Climate Change is Decreasing Risk Aversion Globally

Remy Levin
University of Connecticut

Meng Song*
University of Connecticut

May 6, 2026

Abstract

We estimate the effects of climate change on individual risk preferences, using global cross-sectional preference data linked to personal climate histories. We find that a 1°C increase in mean lifetime temperature is associated with a 0.318-unit decrease in relative risk aversion globally. We replicate the results in panel datasets from four countries containing longitudinal measures of risk preferences. Historical projections indicate that climate change between 1981 and 2026 decreased median global risk aversion by 16%. Our results suggest that climate change may be shaping economic preferences in ways that could erode support for its own mitigation.

Keywords: Climate Change, Risk Preferences, Experience Effects
JEL Codes: D81, Q54

*Levin: Department of Economics, University of Connecticut. Email: remy.levin@uconn.edu. Song: Department of Economics, University of Connecticut. Email: meng.song@uconn.edu. We would like to thank Spencer Cooper, Joshua Graff Zivin, Ulrike Malmendier, Wayne Sandholtz, Jeffrey Sun, Lisa Tarquinio, and Daniela Vidart; seminar participants at the University of Geneva, Brown University, and the University of Massachusetts-Amherst; as well as participants in the Toronto Meeting on the Economics of Climate, Risk Theory Society Annual Meeting, the Association of Environmental and Resource Economists meetings, Normative Economics and Economic Policy Seminar, Toulouse Conference on Energy and Climate, and the UCLA Climate Adaptation Research Symposium for their helpful comments. An earlier version of this paper circulated under the title *The Global Impacts of Climate Change on Risk Preferences.*

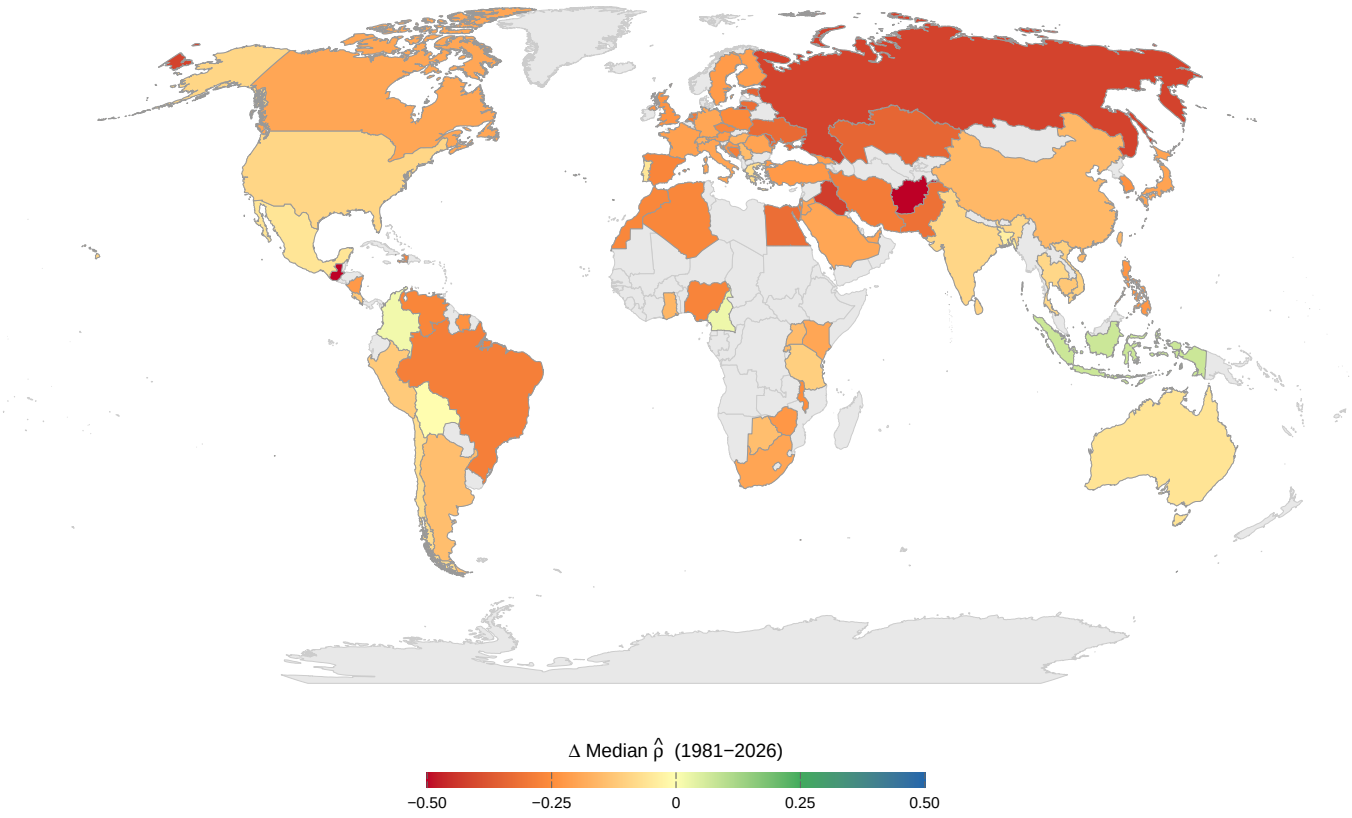
1 Introduction

Climate change is reshaping the world in which we live. Could it also be reshaping us? Standard economic analyses assume the answer is no: they treat climate change as a process that affects the options agents face, but not their preferences over those options. In this paper we provide empirical evidence that the answer is yes: climate change is decreasing individual risk aversion globally. We find that its cumulative effects on risk aversion have been negative in almost every country for which we have data (Figure 1), and that their magnitude is substantial. Historical projections indicate that rising temperatures between 1981 and 2026 decreased global median relative risk aversion by 0.22 points (95% CI: $[0.16, 0.26]$), a 16% decline.

Risk preferences are at the heart of economic behavior across many domains, including investment decisions (Merton (1969), Choukhmane and de Silva (2026)), insurance choices (Mossin (1968), Einav et al. (2012)), and entrepreneurship (Kihlstrom and Laffont (1979), Hvide and Panos (2014)). Our estimates suggest that climate change could be systematically increasing worldwide risk-taking across these domains. For risky behaviors that are known to be impacted by climate change — such as migration (Cattaneo and Peri (2016), Missirian and Schlenker (2017)), agricultural production (Burke and Emerick (2016), Ortiz-Bobea et al. (2021)), and conflict (Burke et al. (2015a), Mach et al. (2019)) — we identify a novel mechanism through which those effects may be operating.

These findings are also relevant to the study and practice of climate change mitigation. Many of the most serious costs of climate change take the form of tail risks, and preferences for climate action necessarily involve the weighing of those risks against the costs of mitigation (Cologna et al. (2025), Dechezleprêtre et al. (2025)). The results indicate that climate change could be driving a behavioral feedback loop, whereby societies become more tolerant of the risks it creates as it progresses, and are therefore slower to act to stop it. They also bear on the Social Cost of Carbon, the cornerstone policy tool that estimates the current welfare cost of emission damages using fixed social preferences. Our findings suggest its methodology

Figure 1: Estimated change in risk aversion due to climate change, 1981 – 2026



Notes: Results based on historical projections from our preferred specification, using the methodology in Section 3.3. Orange and red indicate increased median risk-seeking, blue and green increased median risk aversion. Afghanistan (-0.721) and Guatemala (-0.534) are bottom-coded in the map.

may need to be revised to account for preference adaptation along the warming path.[1]

To establish these findings, our primary analysis combines global cross-sectional preference data with individual climate histories covering the last century. We use preference data

[1]The Social Cost of Carbon (Nordhaus (2017)) measures the present welfare cost of an additional ton of CO_2 emitted. It is used in benefit-cost calculations for climate mitigation policies, and as a benchmark for the optimal level of carbon taxes. In standard implementations, welfare is aggregated using representative-agent utility and discounted with a Ramsey formula, so its value depends on the pure rate of time preference and the curvature of utility with respect to consumption. Under the CRRA utility function the EPA has used, which we also employ in our analysis, that curvature parameter is also the coefficient of relative risk aversion. The SCC therefore depends directly on how the model values consumption loss from climate risk (U.S. Environmental Protection Agency (2023)). Because the chosen parameters substantially affect the estimate (Anthoff et al. (2009), Nath et al. (2023)), there is a long-running debate in the literature about the right values to use, including the proper way to integrate the costs of climate tail risks into the calculation (Weitzman (2007), Weitzman (2014)). But regardless of which values are chosen, current practice treats them as fixed and exogenous. Allowing preferences to endogenously evolve in response to warming would be a qualitative change in the SCC framework.

from the Global Preference Survey (GPS), a landmark study that collected experimentally validated economic preference measures as part of the 2012 wave of the Gallup World Poll. Its sample contains approximately 80,000 subjects across 76 countries, and is representative of 90% of the world's population and GDP (Falk et al. (2018)).

The GPS risk aversion measure is a weighted composite of two sub-scales: a lottery-choice based scale (a 32-bin adaptive staircase measure), and a self-reported risk aversion scale. The staircase measure is required for estimating subjects' risk preference parameters, and for deriving quantitative (and not just qualitative) results. Unfortunately, data from the sub-scales have never been published, and they cannot be readily reconstructed from the public release of the GPS. To overcome this issue, we design an algorithm that successfully decodes the two sub-scales from the composite, by exploiting its latent lattice structure and value frequency distribution. We verify the fidelity of the reconstructed sub-scales by showing that they recover the full latent lattice, closely reproduce the published GPS weights, and precisely reconstruct the original composite.

Using the decoded staircase measure, we structurally estimate the coefficient of relative risk aversion $\hat{\rho}$ for each subject under CRRA utility. We employ a random preference model with a tremble parameter (Apesteguia and Ballester (2018), Harless and Camerer (1994)), combined with broad bracketing over a data-driven income horizon. We then apply classical empirical Bayes shrinkage (Chetty and Hendren (2018)) within-country to the estimates to account for heterogeneity in their precision. This yields the first individual-level, globally-representative distribution of risk preference parameters in the literature.[2] We find that the median global individual is moderately risk averse ($\hat{\rho} = 1.27$ in 2012), with substantial heterogeneity within and across countries. The reduced-form staircase measure and the structural $\hat{\rho}$ serve as the dependent variables in our global analysis.

To link the survey data to spatial climate data, we match GPS survey regions to administrative boundaries from global GIS data using a multistage pipeline. The procedure

[2]Global and country-level risk aversion statistics are reported in Appendix A.2. We recommend that researchers use country medians as summary statistics, as they are less sensitive to extreme outliers.

successfully matches 99% of GPS regions (1,144 of 1,148) to their geographic boundaries. This allows us to analyze the preference data at a spatial resolution and scope at which it has rarely been employed.[3]

We combine three gridded climate datasets to construct a monthly temperature and precipitation record on a 0.5° latitude–longitude grid, from January 1900 through 2025 (Willmott and Matsuura (2001), Harris et al. (2020), Fan and van den Dool (2008)). These data are aggregated to GPS regions using area weighting over their geographic boundaries. For each subject, we then construct lifetime mean temperature, the main independent variable in the analysis, by averaging monthly temperatures in their survey region from birth until the survey. This allows us to study the role of long-run climate experiences, accumulated over decades, rather than the effects of short-run weather fluctuations.

Our baseline specification includes region and cohort (i.e. birth year) fixed effects, as well as controls for gender, language, lifetime mean precipitation, and survey seasonality. Identifying variation comes from differences in warming across birth cohorts within regions, net of time-invariant regional differences and global cohort effects. The identifying assumption is that, conditional on fixed effects and controls, variation in climate experiences is orthogonal to unobserved determinants of risk preferences. We examine the most plausible threats to this assumption from the experience effects literature — macroeconomic conditions, conflict, and natural disasters — and argue that they are unlikely to drive the observed relationship. We also present the results of a specification with region and cohort-by-country fixed effects, the most demanding that can be estimated with these data.

We find consistent negative effects of lifetime mean temperature on risk aversion. In our preferred specification, a 1°C increase in lifetime mean temperature is associated with a 0.318-unit decrease in $\hat{\rho}$ (95% CI: [-0.419, -0.218], $p < 0.01$). The staircase coefficient is also negative in this specification, though not significant. The $\hat{\rho}$ coefficient is marginally significant even in the country-by-cohort specification (-0.177, $p < 0.10$), where identification

[3]The only other paper to analyze the GPS at the region level, to our knowledge, is Sunde et al. (2022), who match 61.3% of GPS regions in 55 countries to the region dataset in Gennaioli et al. (2013).

is restricted to differential warming rates across regions within the same country and cohort. Here the staircase coefficient is also negative and statistically significant (-0.722, $p < 0.05$).

These results are robust to six categories of sensitivity checks: alternative shrinkage regimes, restricting the sample to the interior of the staircase scale, alternative sets of individual controls (including controlling for all other GPS preferences), alternative climate and weather controls (including lifetime climate variability and interview-day temperature anomalies), alternative seasonality controls, and alternative clustering schemes.

To extrapolate from the marginal effect to the cumulative effect of climate change to date, we conduct a historical counterfactual exercise in the spirit of Burke et al. (2015b) and Carleton et al. (2022). We reassign the GPS sample counterfactual lifetime climate experiences as if they had been measured at the same age and in the same region in 1981 and in 2026, then use our preferred estimates to project their risk preferences in those years. The resulting counterfactual preference distributions shift substantially towards risk-seeking over time. At the country level, the projected change in median $\hat{\rho}$ is negative in all but three of 76 countries, with an average country-median change of -0.208. At the global level, we estimate that the median of the globally-representative preference distribution, *ceteris paribus*, was 1.381 in 1981 and 1.161 in 2026, a decrease of 0.22 units or 15.9% over 45 years. Roughly half of the warming over the period occurred between 1981 and 2012, and the other half between 2012 and 2026. Our results suggest that as global warming has accelerated over the last few decades, so has the decrease in risk aversion worldwide.

A concern with the global analysis is that, since the public GPS data does not contain information on subjects' region of birth, our estimates may partly reflect endogenous migration that is plausibly correlated with both climate experiences and risk preferences. To address this concern, and to support our global estimates with evidence from within-person changes in risk preferences, we assemble data from four national panel surveys with longitudinal risk aversion measures: the Indonesian Family Life Survey, the Mexican Family Life Survey, the Chilean Social Protection Survey, and the Japanese Preference Parameters Study. Each of

the four surveys contains multiple measurements of risk aversion for the same subjects taken years apart, using instruments that are similar to the staircase instrument in the GPS. We regress these measures on lifetime mean temperature in subjects' birth region, from birth to measurement in the survey, as well as precipitation experiences and individual and survey year fixed effects.

In each of the four panels we find significant results consistent with the global analysis: increases in lifetime mean temperature lead to decreases in subjects' risk aversion. These results are especially notable given the diversity of this set of countries, which spans low, medium, and high income; very different climates; and different cultures, religions, histories, and demographics. The similarity of the findings across all four bolsters the evidence for both the internal consistency and external consistency of our analysis as a whole.

We also examine heterogeneity in the effects along three dimensions we can measure across settings: age, income, and gender. The effects are broadly stronger for younger individuals, and are generally larger for lower-income individuals, who have the fewest resources to adapt and who are typically most exposed to the physical costs of climate change. The gender patterns are more mixed: in the global data, effects are larger for women. This implies gender convergence in risk preferences under warming, consistent with evidence in Falk and Hermle (2018). In most of the national panels the opposite pattern holds, in line with panel evidence from natural-disaster settings (Hanaoka et al. (2018), Beine et al. (2025)).

Our results contribute to several strands of the literature. First, we contribute to the aforementioned literature on the impacts and mitigation of climate change. At the micro level, the findings may help explain observed climate impacts across multiple domains involving risk-taking behavior. At the macro level, they point to a broad behavioral shift driven by warming that could be dampening international support for its mitigation, given that climate risk perceptions are a major determinant of that support (O'Connor et al. (1999), Zahran et al. (2006), Viscusi and Zeckhauser (2006), Drews and van den Bergh (2016), Cologna et al. (2025)). Our results also bolster a recent theoretical literature on the role of endogenous

preferences in shaping the welfare calculus of the Social Cost of Carbon and carbon taxes (Konc et al. (2021), Mattauch et al. (2022), Besley and Persson (2023), Beck (2026)). We provide direct empirical evidence that economic preferences worldwide are endogenous to warming. Notably, whereas most of this literature assumes that preference endogeneity will increase the welfare costs of climate damages, our results suggest that the opposite may be true, at least with regards to risk preferences under current climate trends.

Second, this paper contributes to the fast-growing literature on the role of lifetime experiences in shaping risk preferences (Malmendier and Wachter (2024)). Prior work has shown that risk aversion responds to lifetime macroeconomic experiences (Malmendier and Nagel (2011), Guiso et al. (2018), Levin and Vidart (2026)), experiences of violence (Voors et al. (2012), Jakiela and Ozier (2019), Brown et al. (2019b)), and natural disasters (Cameron and Shah (2015), Cassar et al. (2017), Hanaoka et al. (2018), Reynaud and Aubert (2019)). Most closely related to our work is the subset of the literature that studies the effects of rainfall shocks on risk preferences, especially Di Falco and Vieider (2022) who find risk preference adaptation to the natural environment in Ethiopia, as well as Liebenehm et al. (2023) who document parallel results in Thailand and Vietnam, and Jaramillo et al. (2025) who do so in Indonesia.

This paper is the first in the literature to show that lifetime increases in average temperatures shape individual risk preferences. This is also the first study to present evidence that lifetime experiences of any kind produce consistent causal effects on risk preferences globally, rather than in particular times and places. Our use of four country panels combined with global cross-sectional data is a significant expansion in scope for the experience effects literature, which is dominated by single panel studies. The results suggest that preference adaptation to lived experience and environmental change may be a more universal phenomenon than has previously been appreciated. As such, our findings broaden the challenge that the experience effects literature poses to economic analyses that employ fixed and exogenous risk preferences.

Third, we provide the most comprehensive quantitative estimates of risk aversion parameters across countries to date. Using lottery-choice data we recover representative distributions of structurally-estimated risk preferences in 76 countries, covering 90% of the world's population and GDP.[4] These data should be of value to researchers conducting microeconomic and macroeconomic analyses in international settings.

The remainder of the paper is organized as follows. Section 2 details the data and empirical methodology of the global analysis. Section 3 presents the global results, including the descriptive characteristics of the global risk preference distribution, the main regression results, and the results of the historical projection exercise. Section 4 covers robustness analyses for the main regression results. In Section 5 we describe the data, methodology, and results of the four national panel analyses. Section 6 presents the results of the heterogeneity analysis in the panels and the global data. Section 7 concludes.

2 Data and methodology

2.1 Global preference data

Our source for risk preference data is the Global Preference Survey (GPS). The GPS is a landmark study that, for the first time, measured risk, time, and social preferences using experimentally validated instruments for a large global sample (Falk et al. (2018), Falk et al. (2022)). The global preference module was attached to the 2012 wave of the Gallup World Poll, and collected data for a sample of approximately 80,000 individuals across 76 countries. The survey's sample is representative for 90% of the world's population and GDP. Data from the GPS has been used in several prominent papers describing the distribution and correlates

[4] Falk et al. (2018) and Bouchouicha and Vieider (2019) provide measures of risk aversion for international samples with similar scopes to ours, but using qualitative instruments that cannot be used for quantitative analyses. Vieider et al. (2015), Rieger et al. (2014), L'Haridon and Vieider (2019), and Meissner et al. (2023) present international quantitative estimates of risk aversion, but for much smaller samples of subjects and countries. Gandelman and Hernández-Murillo (2014) provide a quantitative estimate for a representative sample from 75 countries, but only for a single country-level CRRA parameter, using subjective life satisfaction and income data, and missing major economies that our data covers, such as China and Italy.

of preferences worldwide (Falk et al. (2018), Falk and Hermle (2018), Enke (2019), Becker et al. (2020), Sunde et al. (2022)).

In this paper we use data exclusively from the public release version of the GPS. This dataset is composed of six measured preference indices (risk aversion, patience, positive reciprocity, negative reciprocity, altruism, and trust), three additional individual variables (gender, age, and self-reported subjective math ability), and six survey variables (interview country, region, date, and language, as well as Gallup subject ID and survey weight).

