in Alesina and Gi-

\textsuperscript{3}In detail, Im and Chen (2022) is an internal meta-analysis, Yilmaz and Alper (2019) is a meta-analysis of crowd-sourced data, and the other four are traditional meta-analyses.
uliano (2015), the most widespread definition of culture in the empirical literature combines values and beliefs, referring to culture as “those customary beliefs and values that ethnic, religious, and social groups transmit fairly unchanged from generation to generation” (Guiso et al., 2006). The measurement of culture in terms of values and beliefs has mainly relied on survey questions, whose answers are usually aggregated at the country level (Alesina and Giuliano, 2015). However, when regressions relate these country-level variables to economic outcomes measured through observational data, the problem of reverse causality turns out to be cumbersome (Alesina and Giuliano, 2015; Castellani, 2019). On the other hand, experiments measuring trust, cooperation, altruism, or other constructs have also played a part in accounting for heterogeneous human behavior across societies (Henrich et al., 2001), in spite of the issue of external validity arising from the small convenience samples. For these reasons, the use of country-level cultural indicators in meta-analyses of experimental evidence appears to be an inviting intermediate solution in order to simultaneously mitigate the problems of reverse causality and limited external validity.

Cultural traits that have captured a great deal of interest in the empirical literature are generalized trust, individualism-collectivism, family ties and generalized morality, as well as attitudes toward work and the perception of poverty (Alesina and Giuliano, 2015; Castellani, 2019). In particular, the individualism-collectivism dimension has received the lion’s share of attention not only in Cultural Psychology (Oyserman et al., 2002; Taras et al., 2010) but also in Cultural Economics, where it has proved to be “the central cultural variable that matters for long-run growth” (Gorodnichenko and Roland, 2011). Although alternatives do exist (e.g., the framework by Schwartz (1992), and the GLOBE framework by House et al. (2004)), Alesina and Giuliano (2015) remark that the most popular measure of Individualism has been provided by Hofstede (2001), who defines it as “the degree to which people in a country prefer to act as individuals rather than as members of groups” (Hofstede, 1994). Surveying IBM employees worldwide through the Values Survey Module (VSM), Hofstede (2001) has then developed four country-level cultural dimensions via factor analysis (i.e., individualism-collectivism, power distance, uncertainty avoidance, and masculinity-femininity), adding two more dimensions at a later time (i.e., long-term orientation and indulgence-restraint (Hofstede et al., 2010)). Meta-analyzing nearly 600 empirical studies that explain various organizational outcomes by means of Hofstede’s framework, Taras et al. (2010) find that the predictive power of these cultural dimensions is equal (and sometimes superior) to that one of individual-level variables such as personality traits and demographic characteristics. The same authors also conclude that Hofstede’s framework “is actually more valid at the individual level, rather
than the nation level”. In sum, we view these as sound arguments for (i) the adoption of country-level cultural dimensions (e.g., Hofstede’s) in traditional meta-analyses, for which individual-level measures of cultural values are unobtainable by construction; (ii) the direct measurement of cultural values (e.g., via Hofstede’s VSM) when it comes to meta-studies of crowd-sourced data and internal meta-analyses.

At the same time, we also recognize that using country-level scores from previous research is a second-best solution for a few reasons. First, it assumes persistence of culture, but among scholars there is consensus that culture does change over time (Hofstede, 2001; Taras et al., 2010; Castellani, 2019), with the academic debate mostly focusing on the speed of cultural change. Nonetheless, this issue might be not that serious in light of evidence that “Hofstede’s cultural values do change but in parallel, implying that national cultural distances scores are relatively stable over time” (Beugelsdijk et al., 2017).

Second, it poses identification challenges when it comes to disentangling the effect of culture from other confounding factors. Clearly, the combination of experimental and non-experimental data does not possess the strength of purely experimental data with regard to the chance to make causal statements. However, in the context of meta-analysis, perpetuating the agnostic stance that to date has characterized the literature seems to be the only alternative. For this reason we do believe that this avenue should be pursued, for instance by simultaneously resorting to cultural variables and controls that measure the level of economic development in the country.

Third, it disregards within-country variation in cultural values, which can be extremely large. For example, country-level scores are somewhat uninformative with reference to India, which historically exhibits a number of heterogeneous regional subcultures due to its colonization by people from different societies (Dheer et al., 2015). Consequently, if the characteristics of the country do not reflect those of its inhabitants, projecting national culture onto individuals becomes a problematic practice called ecological fallacy (Brewer and Venaik, 2014). We envision two non-mutually exclusive ways to mitigate this issue. Whenever standard deviations of the country-level scores are not available (e.g., in Hofstede’s framework), the former approach involves the joint use of country-level cultural values and country-level measures of cultural heterogeneity, such as ethno-linguistic-religious Fractionalization à la Alesina et al. (2003). These measures of cultural heterogeneity are more than simple controls in meta-regressions, as they can be meaningfully linked to economic outcomes and answer additional research questions (e.g., Lago-Peñas and Lago-Peñas (2010); Andreoni et al. (2016)). The latter approach for blunting ecological fallacy entails the adoption of multilevel mixed-
effects models, in which individual-level observations are nested in higher-level clusters (e.g., countries and studies) that are treated as random factors. Indeed, as extensively discussed in Fernández-Castilla et al. (2020) hierarchical and non-hierarchical models are increasingly popular across disciplines in order to model dependence at multiple levels in meta-analysis. In this specific case, they acknowledge that individuals belonging to the same country may share some common contextual factors, such as cultural values.4

3 Case studies

We now present three case studies to illustrate the potential role of cross-cultural differences in explaining three specific economic outcomes: risk aversion (subsection 3.1), tax compliance (subsection 3.2), and prosocial behavior (subsection 3.3). The choice of the three case studies is based on individualism-related and fractionalization-related hypotheses that are described in the respective subsections. In any case, the three articles share some common methodological features: (a) they consist in published traditional IPD meta-analyses of experimental evidence; (b) they perform meta-regressions to account for between-study heterogeneity; (c) they do not delve into cultural diversity, at most controlling for the different study location through either country or continent fixed effects in meta-regressions.