2.2 Decoding the GPS risk preference sub-scales

The risk preference measure in the GPS is a composite index of two underlying items: a 32-bin measure derived from a series of five adaptive, hypothetical choices between a sure amount of money and a monetary lottery (the *staircase* measure); and an 11-bin self-reported measure of the subject's propensity to take risks (the *Likert* scale). The staircase measure is a standard experimental measure of risk preferences used in economics, particularly in field settings (Barsky et al. (1997), Hackethal et al. (2023)), while the Likert scale is a commonly used survey instrument to measure attitudes towards risk (Caliendo et al. (2010), Dohmen et al. (2011)). To obtain the composite index (the *combined* measure), each sub-scale bin was normalized to its respective z-score in the empirical distribution after data collection. The two normalized measures were then added together, using weights derived from a validation procedure that tested the power of different linear combinations of the sub-scales to predict risky choices (i.e. revealed preference) in incentivized laboratory experiments amongst a sample of German university students. The combined scale loads at approximately 47.3% on the normalized staircase scale, and 52.7% on the normalized Likert scale.[5]

[5]Comprehensive details on the development process of the GPS and construction of the combined scale are available in the online appendix of Falk et al. (2018), and in Falk et al. (2022). We represent the GPS staircase instrument in Appendix A.1. The monetary amounts in the staircase instrument were adjusted across countries using the ratio of country median wages to the median wage in Germany at the time of the survey. For the Likert instrument, subjects were asked to indicate "in general, how willing or unwilling you are to take risks," on a scale from 0 to 10.

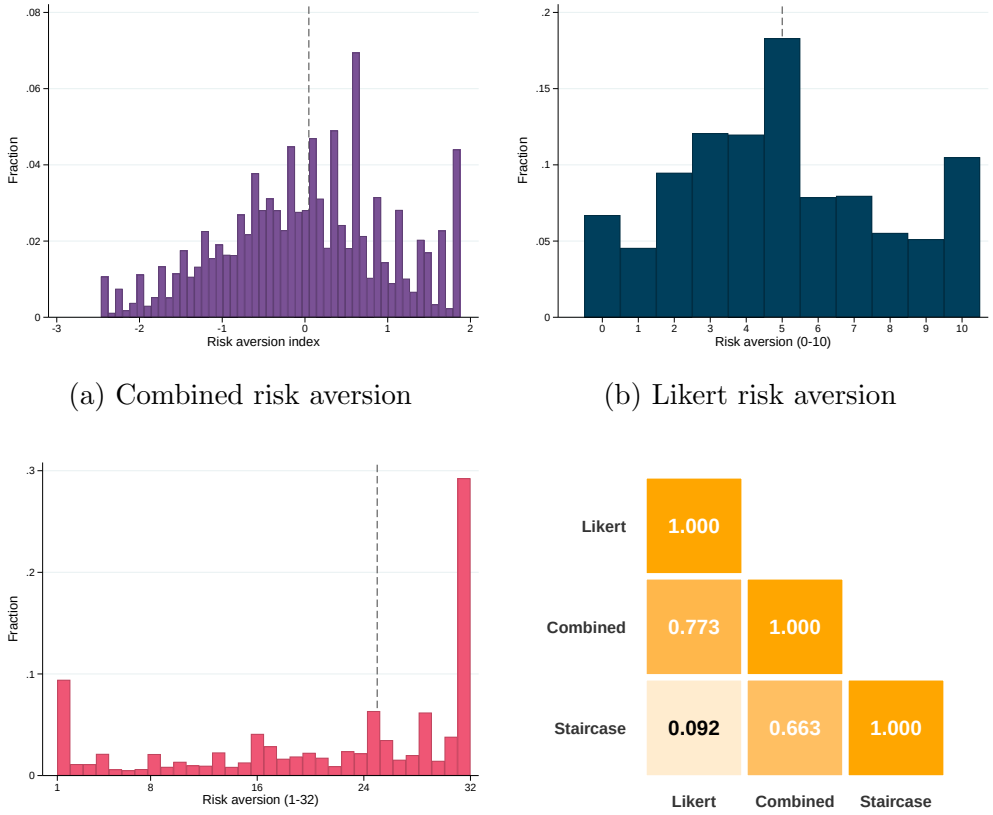
The sub-scale data underlying the GPS preference measures have never been published,[6] nor can they be readily reconstructed from the combined scales, because the sub-scale z-scores are unavailable and the imputation procedure for missing data introduces additional noise.[7] This creates two important limitations. First, a *validation and identification* concern: the underlying sub-scales rely on well-established, carefully translated instruments, whereas the weighting scheme used to aggregate them into a single index was validated in laboratory experiments on a narrow subject pool (a single cohort of students at one German university), and follow-up evidence on its external validity is mixed (Kosfeld and Sharafi (2024), Bittner et al. (2024), Kosfeld et al. (2025)). Access to the unweighted sub-scales would let researchers assess whether findings are driven disproportionately by one component or are artifacts of the aggregation procedure. Second, a limitation of *interpretation and quantification*: because the combined measures are composite indices of qualitatively different sub-scales, their units lack a clear economic interpretation, so that analyses using them can at best estimate only the *sign* of an effect and not its *magnitude*, often the policy-relevant object. Moreover, the composite indices cannot be mapped into structural preference parameters such as risk aversion or discount rates. The staircase components are designed to recover precisely these objects, but without access to them the GPS data cannot be directly integrated with standard economic models or structural empirical analyses.

To address these issues, we design an algorithm for decoding the risk preference sub-scales from the public combined measure alone. Our procedure exploits the latent lattice structure of the composite for complete-case respondents: because z-scoring is linear, one-step changes in the Likert and staircase responses shift the composite by constant step sizes (A and B), so that true (L, S) combinations must lie on a 32×11 grid of at most 352 exact values, with imputed observations falling off the lattice. We identify A and B directly from the data by

[6] Additional individual data from the 2012 World Poll are only available through purchase of institutional access to all Gallup data. Unlike the Gallup World Poll data, however, there is no known legal or contractual restriction on the publication of the GPS sub-scales.

[7] For example, while the combined risk preference scale should, in theory, contain at most 352 (11×32) unique values, the published data contains 735.

Figure 2: Distributions and correlations of GPS risk aversion sub-scales



(a) Combined risk aversion

(b) Likert risk aversion

(c) Staircase risk aversion

(d) Pairwise correlations

Notes: All measures are scaled so that higher values correspond to greater risk aversion, the reverse of the ordering in the GPS. Panel (a) is the combined risk aversion measure in the GPS. Panels (b) and (c) display the decoded sub-scales of the GPS risk preference measure, using the procedure outlined in Section 2.2, and detailed in Appendix B.1. Dashed lines indicate subject medians. Panel (d) reports Spearman rank correlations ($N = 79{,}703$); all correlations are significant at $p < 0.001$.

searching among the most frequent composite values for highly discriminating symmetric distance patterns that imputed values are extremely unlikely to satisfy. We then iteratively expand outward from endpoint anchors to recover a connected component of exactly 352 lattice points. Since the recovered A is substantially larger than B ($A/B \approx 4.26$), we can confidently assign A to the Likert scale (which has fewer bins and therefore larger z-score steps) and B to the staircase. We then propagate coordinate labels along A- and B-steps to assign a unique (L, S) pair to each lattice value, and assign off-lattice (imputed) composites the coordinates of the nearest lattice point.

We verify the decoding in several ways. The recovered lattice contains exactly 352 unique composite values forming a complete 32×11 grid with no duplicate (S, L) assignments, covers 89.6% of respondents, and reconstructs on-lattice composites within floating-point tolerance (max error 9.1×10^{-6}). A placebo test starting from off-lattice points reveals no other substantial connected components. The implied sub-scale weights (≈ 0.477 staircase, ≈ 0.523 Likert) closely match the published GPS weights ($\approx 0.473/0.527$). Rebuilding the composite from the decoded z-scores yields perfect correlation on-lattice ($r = 1.000000$) and near-perfect correlation in the full sample, including nearest-neighbor assignments for off-lattice values ($r = 0.999998$), with 99.4% of observations reconstructed within 0.01 of the original. Complete details of the decoding and verification procedure are available in Appendix B.1.

Histograms and correlations for the combined scale and its sub-scales are in Figure 2, normalized so that higher values indicate greater risk aversion (flipped relative to the GPS scale). As found in other settings, the staircase instrument loads heavily on extreme values, while the Likert instrument is more normally distributed, with clustering at extremes and the midpoint.[8] The combined measure correlates strongly with both sub-scales (staircase: Spearman's $r_s = 0.663$; Likert: $r_s = 0.773$), indicating that neither dominates its variation. However, the two sub-scales are only marginally correlated with each other ($r_s = 0.092$), which is surprising given that both target the same underlying trait and jointly predict risk-taking in experimental validations. The most likely explanations are that either the scales capture different aspects of risk aversion, or that one or both are measured with substantial noise.

In the rest of the paper we focus on the staircase instrument, as it is the one most closely tied to economic theory, and can be used to derive a structural estimate of the underlying risk aversion parameter.

[8]We show that our results are robust to excluding FOSD violators and subjects in end bins in Section 4.

2.3 Structural estimation of the relative risk aversion parameter

Next, we use subject choices in the staircase task to structurally estimate their coefficient of relative risk aversion. Complete details of our procedure are available in Appendix B.2. The first step in this process is recovering country-specific lottery prizes and subject income. The GPS scaled all monetary lottery amounts by country-specific factors k_c equal to a rounded ratio of national median incomes relative to Germany, but the specific factors used were never published. We construct several proxies for these factors using multiple vintages of data from the World Bank Poverty and Inequality Platform (PIP) with different PPP adjustments, as well as published data from Gallup on median household income. We validate these measures against five ground-truth country observations where implementation errors in the GPS inadvertently reveal the true k_c (Falk et al. (2018) online Appendix §AG.1), and select the near-contemporaneous PIP 2011 vintage measure based on overall fit. For subject income, since individual-level income variables are not available in the public release of the GPS, we construct a proxy w using the regional GNI per capita from Chrisendo et al. (2025), matched to GPS survey regions via centroid-in-polygon spatial join (using the matching procedure in Section 2.4), and assign it to all subjects in the region.

As is standard, we assume CRRA utility ($u(c) = c^{1-\rho}/(1-\rho)$) with background income w. Under narrow bracketing ($w = 0$), the staircase maps to a compressed range $\rho \in (-0.1, 0.8)$ that cannot distinguish among moderate-to-high risk aversion (Cox and Sadiraj, 2006). Under broad bracketing ($w > 0$), subjects are assumed to integrate lottery prizes with background income. This yields cardinal risk aversion parameters more consistent with macroeconomics and finance analyses, but leaving the bracketing horizon as a free parameter (Read et al., 1999). We use broad bracketing in our analysis, and take a data-driven approach to the bracketing horizon by estimating the model under five candidates — daily, weekly, monthly, yearly, and a heterogeneous patience-based horizon — and selecting the one that best fits the choice data based on model fit and monotonicity of $\hat{\rho}$.

At each of the 31 staircase boundaries, we solve the CRRA indifference equation for the

ρ_b at which a respondent with background income w is exactly indifferent between the lottery and the safe payment. These boundary values are fed into an RPM+Tremble estimator. This models the five binary choices that each subject makes via the Random Preference Model (Apesteguia and Ballester, 2018) with a tremble parameter (Harless and Camerer, 1994):

$$P(\text{lottery}) = (1-\omega)\,\Phi[(\rho_b - \rho_i)/\sigma] + \omega/2,$$

where σ captures noise in ρ-space and ω is the probability of a random coin-flip choice. The tremble component accommodates FOSD violators and unsolvable-boundary respondents, resulting in coverage of approximately 79,000 subjects (98.5%). We estimate (σ, ω) at the population level via maximum likelihood, and each subject's $\hat{\rho}_i$ is then obtained by maximizing their personal 5-choice likelihood at the optimal $(\hat{\sigma}, \hat{\omega})$.

The daily horizon is the only specification that passes the key construct-validity diagnostic of monotonically increasing mean $\hat{\rho}$ across staircase bins (Apesteguia and Ballester, 2018), with a low tremble estimate ($\hat{\omega} = 0.09$). In contrast, the non-daily horizons exhibit 5–13 monotonicity violations out of 28 adjacent bin pairs, and progressively higher tremble parameters (Figure 15 and Table 27). We therefore adopt daily income bracketing — the most conservative fixed horizon given the evidence for narrow bracketing in the literature (Thaler, 1999; Andersen et al., 2018a) — as our primary specification.

To account for heterogeneous noise across subjects in our estimates, we implement Empirical Bayes shrinkage using the classical Gaussian hierarchical model (James and Stein, 1961; Morris, 1983; Efron and Morris, 1975). Within each country c, the posterior mean shrinks each individual toward the precision-weighted country mean by $B_i = \hat{\sigma}_i^2/(\hat{\sigma}_i^2 + \hat{\tau}_c^2)$, where $\hat{\sigma}_i^2$ is the individual measurement error variance and $\hat{\tau}_c^2$ is the between-individual variance of true preferences estimated via DerSimonian and Laird (1986). This follows the standard approach in the literature on teacher value-added (Kane and Staiger, 2008; Chetty et al., 2014), neighborhood effects (Chetty and Hendren, 2018), and school effectiveness

(Angrist et al., 2017). We verify the Gaussian error model by estimating a robust alternative with a heteroskedastic t-distribution (Lange et al., 1989; Huber and Ronchetti, 2009). The MLE over degrees of freedom for the t selects $\nu^* = 10{,}000$, the normal limit, confirming that the profile-likelihood standard errors fully capture the relevant measurement error structure from the estimation model.[9]

2.4 Matching GPS regions to geographic boundaries

In order to maximize power in our analysis we conduct it at the GPS region level globally. With the exception of Sunde et al. (2022), who study the relationship between patience and income sub-nationally, we are not aware of other papers in the literature that use global sub-national GPS data in their analysis.[10]

In their paper, Sunde et al. (2022) match regions in the GPS to the regional dataset in Gennaioli et al. (2013). This approach has two drawbacks. First, geographic overlap between the two datasets is limited: Sunde et al. (2022) report matching 704 of 1148 regions (61.3%) in 55 countries. Second, since the Gennaioli et al. (2013) database contains no GIS data or region boundaries, it is not possible to readily link it to other, external spatial variables. This makes its use especially limiting for our analysis, which employs gridded climate data.

Our matching procedure instead relies on linking GPS regions to administrative boundary polygons using global GIS data from GADM 4.1.0. We use a two-pass automated pipeline followed by extensive manual crosswalks. We successfully match 1,144 of 1,148 GPS regions, covering over 99% of observations. The procedure is fully deterministic and reproducible. A comprehensive description of it can be found in Appendix B.3.

[9] In Section 4 we show that our main estimates are robust to alternative shrinkage regimes, including no shrinkage, Winsorization, a t-distribution empirical Bayes with constrained ν, and a mixture model that treats excess end-bin respondents as uninformative.

[10] Hanushek et al. (2026) study the relationship between patience and student achievement in the United States and Italy regionally. However, their primary analysis uses Facebook-derived proxies which are validated against the GPS, not GPS micro data directly.

2.5 Lifetime climate experience variables

Our climate data combine three gridded sources to construct a monthly temperature and precipitation record on a $0.5° \times 0.5°$ latitude–longitude grid from January 1900 through 2025. The primary source is the University of Delaware Air Temperature and Precipitation dataset, version 5.01 (Willmott and Matsuura, 2001), which covers 1900–2017. We extend the record through 2024 using CRU TS v4.09 (Harris et al., 2020), and through 2025 using CPC Global Unified daily temperature (Fan and van den Dool, 2008), aggregated to monthly frequency. Splice validity was confirmed over the 2010–2025 overlap period. We aggregate gridded climate to 1,144 GPS survey regions using area weighting.

For each GPS respondent with survey region r and birth year b (computed as survey year minus age), we construct lifetime mean temperature as the average over monthly values from January of year b through December 2011:

$$\overline{T}_{r,b} = \frac{1}{N_b} \sum_{t \in \{\text{Jan}(b),\ldots,\text{Dec}(2011)\}} T_{r,t},$$

where $N_b = 12 \times (2012 - b)$ is the number of months in the experience window. The lifetime standard deviation of temperature, and analogous measures for precipitation, are constructed similarly. We truncate the experience window at December 2011, the last full calendar year before the 2012 GPS survey, rather than extending it to the interview date to avoid conflating lifetime climate trends with short-run weather shocks. Lifetime mean temperature is the main independent variable, and lifetime mean precipitation as a control in our baseline specifications. Complete details on the climate data and the construction process are presented in Appendix B.4.

2.6 Controls and weights

Our analysis uses three additional categories of controls. First, the public GPS data provide subject-level covariates: age, gender, interview language, self-reported math ability, and five

preference indices (patience, positive and negative reciprocity, altruism, trust). Age enters several specifications as a birth-year fixed effect. Gender and language, unlikely to be affected by climate experiences, are included in baseline regressions for precision. Robustness to alternative covariate sets is shown in Section 4.

Second, we construct controls to address noise and potential bias from survey seasonality. If the timing of survey deployment was correlated with local climate — for instance, if hot regions were surveyed in their cold months and cold regions in their hot months, when the effect of long-run heat exposure is potentially magnified — this may attenuate the estimated effects. Conversely, if the correlation between local temperature and seasonal timing is positive, our results might be biased away from zero. We test this directly by regressing each region's mean interview-month temperature percentile on mean annual temperature, finding a significant negative correlation ($r_s = -0.105, p = 0.0002$): colder regions were surveyed in relatively warmer months and vice versa. To absorb this variation, we construct seasonality fixed effects, by interacting survey month with the subject's hemisphere. We include these in the baseline specification.

Third, we control in robustness checks for short-run temperature shocks on the interview day and preceding night. These could affect cognitive and emotional outcomes, including through sleep quality (Obradovich et al., 2018; Mullins and White, 2019; Minor et al., 2022; Romanello et al., 2023), but reflect weather rather than climate variation. For each survey region r, interview date d, and $k \in \{\text{day of, night before}\}$ we compute a z-score anomaly:

$$z_{r,d,k} = \frac{T_{r,d,k} - \bar{T}_{r,d,k}}{\text{SD}(T)_{r,d,k}},$$

where $\bar{T}_{r,d,k}$ and $\text{SD}(T)_{r,d,k}$ are the historical mean and standard deviation of temperature for that day of year in region r, computed over the standard WMO 1981–2010 reference period.

We employ three subject weighting schemes in the analysis. To identify the causal effects of the treatment, we maintain an equal weight for each subject. For results at the country

level, we weight subjects by the survey weights provided by Gallup, in order to derive a representative sample in each country. When estimating effects at the global level, we further weight subjects by the ratio of their country's population in 2012 to its survey sample size. The results of the weighting schemes on the preference distribution and the estimated effects are small, as can be seen in Section 3.1 and Section 4.

2.7 Identification

The aim of our analysis is to estimate the causal effect of climate change on individual risk preferences globally. The ideal experiment would randomly assign climate experiences over decades and measure preferences before and after. In its absence, we exploit between-person variation in experienced lifetime mean temperature linked to a global cross-sectional measure of risk preferences. The lifetime mean captures experiences over the time scale of climate, and generates between-subject variation even within region through non-shared climate experiences and differential weighting of shared experiences by age.

The dominant dimension of variation in lifetime temperature is cross-sectional: some regions are generally hot and others are cold. The effects of these cross-sectional differences are clearly confounded with persistent differences between hot and cold places such as culture, institutions, or development. Cross-sectional differences in climate are also mostly orthogonal to the treatment of interest, which is the *change* in climate. In our first specification we control for cross-sectional differences in climate by including survey region fixed effects. These net out permanent location-based differences, and isolate within-region across-cohort variation that arises from cohorts being born into different segments of their region's climate history.

The region fixed effects specification identifies off the primary component of climate change, a relatively uniform increasing trend in global temperature. However, these estimates may still be confounded with other global trends that are correlated with both lifetime temperature and risk preferences. Based on current evidence in the experience effects literature, there are four major possible confounds: economic growth, violence, natural disasters, and age. The

first three are unlikely to be driving our results. Since we find that risk aversion is negatively correlated with lifetime temperature, confounders would need to be positively correlated with temperature and negatively with risk aversion, or vice versa. Macroeconomic growth has been found to be negatively correlated with both risk aversion (Malmendier and Nagel (2011), Levin and Vidart (2025)) and warming (Dell et al. (2012), Burke et al. (2015b)). Conflict is positively correlated with both (Brown et al. (2019a), Hsiang et al. (2013), Ash and Obradovich (2019)). The most significant disaster-driven changes in risk preferences stem from events like earthquakes, whose occurrence is mostly orthogonal to warming trends (Cameron and Shah (2015), Hanaoka et al. (2018)).

Confounding with subject age is far and away the most serious threat in the region fixed effects specification: young people are on average less risk averse (Dohmen et al. (2017)) and have lived through a hotter climate than their elders. To address this, our preferred specification additionally includes cohort (i.e. birth year) fixed effects. These net out age effects common across all regions. We also estimate a specification with region and cohort-by-country fixed effects, the most demanding we can estimate with our data, which further controls for country-specific age effects.