Following up on the previous section, we utilize the same modus operandi across case studies. After obtaining the original datasets from the authors of the three meta-analyses, we code seven country-level cultural indices and two country-level proxies of economic development as additional regressors, and re-analyze the data. Out of these nine variables, the two of interest are:

• Individualism from the framework of Hofstede (2001), which can potentially range from 0 to 100. The higher (lower) this index, the more individualist (collectivist) the country where the experiment takes place. Individualism is defined as “the degree to which people in a country prefer to act as individuals rather than as members of groups” (Hofstede, 1994);

• ethno-linguistic-religious Fractionalization, namely, the probability that any two citizens of a country belong to a different ethno-linguistic-religious group. In line with other meta-analyses (Lane, 2016; Marini, 2023, 2022; Chowdhury and Marini, 2023), this index 4For example, in our dataset multilevel models with country as a random factor are used in the meta-analysis by Im and Chen (2022).
is constructed by averaging each country’s score for the three separate measures of ethnic, linguistic, and religious fractionalization provided by Alesina et al. (2003).

The other seven variables simply serve as controls in our re-analysis. Five of these variables are the remaining cultural dimensions from Hofstede’s framework: \textit{Power distance}, \textit{Uncertainty avoidance}, \textit{Masculinity}, \textit{Long-term orientation}, and \textit{Indulgence}. The remaining two variables are the \textit{Ease of doing business} indicator of the World Bank and the \textit{Account penetration} rate of financial inclusion of the Global Findex Database, already used as proxies of economic development in the meta-study by Cochard et al. (2021).\footnote{The following cultural indices can vary from 0 to 100: (a) \textit{Power distance}, defined as the extent to which people in a country accept hierarchical differences; (b) \textit{Uncertainty avoidance}, defined as the extent to which people in a country feel threatened by uncertain situations; (c) \textit{Masculinity}, defined as societal preference for achievement, heroism, assertiveness, and material rewards for success; (d) \textit{Long-term orientation}, which characterizes those societies encouraging future-oriented values such as thrift and persistence; (e) \textit{Indulgence}, defined as the extent to which people in a country try to control their desires and impulses. The \textit{Ease of doing business} index ranks countries from 1 to 189, where lower values refer to the most business-friendly economies. Finally, the \textit{Account penetration} rate indicates the percentage of the population aged 15 and over who possess a bank account at a financial institution.}

\subsection{Culture and risk aversion}

Our first case study is the paper “A reconsideration of gender differences in risk attitudes” by Filippin and Crosetto (2016) published in \textit{Management Science}. We therefore focus on risk aversion as first economic outcome of interest, given the existence of two opposed hypotheses linking individualism to risk attitude: the \textit{cushion hypothesis} and the \textit{tough guy hypothesis}. The former predicts that people in collectivist countries are more risk-seeking than those in individualist countries, since their relatively more extended social network could cushion them financially in case of unfavorable events (Hsee and Weber, 1999; Illiashenko, 2019; Illiashenko and Laidroo, 2020). Conversely, the latter posits that individualist societies reward people for personal success and accordingly lead them to take relatively more risks (Breuer et al., 2014; Ashraf et al., 2016; Gaganis et al., 2019; Frijns et al., 2022).

The meta-study by Filippin and Crosetto (2016) investigates gender differences in risk attitude measured through a Multiple Price List (MPL) à la Holt and Laury (2002), which is by far the most popular risk elicitation method in Economics. Table 2 depicts the original version of the task, where subjects are supposed to make multiple choices between paired lotteries. Option A has lower variance than Option B, and always represents the safe choice. While the set of outcomes is unchanged across Problems, the two lotteries are placed in
increasing order of expected value so that, from Problem 5 onwards, a risk-neutral person should switch from Option A to Option B. The higher the *Number of safe choices*, the more risk-averse the subject. Also, switching from Option B to Option A is regarded as inconsistent behavior in the same way as never choosing Option B, since in the last row of the table Option A is strictly dominated. In relation to expected utility theory, inconsistent behaviors are commonly treated as computational errors and have to be handled through structural models that include a stochastic component.

Table 2: Original MPL in Holt and Laury (2002)

<table>
<thead>
<tr>
<th>Problem</th>
<th>Option A</th>
<th>Option B</th>
<th>Exp. payoff difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1/10 of $2.00, 9/10 of $1.60</td>
<td>1/10 of $3.85, 9/10 of $0.10</td>
<td>$1.17</td>
</tr>
<tr>
<td>2</td>
<td>2/10 of $2.00, 8/10 of $1.60</td>
<td>2/10 of $3.85, 8/10 of $0.10</td>
<td>$0.83</td>
</tr>
<tr>
<td>3</td>
<td>3/10 of $2.00, 7/10 of $1.60</td>
<td>3/10 of $3.85, 7/10 of $0.10</td>
<td>$0.50</td>
</tr>
<tr>
<td>4</td>
<td>4/10 of $2.00, 6/10 of $1.60</td>
<td>4/10 of $3.85, 6/10 of $0.10</td>
<td>$0.16</td>
</tr>
<tr>
<td>5</td>
<td>5/10 of $2.00, 5/10 of $1.60</td>
<td>5/10 of $3.85, 5/10 of $0.10</td>
<td>-$0.18</td>
</tr>
<tr>
<td>6</td>
<td>6/10 of $2.00, 4/10 of $1.60</td>
<td>6/10 of $3.85, 4/10 of $0.10</td>
<td>-$0.51</td>
</tr>
<tr>
<td>7</td>
<td>7/10 of $2.00, 3/10 of $1.60</td>
<td>7/10 of $3.85, 3/10 of $0.10</td>
<td>-$0.85</td>
</tr>
<tr>
<td>8</td>
<td>8/10 of $2.00, 2/10 of $1.60</td>
<td>8/10 of $3.85, 2/10 of $0.10</td>
<td>-$1.18</td>
</tr>
<tr>
<td>9</td>
<td>9/10 of $2.00, 1/10 of $1.60</td>
<td>9/10 of $3.85, 1/10 of $0.10</td>
<td>-$1.52</td>
</tr>
<tr>
<td>10</td>
<td>10/10 of $2.00, 0/10 of $1.60</td>
<td>10/10 of $3.85, 0/10 of $0.10</td>
<td>-$1.85</td>
</tr>
</tbody>
</table>

Filippin and Crosetto (2016) implement a three-part meta-analysis that first evaluates effect sizes separately paper by paper, then performs IPD meta-regressions, and finally relies on structural equation modeling. In conclusion, the authors find that women are significantly more risk-averse than men, even though the overall effect size turns out to be small. It is worth noting that, to better suit the different statistical techniques, the three parts of their meta-analysis are based on different sets of studies. For the purposes of our paper we consider the part of the analysis exploiting IPD meta-regressions (Table 9 of their article), which include 5,807 datapoints from 48 studies conducted in 16 countries. These observations originate from consistent subjects that face paired lotteries with increasing probabilities and fixed outcomes. Indeed, on the one hand the meta-study by Filippin and Crosetto (2016) completely leaves out tasks using an outcome scale and fixed probabilities, on the other hand inconsistent behaviors are modeled only at a later stage via structural model estimation. Nevertheless, their meta-study also features modified versions of the task with a number of Problems different from 10, adjusting the *Number of safe choices* accordingly.6

---

6For example, in case one makes 15 safe choices in a 20-row table, Filippin and Crosetto (2016) assign the
After receiving our data request in September 2022, Filippin and Crosetto (2016) shared their dataset and the script to reproduce the meta-regressions displayed in Table 9 of their article. Given that culture is not examined by Filippin and Crosetto (2016), we match the included studies with the countries where they were conducted, and then code the nine aforementioned country-level variables. This step leads to the exclusion of 11 observations from Israel (Yechiam and Hochman, 2013) since no Indulgence score is available for this country, resulting in the final sample of 5,796 datapoints from 47 studies and 15 countries.\(^7\) LOWESS smoother (a) in Figure 2 provides some initial insights into the relationship between individualism and risk aversion, suggesting that it may be non-linear.\(^8\) Also, it is worth mentioning that the great bulk of the observations (91.8%) stem from individualist countries, for which the Individualism score is greater than 50.

Next, in Table 3 we conduct IPD meta-regressions with the Number of safe choices as dependent variable. Column (1) faithfully reproduces the second specification in Table 9 of Filippin and Crosetto (2016), namely, the authors’ “preferred specification”. This model performs an OLS estimation with robust standard errors and finds gender differences in risk attitude, with women being significantly more risk-averse than men \((p < 0.001)\). In column (2) we extend the same linear regression model by simply adding the nine country-level variables. In doing so, along with an increase in the Adjusted R-squared from 1.9% to 3.1% we observe that the coefficient of Individualism is not significant \((p = 0.619)\). However, building on the insights from Figure 2, in specification (3) we include the squared term of Individualism. At this point, the coefficient of the linear term appears to be positive and significant \((p < 0.001)\), with the negative coefficient of the squared term unveiling an inverted U-shaped relationship between individualism and risk aversion \((p < 0.001)\). These results provide mixed evidence in relation to the two underlying hypotheses. As a robustness check, column (4) implements a multilevel mixed-effects (MME) model with standard errors clustered at both the study and the country level. Note that we use a non-hierarchical model with crossed random effects (also known as cross-classified random-effects model (Fernández-Castilla et al., 2020)) because the two random factors are not nested within each other, but rather crossed.\(^9\) Also, the slope value 7.5 to the variable Number of safe choices. In sum, the latter variable can assume values between 0 and 9, given that consistent behavior involves choosing Option B at Problem 10.

\(^7\)The 15 included countries are listed in Figure A1 in Appendix A, where a horizontal bar chart displays the average Number of safe choices by country.

\(^8\)Note that, differently from the other two case studies, here fractionalization serves just as a control and we therefore refrain from commenting on it.

\(^9\)Indeed, studies are not nested within countries since one study (Grijalva et al., 2011) administers the MPL in multiple countries, nor are countries nested within studies because seven countries (USA, Canada, France,
Figure 2: Relationship between cultural variables and risk aversion

LOWESS smoothers - Risk aversion

(a) Observations = 5,796

(b) Observations = 5,796

on Individualism is allowed to vary between studies in order to reflect the idea that some studies might endorse the cushion hypothesis, while other studies might support the tough guy hypothesis. This is really the case, as revealed by a conservative LR test comparing the model without the random slope against the model with the random slope. Indeed, the p-value of the LR test amounts to 0.001, indicating that the slope on Individualism does vary between studies. Since the inverted U-shaped relationship between individualism and risk aversion is confirmed by the MME model ($p = 0.005$ and $p = 0.009$ for Individualism and its square, respectively), this is further evidence of between-study heterogeneity in findings on the role of culture. LOWESS smoother (a) in Figure 2 leads us to conclude that these results are driven by individualist countries, which provide the large majority of observations (91.8%).

Germany, Netherlands, Spain, Denmark) host multiple studies.

The test is conservative given that the standard deviation of the random slope under the null hypothesis is equal to 0, and therefore on the boundary of the parameter space.
Result 1. There is an inverted U-shaped relationship between individualism and risk aversion. Both the cushion hypothesis and the tough guy hypothesis are supported.