To bolster the evidence for our global estimates, we also assemble data from four panel surveys containing longitudinal elicited measures of risk preferences, for the same individuals years apart. This allows us to estimate the effects of within-person changes in climate experiences on risk preferences, and to address other concerns in the global analysis such as endogenous migration. We discuss these further in Section 5.

2.8 Empirical specifications

We estimate three specifications, each for two left hand side variables: (1) staircase risk aversion from the GPS, and (2) the structurally estimated $\hat{\rho}$ based on it. We label both collectively RA below. For subject i, in region r, born in year b, the region fixed effects

specification is

$$RA_{i,r,b} = \beta \bar{T}_{r,b} + \gamma \bar{P}_{r,b} + \delta X_i + \alpha_{hm} + \alpha_r + \epsilon_{i,r,b}, \tag{1}$$

where $\bar{T}_{r,b}$ is mean monthly temperature in region r from January of year b to December 2011, $\bar{P}_{r,b}$ is the equivalent measure for precipitation, X_i are individual controls (gender and interview language dummies in the baseline), α_{hm} is a hemisphere by interview month fixed effect controlling for seasonality, and α_r is the region fixed effect. For the main specification we add a cohort fixed effect α_b:

$$RA_{i,r,b} = \beta \bar{T}_{r,b} + \gamma \bar{P}_{r,b} + \delta X_i + \alpha_{hm} + \alpha_r + \alpha_b + \epsilon_{i,r,b}, \tag{2}$$

while in the third specification we instead add a country-by-cohort fixed effect α_{cb}:

$$RA_{i,r,b} = \beta \bar{T}_{r,b} + \gamma \bar{P}_{r,b} + \delta X_i + \alpha_{hm} + \alpha_r + \alpha_{cb} + \epsilon_{i,r,b}. \tag{3}$$

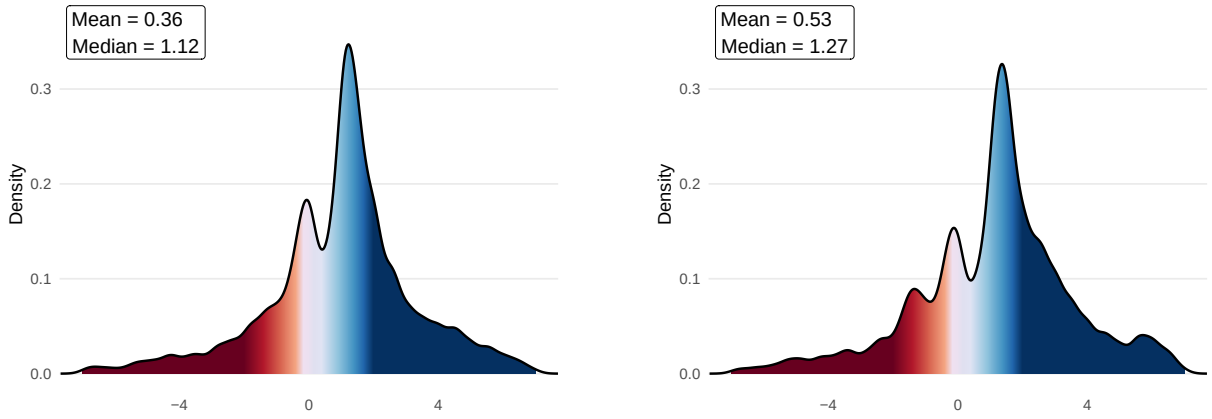
Standard errors for all specifications are clustered two-ways by region and birth year.

3 Results

3.1 The global risk preference distribution

In this section we present the results of our structural estimation of global risk preferences. These include three distributions of the CRRA $\hat{\rho}$ parameter: (1) the *unweighted* distribution; (2) the distribution weighted by the Gallup-survey weights, so it is *nationally-representative*; and (3) the nationally-representative distribution additionally weighted by countries' population sizes (relative to their survey sample size), to yield a *globally-representative* preference distribution. We use the unweighted distribution for causal inference (Section 3.2), the nationally-representative distribution when reporting cross-country results, and the globally-representative distribution for the counterfactual analysis (Section 3.3).

Figure 3: Global Distribution of Relative Risk Aversion, 2012



(a) Unweighted $\hat{\rho}$

(b) Globally-representative $\hat{\rho}$

Notes: This figure displays the global cross-sectional distribution of our individual estimates of CRRA coefficient $\hat{\rho}$. Red indicates risk-seeking ($\hat{\rho} < 0$), blue indicates risk aversion ($\hat{\rho} > 0$). Panel (a) displays the kernel density of the unweighted $\hat{\rho}$ distribution. Panel (b) displays the kernel density of the globally-representative distribution, which is result of weighting subjects by national Gallup survey weights and by the ratio of their country's population in 2012 to the country's Gallup World Poll sample size. Both density plots are trimmed to the range $\hat{\rho} \in [-7, 7]$ (comprising $\approx 95\%$ of the sample) for visual clarity. Reported means and medians are computed from the untrimmed distributions.

Kernel density plots of the estimated preference distributions are presented in Figure 3. The unweighted distribution (Panel a) has a median of 1.12 and a mean of 0.36, with an interquartile range of -0.33 to 2.10, and an interdecile range of -2.82 to 3.99. The distribution is left skewed, with a long tail of substantial risk-seeking estimates. It has a primary peak at $\rho = 1.34$ (density 0.288) and a secondary peak at risk neutrality $\rho = 0.03$ (density 0.152). The nationally-representative distribution is virtually identical, so we do not present it here. The globally-representative distribution (Panel b) has a median of 1.27, a mean of 0.53, with IQR $= [-0.35, 2.59]$ and IDR $= [-2.87, 4.54]$. Relative to the unweighted distribution, this distribution is shifted rightward, reflecting the fact that some of the world's most populous countries — especially India, Indonesia, Bangladesh, and the Philippines — have above average risk aversion.

Maps of nationally-representative average risk aversion by country are presented in Figure 4. Panel (a) displays country medians and panel (b) country means. Several patterns are evident. Across almost all countries, the median individual is moderately risk averse, with

the exception of a cluster of countries in the Middle East and Africa exhibiting median risk neutrality or risk-seeking. South and South East Asia contain a cluster of especially risk averse countries. Turning to the mean, we see a nearly-uniform global shift towards risk-seeking, indicating substantial left skew in almost every country. We can also see that some country pairs which have very similar median risk aversion, such as the United and China, or Mexico and Colombia, have very different means, indicating that there are differential patterns of preference heterogeneity between countries. We provide complete distributional statistics by country in Appendix A.2.[11]

3.2 The effects of climate change on risk aversion

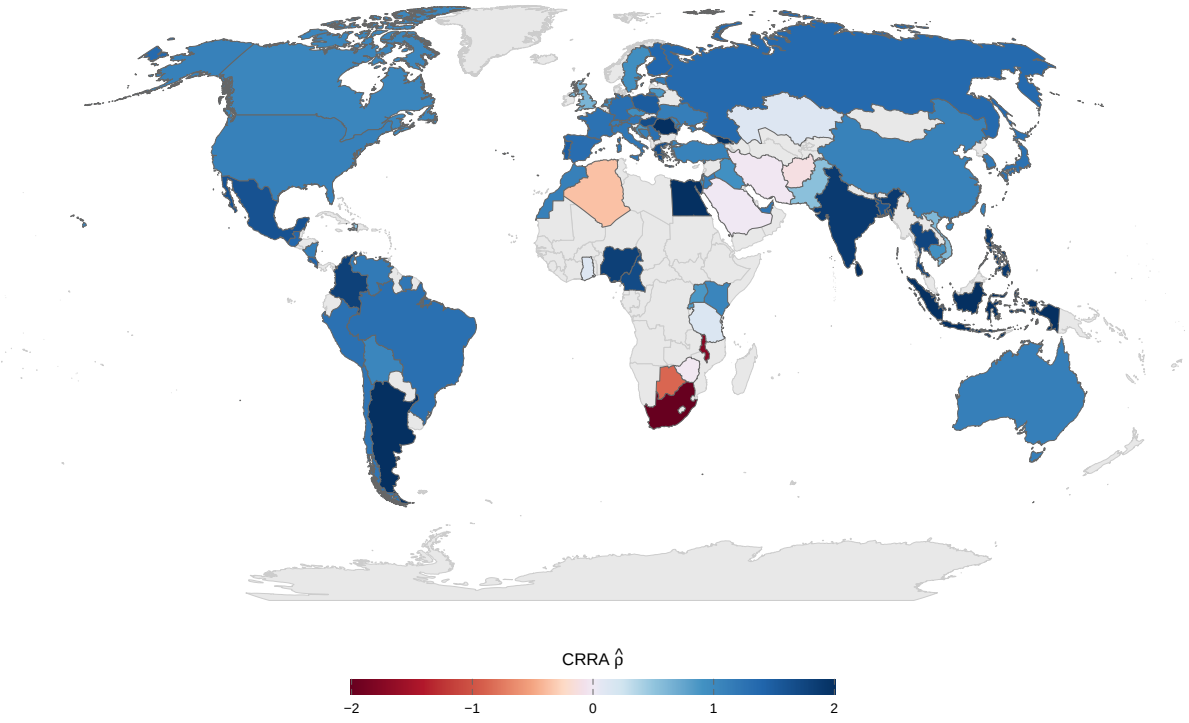
Table 1 presents our main results. Each column reports the coefficient on lifetime mean temperature from a regression of risk aversion on experienced climate, with progressively more demanding fixed effects specifications. Panel A shows the results for the raw staircase measure (1–32 scale, normalized so higher values indicate greater risk aversion), a reduced-form ordinal measure of risky lottery choices. Panel B displays the results for the structurally estimated CRRA coefficient $\hat{\rho}$. Specifications for the three columns are detailed in Section 2.8. All specifications include region fixed effects, as well as controls for gender, interview language dummies, lifetime mean precipitation, and hemisphere-by-interview-month seasonality fixed effects. Standard errors are clustered two-way by survey region and birth year.

The results are consistent across panels and measures: higher lifetime mean temperature is associated with lower risk aversion. In the region fixed effects specification (Column 1), a 1°C increase in lifetime mean temperature is associated with a 1.88-point reduction in staircase risk aversion ($p < 0.01$) and a 0.622-unit decrease in $\hat{\rho}$ ($p < 0.01$).

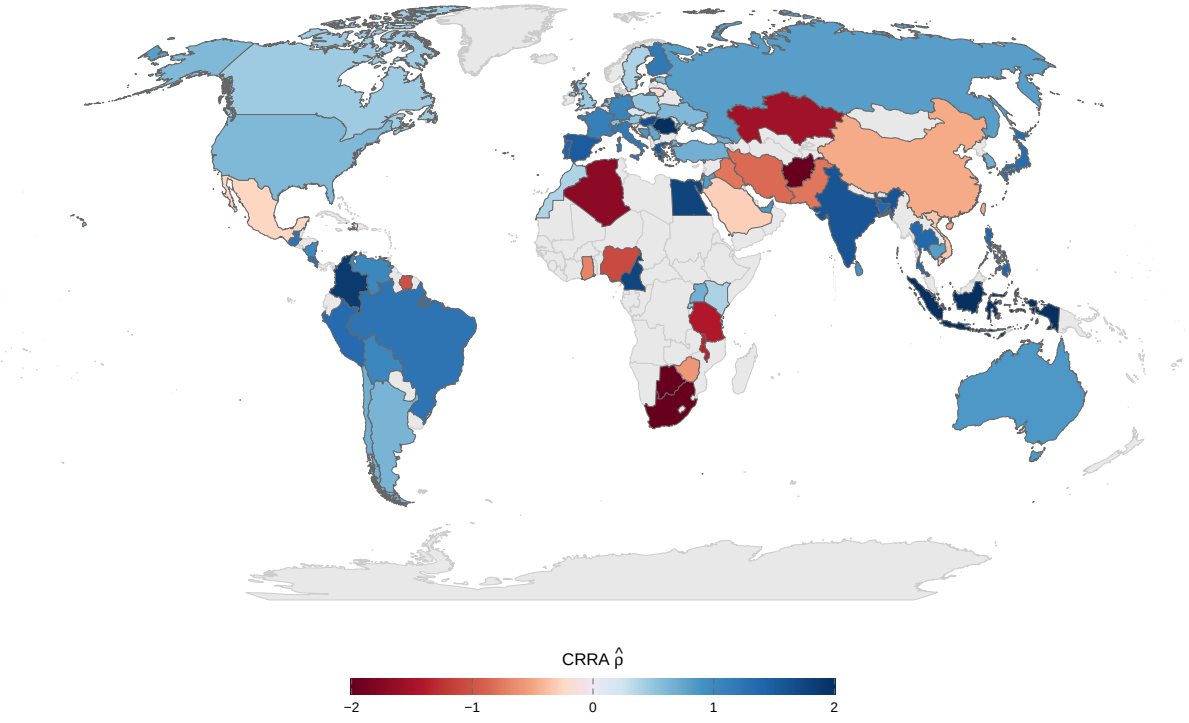
In our preferred specification (Column 2), we add cohort fixed effects, absorbing both cross-region differences in climate and generational shifts in risk tolerance (due to age or

[11]We recommend that researchers employing these data in applications use the country medians for the average country risk aversion rather than the means, since the means are sensitive to extreme values of $\hat{\rho}$. For similar reasons, we recommend the interdecile range for cross-country estimates of preference heterogeneity.

Figure 4: Distribution of Relative Risk Aversion by Country, 2012



(a) Median $\hat{\rho}$ by country



(b) Mean $\hat{\rho}$ by country

other factors) that are common across all regions. Even with this substantial reduction in identifying variation, the $\hat{\rho}$ coefficient remains negative and precisely estimated: $\hat{\beta} = -0.318$ ($p < 0.01$). This implies that a 1°C increase in lifetime mean temperature is associated with a 0.318 (95% CI: [-0.419,-0.218]) unit decrease in risk aversion $\hat{\rho}$. The staircase coefficient is also negative in this specification, but is imprecise (-0.196, SE $= 0.310$).

In Column 3 we replace the cohort fixed effects with country-by-cohort fixed effects. These absorb not only global generational trends but also country-specific cohort effects. Under this specification $\hat{\rho}$ coefficient remains negative (-0.177, $p < 0.10$), though less precisely estimated, while the staircase coefficient is grows larger and becomes significant at conventional levels (-0.722, $p < 0.05$).

Although it is not the main focus of the paper, we can also estimate the effects of changes in lifetime precipitation (per cm) on risk preferences from the same regression (Appendix A.3). $\hat{\beta}_P$ from the preferred specification is equal to -0.087 ($p < 0.01$).

3.3 Historical projections of the risk preference distribution

Our main results tell us about the average *marginal effect* of climate change on risk preferences. In this section, we use these estimates to extrapolate to its *total effect*. Our method is in the spirit of work in the economics of climate that projects impacts into the future based on present day estimates (Burke et al. (2015b), Carleton et al. (2022)). In our case we use our estimates and known realizations of global climate and project its effects into the near past, specifically the period from 1981 to 2026.[12] Note that this is a partial equilibrium exercise, both because we are projecting the estimated relationship for 2012 to other periods where it might differ, and because we are holding constant characteristics, locations, and weighting schemes across the counterfactual samples.

[12]We focus on this period for three reasons. First, 1980 is a common benchmark in the climate change literature, since most warming has occurred since then. Second, this period is symmetric in warming with respect to our preference data, since approximately the same amount of warming has taken place between 1981 and 2021 as between 2012 and 2026. Third, our method involves reconstructing counterfactual lifetime experiences for subjects as if they were measured in different years, and 1981 is the earliest year for which we have sufficient historical climate data to cover the oldest subjects in the sample.

Table 1: Lifetime Mean Temperature and Risk Aversion

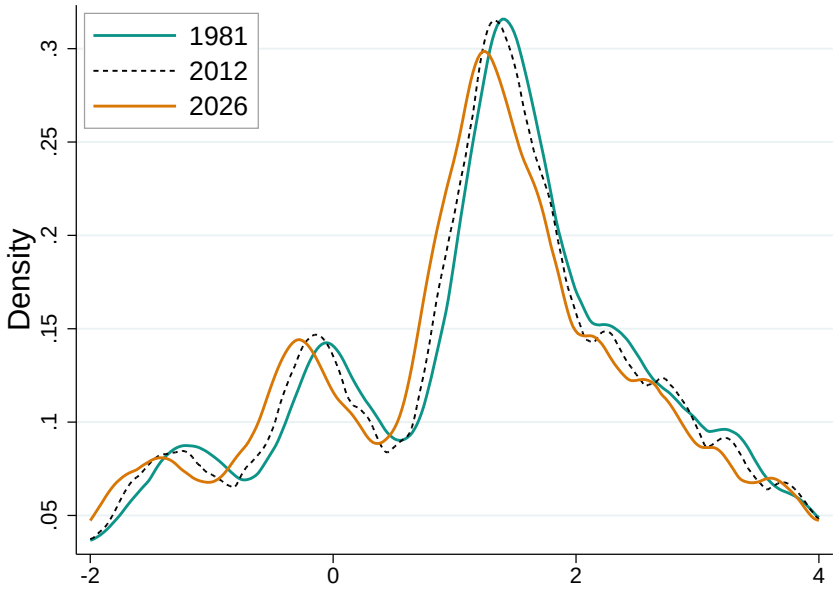
	(1)	(2)	(3)
Panel A: Staircase Risk Aversion			
Mean Lifetime Temp	-1.883***	-0.196	-0.722**
	(0.608)	(0.310)	(0.354)
R^2	0.153	0.159	0.212
Panel B: CRRA $\hat{\rho}$			
Mean Lifetime Temp	-0.622***	-0.318***	-0.177*
	(0.075)	(0.051)	(0.100)
R^2	0.124	0.126	0.177
N	78,221	78,221	77,828
Region FE	X	X	X
Cohort FE		X	
Country × Cohort FE			X

Notes: The dependent variable in panel A is the GPS staircase lottery measure (1–32 scale) normalized so higher numbers mean more risk aversion. The dependent variable in Panel B is the CRRA $\hat{\rho}$ estimated from the staircase measure using the procedure in Section 2.3. Empirical specifications for the three columns can be found in Section 2.8. Regressions are at the survey region level. All specifications control for gender, language dummies, lifetime mean precipitation, and hemisphere × interview-month seasonality fixed effects. Standard errors, clustered two-way by region and birth year, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

We assign to the sample of subjects in the GPS counterfactual lifetime climate experiences as if they were measured at the same age, in the same location, at a different point in time. For example, a 40-year-old measured in 2012, who experienced climate during 1972–2011, is assigned climate experiences in the same region 1941–1980 and 1986–2025. We then use the estimates from our preferred specification (Table 1, Panel B, Column 2) to predict their preferences in 1981 and 2026, using their measured preferences in 2012 as a starting point.

Formally, for each respondent with birth year b surveyed in 2012 in region r, we define shifted birth years $b' = b + (t - 2012)$ for endpoint years $t \in \{1981, 2026\}$. The lifetime experience window for mean temperature T and precipitation P runs from January of year b'

Figure 5: Estimated change in global $\hat{\rho}$ distribution due to climate change, 1981 – 2026



through December of year $t-1$. The projected preferences $\tilde{\rho}$ are

$$\tilde{\rho}^t_{i,r,b} = \hat{\rho}^{2012}_i + \hat{\beta}_T \cdot T^t_{i,r,b'} + \hat{\beta}_P \cdot P^t_{i,r,b'} \tag{4}$$

$$= \hat{\rho}^{2012}_i - 0.318 \cdot T^t_{i,r,b'} - 0.087 \cdot P^t_{i,r,b'}. \tag{5}$$

Results for the unweighted distribution of $\hat{\rho}$ are presented in Figure 5, trimmed to $\hat{\rho} = [-2,4]$ for clarity. As expected, the preference distribution shifts to the left, towards risk-seeking, over time. Its median and mean in 1981 are 0.44 and 1.204; 0.36 and 1.12 in 2012; and 0.23 and 1.00 in 2026.

Figure 1 and Table 7 present the results for country median $\tilde{\rho}$ changes, using the nationally-representative $\hat{\rho}$ distribution. These range from -0.721 (Afghanistan) to 0.072 (Indonesia), with an average country median of -0.208. Only three countries — Colombia (0.018), Cameroon (0.024), and Indonesia — experience increases in risk aversion due to climate change over the period. We also estimate the overall change in global risk aversion, using the globally-representative preference distribution. Its median is 1.27 in 2012, while its

counterfactual median is 1.381 in 1981 and 1.161 in 2026. This represents a 0.22 unit or 15.9% reduction in risk aversion over the 45 year period.