### Table 3: Explaining risk aversion through cultural values

<table>
<thead>
<tr>
<th>Dependent variable: Number of safe choices</th>
<th>(1) Filippin and Crosetto (OLS)</th>
<th>(2) Extended OLS model</th>
<th>(3) Extended OLS model with squared terms</th>
<th>(4) Extended MME model with squared terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>0.326*** (0.050)</td>
<td>0.318*** (0.049)</td>
<td>0.333*** (0.049)</td>
<td>0.304*** (0.048)</td>
</tr>
<tr>
<td>Realmoney</td>
<td>0.013*** (0.002)</td>
<td>0.021*** (0.003)</td>
<td>0.017*** (0.004)</td>
<td>0.005 (0.007)</td>
</tr>
<tr>
<td>Realmoney / 100</td>
<td>-0.004*** (0.001)</td>
<td>-0.008*** (0.001)</td>
<td>-0.010*** (0.001)</td>
<td>-0.004 (0.003)</td>
</tr>
<tr>
<td>Exchange / 100</td>
<td>0.011 (0.009)</td>
<td>0.049*** (0.012)</td>
<td>0.104*** (0.018)</td>
<td>0.079*** (0.022)</td>
</tr>
<tr>
<td>Randomorder</td>
<td>0.360** (0.128)</td>
<td>0.856*** (0.209)</td>
<td>0.947*** (0.210)</td>
<td>1.407** (0.412)</td>
</tr>
<tr>
<td>Individualism</td>
<td>0.002 (0.004)</td>
<td>0.156*** (0.032)</td>
<td>0.147** (0.052)</td>
<td></td>
</tr>
<tr>
<td>Individualism^2</td>
<td>-14.1e^{-4***} (2.9e^{-4})</td>
<td>-12.6e^{-4**} (4.8e^{-4})</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fractionalization</td>
<td>-0.866 (0.516)</td>
<td>-2.249 (2.807)</td>
<td>3.811 (4.867)</td>
<td></td>
</tr>
<tr>
<td>Fractionalization^2</td>
<td>-0.999 (3.366)</td>
<td>-8.950 (5.673)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>5.303*** (0.039)</td>
<td>3.567*** (0.913)</td>
<td>1.613*** (2.643)</td>
<td>-1.255 (3.851)</td>
</tr>
</tbody>
</table>

(1), (2), and (3): coefficient estimates from OLS regression models, with robust standard errors in parentheses. (4): coefficient estimates from multilevel mixed-effects (MME) model, with standard errors clustered at both the study and the country level in parentheses. This is a non-hierarchical model with crossed random effects, where the slope on Individualism is allowed to vary between studies. In models (2), (3), and (4) the 11 observations from Israel are removed since no Indulgence score is available for this country. The label “Cultural & ED controls” includes Power distance, Uncertainty avoidance, Masculinity, Long-term orientation, Indulgence, Ease of doing business, and Account penetration.

*** p-value < 0.001.
** p-value < 0.01.
* p-value < 0.05.
3.2 Culture and tax compliance

Our second case study is the paper “40 years of tax evasion games: a meta-analysis” by Alm and Malézieux (2021) published in Experimental Economics. We therefore focus on tax compliance as second economic outcome of interest, given the existence of two opposed theories linking individualism to tax compliance. One theory claims that collectivist societies’ concern for the in-group is so pervasive that it can override written laws, which in these countries are considered as group-specific instead of universal (Tsakumis et al., 2007; Richardson, 2008). For this reason, individualist countries end up with more powerful economies as compared with collectivist countries, as well as with stricter regulatory systems that discourage tax evasion. In a few words, people in individualist countries are described as more compliant than those in collectivist countries. Conversely, the institutional anomie theory argues that individualism causes the erosion of moral codes because of its focus on gaining materialistic rewards at the expense of ethical considerations (Messner and Rosenfeld, 2001; Bame-Aldred et al., 2013). Hence, according to this theory people in individualist countries are expected to be less compliant and more prone to deviant behaviors such as tax evasion. Furthermore, there is well-established evidence for a negative link between culturally fractionalized societies and tax compliance (Lassen, 2007; Lago-Peñas and Lago-Peñas, 2010). The most popular explanation is that cultural heterogeneity is negatively associated with trust and, if people do not trust others to comply with tax laws, they prefer to evade taxes themselves. A low level of trust also implies inefficient public goods provision, and “according to the quid pro quo argument, tax compliance depends in part on taxpayers receiving sufficient public goods in return for taxes paid”.

The meta-study by Alm and Malézieux (2021) examines the impact of public policy, experimental design, and individual-level variables on tax evasion choices elicited through a Tax Evasion Game (TEG) à la Friedland et al. (1978), arguably the most common method to measure tax compliance in the lab. The original version of the game is built on the economic model of taxpayer behavior of Allingham and Sandmo (1972) and works as follows. After being endowed with a fixed amount of income, in each round of the game a subject has to decide how much of it to report to tax authorities. Differently from underreported income, on which no taxes are paid, reported income is taxed at a certain rate. Also, the subject has to face the risk of being audited with fixed probability, and consequently fined at a fixed rate on each experimental currency unit that was underreported. In sum, the individual rationally weighs the expected pros and cons of cheating, thus maximizing the expected utility of the tax
evasion gamble. The model and the TEG have been updated over the years to incorporate additional variables expected to account for tax compliance, such as ethical considerations (Eisenhauer et al., 2011) and the redistribution of taxes to participants.

To avoid dropping the great bulk of the observations due to missing values and multicollinearity problems, Alm and Malézieux (2021) conduct separate IPD meta-regressions of tax compliance on public policy, experimental design, and individual-level variables, further discriminating between extensive and intensive margin of compliance. Among other things, the authors find ubiquitous negative impact of flat-rate tax systems, positive impact of tax redistribution and investment of taxes in real-life public goods, as well as gender differences, with women evading less than men. Therefore, also in this case the various models are based on different sets of studies. For the sake of comparison with the other case studies, here we consider the OLS estimation performed by Alm and Malézieux (2021) in Table 16 of their Appendix, where the extensive margin is not distinguished from the intensive margin of compliance, and the dependent variable is the Compliance rate ranging from 0 to 1.\footnote{The compliance rate is defined as the ratio between reported income and total income.} In particular, we focus on the fourth specification of their table (i.e., the analysis of the effects of public policy regressors) because it includes the highest number of observations (163,123), datasets (48), and countries (15).\footnote{Note that, in this case study, one dataset is exploited by six articles.} After receiving our data request, Alm and Malézieux (2021) provided a list of contributors to whom we then sent authorization requests to reuse data, as the latter were not publicly available in most cases. The collection of authorizations started in February 2023 and ended in March 2023. After three reminders all the contributors gave express consent, so that Alm and Malézieux (2021) shared the whole meta-dataset with us in March 2023. The included datapoints are already matched with the respective countries, since in their meta-regressions Alm and Malézieux (2021) use country fixed effects to control for the different study location. Although the choice of evading taxes can be mostly viewed as individual decision making, this is not the case for a noticeable portion of the included TEGs (e.g., 21.6\% of the considered data stem from TEGs where taxes are redistributed). To obtain the purest evidence, we therefore decide to re-analyze only the observations from Round 1 of the TEGs, which amount to 11,101.\footnote{This decision is also based on computational reasons in the case of the crossed-effects model, which we could not fit otherwise. Also, the 15 included countries are listed in Figure A2 in Appendix A, where a horizontal bar chart displays the average Compliance rate by country.}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure3}
\caption{LOWESS smoother (a) in Figure 3 anticipates that individualism may be negatively related to tax compliance, with the large majority of the datapoints (92.9\%) referring to individualist countries. The sample exhibits low variability}
\end{figure}
with regard to fractionalization too, as 85.9% of the data are gathered in relatively less fractionalized countries, where the Fractionalization score is lower than 0.5 and the relationship between fractionalization and tax compliance appears to be non-linear.