4 Robustness

We subject our main estimates to six categories of robustness checks, each targeting a specific concern about the empirical specification. Here we summarize the results, focusing in the narrative on Columns 2 (region + cohort fixed effects) and 3 (region + country×cohort fixed effects) for both the staircase measure and the structurally estimated $\hat{\rho}$.

In Appendix A.5 we examine the robustness of the $\hat{\rho}$ estimates to **alternative shrinkage regimes**. Our baseline uses normal empirical Bayes shrinkage to reduce noise in individual-level CRRA estimates. When we instead use the unshrunk raw estimates, the Column 2 coefficient on $\hat{\rho}$ grows to -0.505 and the Column 3 coefficient to -0.851 ($p < 0.01$). Winsorizing at the 2.5th/97.5th percentiles without shrinkage yields qualitatively consistent results (-0.474 and -0.808). Both the t-distribution EB ($\nu = 5$) and the mixture model shrinkage produce larger coefficients (-0.861 and -0.515 in Column 2, both $p < 0.01$). The pattern is consistent with shrinkage compressing the dependent variable and mechanically attenuating the regression coefficient: all alternative CRRA measures yield effects that are at least as large as the baseline, and all preserve the sign and significance.

Appendix A.6 presents the results of **restricting the sample to subjects in the interior of the staircase scale**, which is censored at both extremes. Subjects who always chose the lottery (bins 1–2) make first-order stochastically dominated choices. Those who always chose the safe option (bin 32) are at the maximum risk aversion boundary. In Column 2, excluding FOSD violators attenuates both estimates modestly (staircase: -0.308; $\hat{\rho}$: -0.227, $p < 0.01$), while excluding the most risk-averse subjects amplifies them (staircase: -0.243; $\hat{\rho}$: -0.391, $p < 0.01$). Excluding both boundary groups yields intermediate estimates (-0.228 and -0.341, $p < 0.01$). Column 3 shows a similar pattern for the staircase (-0.935

to -0.352 across exclusions) and for $\hat{\rho}$ (-0.156 to -0.277). The sign is preserved across all specifications, indicating that the result is not driven by subjects making extreme or dominated lottery choices.

In Appendix A.7 we show robustness of the results to **including alternative sets of individual controls** from the GPS. Adding subjective math ability to the baseline controls leaves both Column 2 estimates virtually unchanged (staircase: -0.206; $\hat{\rho}$: -0.319, $p < 0.01$), as does the further inclusion of all five GPS preference indices — patience, trust, altruism, and positive and negative reciprocity (staircase: -0.186; $\hat{\rho}$: -0.321, $p < 0.01$). The Column 3 estimates are similarly stable, and in fact become more precisely estimated with the full set of preference controls (staircase: -0.886, $p < 0.05$; $\hat{\rho}$: -0.225, $p < 0.05$). The insensitivity to these controls indicates that the temperature–risk aversion relationship is not mediated by numeracy or correlated shifts in other preference domains. Removing gender and language controls reduces precision in Column 2 but leaves the Column 3 staircase estimate significant (-0.607, $p < 0.05$), suggesting that the within-country variation driving Column 3 is less dependent on demographic controls for identification.

Appendix A.8 demonstrates robustness to **alternate climate and weather controls**. Adding interview-day temperature anomalies (the day-of-interview mean temperature and the night-before minimum temperature, both as z-scores relative to the region's 1981–2010 climatology) barely moves the estimates: Column 2 $\hat{\rho}$ goes from -0.318 to -0.317, and the staircase from -0.196 to -0.196. Further adding the lifetime standard deviations of temperature and precipitation yields -0.314 and -0.182 in Column 2. This confirms that the estimated effect of lifetime mean temperature is not confounded by acute weather shocks or by lifetime climate variability. Dropping lifetime mean precipitation attenuates the $\hat{\rho}$ estimate slightly (-0.292, $p < 0.01$), reflecting modest correlation between temperature and precipitation trends, and weakens the staircase estimate (-0.123, not significant).

In Appendix A.9 we show robustness of the results to **alternative seasonality controls**. Our baseline absorbs hemisphere $\times$ interview-month fixed effects. Dropping these entirely

leaves the Column 2 $\hat{\rho}$ estimate qualitatively consistent (-0.255) but insignificant, while the staircase estimate attenuates to zero (0.004). Replacing them with country $\times$ interview-month FEs yields Column 2 estimates of -0.696 ($p < 0.05$, staircase) and -0.125 (not significant, $\hat{\rho}$), as these finer controls absorb a larger share of the cross-cohort variation. The Column 3 estimates are notably more stable across all seasonality specifications, with the staircase ranging from -0.668 to -0.732 (all $p < 0.05$ or better) and $\hat{\rho}$ from -0.176 to -0.204 (all $p < 0.10$). This is consistent with the within-country identifying variation in Column 3 being less sensitive to the treatment of interview timing than the cross-country variation that enters Column 2.

Finally, Appendix A.10 shows robustness to **alternative clustering**. The point estimates are mechanically invariant to the clustering choice; only the standard errors change. Our baseline two-way clustering by region and birth year is the most conservative specification. Clustering by country alone yields the tightest Column 2 standard errors for $\hat{\rho}$ (SE = 0.052), while region $\times$ cohort clustering produces the widest (SE = 0.100). The $\hat{\rho}$ coefficient is significant at the 1% level under all four alternative clustering schemes in Column 2. The staircase coefficient remains insignificant in Column 2 regardless of clustering, while the Column 3 staircase estimate (-0.722) ranges from marginally significant ($p < 0.10$, country clustering) to significant at the 5% level (baseline and region $\times$ cohort clustering).

5 National panel analyses

The analysis to this point has had the advantage of being global in scope. It is limited, however, in that we are estimating the effects of climate change on preference changes from cross-sectional preference data. In addition, since the public GPS data does not contain information on subjects' region of birth, it is possible that the estimates may reflect endogenous migration, plausibly correlated with both climate and risk preferences, rather than the causal effects we are after.

To support the global estimates, in this section we report the results from analyses of panel surveys from four countries: Indonesia, Mexico, Chile, and Japan. All four data sets contain repeat elicited measures of risk aversion for the same individuals years apart. These measures involve selections between hypothetical monetary lotteries, similar to the staircase measure in the GPS. Because of their panel structure we can derive within-subject estimates in these data using individual fixed effects. The panels also contain information on subjects' location of birth, allowing us to control for endogenous migration by using climate experiences in a location that subjects do not themselves choose.

5.1 Data sources

Our data source for **Indonesia** is the Indonesian Family Life Survey (IFLS) (Strauss et al., 2009, 2016). The IFLS is a longitudinal study administered by the RAND corporation in 13 provinces in Indonesia in five waves, starting in 1993. IFLS4 (2007-2008) and IFLS5 (2014) contain staircase measures of risk aversion, using three adaptive choices over income lotteries, where the safe amount is fixed and the 50/50 lottery amounts vary. This yields a measure with five bins, ordered so that higher numbers correspond to more risk aversion. The risk aversion instrument in Indonesia is identical in the two IFLS waves. We restrict the analysis to subjects in these two waves with recorded responses for the risk aversion measure in both, as well as information on birth province and birth-year. The result is a sample of 21,057 subjects, each observed twice, for a total of 42,114 subject $\times$ wave observations.

In **Mexico** we use data from the Mexican Family Life Survey (MxFLS), a longitudinal study administered in 16 states in three waves starting in 2002, and managed by the Iberoamerican University (UIA) and the Center for Economic Research and Teaching (CIDE). MxFLS2 (2005-2006) (Rubalcava and Teruel (2008)) and MxFLS3 (2009-2012) (Rubalcava and Teruel (2013)) contain similar staircase income lottery instruments to those in the IFLS. In MxFLS2 the instrument has five bins, while in MxFLS3 it has seven bins, which we collapse to five using the procedure in Brown et al. (2019a). Our final sample contains 12,294

subjects, each observed twice, for a total of 24,588 subject × wave observations.

The data for **Chile** comes from the Chilean Social Protection Survey (EPS), a national longitudinal panel survey launched in 2002 through the Ministry of Labor and Social Welfare, in collaboration with the University of Chile (Bravo et al. (2008)). The EPS waves in 2004, 2006, and 2009 contain identical hypothetical job choice questions, where subjects choose between a position with a guaranteed income, and one with a risky income. The instrument employs a staircase design with two layers, resulting in a measure with four bins. Our final sample contains 15,496 subjects, each appearing in 2.62 waves on average, for a total of 40,672 subject × wave observations.

For **Japan** we use data from the Preference Parameters Study (PPS), a Japanese household panel survey administered by the Institute of Social and Economic Research (ISER) at Osaka University. The survey was conducted annually from 2003-2013, 2016-2018, and 2021-2024, and contains a rich set of elicited preference measures. In the current analysis we use the "speed lottery" measure employed by Hanaoka et al. (2018): subjects are asked to imagine that they have a chance to buy a lottery ticket with a 50% chance of winning ¥100,000, and given a multiple price list of seven possible amounts that they would be willing to pay for it, ranging from ¥10 to ¥50,000. Using choices in this instrument, we construct an ordinal risk aversion measure with 9 bins. We use data from all ten PPS waves from 2011 to 2024, in which the speed lottery instrument is identical. For subjects' location we use their prefecture at age 15. Our final sample consists of 4,980 subjects, each appearing in 6 waves on average, for a total of 29,949 subject × wave observations.

The language and distributions of the risk aversion instruments in all panels are available in Appendix A.11.

5.2 Empirical specification

The specification in the panel surveys closely mirrors the global specifications containing cohort fixed effects, but substituting for them with individual fixed effects. We construct

climate experience measures using the same data and method as in the global analysis, described in Section 2.5, based on subjects' birth location (or location at 15 in Japan). For each risk aversion measure RA and subject i, measured in the survey in year t, we estimate

$$RA_{i,t} = \beta \bar{T}_{i,t} + \gamma \bar{P}_{i,t} + \alpha_i + \alpha_t + \epsilon_{i,t}, \tag{6}$$

where $\bar{T}_{i,t}$ is mean monthly temperature in subject i's region of birth, starting in January of their birth year and concluding in December of the year $t-1$. $\bar{P}_{i,t}$ is the equivalent measure for precipitation, α_i is an individual fixed effect, and α_t is a survey wave year (time) fixed effect. Standard errors are clustered at the birth-region by birth-year level.

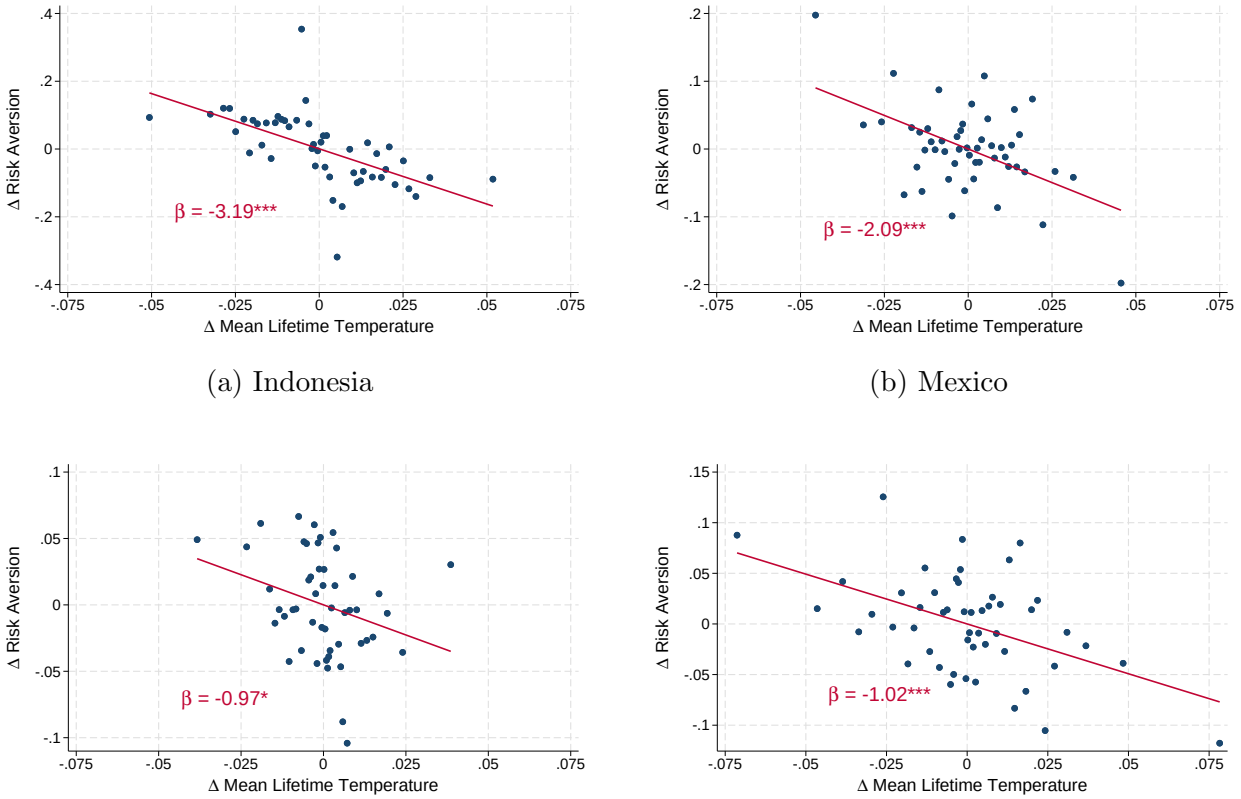
5.3 Results

Figure 6 presents binned scatterplots of within-person changes in risk aversion against within-person changes in lifetime mean temperature, separately for each of the four national panels. The full regression table is also available in Appendix A.13. Each plot residualizes both variables on individual fixed effects, year fixed effects, and lifetime mean precipitation, so that the slope of the fitted line equals the coefficient $\hat{\beta}$ from the full panel regression.

In all four countries, the relationship is negative and significant: individuals who experience larger increases in lifetime temperature between survey waves exhibit larger decreases in risk aversion. In Indonesia, a one-degree-Celsius increase in lifetime mean temperature is associated with a 3.19-unit decrease in risk aversion ($p < 0.01$), on a five-bin scale. In Mexico, the corresponding estimate is -2.09 ($p < 0.01$) on the harmonized five-bin scale (note that coefficient magnitudes are not comparable across surveys, because of the different instruments they use). In Chile, the estimate is -0.97 ($p = 0.061$) on a four-bin scale, marginally significant. In Japan, using the nine-bin speed lottery measure across ten waves spanning 2011–2024, the estimate is -1.02 ($p < 0.01$).[13] We also see broad consistency with

[13] Hanaoka et al. (2018), use a structural estimate of risk aversion based on the speed lottery to show that the Great Eastern Japan Earthquake in 2011 changed the risk preferences of subjects between the 2011 and

Figure 6: Change in lifetime temperature and change in risk aversion in four national panels



(a) Indonesia

(b) Mexico

(c) Chile

(d) Japan

Notes: Each panel plots means of risk aversion against lifetime mean temperature, both residualized on individual fixed effects, wave fixed effects, and lifetime mean precipitation. Effect sizes not comparable across panels. The slope of the fitted line equals the coefficient on lifetime temperature from the full panel regression. 50 quantile bins. Standard errors clustered at birth region × birth year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

the precipitation results in the main analysis, with three of the four estimates significant and negative, except in Chile, where the precipitation coefficient is significant and positive (Appendix A.12).

The consistency of the sign of the effects across the four countries is notable. Indonesia, Mexico, Chile, and Japan are highly diverse along every margin we might expect to be muddying the global estimates. They are at different development levels (low for Indonesia, middle-income for Mexico and Chile, high income for Japan); demographically diverse; geographically and climactically disparate; and have very distinct cultures, religions, and

2012 waves. As a robustness check, in Appendix A.14 we estimate our results excluding these two waves of the survey, and find that the temperature effects become stronger.

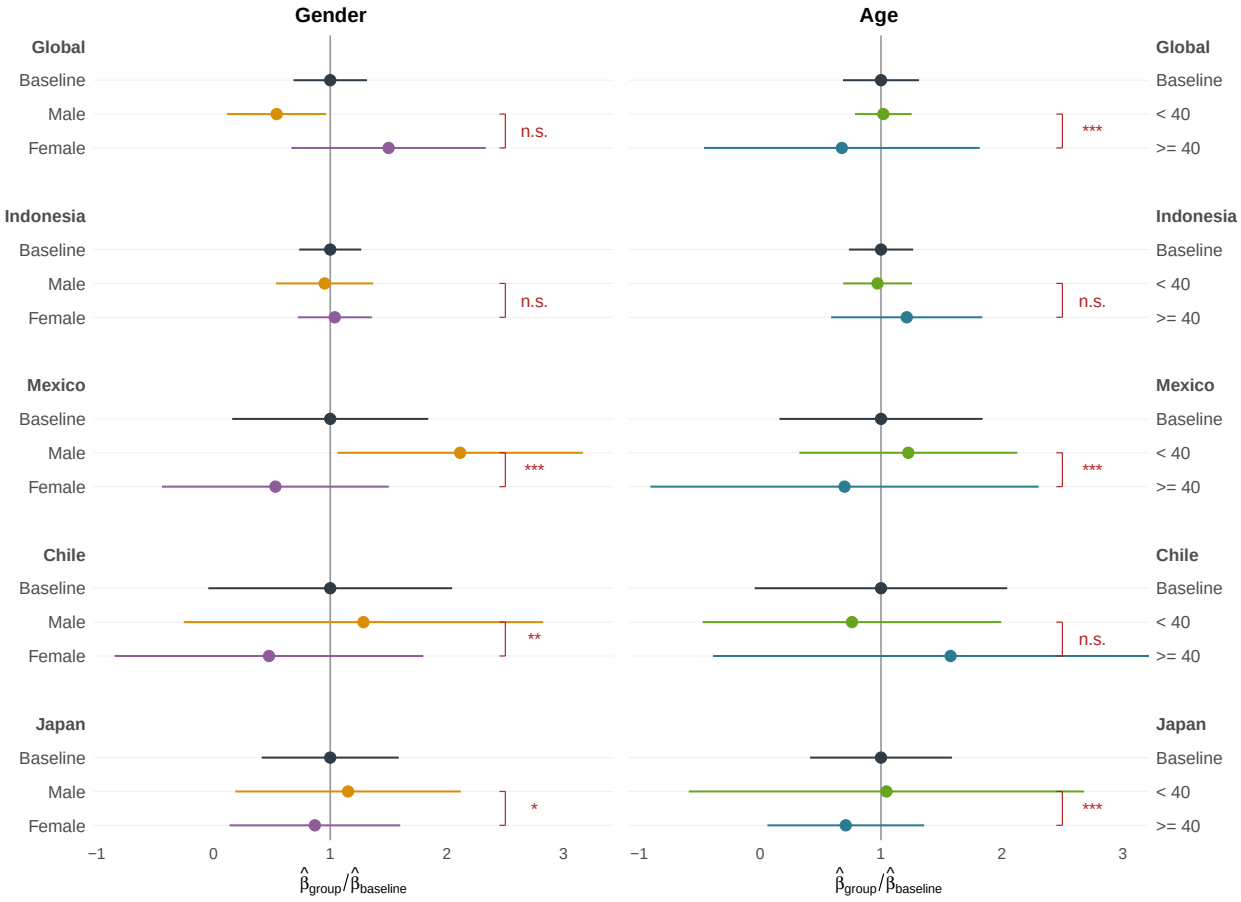
histories. The fact that the demanding within-person specification for the effects is consistent across these settings supports the finding that rising temperatures have consistent results across many diverse environments, as our global results indicate.

6 Heterogeneity of effects globally and within-country

In this section we examine heterogeneity in the lifetime temperature and risk aversion relationship in our global and country-panel samples. In the global data we use the lifetime mean temperature coefficient for $\hat{\rho}$ from our preferred specification, Table 1, Panel B, Column 2. In the country panels we use the lifetime mean temperature coefficient for the reduced-form measures from Figure 6. Our focus is on three dimensions of heterogeneity which are known correlates of risk aversion, and which we can measure consistently across settings: gender, age, and income.

The global analysis and the country panels appear to tell different stories about the effects by gender (Figure 7). In the global data, the point estimate for women ($\hat{\beta} = -0.477$, 95% CI: $[-0.742, -0.212]$) is nearly three times larger in magnitude than for men (-0.172, $[-0.307, -0.037]$), though the difference is not statistically significant ($p = 0.42$). In contrast, men show significantly larger effects than women in Mexico ($p < 0.01$) and Chile ($p < 0.05$), with a marginal difference in the same direction in Japan ($p = 0.054$), while Indonesia shows no gender heterogeneity ($p = 0.19$). Notably, although the results appear in conflict with each other, they are consistent with the current literature on differential preference changes. Since men are on average more risk-seeking than women, the global results suggest preference convergence under climate change, in line with Falk and Hermle (2018) evidence of gender preference convergence under the process of development. The panel results, meanwhile, suggest preference divergence, which is consistent with the findings in Hanaoka et al. (2018) about the differential effects of a natural disaster on the risk preferences of men and women. More work is clearly needed on this front.

Figure 7: Heterogeneity of estimates by gender and age



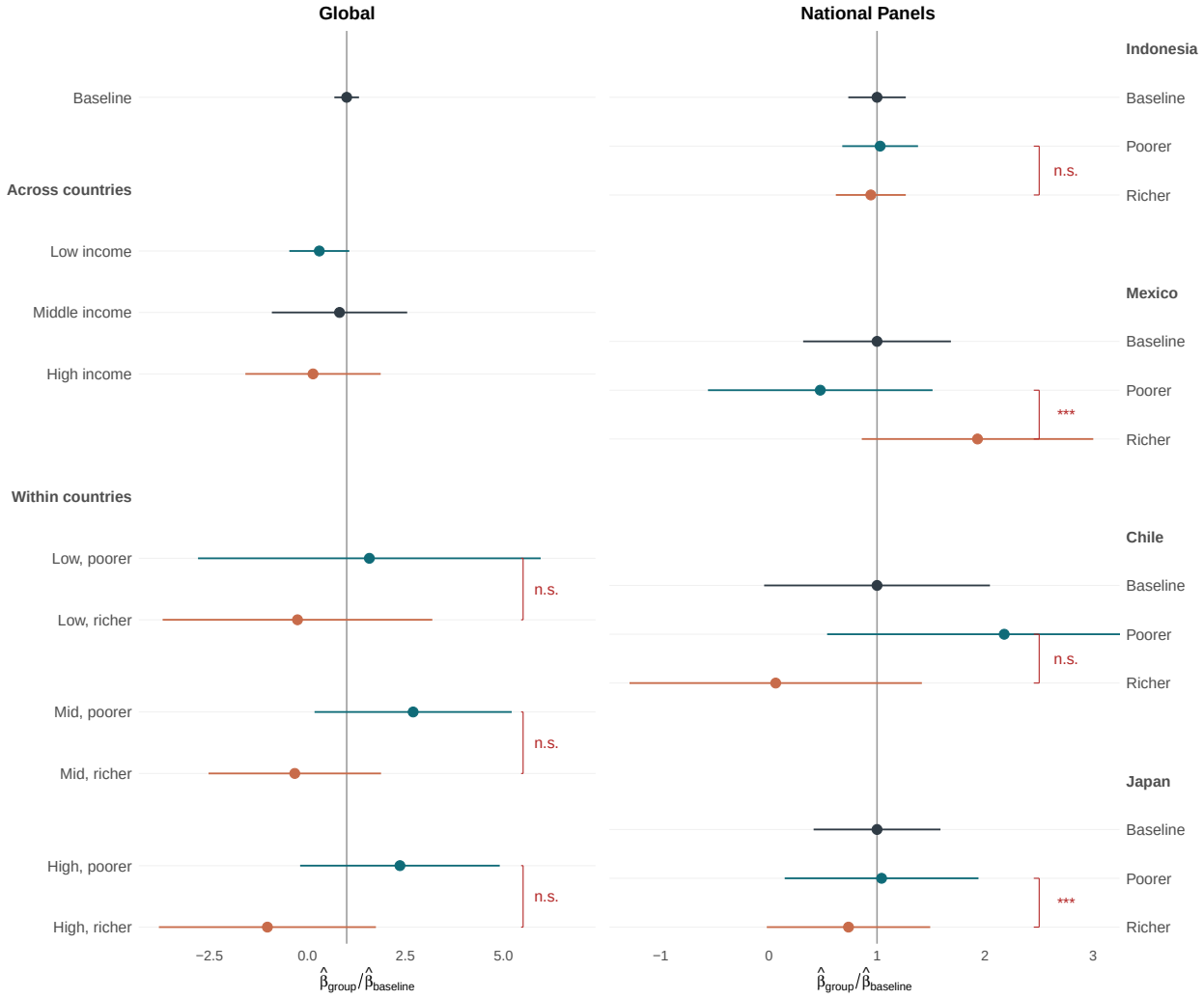
Notes: Estimates are regression coefficients from Table 1, Panel B, Column 2 (global) and Figure 6 (country panels), by subgroup. Age in panels is at midpoint between subject's first and last appearance. Coefficients and CIs are normalized by the baseline, so values greater than 1 indicate a larger effect size in absolute value. Significance is for differences between sub-groups. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$, n.s. $p > 0.10$.

The differential effects of age (Figure 7) are more consistent across settings. Splitting the global sample at the median age of 40, younger respondents show a significantly larger effect ($\hat{\beta} = -0.324$, $[-0.399, -0.248]$) than older respondents (-0.215, $[-0.577, 0.148]$; interaction $p < 0.01$). In the panels, Mexico ($p < 0.01$) and Japan ($p < 0.001$) show significantly larger effects for younger individuals at the panel midpoint, while Indonesia and Chile show no significant age heterogeneity. These results are consistent with younger individuals being more responsive to temperature changes, but interpretation requires some care given the fixed effects structure. Although global temperature has been steadily increasing over time, the cohort (global) and individual (panels) fixed effects should absorb variation from secular trends in the independent variables. More likely explanations of the differences in the marginal effects between the young and the old is that the young might be more plastic in their preferences, and/or that the effects of temperature on risk aversion could have a non-linear component, where additional warming amplifies the effects of prior warming.

Figure 8 presents results on heterogeneity by income. In the global data, we examine this at two levels: (1) by splitting the sample of countries into development categories using the 2012 World Bank PIP median daily per capita consumption (low <\$5, middle \$5–20, high >\$20/day); and (2) within each of these categories, splitting the sample by whether its subjects' regional GNI per capita (from the Chrisendo et al. (2025) sub-national dataset) is above or below its country's median regional GNI per capita across regions in the sample. The coarser split does not result in a clear income gradient: middle income countries exhibit the strongest effects, followed by lower-income countries, and then high income countries. The finer splits tell a much clearer story: across all three country-income bins, subjects in poorer regions show stronger effects than subjects in richer regions, although in all three cases the differences are not statistically significant.

In the country panels, we calculate for each subject their mean household income across all panel appearances, and then split them based on the median of these across subjects in the

Figure 8: Heterogeneity of estimates by income



Notes: Estimates are regression coefficients from Table 1, Panel B, Column 2 (global) and Figure 6 (country panels), by subgroup. The left panel displays a 3-way split in the global data based on subjects' country income group (low/middle/high, based on PIP median daily consumption in 2011 PPP$), and a further binary split within each group based on whether subjects' region is above or below median region GNI per capita among regions in their country that are in the sample. The right panel splits subjects on within-sample median income in the national panels, where subject income is the mean of income across appearances in the panel. Coefficients and CIs are normalized by the baseline, so values greater than 1 indicate a larger effect size in absolute value. Significance is for differences between sub-groups. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$, n.s. $p > 0.10$.

sample.[14] In Japan and Chile we see a similar pattern to the one in the global analysis, with poorer subjects having larger effects, though only Japan's difference ($p < 0.01$) is statistically significant, while Chile's is merely suggestive ($p = 0.16$). In Indonesia the pattern is nominally the same, but the difference is far from significant ($p = 0.94$). Mexico is a notable exception: the effect is concentrated among richer households ($p < 0.001$), the opposite of the pattern in other settings.

Taken together, the results across settings suggest that the effects of lifetime temperature on risk preferences are likely stronger for lower-income individuals, with some possible exceptions like Mexico. These are exactly the individuals that have the fewest resources at their disposal to adapt to environmental change, and who typically bear the brunt of climate change overall.

7 Conclusion

In October of 1943, the British Parliament was debating how to rebuild the House of Commons, destroyed in the Blitz two years prior. Its old chamber had been small and crowded, with government and opposition facing each other across a narrow floor. Reformers wanted to modernize: a larger chamber, semi-circular in the continental style, with a seat for every member. Opening the debate, Winston Churchill rose to defend the old design. A smaller room made disagreements closer, sharper, more human in scale. It made Parliament a living institution, he argued, one with the resilience and stability to govern through crisis. The physical form of the old chamber had molded Parliament's very habit of mind, and with it the character of British democracy. "We shape our buildings," Churchill concluded, "and afterwards our buildings shape us."

Anthropogenic climate change is the most significant transformation to our natural environment that humanity has ever made — our grandest building yet. It would be hubris

[14]The Indonesia, Mexico, and Japan surveys elicit total household income directly. In Chile, total monthly household income is constructed by aggregating labor income, pensions, government transfers, private transfers, and capital income across the 2004, 2006, and 2009 waves.

to expect that such a sweeping change to our habitat will not have consequences for our own nature. In fact, we have significant evidence from evolutionary biology that these kinds of reciprocal dynamics are common, including in humans. This is the core precept of the theory of niche construction (Odling-Smee et al. (1996)). Organisms do not merely adapt to environments. They alter them, and those alterations then become part of the selective landscape they face. In humans this mechanism interacts with cultural and cognitive evolution (Laland et al. (2001), Kendal et al. (2011)), and it is likely to be amplified by the scale of the changes that we are bringing about.

The most significant open question in our analysis is the mechanism through which long-run warming is affecting risk aversion. Here the scope of our findings is instructive: whichever mechanisms are in operation, they must be affecting human populations globally. This suggests to us that the most likely candidates are physical or neurological changes induced by warming, consistent with evidence on these effects in animals (Moiroux et al. (2015), Gutiérrez et al. (2023)), and in the related domain of conflict, impulsivity, and aggression in humans (Burke et al. (2015a), Tiihonen et al. (2017)). The fact that the effects are generally larger for lower-income individuals in our data, who are most exposed to the physical costs of climate change, is also supportive of this hypothesis. Future work should endeavor to answer this question more fully.

Another important area for future work is the implications of our results for the analysis of climate policy. On the positive side, work on climate politics that treats risk preferences as fixed may misjudge the trajectory of public support for mitigation, since the warming these policies aim to address may itself reduce demand for them. On the normative side, the issue is deeper. If climate change alters the preferences used to evaluate climate damages, welfare analyses must confront the question of which preferences should be given normative authority — those that prevail before the damages occur, those that emerge after individuals adapt to them, or those that evolve along the warming path. Although these questions are beyond the scope of the current paper, our findings indicate they can no longer be set aside.

References

Steffen Andersen, James C. Cox, Glenn W. Harrison, Morten I. Lau, E. Elisabet Rutström, and Vjollca Sadiraj. Asset Integration and Attitudes to Risk: Theory and Evidence. *Review of Economics and Statistics*, 100(5):816–830, 2018a. doi: 10.1162/rest_a_00746. (Cited on page 15.)

Steffen Andersen, Glenn W. Harrison, Morten Igel Lau, and E. Elisabet Rutström. Multi-attribute Utility Theory, Intertemporal Utility, and Correlation Aversion. *International Economic Review*, 59(3):537–555, 2018b. doi: 10.1111/iere.12279. (Cited on page 103.)

Joshua D. Angrist, Peter D. Hull, Parag A. Pathak, and Christopher R. Walters. Inputs and the Education Production Function. *American Economic Review*, 107(5):132–136, 2017. doi: 10.1257/aer.p20171038. (Cited on pages 16, 109, and 111.)

David Anthoff, Richard S. J. Tol, and Gary W. Yohe. Risk aversion, time preference, and the social cost of carbon. *Environmental Research Letters*, 4(2):024002, April 2009. ISSN 1748-9326. doi: 10.1088/1748-9326/4/2/024002. URL https://iopscience.iop.org/article/10.1088/1748-9326/4/2/024002. (Cited on page 3.)

Jose Apesteguia and Miguel A Ballester. Monotone Stochastic Choice Models: The Case of Risk and Time Preferences. *Journal of Political Economy*, 126(1):74–106, 2018. doi: 10.1086/695504. (Cited on pages 4, 15, 105, and 106.)

Konstantin Ash and Nick Obradovich. Climatic Stress, Internal Migration, and Syrian Civil War Onset. *Journal of Conflict Resolution*, pages 1–29, 2019. ISSN 15528766. doi: 10.1177/0022002719864140. (Cited on page 20.)

Robert B. Barsky, F. Thomas Juster, Miles S. Kimball, and Matthew D. Shapiro. Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study. *The Quarterly Journal of Economics*, 112(2):537–579, 1997. ISSN 00335533. doi: 10.1162/003355397555280. ISBN: 00335533. (Cited on page 10.)

Lukas Beck. Endogenous preferences, environmental economics, and welfare. *Politics, Philosophy & Economics*, 2026. ISSN 1470-594X. doi: 10.1177/1470594X251412790. URL https://journals.sagepub.com/doi/10.1177/1470594X251412790. (Cited on page 8.)

Anke Becker, Benjamin Enke, and Armin Falk. Ancient Origins of the Global Variation in Economic Preferences. *AEA Papers and Proceedings*, (110), 2020. (Cited on page 10.)

Michel Beine, Gary Charness, Arnaud Dupuy, and Majlinda Joxhe. A natural experiment on earthquakes and risk preferences. *European Economic Review*, 179:105108, 2025. ISSN 0014-2921. doi: 10.1016/j.euroecorev.2025.105108. URL https://www.sciencedirect.com/science/article/abs/pii/S0014292125001588. (Cited on page 7.)

Timothy Besley and Torsten Persson. The Political Economics of Green Transitions. *The Quarterly Journal of Economics*, 138(3):1863–1906, August 2023. ISSN 0033-5533. doi: 10.1093/qje/qjad006. URL https://academic.oup.com/qje/article/138/3/1863/7000849. (Cited on page 8.)

Anika Bittner, Alexander Haering, Timo Heinrich, and Thomas Mayrhofer. Revalidating a survey instrument for measuring risk preferences. *Economics Letters*, 242:111851, September 2024. ISSN 01651765. doi: 10.1016/j.econlet.2024.111851. URL https://linkinghub.elsevier.com/retrieve/pii/S0165176524003355. (Cited on page 11.)

Ranoua Bouchouicha and Ferdinand M. Vieider. Growth, entrepreneurship, and risk-tolerance: a risk-income paradox. *Journal of Economic Growth*, 24(3):257–282, 2019. ISSN 1381-4338. doi: 10.1007/s10887-019-09168-0. URL https://link.springer.com/article/10.1007/s10887-019-09168-0. (Cited on page 9.)

David Bravo, Jere R. Behrman, Olivia S. Mitchell, Petra E. Todd, and Javiera Vásquez. Encuesta de protección social 2006: Presentación general y principales resultados. Technical report, Department of Economics, University of Chile, 2008. (Cited on page 32.)

Ryan Brown, Verónica Montalva, Duncan Thomas, and Andrea Velásquez. Impact of Violent Crime on Risk Aversion: Evidence from the Mexican Drug War. *Review of Economics and Statistics*, 2019a. URL http://www.nber.org/papers/w23181. (Cited on pages 20 and 31.)

Ryan Brown, Verónica Montalva, Duncan Thomas, and Andrea Velásquez. Impact of violent crime on risk aversion: Evidence from the mexican drug war. *Review of Economics and Statistics*, 101(5):892–904, 2019b. ISSN 15309142. doi: 10.1162/rest_a_00788. (Cited on page 8.)

Marshall Burke and Kyle Emerick. Adaptation to climate change: Evidence from US agriculture. *American Economic Journal: Economic Policy*, 8(3):106–140, 2016. ISSN 1945774X. doi: 10.1257/pol.20130025. (Cited on page 2.)

Marshall Burke, Solomon M Hsiang, and Edward Miguel. Climate and Conflict. *Annual Review of Economics*, 7:577–615, 2015a. doi: 10.1146/annurev-economics-080614-115430. (Cited on pages 2 and 40.)

Marshall Burke, Solomon M. Hsiang, and Edward Miguel. Global non-linear effect of temperature on economic production. *Nature*, 527(7577):235–239, 2015b. doi: 10.1038/nature15725. (Cited on pages 6, 20, and 25.)

Marco Caliendo, Frank Fossen, and Alexander Kritikos. The Impact of Risk Attitudes on Entrepreneurial Survival. *Journal of Economic Behavior & Organization*, 76(1):45–63, October 2010. (Cited on page 10.)

Lisa Cameron and Manisha Shah. Risk-taking behaviour in the wake of natural disasters. *Journal of Human Resources*, 50(2):484–515, 2015. (Cited on pages 8 and 20.)

Tamma Carleton, Amir Jina, Michael Delgado, Michael Greenstone, Trevor Houser, Solomon Hsiang, Andrew Hultgren, Robert E. Kopp, Kelly E. McCusker, Ishan Nath, James Rising, Ashwin Rode, Hee Kwon Seo, Arvid Viaene, Jiacan Yuan, and Alice Tianbo Zhang. Valuing the Global Mortality Consequences of Climate Change Accounting for Adaptation Costs and Benefits. *The Quarterly Journal of Economics*, 137(4):2037–2105, 2022. doi: 10.1093/qje/qjac020. (Cited on pages 6 and 25.)

Alessandra Cassar, Andrew Healy, and Carl V O N Kessler. Trust, Risk, and Time Preferences After a Natural Disaster : Experimental Evidence from Thailand. *World Development*, 94:90–105, 2017. ISSN 0305-750X. doi: 10.1016/j.worlddev.2016.12.042. URL http://dx.doi.org/10.1016/j.worlddev.2016.12.042. (Cited on page 8.)

Cristina Cattaneo and Giovanni Peri. The migration response to increasing temperatures. *Journal of Development Economics*, 122:127–146, September 2016. ISSN 0304-3878. doi: 10.1016/j.jdeveco.2016.05.004. URL https://www.sciencedirect.com/science/article/pii/S0304387816300372. (Cited on page 2.)

Raj Chetty and Nathaniel Hendren. The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects. *Quarterly Journal of Economics*, 133(3):1107–1162, 2018. doi: 10.1093/qje/qjy007. (Cited on pages 4, 15, 109, and 111.)

Raj Chetty, John N. Friedman, and Jonah E. Rockoff. Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates. *American Economic Review*, 104(9): 2593–2632, 2014. doi: 10.1257/aer.104.9.2593. (Cited on pages 15, 109, and 111.)

Taha Choukhmane and Tim de Silva. What Drives Investors' Portfolio Choices? Separating Risk Preferences from Frictions. *The Journal of Finance*, 81(1), February 2026. ISSN 0022-1082. doi: 10.1111/jofi.70013. URL https://onlinelibrary.wiley.com/doi/10.1111/jofi.70013. (Cited on page 2.)

Daniel Chrisendo, Venla Niva, Roman Hoffman, Sina Masoumzadeh Sayyar, Juan Rocha, Vilma Sandström, Frederick Solt, and Matti Kummu. Rising income inequality across half of global population and socioecological implications. *Nature sustainability*, 8:1601–1613, December 2025. (Cited on pages 14, 37, and 100.)

Viktoria Cologna, Simona Meiler, Chahan M. Kropf, Samuel Lüthi, Niels G. Mede, David N. Bresch, Oscar Lecuona, Sebastian Berger, John Besley, Cameron Brick, Marina Joubert, Edward W. Maibach, Sabina Mihelj, Naomi Oreskes, Mike S. Schäfer, Sander van der Linden, and TISP Consortium. Extreme weather event attribution predicts climate policy support across the world. *Nature Climate Change*, 15(7):725–735, July 2025. ISSN 1758-6798. doi: 10.1038/s41558-025-02372-4. URL https://www.nature.com/articles/s41558-025-02372-4. (Cited on pages 2 and 7.)

James C. Cox and Vjollca Sadiraj. Small- and Large-Stakes Risk Aversion: Implications of Concavity Calibration for Decision Theory. *Games and Economic Behavior*, 56(1):45–60, 2006. doi: 10.1016/j.geb.2005.08.001. (Cited on page 14.)

Antoine Dechezleprêtre, Adrien Fabre, Tobias Kruse, Bluebery Planterose, Ana Sanchez Chico, and Stefanie Stantcheva. Fighting Climate Change: International Attitudes toward Climate Policies. *American Economic Review*, 115(4):1258–1300, April 2025. ISSN 0002-8282. doi: 10.1257/aer.20230501. URL https://www.aeaweb.org/articles?id=10.1257/aer.20230501. (Cited on page 2.)

Melissa Dell, Benjamin F. Jones, and Benjamin A. Olken. Temperature shocks and economic growth: Evidence from the last half century. *American Economic Journal: Macroeconomics*, 4(3):66–95, 2012. ISSN 19457707. doi: 10.1257/mac.4.3.66. (Cited on page 20.)