Figure 3: Relationship between cultural variables and tax compliance

Table 4 shows IPD meta-regressions with the Compliance rate as dependent variable. In column (1) we reproduce the fourth specification in Table 16 of Alm and Malézieux (2021), which estimates the impact of public policy variables by means of a linear regression model with standard errors clustered at the individual level. This model is equipped with round, country, study, and year fixed effects to control for possible confounding factors, and among other things it finds negative effects of amnesties, flat tax systems, and tax rate on compliance ($p < 0.001$ in all three cases). Specification (2) includes only the observations from Round 1 and adds the two cultural variables of interest, thus extending the previous model. Also, country fixed effects are removed. It turns out that participants from individualist countries
are significantly less compliant than those from collectivist countries ($p < 0.001$), as indicated by the negative coefficient of Individualism. Similarly, there is evidence that Fractionalization significantly decreases the compliance rate ($p < 0.001$). Nevertheless, this result is not robust to the inclusion of controls in model (3), where on the contrary the coefficient of Fractionalization is positive and the related test misses significance ($p = 0.051$). As the relationship between fractionalization and tax compliance appears to be non-linear, it would be appropriate to use the square of Fractionalization as additional regressor to shed more light, but unfortunately this is not feasible because of multicollinearity problems. Instead, individualism continues to be associated with significantly lower levels of tax compliance in both column (3) and column (4) ($p = 0.001$ in both cases), where we perform a non-hierarchical MME model with standard errors clustered at both the study and the country level. Our findings therefore endorse the institutional anomie theory. Also, a conservative LR test comparing the MME model against the linear model proves that the clustering adds no value to the analysis ($p = 1.000$).

**Result 2.** There is evidence of a negative link between individualism and tax compliance.

### 3.3 Culture and prosocial behavior

Our third case study is the paper “Are women more generous than men? A meta-analysis” by Bilén et al. (2021) published in the *Journal of the Economic Science Association*. We therefore focus on prosocial behavior as third economic outcome of interest, given the existence of mixed findings linking individualism to prosocial behavior. On the one hand, there is evidence that collectivism positively relates to formal volunteering, in line with the hypotheses that “societal collectivism is synonymous with actions that are directed to maximize the well-being of society” (Parboteeah et al., 2004), and individualism is associated with the pursuit of self-interest rather than group interest (Earley, 1989). On the other hand, the most recent school of thought points out that “individuals behave in a prosocial manner because it serves their own purposes” (Luria et al., 2015), and such self-serving behavior arises mostly in individualist countries. A positive impact of individualism on self-interested giving, which is commonly

---

14 The findings related to cultural variables remain unchanged if we run linear regressions including data from all rounds. These robustness checks are available upon request.

19
Table 4: Explaining tax compliance through cultural values

<table>
<thead>
<tr>
<th></th>
<th>(1) Alm and Malézieux (OLS)</th>
<th>(2) Extended OLS model</th>
<th>(3) Extended OLS model with controls</th>
<th>(4) Extended MME model with controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random audit</td>
<td>-0.008</td>
<td>-0.137**</td>
<td>-0.135**</td>
<td>-0.135**</td>
</tr>
<tr>
<td></td>
<td>(0.050)</td>
<td>(0.047)</td>
<td>(0.047)</td>
<td>(0.049)</td>
</tr>
<tr>
<td>Audit probability</td>
<td>0.002</td>
<td>-0.237***</td>
<td>-0.245***</td>
<td>-0.245**</td>
</tr>
<tr>
<td></td>
<td>(0.034)</td>
<td>(0.067)</td>
<td>(0.067)</td>
<td>(0.086)</td>
</tr>
<tr>
<td>Fine size</td>
<td>0.006</td>
<td>0.006</td>
<td>0.008</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.014)</td>
<td>(0.014)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>Audit * Fine</td>
<td>-0.233***</td>
<td>-0.255***</td>
<td>-0.241***</td>
<td>-0.241**</td>
</tr>
<tr>
<td></td>
<td>(0.033)</td>
<td>(0.064)</td>
<td>(0.064)</td>
<td>(0.076)</td>
</tr>
<tr>
<td>Amnesty</td>
<td>-0.313***</td>
<td>-0.337**</td>
<td>-0.337</td>
<td>-0.337</td>
</tr>
<tr>
<td></td>
<td>(0.029)</td>
<td>(0.121)</td>
<td>(0.121)</td>
<td>(0.226)</td>
</tr>
<tr>
<td>Flat tax</td>
<td>-0.127***</td>
<td>-0.195*</td>
<td>-0.195*</td>
<td>-0.195*</td>
</tr>
<tr>
<td></td>
<td>(0.025)</td>
<td>(0.079)</td>
<td>(0.079)</td>
<td>(0.088)</td>
</tr>
<tr>
<td>Tax rate</td>
<td>-0.175***</td>
<td>-0.406**</td>
<td>-0.403**</td>
<td>-0.403**</td>
</tr>
<tr>
<td></td>
<td>(0.032)</td>
<td>(0.142)</td>
<td>(0.142)</td>
<td>(0.160)</td>
</tr>
<tr>
<td>Individualism</td>
<td>-29.6e^{-4}***</td>
<td>-0.003**</td>
<td>-0.003**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.4e^{-2})</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>Fractionalization</td>
<td>-0.274***</td>
<td>0.236</td>
<td>0.236</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.049)</td>
<td>(0.121)</td>
<td>(0.126)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.327***</td>
<td>1.420***</td>
<td>0.135</td>
<td>0.135</td>
</tr>
<tr>
<td></td>
<td>(0.069)</td>
<td>(0.105)</td>
<td>(0.451)</td>
<td>(0.428)</td>
</tr>
</tbody>
</table>