Rebecca DerSimonian and Nan Laird. Meta-Analysis in Clinical Trials. *Controlled Clinical Trials*, 7(3):177–188, 1986. doi: 10.1016/0197-2456(86)90046-2. (Cited on pages 15 and 108.)

Salvatore Di Falco and Ferdinand M. Vieider. Environmental Adaptation of Risk Preferences. *The Economic Journal*, 132(648):2737–2766, 2022. ISSN 0013-0133. doi: 10.1093/ej/ueac030. URL https://academic.oup.com/ej/article/132/648/2737/6581312. (Cited on page 8.)

Thomas Dohmen, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G. Wagner. Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3):522–550, 2011. ISSN 15424766. doi: 10.1111/j.1542-4774.2011.01015.x. ISBN: 1542-4766. (Cited on page 10.)

Thomas Dohmen, Armin Falk, Bart H. H. Golsteyn, David Huffman, and Uwe Sunde. Risk Attitudes Across The Life Course. *The Economic Journal*, 127(605):F95–F116, 2017. ISSN 00130133. doi: 10.1111/ecoj.12322. URL http://doi.wiley.com/10.1111/ecoj.12322. ISBN: 3143388382. (Cited on page 20.)

Stefan Drews and Jeroen C. J. M. van den Bergh. What explains public support for climate policies? A review of empirical and experimental studies. *Climate Policy*, 16(7):855–876, 2016. ISSN 1469-3062. doi: 10.1080/14693062.2015.1058240. URL https://www.tandfonline.com/doi/full/10.1080/14693062.2015.1058240. (Cited on page 7.)

Bradley Efron and Carl Morris. Data Analysis Using Stein's Estimator and its Generalizations. *Journal of the American Statistical Association*, 70(350):311–319, 1975. doi: 10.1080/01621459.1975.10479864. (Cited on pages 15 and 108.)

Liran Einav, Amy Finkelstein, Iuliana Pascu, and Mark R Cullen. How general are risk preferences? Choices under uncertainty in different domains. *American Economic Review*, 102(6):1–29, 2012. (Cited on page 2.)

Benjamin Enke. Kinship, cooperation, and the evolution of moral systems. *Quarterly Journal of Economics*, 134(2):953–1019, 2019. ISSN 15314650. doi: 10.1093/qje/qjz001. (Cited on page 10.)

Armin Falk and Johannes Hermle. Relationship of gender differences in preferences to economic development and gender equality. *Science*, 362(6412):eaas9899, 2018. ISSN 0036-8075. doi: 10.1126/science.aas9899. URL http://www.sciencemag.org/lookup/doi/10.1126/science.aas9899. (Cited on pages 7, 10, and 35.)

Armin Falk, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde. Global Evidence on Economic Preferences. *Quartely Journal of Economics*, (May):1 – 80, 2018. doi: https://doi.org/10.1093/qje/qjy013. (Cited on pages 4, 9, 10, 14, 53, 93, and 98.)

Armin Falk, Thomas Dohmen, David Huffman, and Uwe Sunde. The preference survey module: A validated instrument for measuring risk, time, and social preferences. *Management Science*, 69(4):1935–1950, 2022. (Cited on pages 9 and 10.)

Yun Fan and Huug van den Dool. A global monthly land surface air temperature analysis for 1948-present. *Journal of Geophysical Research*, 113:D01103, 2008. doi: 10.1029/2007JD008470. (Cited on pages 5, 17, 120, and 121.)

Néstor Gandelman and Rubén Hernández-Murillo. Risk Aversion at the Country Level. *Federal Reserve Bank of St. Louis Review*, 97(1):53–66, 2014. (Cited on page 9.)

Nicola Gennaioli, Rafael La Porta, Florencio Lopez-de Silanes, and Andrei Shleifer. Human Capital and Regional Development. *The Quarterly Journal of Economics*, 128(1):105–164, February 2013. ISSN 0033-5533, 1531-4650. doi: 10.1093/qje/qjs050. URL https://academic.oup.com/qje/article/128/1/105/1840182. (Cited on pages 5 and 16.)

Luigi Guiso, Paola Sapienza, and Luigi Zingales. Time varying risk aversion. *Journal of Financial Economics*, 128(3):403–421, 2018. ISSN 0304-405X. doi: 10.1016/j.jfineco.2018.02.007. URL https://doi.org/10.1016/j.jfineco.2018.02.007. (Cited on page 8.)

Jorge S. Gutiérrez, Teresa Catry, María Espinosa-Colín, José A. Masero, and José Pedro Granadeiro. Heat stress increases risk taking in foraging shorebirds. *Functional Ecology*, 37(4):1005–1019, 2023. ISSN 0269-8463. doi: 10.1111/1365-2435.14288. URL https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2435.14288. (Cited on page 40.)

Andreas Hackethal, Michael Kirchler, Christine Laudenbach, Michael Razen, and Annika Weber. On the role of monetary incentives in risk preference elicitation experiments. *Journal of Risk and Uncertainty*, 66:189–213, 2023. (Cited on page 10.)

Chie Hanaoka, Hitishi Shigeoka, and Yasutora Watanabe. Do Risk Preferences Change? Evidence from the Great Eastern Japan Earthquake. *American Economic Journal: Applied Economics*, 10(2)(2):298–330, 2018. ISSN 1945-7782. doi: 10.1257/app.20170048. ISBN: w21400. (Cited on pages 7, 8, 20, 32, 33, 35, and 92.)

Eric A Hanushek, Lavinia Kinne, Pietro Sancassani, and Ludger Woessmann. Patience and Subnational Differences in Human Capital: Regional Analysis with Facebook Interests. *The Economic Journal*, 136(673):335–350, January 2026. ISSN 0013-0133, 1468-0297. doi: 10.1093/ej/ueaf045. URL https://academic.oup.com/ej/article/136/673/335/8168471. (Cited on page 16.)

David W. Harless and Colin F. Camerer. The Predictive Utility of Generalized Expected Utility Theories. *Econometrica*, 62(6):1251–1289, 1994. doi: 10.2307/2951750. (Cited on pages 4, 15, and 105.)

Ian Harris, Timothy J. Osborn, Phil Jones, and David Lister. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data*, 7:109, 2020. doi: 10.1038/s41597-020-0453-3. (Cited on pages 5, 17, and 120.)

Glenn W. Harrison and E Elisabet Rutstrom. Risk Aversion in the Laboratory. In *Risk Aversion in Experiments*, volume 12, pages 41–196. 2008. doi: 10.1016/S0193-2306(08)00003-3. Issue: 08. (Cited on page 105.)

John D. Hey and Chris Orme. Investigating Generalizations of Expected Utility Theory Using Experimental Data. *Econometrica*, 62(6):1291–1326, 1994. doi: 10.2307/2951750. (Cited on page 105.)

Solomon M. Hsiang, Marshall Burke, and Edward Miguel. Quantifying the influence of climate on human conflict. *Science*, 341(6151), 2013. ISSN 10959203. doi: 10.1126/science.1235367. (Cited on page 20.)

Peter J. Huber and Elvezio M. Ronchetti. *Robust Statistics: Theory and Methods*. Wiley, Hoboken, NJ, 2nd edition, 2009. ISBN 978-0-470-12990-6. (Cited on pages 16 and 109.)

Hans K. Hvide and Georgios A. Panos. Risk tolerance and entrepreneurship. *Journal of Financial Economics*, 111(1):200–223, January 2014. ISSN 0304-405X. doi: 10.1016/j.jfineco.2013.06.001. URL https://www.sciencedirect.com/science/article/pii/S0304405X13001748. (Cited on page 2.)

Pamela Jakiela and Owen Ozier. The Impact of Violence on Individual Risk Preferences: Evidence from a Natural Experiment. *Review of Economics and Statistics*, 101(3):547–559, 2019. ISSN 0034-6535. doi: 10.1162/rest_a_00763. URL https://direct.mit.edu/rest/article/101/3/547/58515/The-Impact-of-Violence-on-Individual-Risk. (Cited on page 8.)

W. James and Charles Stein. Estimation with Quadratic Loss. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 361–379. University of California Press, 1961. (Cited on pages 15 and 108.)

Paula Jaramillo, Daniel LaFave, and Lindsey K. Novak. Extreme Weather Events Impact Risk Tolerance and Time Preferences. *The World Bank Economic Review*, 2025. ISSN 0258-6770. doi: 10.1093/wber/lhaf018. URL https://academic.oup.com/wber/advance-article-abstract/doi/10.1093/wber/lhaf018/8242438. Advance article. (Cited on page 8.)

Thomas J. Kane and Douglas O. Staiger. Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation. Working Paper 14607, National Bureau of Economic Research, 2008. (Cited on pages 15, 109, and 111.)

Jeremy Kendal, Jamshid J. Tehrani, and John Odling-Smee. Human niche construction in interdisciplinary focus. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1566):785–792, 2011. ISSN 0962-8436. doi: 10.1098/rstb.2010.0306. URL https://royalsocietypublishing.org/doi/10.1098/rstb.2010.0306. (Cited on page 40.)

Richard E. Kihlstrom and Jean-Jacques Laffont. A General Equilibrium Entrepreneurial Theory of Firm Formation Based on Risk Aversion. *Journal of Political Economy*, 87(4):719–748, 1979. ISSN 0022-3808. doi: 10.1086/260790. (Cited on page 2.)

Théo Konc, Ivan Savin, and Jeroen C. J. M. van den Bergh. The social multiplier of environmental policy: Application to carbon taxation. *Journal of Environmental Economics and Management*, 105:102396, 2021. ISSN 0095-0696. doi: 10.1016/j.jeem.2020.102396. URL https://www.sciencedirect.com/science/article/pii/S0095069620301194. (Cited on page 8.)

Michael Kosfeld and Zahra Sharafi. The Preference Survey Module: evidence on social preferences from Tehran. *Journal of the Economic Science Association*, 10(1):152–164, June 2024. ISSN 2199-6784, 2199-6776. doi: 10.1007/s40881-023-00151-5. URL https://www.cambridge.org/core/product/identifier/S2199678424003292/type/journal_article. (Cited on page 11.)

Michael Kosfeld, Zahra Sharafi, Maíra Sontag González, and Na Zou. Measuring Economic Preferences with Surveys and Behavioral Experiments. *Working Paper*, 2025. (Cited on page 11.)

Kevin N. Laland, John Odling-Smee, and Marcus W. Feldman. Cultural niche construction and human evolution. *Journal of Evolutionary Biology*, 14(1):22–33, 2001. ISSN 1010-061X. doi: 10.1046/j.1420-9101.2001.00262.x. URL https://academic.oup.com/jeb/article-abstract/14/1/22/7323262. (Cited on page 40.)

Kenneth L. Lange, Roderick J. A. Little, and Jeremy M. G. Taylor. Robust Statistical Modeling Using the t Distribution. *Journal of the American Statistical Association*, 84 (408):881–896, 1989. doi: 10.1080/01621459.1989.10478852. (Cited on pages 16 and 109.)

Remy Levin and Daniela Vidart. Risk-Taking Adaptation to Macroeconomic Experiences. *Journal of Political Economy Microeconomics*, 2025. URL https://www.dropbox.com/scl/fi/3bua4xsyukbl6uq60nlfn/Risk_adaptation_macro_2025_02.pdf?dl=0&rlkey=bye0gqgfot1crlhtc41ls5cva&st=v7h1x8lc. (Cited on page 20.)

Remy Levin and Daniela Vidart. Risk-Taking Adaptation to Macroeconomic Experiences. *Forthcoming, Journal of Political Economy Microeconomics*, 2026. (Cited on page 8.)

Olivier L'Haridon and Ferdinand M Vieider. All over the map: A worldwide comparison of risk preferences. *Quantitative Economics*, 10:185–215, 2019. doi: 10.3982/QE898. (Cited on page 9.)

Sabine Liebenehm, Ingmar Schumacher, and Eric Strobl. Rainfall shocks and risk aversion: Evidence from Southeast Asia. *American Journal of Agricultural Economics*, 106(1): 145–176, 2023. ISSN 0002-9092. doi: 10.1111/ajae.12403. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/ajae.12403. (Cited on page 8.)

Katharine J. Mach, Caroline M. Kraan, W. Neil Adger, Halvard Buhaug, Marshall Burke, James D. Fearon, Christopher B. Field, Cullen S. Hendrix, Jean-Francois Maystadt, John O'Loughlin, Philip Roessler, Jürgen Scheffran, Kenneth A. Schultz, and Nina von Uexkull. Climate as a risk factor for armed conflict. *Nature*, 571(7764):193–197, July 2019. ISSN

1476-4687. doi: 10.1038/s41586-019-1300-6. URL https://www.nature.com/articles/s41586-019-1300-6. (Cited on page 2.)

Ulrike Malmendier and Stefan Nagel. Depression Babies: Do Macroeconomic Experiences Affect Risk Taking? *The Quarterly Journal of Economics*, 126(1):373–416, 2011. ISSN 0033-5533. doi: 10.1093/qje/qjq004. ISBN: 00335533. (Cited on pages 8 and 20.)

Ulrike Malmendier and Jessica A. Wachter. Memory of Past Experiences and Economic Decisions. In Michael J. Kahana and Anthony D. Wagner, editors, *The Oxford Handbook of Human Memory, Two Volume Pack: Foundations and Applications*, pages 2228–2266. Oxford University Press, Oxford, 2024. ISBN 978-0-19-091798-2. doi: 10.1093/oxfordhb/9780190917982.013.78. URL https://academic.oup.com/edited-volume/57928/chapter-abstract/475482667. (Cited on page 8.)

Linus Mattauch, Cameron Hepburn, Fiona Spuler, and Nicholas Stern. The economics of climate change with endogenous preferences. *Resource and Energy Economics*, 69: 101312, 2022. ISSN 0928-7655. doi: 10.1016/j.reseneeco.2022.101312. URL https://www.sciencedirect.com/science/article/pii/S092876552200029X. (Cited on page 8.)

Thomas Meissner, Xavier Gassmann, Corinne Faure, and Joachim Schleich. Individual characteristics associated with risk and time preferences: A multi country representative survey. *Journal of Risk and Uncertainty*, 66(1):77–107, 2023. ISSN 0895-5646. doi: 10.1007/s11166-022-09383-y. URL https://link.springer.com/article/10.1007/s11166-022-09383-y. (Cited on page 9.)

Robert C. Merton. Lifetime Portfolio Selection under Uncertainty : The Continuous-Time Case. *The Review of Economics and Statistics*, 51(3):247–257, 1969. (Cited on page 2.)

Kelton Minor, Andreas Bjerre-Nielsen, Sigga Svala Jonasdottir, Sune Lehmann, and Nick Obradovich. Rising temperatures erode human sleep globally. *One Earth*, 5(5):534–549, May 2022. ISSN 25903322. doi: 10.1016/j.oneear.2022.04.008. URL https://linkinghub.elsevier.com/retrieve/pii/S2590332222002093. (Cited on page 18.)

Anouch Missirian and Wolfram Schlenker. Asylum applications respond to temperature fluctuations. *Science*, 358(6370):1610–1614, December 2017. ISSN 0036-8075. doi: 10.1126/science.aao0432. URL https://www.science.org/doi/10.1126/science.aao0432. (Cited on page 2.)

Joffrey Moiroux, Guy Boivin, and Jacques Brodeur. Temperature influences host instar selection in an aphid parasitoid: support for the relative fitness rule. *Biological Journal of the Linnean Society*, 115(4):792–801, 2015. ISSN 0024-4066. doi: 10.1111/bij.12545. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/bij.12545. (Cited on page 40.)

Carl Morris. Empirical Bayes Estimates of Age-Standardized Relative Risks for Use in Disease Mapping. *Biometrics*, 39(1):29–38, 1983. doi: 10.2307/2530801. (Cited on pages 15 and 108.)

Jan Mossin. Aspects of Rational Insurance Purchasing. *Journal of Political Economy*, 76 (4, Part 1):553–568, July 1968. ISSN 0022-3808. doi: 10.1086/259427. URL https://www.journals.uchicago.edu/doi/10.1086/259427. (Cited on page 2.)

Jamie T. Mullins and Corey White. Temperature and Mental Health: Evidence from the Spectrum of Mental Health Outcomes. *Journal of Health Economics*, 68:102240, 2019. doi: 10.1016/j.jhealeco.2019.102240. (Cited on page 18.)

Ishan Nath, Kelly McCusker, Ian Bolliger, Tamma Carleton, Michael Delgado, Michael Greenstone, Trevor Houser, Solomon Hsiang, Andrew Hultgren, Amir Jina, Robert Kopp, James Rising, and Ashwin Rode. The Welfare Economics of a Data-Driven Social Cost of Carbon. *Working Paper*, 2023. (Cited on page 3.)

William D. Nordhaus. Revisiting the social cost of carbon. *Proceedings of the National Academy of Sciences*, 114(7):1518–1523, February 2017. ISSN 0027-8424. doi: 10.1073/pnas.1609244114. URL https://www.pnas.org/doi/10.1073/pnas.1609244114. (Cited on page 3.)

Nick Obradovich, Robyn Migliorini, Martin P. Paulus, and Iyad Rahwan. Empirical Evidence of Mental Health Risks Posed by Climate Change. *Proceedings of the National Academy of Sciences*, 115(43):10953–10958, 2018. doi: 10.1073/pnas.1801528115. (Cited on page 18.)

Robert E. O'Connor, Richard J. Bord, and Ann Fisher. Risk Perceptions, General Environmental Beliefs, and Willingness to Address Climate Change. *Risk Analysis*, 19(3): 461–471, 1999. ISSN 0272-4332. doi: 10.1111/j.1539-6924.1999.tb00421.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1539-6924.1999.tb00421.x. (Cited on page 7.)

F. John Odling-Smee, Kevin N. Laland, and Marcus W. Feldman. Niche Construction. *The American Naturalist*, 147(4):641–648, 1996. ISSN 0003-0147. doi: 10.1086/285870. URL https://www.journals.uchicago.edu/doi/abs/10.1086/285870. (Cited on page 40.)

Ariel Ortiz-Bobea, Toby R. Ault, Carlos M. Carrillo, Robert G. Chambers, and David B. Lobell. Anthropogenic climate change has slowed global agricultural productivity growth. *Nature Climate Change*, 11(4):306–312, April 2021. ISSN 1758-6798. doi: 10.1038/s41558-021-01000-1. URL https://www.nature.com/articles/s41558-021-01000-1. (Cited on page 2.)

Matthew Rabin. Risk Aversion and Expected-Utility Theory : A Calibration Theorem. *Econometrica*, 68(5):1281–1292, 2000. (Cited on page 103.)

Daniel Read, George Loewenstein, and Matthew Rabin. Choice Bracketing. *Journal of Risk and Uncertainty*, 19(1-3):171–197, 1999. doi: 10.1023/A:1007879411489. (Cited on page 14.)

Arnaud Reynaud and Cécile Aubert. Does flood experience modify risk preferences? Evidence from an artefactual field experiment in Vietnam. *The Geneva Risk and Insurance Review*, 45(1):36–74, 2019. ISSN 1554-964X. doi: 10.1057/s10713-019-00044-w. URL https://link.springer.com/article/10.1057/s10713-019-00044-w. (Cited on page 8.)

Marc Oliver Rieger, Mei Wang, and Thorsten Hens. Risk preferences around the world. *Management Science*, 61(3):637–648, 2014. ISSN 15265501. doi: 10.1287/mnsc.2013.1869. (Cited on page 9.)