Cultural & ED controls    No                          No                          Yes                          Yes
Round FE                  Yes                         No                          No                          No
Country FE                Yes                         No                          No                          No
Study FE                  Yes                         Yes                         Yes                         Yes
Year FE                   Yes                         Yes                         Yes                         Yes
R-squared (%)             8.836                       9.437                      9.723                      -
Adj. R-squared (%)        8.769                       8.978                      9.233                      -
LR $\chi^2$ vs. linear model -                          -                          1.8e^{-12}
No. of observations       163,123                     11,101                     11,101                     11,101

(1), (2), and (3): coefficient estimates from OLS regression models, with standard errors clustered at the individual level in parentheses. (4): coefficient estimates from multilevel mixed-effects (MME) model, with standard errors clustered at both the study and the country level in parentheses. This is a non-hierarchical model with crossed random effects, where the slope on Individualism is allowed to vary between studies. In models (2), (3), and (4) we include only the observations from Round 1. The label “Cultural & ED controls” includes Power distance, Uncertainty avoidance, Masculinity, Long-term orientation, Indulgence, Ease of doing business, and Account penetration.

*** $p$-value < 0.001.
** $p$-value < 0.01.
* $p$-value < 0.05.
known as warm-glow giving (Andreoni, 1990), is thus expected to take place through the moderation of social norms (Cai et al., 2022). Hence, “individualism can bring a community together rather than dividing it” (Kemmelmeier et al., 2006). Moreover, there is solid evidence for a negative link between culturally fractionalized societies and charitable donations (Okten and Osili, 2004; Andreoni et al., 2016). One possible explanation is that cultural heterogeneity implies different tastes and agendas between groups when it comes to allocating private and public funds, thereby decreasing fundraising and contributions to specific recipients.

The meta-study by Bilén et al. (2021) investigates gender differences in prosocial behavior measured through a Dictator Game (DG) à la Forsythe et al. (1994), where altruistic reasons are known to play a role (Chai et al., 2011). In the standard DG, a subject serving as the dictator is supposed to decide how to split a windfall endowment with another individual that acts as a recipient. In a different version of the DG, the role of the recipient is played by a charity. Bilén et al. (2021) meta-analyze both versions of the DG, including studies with monetary incentives where (i) participants play anonymously, (ii) dictators do not have the option to take money from recipients, and (iii) the price of giving is equal to 1. The authors carry out IPD meta-regressions of the Share donated by the dictator on a number of regressors that include a gender dummy, thereby finding that women on average give more than do men in both the standard and the charity DG. For the purposes of our paper we consider the OLS estimation performed in Table 3 of their article, which is based on 15,016 observations from 53 studies.

After receiving our data request, Bilén et al. (2021) provided a list of contributors to whom we then sent authorization requests to reuse data, as the latter were not publicly available in most cases. The collection of authorizations started in June 2023 and ended in July 2023. After two reminders all the contributors gave express consent except for Bachke et al. (2017), so that Bilén et al. (2021) shared their meta-dataset with us in July 2023 by removing the 189 observations from Bachke et al. (2017). The included datapoints are already matched with the respective countries, since in their meta-regressions Bilén et al. (2021) use continent fixed effects to control for the different study location. 10 of the included experimental studies are conducted online with the chance to recruit participants from multiple countries, which consequently leads to a highly fragmented meta-dataset from the viewpoint of nationality. For example, the study by Dreber et al. (2013) per se features observations from at least 38 countries, and only two of these countries are associated with more than 10 observations. We therefore decide to re-analyze only the data from countries associated with more than 30 observations. Moreover, our re-analysis drops 45 observations from online experiments for
which information about the country of origin is missing, thus finally considering a restricted dataset of 14,689 datapoints from 52 studies and 19 countries.\textsuperscript{15} LOWESS smoother (a) in Figure 4 suggests that the relationship between individualism and prosocial behavior may be non-linear, especially with regard to the portion of observations coming from individualist countries (79.9%). Conversely, LOWESS smoother (b) anticipates that fractionalization may be negatively related to prosocial behavior, with highly and lowly fractionalized countries being equally represented (49.4% vs. 50.6%, respectively).

Figure 4: Relationship between cultural variables and prosocial behavior

Table 5 shows IPD meta-regressions with the \textit{Share donated} as dependent variable. In column (1) we reproduce the fourth specification in Table 3 of Bilén et al. (2021), which estimates the gender gap in prosocial behavior by means of a linear regression model with

\textsuperscript{15}The 19 included countries are listed in Figure A3 in Appendix A, where a horizontal bar chart displays the average \textit{Share donated} by country.
standard errors clustered at the condition level. In addition to other controls, this model is equipped with continent fixed effects (i.e., Region dummies, according to the original wording) and finds gender differences in prosocial behavior, with women’s donations being significantly higher than men’s donations ($p = 0.001$). Specification (2) simply considers the aforementioned restricted dataset and removes continent fixed effects, showing that the results are unchanged. In column (3) we extend the previous model by adding country-level Individualism and Fractionalization, alongside the other country-level variables measuring cultural values and economic development.\footnote{Please note that neither Long-term orientation nor Indulgence scores are available for Sierra Leone. Consequently, in Table 5 we decide to drop these two controls rather than the 930 observations from this non-WEIRD country.} Participants from individualist countries donate a significantly lower share of their endowment ($p < 0.001$), as the related regression coefficient turns out to be negative. However, the coefficient of Fractionalization is a long way off from statistical significance ($p = 0.312$) and accordingly we are unable to support the initial insights from LOWESS smoother (b). These results are robust to the use of a non-hierarchical MME model with standard errors clustered at both the study and the country level in column (4) ($p = 0.001$ and $p = 0.100$, respectively), where the slope on Individualism is allowed to vary between studies in order to reflect previous mixed evidence. In this regard, we use a conservative LR test comparing the model without the random slope against the model with the random slope, and are able to reject the null hypothesis that the between-study variation is zero ($p = 0.041$). Hence, since the point estimate of the effect of Individualism is significantly negative (-0.004), we are led to conclude that the majority of studies find a negative relationship between individualism and prosociality, but a few studies also support the idea that donations are higher in individualist countries out of warm glow.