Marina Romanello, Claudia Di Napoli, Carole Green, Harry Kennard, Pete Lampard, Daniel Scamman, Maria Walawender, Zakari Ali, Nadia Ameli, Sonja Ayeb-Karlsson, Paul J Beggs, Kristine Belesova, Lea Berrang Ford, Kathryn Bowen, Wenjia Cai, Max Callaghan, Diarmid Campbell-Lendrum, Jonathan Chambers, Troy J Cross, Kim R Van Daalen, Carole Dalin, Niheer Dasandi, Shouro Dasgupta, Michael Davies, Paula Dominguez-Salas, Robert Dubrow, Kristie L Ebi, Matthew Eckelman, Paul Ekins, Chris Freyberg, Olga Gasparyan, Georgiana Gordon-Strachan, Hilary Graham, Samuel H Gunther, Ian Hamilton, Yun Hang, Risto Hänninen, Stella Hartinger, Kehan He, Julian Heidecke, Jeremy J Hess, Shih-Che Hsu, Louis Jamart, Slava Jankin, Ollie Jay, Ilan Kelman, Gregor Kiesewetter, Patrick Kinney, Dominic Kniveton, Rostislav Kouznetsov, Francesca Larosa, Jason K W Lee, Bruno Lemke, Yang Liu, Zhao Liu, Melissa Lott, Martín Lotto Batista, Rachel Lowe, Maquins Odhiambo Sewe, Jaime Martinez-Urtaza, Mark Maslin, Lucy McAllister, Celia McMichael, Zhifu Mi, James Milner, Kelton Minor, Jan C Minx, Nahid Mohajeri, Natalie C Momen, Maziar Moradi-Lakeh, Karyn Morrissey, Simon Munzert, Kris A Murray, Tara Neville, Maria Nilsson, Nick Obradovich, Megan B O'Hare, Camile Oliveira, Tadj Oreszczyn, Matthias Otto, Fereidoon Owfi, Olivia Pearman, Frank Pega, Andrew Pershing, Mahnaz Rabbaniha, Jamie Rickman, Elizabeth J Z Robinson, Joacim Rocklöv, Renee N Salas, Jan C Semenza, Jodi D Sherman, Joy Shumake-Guillemot, Grant Silbert, Mikhail Sofiev, Marco Springmann, Jennifer D Stowell, Meisam Tabatabaei, Jonathon Taylor, Ross Thompson, Cathryn Tonne, Marina Treskova, Joaquin A Trinanes, Fabian Wagner, Laura Warnecke, Hannah Whitcombe, Matthew Winning, Arthur Wyns, Marisol Yglesias-González, Shihui Zhang, Ying Zhang, Qiao Zhu, Peng Gong, Hugh Montgomery, and Anthony Costello. The 2023 report of the Lancet Countdown on health and climate change: the imperative for a health-centred response in a world facing irreversible harms. *The Lancet*, 402 (10419):2346–2394, December 2023. ISSN 01406736. doi: 10.1016/S0140-6736(23)01859-7. URL https://linkinghub.elsevier.com/retrieve/pii/S0140673623018597. (Cited on page 18.)

Luis Rubalcava and Graciela Teruel. Mexican Family Life Survey, Second Wave. Technical report, 2008. URL http://www.ennvih-mxfls.org. (Cited on page 31.)

Luis Rubalcava and Graciela Teruel. Mexican Family Life Survey, Third Wave. Technical report, 2013. URL http://www.ennvih-mxfls.org. (Cited on page 31.)

John Strauss, Kathleen Beegle, Bondan Sikoki, Agus Dwiyanto, Yulia Herawati, and Firman Witoelar. The Fourth Wave of the Indonesia Family Life Survey: Overview and Field Report. *RAND Working Paper*, WR-675/1-N(April 2009):i–82, 2009. (Cited on page 31.)

John Strauss, Firman Witoelar, and Sikoki Bondan. The Fifth Wave of the Indonesia Family Life Survey: Overview and Field Report. Technical report, 2016. Volume: 1 Issue: March. (Cited on page 31.)

Uwe Sunde, Thomas Dohmen, Benjamin Enke, Armin Falk, David Huffman, and Gerrit Meyerheim. Patience and Comparative Development. *The Review of Economic Studies*, 89 (5):2806–2840, September 2022. ISSN 0034-6527, 1467-937X. doi: 10.1093/restud/rdab084. URL https://academic.oup.com/restud/article/89/5/2806/6447525. (Cited on pages 5, 10, and 16.)

Richard H. Thaler. Mental Accounting Matters. *Journal of Behavioral Decision Making*, 12(3): 183–206, 1999. doi: 10.1002/(SICI)1099-0771(199909)12:3⟨183::AID-BDM318⟩3.0.CO;2-F. (Cited on page 15.)

Jari Tiihonen, Pirjo Halonen, Laura Tiihonen, Hannu Kautiainen, Markus Storvik, and James Callaway. The Association of Ambient Temperature and Violent Crime. *Scientific Reports*, 7(1):6543, 2017. ISSN 2045-2322. doi: 10.1038/s41598-017-06720-z. URL https://www.nature.com/articles/s41598-017-06720-z. (Cited on page 40.)

U.S. Environmental Protection Agency. Report on the Social Cost of Greenhouse Gases: Estimates Incorporating Recent Scientific Advances. Supplementary Material for the Regulatory Impact Analysis EPA-HQ-OAR-2021-0317, U.S. Environmental Protection Agency, National Center for Environmental Economics; Climate Change Division, Office of Air and Radiation, Washington, DC, November 2023. URL https://www.epa.gov/system/files/documents/2023-12/epa_scghg_2023_report_final.pdf. (Cited on page 3.)

Ferdinand M. Vieider, Mathieu Lefebvre, Ranoua Bouchouicha, Thorsten Chmura, Rustamdjan Hakimov, Michal Krawczyk, and Peter Martinsson. Common Components of Risk and Uncertainty Attitudes Across Contexts and Domains: Evidence from 30 Countries. *Journal of the European Economic Association*, 13(3):421–452, 2015. ISSN 1542-4766. doi: 10.1111/jeea.12102. URL https://academic.oup.com/jeea/article-abstract/13/3/421/2319769. (Cited on page 9.)

W. Kip Viscusi and Richard J. Zeckhauser. The Perception and Valuation of the Risks of Climate Change: A Rational and Behavioral Blend. *Climatic Change*, 77(1-2):151–177, 2006. ISSN 0165-0009. doi: 10.1007/s10584-006-9075-9. URL https://link.springer.com/article/10.1007/s10584-006-9075-9. (Cited on page 7.)

Maarten J. Voors, Eleonora E. M. Nillesen, Philip Verwimp, Erwin H. Bulte, Robert Lensink, and Daan P. Van Soest. Violent Conflict and Behavior: A Field Experiment in Burundi. *American Economic Review*, 102(2):941–964, 2012. ISSN 0002-8282. doi: 10.1257/aer.102.2.941. URL https://www.aeaweb.org/articles?id=10.1257/aer.102.2.941. (Cited on page 8.)

Christopher R. Walters. The Use of Connected Partners to Estimate Selection: Empirical Bayes and Jackknife Methods for Research Designs with Grouped Data. *Econometrica*, 92 (5):1419–1456, 2024. doi: 10.3982/ECTA20793. (Cited on page 109.)

Martin L. Weitzman. A Review of The Stern Review on the Economics of Climate Change. *Journal of Economic Literature*, 45(3):703–724, September 2007. ISSN 0022-0515. doi: 10.1257/jel.45.3.703. URL https://www.aeaweb.org/articles?id=10.1257/jel.45.3.703. (Cited on page 3.)

Martin L. Weitzman. Fat Tails and the Social Cost of Carbon. *American Economic Review*, 104(5):544–546, May 2014. ISSN 0002-8282. doi: 10.1257/aer.104.5.544. URL https://www.aeaweb.org/articles?id=10.1257/aer.104.5.544. (Cited on page 3.)

Leonie Wenz, Robert D. Carr, Noah Kögel, Maximilian Kotz, and Anders Levermann. DOSE – Global data set of reported sub-national economic output. *Scientific Data*, 10:425, 2023. doi: 10.1038/s41597-023-02323-8. (Cited on page 101.)

Cort J. Willmott and Kenji Matsuura. Terrestrial Air Temperature and Precipitation: Monthly and Annual Time Series (1950-1999), Version 1.02, 2001. URL http://climate.geog.udel.edu/~climate/. (Cited on pages 5, 17, and 120.)

Sammy Zahran, Samuel D. Brody, Himanshu Grover, and Arnold Vedlitz. Climate Change Vulnerability and Policy Support. *Society & Natural Resources*, 19(9):771–789, 2006. ISSN 0894-1920. doi: 10.1080/08941920600835528. URL https://www.tandfonline.com/doi/abs/10.1080/08941920600835528. (Cited on page 7.)

A Empirical Appendix

A.1 GPS staircase measure

Table 2: Staircase elicitation of willingness to take risks in the Global Preference Survey

Round r	Step Δ_r	Safe amounts x_r offered (euros)	# nodes at r
1	80	160	1
2	40	80, 240	2
3	20	40, 120, 200, 280	4
4	10	20, 60, 100, 140, 180, 220, 260, 300	8
5	—	10, 30, 50, . . . , 270, 290, 310	16

Notes. In each round $r = 1, \ldots, 5$ the respondent chooses between (A) a 50:50 lottery paying 300 or 0 euros, or (B) a sure payment of x_r euros. Choosing the lottery in round r raises the next-round safe amount by Δ_r; choosing the safe option lowers it by Δ_r. The step size halves each round, producing a binary tree with $2^5 = 32$ terminal nodes. The five-choice sequence maps one-to-one onto an integer "willingness to take risks" score: 1 corresponds to choosing B at every round (most risk averse) and 32 to choosing A at every round (most risk tolerant). In our analysis we reverse this ordering so that 1 indicates the most risk-seeking bin and 32 the most risk-averse bin. A visual representation of the instrument is available in the appendix of Falk et al. (2018).

A.2 Information on Estimated Global Preference Distribution

Survey-weighted distributions of CRRA $\hat{\rho}$ (estimated using the procedure in Section 2.3) by country, ordered by median. Global distributions shown at top of first table.

Table 3: $\hat{\rho}$ Distribution by Country (Global + Ranks 1–25)

Rank	Country		Median	Mean	SD	Min	p10	p25	p75	p90	Max
Global distributions											
	Unweighted	79275	1.12	0.36	4.69	-58.83	-2.82	-0.33	2.10	3.99	34.29
	Country-representative	79275	1.12	0.38	4.57	-58.83	-2.79	-0.33	2.10	3.97	34.29
	Globally-representative	79275	1.27	0.53	5.02	-58.83	-2.87	-0.35	2.59	4.54	34.29
Country-representative distributions (most risk-averse first), 1–25											
1	Egypt	1019	3.11	1.77	5.24	-33.13	-5.74	0.63	4.11	6.07	10.53
2	Indonesia	995	2.63	2.49	3.75	-37.50	-1.35	1.99	3.46	7.42	12.46
3	Romania	952	2.21	2.06	9.20	-58.83	-3.47	-0.04	3.37	7.08	34.29
4	Argentina	988	2.03	0.63	5.10	-24.56	-2.29	-0.32	2.26	5.05	7.78
5	Sri Lanka	996	2.01	0.91	4.60	-26.13	-2.55	-0.61	2.57	5.05	7.39
6	Georgia	986	1.96	0.69	5.99	-33.56	-4.68	-1.98	3.61	5.58	11.09
7	Philippines	1000	1.88	1.41	4.13	-29.39	-2.57	0.49	2.70	5.54	8.90
8	India	2485	1.88	1.58	5.12	-39.59	-1.12	0.64	3.48	5.93	19.32
9	Nigeria	995	1.82	-1.08	9.49	-49.05	-12.39	-5.35	4.84	6.73	21.09
10	Colombia	996	1.80	1.89	3.18	-23.73	-0.32	1.57	2.61	5.51	7.66
11	Thailand	999	1.74	1.38	4.14	-19.29	-0.30	1.13	2.61	5.40	6.60
12	Cameroon	1000	1.69	1.76	2.71	-24.27	-0.32	1.54	2.36	4.22	8.05
13	Hungary	1004	1.67	1.66	4.29	-30.23	-1.10	1.13	4.07	5.06	10.33
14	Greece	1000	1.66	1.57	3.72	-23.25	-1.03	0.94	3.59	5.29	7.11
15	Mexico	998	1.61	-0.26	6.54	-36.40	-9.61	-0.86	3.03	5.09	11.76
16	Bangladesh	993	1.61	1.39	6.17	-38.75	-3.70	-0.11	3.58	8.66	13.14
17	Poland	992	1.51	0.49	5.07	-30.94	-3.25	-0.82	2.61	4.90	10.42
18	Portugal	989	1.49	1.42	2.23	-17.99	-0.09	1.43	1.93	3.58	6.29
19	Costa Rica	999	1.41	1.29	2.59	-16.35	-0.08	1.38	1.41	3.87	5.47
20	Guatemala	994	1.39	1.24	3.61	-29.27	-1.34	1.12	2.75	4.55	9.02
21	Finland	999	1.38	1.20	2.13	-11.06	-0.74	0.56	2.10	3.81	4.39
22	Russia	1447	1.36	0.81	4.80	-28.33	-1.90	-0.32	2.61	4.54	9.07
23	Spain	1000	1.32	1.51	2.05	-16.97	-0.07	1.07	2.37	3.99	5.91
24	Germany	838	1.30	1.06	2.92	-18.68	-0.51	0.48	2.11	3.59	5.64
25	Brazil	998	1.29	1.24	2.88	-21.60	-0.86	0.63	2.11	4.55	6.51

Notes: All country statistics are survey-weighted using Gallup probability weights. "Country-representative" uses survey weights; "Globally-representative" uses population weights ($w_i \times \text{Pop}_c / \sum_{i \in c} w_i$). Countries ordered by survey-weighted median $\hat{\rho}$.

Table 4: $\hat{\rho}$ Distribution by Country (Ranks 26–50)

Rank	Country		Median	Mean	SD	Min	p10	p25	p75	p90	Max
26	Serbia	1011	1.29	0.74	3.08	-21.28	-1.44	-0.33	1.32	3.99	6.38
27	Peru	988	1.28	1.34	2.77	-20.21	-0.60	0.16	2.54	4.78	6.40
28	France	982	1.27	1.12	2.50	-17.27	-0.20	0.47	1.61	3.42	6.15
29	Japan	996	1.26	1.32	2.10	-14.18	-0.14	0.65	2.43	3.58	4.90
30	Chile	1003	1.23	0.68	4.60	-25.26	-2.20	-0.03	2.11	4.97	7.60
31	Austria	996	1.22	0.53	2.99	-17.72	-2.25	-0.01	1.95	3.10	5.93
32	United Arab Emirates	999	1.22	0.86	3.43	-16.01	-0.83	0.01	2.05	4.50	4.71
33	Italy	993	1.21	1.20	2.48	-16.55	-0.27	0.97	1.97	3.57	5.81
34	Suriname	495	1.20	-1.03	5.69	-29.89	-6.09	-3.53	1.66	2.11	8.56
35	South Korea	982	1.18	0.68	2.63	-16.03	-1.58	-0.01	1.88	3.09	5.00
36	Venezuela	996	1.18	1.02	3.61	-30.92	-2.18	-0.04	2.74	3.60	9.55
37	Switzerland	820	1.15	0.57	3.47	-16.38	-1.75	-0.01	1.97	3.59	5.29
38	Estonia	998	1.14	0.56	3.41	-21.22	-1.79	-0.12	1.61	3.10	6.93
39	Nicaragua	995	1.13	1.01	1.72	-14.89	0.18	1.10	1.13	1.98	4.53
40	Ukraine	982	1.13	0.60	3.58	-20.68	-1.82	-0.53	1.94	4.05	6.17
41	Jordan	996	1.13	0.82	2.86	-18.15	-0.33	0.63	1.46	3.35	5.24
42	Australia	999	1.12	0.86	2.28	-17.20	-0.79	0.21	1.87	3.08	6.44
43	Morocco	991	1.11	0.38	2.91	-16.35	-0.36	-0.02	1.13	1.88	4.79
44	Bosnia Herzegovina	991	1.11	0.87	1.91	-11.89	-0.29	0.39	1.46	2.58	3.84
45	United States	1067	1.10	0.60	3.06	-16.66	-1.06	-0.01	1.96	3.09	8.14
46	China	2549	1.09	-0.48	5.35	-36.97	-4.22	-1.51	1.62	3.41	12.37
47	Turkey	999	1.09	0.68	3.14	-26.09	-1.32	-0.22	1.61	3.59	8.07
48	Moldova	990	1.08	0.25	3.88	-24.68	-1.68	-0.60	1.61	3.51	7.42
49	Kenya	998	1.06	0.40	4.78	-36.79	-2.48	-0.88	1.95	5.03	11.65
50	Canada	1000	1.04	0.45	3.11	-17.50	-1.92	-0.05	1.61	3.09	5.20

Table 5: $\hat{\rho}$ Distribution by Country (Ranks 51–76)

Rank	Country		Median	Mean	SD	Min	p10	p25	p75	p90	Max
51	Bolivia	990	1.04	1.02	2.12	-13.60	-0.36	0.47	1.90	3.35	4.45
52	Netherlands	996	1.00	0.57	3.11	-19.87	-1.31	-0.10	1.61	3.38	6.07
53	Israel	999	0.96	0.49	3.26	-17.29	-2.30	-0.42	1.95	3.87	5.48
54	Czech Republic	945	0.96	0.13	3.84	-27.86	-3.22	-0.60	1.89	3.41	9.05
55	Sweden	997	0.96	0.39	3.03	-18.28	-2.29	-0.08	1.87	3.09	5.42
56	Cambodia	999	0.93	0.81	2.70	-20.80	-0.18	0.63	1.38	3.09	6.72
57	Iraq	965	0.93	-0.77	6.32	-31.15	-6.27	-2.86	2.56	3.87	8.97
58	Lithuania	987	0.88	-0.17	4.96	-25.23	-2.31	-0.96	1.94	3.10	8.01
59	Croatia	944	0.88	-0.07	4.55	-23.44	-2.82	-0.85	1.92	3.60	6.39
60	Uganda	1000	0.87	0.67	1.97	-14.87	-0.11	0.05	0.89	2.54	4.59
61	Rwanda	1000	0.86	0.46	1.67	-18.78	-0.58	-0.52	0.87	1.43	7.53
62	United Kingdom	1013	0.63	0.46	2.18	-13.11	-0.79	0.09	1.60	1.96	4.62
63	Vietnam	977	0.61	-0.34	3.32	-17.78	-2.65	-1.33	1.11	2.51	4.64
64	Haiti	503	0.55	-0.10	3.20	-19.05	-1.40	-0.38	0.66	2.51	5.06
65	Pakistan	1000	0.54	-0.77	4.48	-15.41	-1.43	-0.35	0.71	1.98	3.96
66	Tanzania	999	0.14	-1.42	4.07	-43.65	-3.07	-2.78	0.28	0.63	9.65
67	Kazakhstan	967	0.13	-1.54	6.94	-31.13	-6.69	-4.86	2.58	3.46	10.35
68	Ghana	1000	0.13	-0.69	3.56	-23.61	-2.34	-1.08	0.50	1.61	6.36
69	Saudi Arabia	1035	-0.01	-0.30	3.75	-18.83	-1.78	-0.58	0.98	2.36	5.57
70	Iran	2463	-0.02	-0.84	4.60	-24.22	-4.29	-1.33	1.11	2.60	6.76
71	Zimbabwe	1000	-0.03	-0.59	3.64	-22.78	-2.32	-0.57	0.59	1.87	5.54
72	Afghanistan	997	-0.12	-2.41	6.81	-22.04	-19.53	-1.88	1.44	2.07	5.75
73	Algeria	1022	-0.37	-1.70	5.71	-27.09	-4.86	-4.02	1.12	2.40	7.52
74	Botswana	1000	-0.87	-3.63	9.55	-45.44	-10.26	-9.31	2.11	4.11	14.71
75	Malawi	1000	-1.76	-1.35	3.33	-22.04	-2.84	-2.03	0.06	0.13	5.44
76	South Africa	999	-4.38	-5.31	8.22	-41.48	-15.94	-8.45	0.66	2.37	9.71

A.3 Precipitation results: Global

Table 6: Lifetime Mean Precipitation and Risk Aversion

	(1)	(2)	(3)
Panel A: Staircase Risk Aversion			
Mean Lifetime Precip	-0.128	-0.238	-0.469**
	(0.220)	(0.158)	(0.221)
R^2	0.153	0.159	0.212
Panel B: CRRA $\hat{\rho}$			
Mean Lifetime Precip	-0.054	-0.087***	-0.137**
	(0.046)	(0.031)	(0.067)
R^2	0.124	0.126	0.177
N	78,221	78,221	77,828
Region FE	X	X	X
Cohort FE		X	
Country × Cohort FE			X

Notes: The dependent variable in panel A is the GPS staircase lottery measure (1–32 scale) normalized so higher numbers mean more risk aversion. The dependent variable in Panel B is the CRRA $\hat{\rho}$ estimated from the staircase measure using the procedure in Section 2.3. Empirical specifications for the three columns can be found in Section 2.8. All specifications control for gender, language dummies, lifetime mean temperature, and hemisphere×interview-month seasonality fixed effects. Standard errors, clustered two-way by region and birth year, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.4 Projected change in median risk preferences