**Result 3.** While the majority of studies find a negative relationship between individualism and prosocial behavior, a few studies also provide opposite findings.

### 4 Discussion and conclusions

This manuscript is a methodological work on the state of research using meta-analytic procedures in Economic Psychology, with a focus on the investigation of cross-cultural differences. We review published meta-analyses and introduce a new classification thereof by data source,
Table 5: Explaining prosocial behavior through cultural values

<table>
<thead>
<tr>
<th></th>
<th>(1) Bilén et al. (OLS)</th>
<th>(2) Bilén et al. (OLS) restricted dataset</th>
<th>(3) Extended OLS model, restricted dataset</th>
<th>(4) Extended MME model, restricted dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>0.020**</td>
<td>0.022***</td>
<td>0.022**</td>
<td>0.021***</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.006)</td>
<td>(0.006)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Charity DG</td>
<td>0.152***</td>
<td>0.166***</td>
<td>0.142***</td>
<td>0.128***</td>
</tr>
<tr>
<td></td>
<td>(0.034)</td>
<td>(0.038)</td>
<td>(0.028)</td>
<td>(0.030)</td>
</tr>
<tr>
<td>Charity DG * Female</td>
<td>0.099***</td>
<td>0.090***</td>
<td>0.097***</td>
<td>0.098***</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.017)</td>
<td>(0.017)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>Individualism</td>
<td></td>
<td>-0.004***</td>
<td>-0.004**</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>Fractionalization</td>
<td></td>
<td>0.084</td>
<td>0.168</td>
<td>0.083</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.102)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.491***</td>
<td>0.245***</td>
<td>1.101***</td>
<td>1.107***</td>
</tr>
<tr>
<td></td>
<td>(0.054)</td>
<td>(0.017)</td>
<td>(0.114)</td>
<td>(0.166)</td>
</tr>
</tbody>
</table>

| Cultural & ED controls | No | No | Yes | Yes |
| Treatment controls     | Yes| Yes| Yes | Yes |
| Continent FE           | Yes| No | No  | No  |
| R-squared (%)          | 12.556 | 10.439 | 13.796 | - |
| Adj. R-squared (%)     | 12.414 | 10.336 | 13.655 | - |
| LR \(\chi^2\) vs. no random slope | - | - | - | 4.180* |

No. of observations 14,827 14,689 14,689 14,689

(1), (2), and (3): coefficient estimates from OLS regression models, with standard errors clustered at the condition level in parentheses. (4): coefficient estimates from multilevel mixed-effects (MME) model, with standard errors clustered at both the study and the country level in parentheses. This is a non-hierarchical model with crossed random effects, where the slope on Individualism is allowed to vary between studies. The label “Cultural & ED controls” includes Power distance, Uncertainty avoidance, Masculinity, Ease of doing business, and Account penetration. The label “Individual controls” includes Student characteristics and Age. The label “Treatment controls” includes Double-blind, Setting characteristics, Random payment, and Partitioning of endowment.

*** p-value < 0.001.
** p-value < 0.01.
* p-value < 0.05.
describing how the different categories relate to the study of cross-cultural differences. We also
discuss related opportunities and challenges, proposing a sustainable methodological approach
that is then implemented in three case studies.

After resorting to the search engines of seven representative field journals, we are able to
locate more than a hundred meta-analyses that can be sorted into four categories: traditional,
crowd-sourced, internal, and organizational meta-analyses. The four types of meta-analyses
are associated with different goals. Among the four types, we consider traditional meta-
analyses as the natural tool to account for heterogeneity in terms of geographic differences.
Meta-analyses of crowd-sourced data are also suitable for this purpose, but not yet popular
in the experimental community. We identify lower potential, instead, with respect to internal
and organizational meta-analyses. The great bulk of meta-studies consist in traditional and
internal meta-analyses, whose yearly number has rapidly increased in recent times. Neverthe-
less, we are unable to detect a similar increase in the number of meta-analyses exploring the
role of geographic differences in the location of data collection. We speculate that scholars
refrain from combining experimental and non-experimental data because of the identification
challenges posed by this approach, when it comes to disentangling the effect of culture from
other confounding factors. However, in the context of meta-analysis, perpetuating the agnostic
stance that to date has characterized the literature seems to be the only alternative. For this
reason we do believe that this approach should be pursued through meta-regression analysis.
In particular, we recommend (a) the adoption of country-level cultural dimensions in tradi-
tional meta-analyses, for which individual-level measures of cultural values are unobtainable
by construction; (b) the direct measurement of cultural values when it comes to meta-studies
of crowd-sourced data and internal meta-analyses. In this regard we identify some advant-
ages related to the use of country-level cultural variables in meta-analyses of experimental
evidence: (i) it is an intermediate solution in order to simultaneously mitigate the problems
of reverse causality and limited external validity, which affect the empirical studies based
on purely observational and experimental data, respectively; (ii) it allows to answer culture-
related research questions not posed by the included studies, which are generally conducted
in a single country; (iii) it is an opportunity to settle controversies arising from conflicting
claims, such as those about the effect of individualism on certain economic outcomes. We
also discuss the risk of ecological fallacy as main limitation, thus recommending the use of
country-level measures of cultural heterogeneity and the adoption of multilevel mixed-effects
models.

At the same time, we make the case that this approach should be backed by relevant
literature, and we therefore choose three case studies on the basis of hypotheses that link cultural values and heterogeneity to three specific economic outcomes: risk aversion, tax compliance, and prosocial behavior. In each case study, after obtaining the dataset of a published IPD meta-analysis that did not delve into cultural diversity, we code country-level individualism and fractionalization in addition to other controls, and re-analyze the data by means of OLS estimation and non-hierarchical multilevel models. First, we find an inverted U-shaped relationship between individualism and risk aversion that on the one hand supports the tough guy hypothesis in relatively more individualist countries, but on the other hand tallies with the cushion hypothesis in relatively less individualist countries. Given that the latter are usually non-WEIRD countries where field instead of lab experiments are conducted, one possible explanation for this result pertains to differences in risk-taking behavior between student subjects and more experienced participants (Huber and König-Kersting, 2022). Second, there is evidence of a negative link between individualism and tax compliance that endorses the institutional anomie theory, namely, the hypothesis that individualism causes the erosion of moral codes because of its focus on gaining materialistic rewards at the expense of ethical considerations. However, despite the large sample size, this finding should be treated with caution in light of the fact that the sample exhibits low variability with regard to individualism scores. Indeed, in this case the original meta-study aimed to include only articles based on lab experiments. Third, we find a negative relationship between individualism and prosocial behavior, but a minority of studies also provide opposite findings. In other words, our results mostly support the hypothesis that collectivism (rather than individualism out of warm glow) leads to maximize the well-being of society. In this case we can also exclude the possibility that differences in the data collection method explain this finding, since the original meta-regressions already control for them.

Finally, we point out that the main results of the three case studies remain unchanged after our re-analysis, and there is no evidence that cultural heterogeneity in the form of fractionalization plays a role in explaining the three economic outcomes of interest. We also believe that the investigation of cross-cultural differences in meta-analyses and the external validity thereof would greatly benefit from further experimental research in non-WEIRD countries, and we therefore call for more studies in this direction.
References


Polman, E. and Wu, K. (2020). Decision making for others involving risk: A review and meta-


Appendix A. Supplementary material

Figure A1: Average number of safe choices by country

<table>
<thead>
<tr>
<th>Country</th>
<th>Risk aversion by country</th>
</tr>
</thead>
<tbody>
<tr>
<td>India</td>
<td>6.351</td>
</tr>
<tr>
<td>Canada</td>
<td>6.203</td>
</tr>
<tr>
<td>Colombia</td>
<td>6.122</td>
</tr>
<tr>
<td>Denmark</td>
<td>6.023</td>
</tr>
<tr>
<td>China</td>
<td>5.601</td>
</tr>
<tr>
<td>South Africa</td>
<td>5.571</td>
</tr>
<tr>
<td>Germany</td>
<td>5.568</td>
</tr>
<tr>
<td>Netherlands</td>
<td>5.552</td>
</tr>
<tr>
<td>France</td>
<td>5.509</td>
</tr>
<tr>
<td>USA</td>
<td>5.498</td>
</tr>
<tr>
<td>Austria</td>
<td>5.481</td>
</tr>
<tr>
<td>Mexico</td>
<td>5.182</td>
</tr>
<tr>
<td>Morocco</td>
<td>5.028</td>
</tr>
<tr>
<td>Greece</td>
<td>5.018</td>
</tr>
<tr>
<td>Spain</td>
<td>4.994</td>
</tr>
</tbody>
</table>
Figure A2: Average compliance rate by country

![Tax compliance by country chart](chart.png)

<table>
<thead>
<tr>
<th>Country</th>
<th>Compliance Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albania</td>
<td>0.956</td>
</tr>
<tr>
<td>Canada</td>
<td>0.855</td>
</tr>
<tr>
<td>Romania</td>
<td>0.719</td>
</tr>
<tr>
<td>Austria</td>
<td>0.690</td>
</tr>
<tr>
<td>France</td>
<td>0.682</td>
</tr>
<tr>
<td>Taiwan</td>
<td>0.675</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>0.623</td>
</tr>
<tr>
<td>USA</td>
<td>0.613</td>
</tr>
<tr>
<td>Italy</td>
<td>0.596</td>
</tr>
<tr>
<td>Netherlands</td>
<td>0.591</td>
</tr>
<tr>
<td>UK</td>
<td>0.589</td>
</tr>
<tr>
<td>Australia</td>
<td>0.547</td>
</tr>
<tr>
<td>Germany</td>
<td>0.546</td>
</tr>
<tr>
<td>Sweden</td>
<td>0.534</td>
</tr>
<tr>
<td>Denmark</td>
<td>0.284</td>
</tr>
</tbody>
</table>
Figure A3: Average share donated by country

Prosocial behavior by country

Armenia: 0.573
Colombia: 0.553
Norway: 0.448
Spain: 0.428
Sweden: 0.375
Netherlands: 0.342
Denmark: 0.341
New Zealand: 0.332
UK: 0.299
Germany: 0.289
USA: 0.287
India: 0.292
Bosnia and Herzegovina: 0.290
Switzerland: 0.254
Sierra Leone: 0.253
Austria: 0.251
France: 0.222
Japan: 0.183
Australia: 0.142
<table>
<thead>
<tr>
<th>Year</th>
<th>Authors</th>
<th>Title</th>
<th>Paper Number</th>
<th>University</th>
<th>DOI Link</th>
</tr>
</thead>
<tbody>
<tr>
<td>2023-08</td>
<td>Fumarco, L., Harrell, B., Button, P., Schwegman, D., Dils, E.</td>
<td>Gender Identity, Race, and Ethnicity-based Discrimination in Access to Mental Health Care: Evidence from an Audit Correspondence Field Experiment.</td>
<td>2023-08</td>
<td>Brno: Masaryk University.</td>
<td><a href="https://doi.org/10.5817/WP_MUNI_ECON_2023-08">https://doi.org/10.5817/WP_MUNI_ECON_2023-08</a></td>
</tr>
<tr>
<td>2023-03</td>
<td>Mikula, Š., Reggiani, T., Sabatini, F.</td>
<td>The long-term impact of religion on social capital: lessons from post-war Czechoslovakia.</td>
<td>2023-03</td>
<td>Brno: Masaryk University.</td>
<td><a href="https://doi.org/10.5817/WP_MUNI_ECON_2023-03">https://doi.org/10.5817/WP_MUNI_ECON_2023-03</a></td>
</tr>
<tr>
<td>2023-02</td>
<td>Clò, S., Reggiani, T., Ruberto, S.</td>
<td>Consumption feedback and water saving: An experiment in the metropolitan area of Milan.</td>
<td>2023-02</td>
<td>Brno: Masaryk University.</td>
<td><a href="https://doi.org/10.5817/WP_MUNI_ECON_2023-02">https://doi.org/10.5817/WP_MUNI_ECON_2023-02</a></td>
</tr>
</tbody>
</table>


ISSN electronic edition 2571-130X