Table 7: Change in median $\tilde{\rho}$ by country, 1981–2026

Country	'81	'26	Δ	Country	'81	'26	Δ
Afghanistan	0.052	-0.669	-0.721	Jordan	1.136	0.932	-0.204
Guatemala	1.696	1.162	-0.534	Japan	1.361	1.157	-0.204
Iraq	1.208	0.781	-0.427	Romania	2.232	2.028	-0.203
Russia	1.672	1.254	-0.417	Italy	1.258	1.057	-0.200
Moldova	1.167	0.808	-0.359	UAE	1.267	1.067	-0.200
Estonia	1.303	0.952	-0.351	South Africa	-4.299	-4.499	-0.200
Kazakhstan	0.317	-0.024	-0.341	Germany	1.410	1.213	-0.198
Ukraine	1.273	0.940	-0.333	Canada	1.142	0.945	-0.197
Egypt	3.308	2.985	-0.323	Kenya	1.120	0.923	-0.197
Lithuania	1.088	0.766	-0.323	Saudi Arabia	0.128	-0.067	-0.196
Pakistan	0.619	0.304	-0.315	Switzerland	1.200	1.023	-0.177
Netherlands	1.178	0.883	-0.295	Hungary	1.786	1.611	-0.174
Iran	0.097	-0.195	-0.292	Ghana	0.166	0.005	-0.162
Brazil	1.500	1.216	-0.283	China	1.100	0.942	-0.158
Bosnia Herz.	1.166	0.884	-0.282	Rwanda	0.971	0.817	-0.154
Haiti	0.707	0.435	-0.272	Uganda	0.933	0.783	-0.150
Nigeria	1.921	1.649	-0.272	Serbia	1.273	1.123	-0.149
Israel	1.101	0.831	-0.271	Botswana	-0.791	-0.935	-0.144
Spain	1.481	1.213	-0.268	Argentina	2.088	1.944	-0.144
Venezuela	1.389	1.122	-0.267	Cambodia	0.966	0.840	-0.125
Czech Republic	1.087	0.821	-0.266	Costa Rica	1.475	1.353	-0.122
Algeria	-0.324	-0.590	-0.265	Peru	1.426	1.308	-0.118
Morocco	1.169	0.905	-0.265	Tanzania	0.184	0.077	-0.107
Poland	1.625	1.368	-0.257	Vietnam	0.445	0.343	-0.102
Malawi	-1.565	-1.816	-0.251	Sri Lanka	2.024	1.924	-0.099
South Korea	1.299	1.056	-0.242	United States	1.103	1.011	-0.092
Suriname	1.218	0.979	-0.238	India	1.876	1.786	-0.091
United Kingdom	0.815	0.576	-0.238	Thailand	1.741	1.652	-0.088
Croatia	1.030	0.797	-0.233	Greece	1.602	1.524	-0.078
Philippines	1.855	1.625	-0.230	Chile	1.291	1.214	-0.077
Zimbabwe	0.152	-0.076	-0.229	Australia	1.111	1.050	-0.060
Turkey	1.142	0.916	-0.226	Mexico	1.569	1.511	-0.058
Nicaragua	1.248	1.025	-0.224	Portugal	1.551	1.495	-0.056
Austria	1.247	1.025	-0.222	Bangladesh	1.598	1.562	-0.036
Global	**1.381**	**1.161**	**-0.220**	Bolivia	1.025	1.020	-0.005
Sweden	1.042	0.823	-0.219	Colombia	1.852	1.870	0.018
Finland	1.437	1.223	-0.214	Cameroon	1.678	1.702	0.024
Georgia	2.116	1.905	-0.211	Indonesia	2.569	2.641	0.072
France	1.380	1.170	-0.210				

Notes: Results from analysis in Section 3.3. Country medians derived from nationally-representative preference distributions; global medians are derived from the globally-representative preference distributions.

A.5 Robustness: alternative shrinkage

Table 8: Lifetime Mean Temperature and Risk Aversion: Alternative $\hat{\rho}$ shrinkage

	(1)	(2)	(3)
Panel A: CRRA $\hat{\rho}$ (Baseline — Normal EB)			
Mean Lifetime Temp	-0.622***	-0.318***	-0.177*
	(0.075)	(0.051)	(0.100)
R^2	0.124	0.126	0.177
Panel B: CRRA $\hat{\rho}$ (Unshrunk)			
Mean Lifetime Temp	-2.147***	-0.505*	-0.851***
	(0.545)	(0.259)	(0.316)
R^2	0.139	0.144	0.195
Panel C: CRRA $\hat{\rho}$ (Winsorized 2.5/97.5)			
Mean Lifetime Temp	-2.085***	-0.474*	-0.808***
	(0.534)	(0.251)	(0.288)
R^2	0.141	0.145	0.197
Panel D: CRRA $\hat{\rho}$ (t-dist EB, ν=5)			
Mean Lifetime Temp	-1.237***	-0.861***	-0.052
	(0.141)	(0.173)	(0.101)
R^2	0.499	0.501	0.534
Panel E: CRRA $\hat{\rho}$ (Mixture Model)			
Mean Lifetime Temp	-0.631***	-0.515***	-0.059
	(0.084)	(0.096)	(0.065)
R^2	0.227	0.228	0.279
N	78,221	78,221	77,828
Region FE	X	X	X
Cohort FE		X	
Country × Cohort FE			X

Notes: Each panel uses a different CRRA $\hat{\rho}$ estimate as the dependent variable. Panel A: Normal empirical Bayes shrinkage (baseline). Panel B: unshrunk raw estimate. Panel C: winsorized at the 2.5th/97.5th percentiles (no shrinkage). Panel D: t-distribution empirical Bayes ($\nu = 5$). Panel E: mixture model shrinkage. Empirical specifications for the three columns can be found in Section 2.8. Regressions are at the subject region level. All specifications control for gender, language dummies, lifetime mean precipitation, and hemisphere×interview-month seasonality fixed effects. Standard errors, clustered two-way by region and birth year, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.6 Robustness: sample restrictions by staircase boundary cases

Table 9: Lifetime Mean Temperature and Risk Aversion (Excluding FOSD Violators)

	(1)	(2)	(3)
Panel A: Staircase Risk Aversion			
Mean Lifetime Temp	-2.166***	-0.308	-0.935**
	(0.624)	(0.383)	(0.393)
R^2	0.163	0.171	0.231
Panel B: CRRA $\hat{\rho}$			
Mean Lifetime Temp	-0.605***	-0.227***	-0.197
	(0.091)	(0.028)	(0.128)
R^2	0.107	0.109	0.170
N	69,929	69,929	69,504
Region FE	X	X	X
Cohort FE		X	
Country × Cohort FE			X

Notes: The dependent variable in panel A is the GPS staircase lottery measure (1–32 scale) normalized so higher numbers mean more risk aversion. The dependent variable in Panel B is the CRRA $\hat{\rho}$ estimated from the staircase measure using the procedure in Section 2.3. Empirical specifications for the three columns can be found in Section 2.8. Regressions are at the subject region level. Sample excludes subjects in staircase bins 31–32 (FOSD-dominated choices: always chose the lottery). All specifications control for gender, language dummies, lifetime mean precipitation, and hemisphere×interview-month seasonality fixed effects. Standard errors, clustered two-way by region and birth year, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 10: Lifetime Mean Temperature and Risk Aversion (Excluding Most Risk-Averse)

	(1)	(2)	(3)
Panel A: Staircase Risk Aversion			
Mean Lifetime Temp	-0.879***	-0.243	-0.352
	(0.254)	(0.222)	(0.449)
R^2	0.186	0.189	0.259
Panel B: CRRA $\hat{\rho}$			
Mean Lifetime Temp	-0.685***	-0.391***	-0.156
	(0.065)	(0.084)	(0.124)
R^2	0.147	0.149	0.217
N	55,417	55,417	54,985
Region FE	X	X	X
Cohort FE		X	
Country × Cohort FE			X

Notes: The dependent variable in panel A is the GPS staircase lottery measure (1–32 scale) normalized so higher numbers mean more risk aversion. The dependent variable in Panel B is the CRRA $\hat{\rho}$ estimated from the staircase measure using the procedure in Section 2.3. Empirical specifications for the three columns can be found in Section 2.8. Regressions are at the subject region level. Sample excludes subjects in staircase bin 1 (always chose the safe option). All specifications control for gender, language dummies, lifetime mean precipitation, and hemisphere×interview-month seasonality fixed effects. Standard errors, clustered two-way by region and birth year, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 11: Lifetime Mean Temperature and Risk Aversion (Excluding All Boundary Cases)

	(1)	(2)	(3)
Panel A: Staircase Risk Aversion			
Mean Lifetime Temp	-1.290***	-0.228	-0.656*
	(0.328)	(0.146)	(0.335)
R^2	0.134	0.138	0.222
Panel B: CRRA $\hat{\rho}$			
Mean Lifetime Temp	-0.773***	-0.341***	-0.277
	(0.088)	(0.095)	(0.205)
R^2	0.107	0.110	0.192
N	47,111	47,111	46,624
Region FE	X	X	X
Cohort FE		X	
Country × Cohort FE			X

Notes: The dependent variable in panel A is the GPS staircase lottery measure (1–32 scale) normalized so higher numbers mean more risk aversion. The dependent variable in Panel B is the CRRA $\hat{\rho}$ estimated from the staircase measure using the procedure in Section 2.3. Empirical specifications for the three columns can be found in Section 2.8. Regressions are at the subject region level. Sample excludes subjects in staircase bins 1, 31, and 32 (both boundary groups: always safe and FOSD-dominated). All specifications control for gender, language dummies, lifetime mean precipitation, and hemisphere×interview-month seasonality fixed effects. Standard errors, clustered two-way by region and birth year, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.7 Robustness: alternative individual controls

Table 12: Lifetime Mean Temperature and Risk Aversion (No Gender or Language Controls)

	(1)	(2)	(3)
Panel A: Staircase Risk Aversion			
Mean Lifetime Temp	-1.166***	-0.097	-0.607**
	(0.428)	(0.257)	(0.268)
R^2	0.148	0.154	0.209
Panel B: CRRA $\hat{\rho}$			
Mean Lifetime Temp	-0.389***	-0.171	-0.135
	(0.136)	(0.147)	(0.104)
R^2	0.120	0.122	0.175
N	78,221	78,221	77,828
Region FE	X	X	X
Cohort FE		X	
Country × Cohort FE			X

Notes: The dependent variable in panel A is the GPS staircase lottery measure (1–32 scale) normalized so higher numbers mean more risk aversion. The dependent variable in Panel B is the CRRA $\hat{\rho}$ estimated from the staircase measure using the procedure in Section 2.3. Empirical specifications for the three columns can be found in Section 2.8. Regressions are at the subject region level. All specifications control for lifetime mean precipitation and hemisphere×interview-month seasonality fixed effects. Gender and language controls are *not* included. Standard errors, clustered two-way by region and birth year, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 13: Lifetime Mean Temperature and Risk Aversion (Adding Subjective Math Ability)

	(1)	(2)	(3)
Panel A: Staircase Risk Aversion			
Mean Lifetime Temp	-1.797***	-0.206	-0.745**
	(0.591)	(0.317)	(0.363)
R^2	0.156	0.161	0.215
Panel B: CRRA $\hat{\rho}$			
Mean Lifetime Temp	-0.594***	-0.319***	-0.186*
	(0.071)	(0.053)	(0.103)
R^2	0.126	0.128	0.180
N	77,233	77,233	76,841
Region FE	X	X	X
Cohort FE		X	
Country × Cohort FE			X

Notes: The dependent variable in panel A is the GPS staircase lottery measure (1–32 scale) normalized so higher numbers mean more risk aversion. The dependent variable in Panel B is the CRRA $\hat{\rho}$ estimated from the staircase measure using the procedure in Section 2.3. Empirical specifications for the three columns can be found in Section 2.8. Regressions are at the subject region level. All specifications control for gender, language dummies, lifetime mean precipitation, hemisphere×interview-month seasonality FEs, and subjective math ability. Standard errors, clustered two-way by region and birth year, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 14: Lifetime Mean Temperature and Risk Aversion (Adding Math + All Preference Controls)

	(1)	(2)	(3)
Panel A: Staircase Risk Aversion			
Mean Lifetime Temp	-1.411***	-0.186	-0.886**
	(0.490)	(0.296)	(0.348)
R^2	0.172	0.176	0.230
Panel B: CRRA $\hat{\rho}$			
Mean Lifetime Temp	-0.511***	-0.321***	-0.225**
	(0.057)	(0.056)	(0.100)
R^2	0.131	0.133	0.186
N	74,799	74,799	74,408
Region FE	X	X	X
Cohort FE		X	
Country × Cohort FE			X

Notes: The dependent variable in panel A is the GPS staircase lottery measure (1–32 scale) normalized so higher numbers mean more risk aversion. The dependent variable in Panel B is the CRRA $\hat{\rho}$ estimated from the staircase measure using the procedure in Section 2.3. Empirical specifications for the three columns can be found in Section 2.8. Regressions are at the subject region level. All specifications control for gender, language dummies, lifetime mean precipitation, hemisphere×interview-month seasonality FEs, subjective math ability, patience, trust, altruism, positive reciprocity, and negative reciprocity. Standard errors, clustered two-way by region and birth year, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.8 Robustness: alternative climate controls

Table 15: Lifetime Mean Temperature and Risk Aversion (No Precipitation Control)

	(1)	(2)	(3)
Panel A: Staircase Risk Aversion			
Mean Lifetime Temp	-1.842***	-0.123	-0.474
	(0.594)	(0.291)	(0.493)
R^2	0.153	0.159	0.212
Panel B: CRRA $\hat{\rho}$			
Mean Lifetime Temp	-0.605***	-0.292***	-0.104
	(0.082)	(0.059)	(0.107)
R^2	0.124	0.126	0.177
N	78,221	78,221	77,828
Region FE	X	X	X
Cohort FE		X	
Country × Cohort FE			X

Notes: The dependent variable in panel A is the GPS staircase lottery measure (1–32 scale) normalized so higher numbers mean more risk aversion. The dependent variable in Panel B is the CRRA $\hat{\rho}$ estimated from the staircase measure using the procedure in Section 2.3. Empirical specifications for the three columns can be found in Section 2.8. Regressions are at the subject region level. All specifications control for gender, language dummies, and hemisphere×interview-month seasonality fixed effects. Lifetime mean precipitation is *not* included. Standard errors, clustered two-way by region and birth year, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 16: Lifetime Mean Temperature and Risk Aversion (Adding Interview-Day Anomalies)

	(1)	(2)	(3)
Panel A: Staircase Risk Aversion			
Mean Lifetime Temp	-1.875***	-0.196	-0.705*
	(0.609)	(0.312)	(0.362)
R^2	0.154	0.160	0.214
Panel B: CRRA $\hat{\rho}$			
Mean Lifetime Temp	-0.617***	-0.317***	-0.177*
	(0.076)	(0.052)	(0.103)
R^2	0.125	0.127	0.178
N	77,254	77,254	76,865
Region FE	X	X	X
Cohort FE		X	
Country × Cohort FE			X

Notes: The dependent variable in panel A is the GPS staircase lottery measure (1–32 scale) normalized so higher numbers mean more risk aversion. The dependent variable in Panel B is the CRRA $\hat{\rho}$ estimated from the staircase measure using the procedure in Section 2.3. Empirical specifications for the three columns can be found in Section 2.8. Regressions are at the subject region level. All specifications control for gender, language dummies, lifetime mean precipitation, hemisphere×interview-month seasonality FEs, interview-day mean temperature anomaly (z-score), and night-before minimum temperature anomaly (z-score). Standard errors, clustered two-way by region and birth year, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 17: Lifetime Mean Temperature and Risk Aversion (Adding Anomalies + Lifetime SD)

	(1)	(2)	(3)
Panel A: Staircase Risk Aversion			
Mean Lifetime Temp	-1.782***	-0.182	-0.741*
	(0.617)	(0.288)	(0.402)
R^2	0.154	0.160	0.214
Panel B: CRRA $\hat{\rho}$			
Mean Lifetime Temp	-0.593***	-0.314***	-0.132
	(0.075)	(0.050)	(0.123)
R^2	0.125	0.127	0.178
N	77,254	77,254	76,865
Region FE	X	X	X
Cohort FE		X	
Country × Cohort FE			X

Notes: The dependent variable in panel A is the GPS staircase lottery measure (1–32 scale) normalized so higher numbers mean more risk aversion. The dependent variable in Panel B is the CRRA $\hat{\rho}$ estimated from the staircase measure using the procedure in Section 2.3. Empirical specifications for the three columns can be found in Section 2.8. Regressions are at the subject region level. All specifications control for gender, language dummies, lifetime mean precipitation, hemisphere×interview-month seasonality FEs, interview-day temperature anomaly (z-score), night-before Tmin anomaly (z-score), lifetime temperature SD, and lifetime precipitation SD. Standard errors, clustered two-way by region and birth year, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.9 Robustness: alternative seasonality fixed effects

Table 18: Lifetime Mean Temperature and Risk Aversion (No Seasonality FEs)

	(1)	(2)	(3)
Panel A: Staircase Risk Aversion			
Mean Lifetime Temp	-0.943**	0.004	-0.729**
	(0.384)	(0.385)	(0.345)
R^2	0.149	0.156	0.212
Panel B: CRRA $\hat{\rho}$			
Mean Lifetime Temp	-0.427***	-0.255	-0.193*
	(0.143)	(0.175)	(0.097)
R^2	0.123	0.125	0.177
N	78,494	78,494	78,100
Region FE	X	X	X
Cohort FE		X	
Country × Cohort FE			X

Notes: The dependent variable in panel A is the GPS staircase lottery measure (1–32 scale) normalized so higher numbers mean more risk aversion. The dependent variable in Panel B is the CRRA $\hat{\rho}$ estimated from the staircase measure using the procedure in Section 2.3. Empirical specifications for the three columns can be found in Section 2.8. Regressions are at the subject region level. All specifications control for gender, language dummies, and lifetime mean precipitation. Hemisphere×interview-month seasonality fixed effects are *not* included. Standard errors, clustered two-way by region and birth year, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 19: Lifetime Mean Temperature and Risk Aversion (Country × Month FEs)

	(1)	(2)	(3)
Panel A: Staircase Risk Aversion			
Mean Lifetime Temp	-3.411***	-0.696**	-0.732**
	(0.497)	(0.346)	(0.351)
R^2	0.156	0.160	0.213
Panel B: CRRA $\hat{\rho}$			
Mean Lifetime Temp	-0.738***	-0.125	-0.197**
	(0.123)	(0.088)	(0.098)
R^2	0.125	0.127	0.178
N	78,493	78,493	78,099
Region FE	X	X	X
Cohort FE		X	
Country × Cohort FE			X

Notes: The dependent variable in panel A is the GPS staircase lottery measure (1–32 scale) normalized so higher numbers mean more risk aversion. The dependent variable in Panel B is the CRRA $\hat{\rho}$ estimated from the staircase measure using the procedure in Section 2.3. Empirical specifications for the three columns can be found in Section 2.8. Regressions are at the subject region level. All specifications control for gender, language dummies, and lifetime mean precipitation. Hemisphere×month FEs are replaced with country×interview-month FEs (∼912 groups). Standard errors, clustered two-way by region and birth year, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.10 Robustness: alternative clustering

Table 20: Lifetime Mean Temperature and Risk Aversion (Clustering: Region × Cohort)

	(1)	(2)	(3)
Panel A: Staircase Risk Aversion			
Mean Lifetime Temp	-1.883***	-0.196	-0.722**
	(0.203)	(0.198)	(0.358)
R^2	0.153	0.159	0.212
Panel B: CRRA $\hat{\rho}$			
Mean Lifetime Temp	-0.622***	-0.318***	-0.177
	(0.082)	(0.100)	(0.121)
R^2	0.124	0.126	0.177
N	78,221	78,221	77,828
Region FE	X	X	X
Cohort FE		X	
Country × Cohort FE			X

Notes: The dependent variable in panel A is the GPS staircase lottery measure (1–32 scale) normalized so higher numbers mean more risk aversion. The dependent variable in Panel B is the CRRA $\hat{\rho}$ estimated from the staircase measure using the procedure in Section 2.3. Empirical specifications for the three columns can be found in Section 2.8. Regressions are at the subject region level. All specifications control for gender, language dummies, lifetime mean precipitation, and hemisphere×interview-month seasonality fixed effects. Standard errors are clustered by region×cohort (single-way). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 21: Lifetime Mean Temperature and Risk Aversion (Clustering: Region Only)

	(1)	(2)	(3)
Panel A: Staircase Risk Aversion			
Mean Lifetime Temp	-1.883***	-0.196	-0.722**
	(0.608)	(0.310)	(0.367)
R^2	0.153	0.159	0.212
Panel B: CRRA $\hat{\rho}$			
Mean Lifetime Temp	-0.622***	-0.318***	-0.177*
	(0.086)	(0.078)	(0.104)
R^2	0.124	0.126	0.177
N	78,221	78,221	77,828
Region FE	X	X	X
Cohort FE		X	
Country × Cohort FE			X

Notes: The dependent variable in panel A is the GPS staircase lottery measure (1–32 scale) normalized so higher numbers mean more risk aversion. The dependent variable in Panel B is the CRRA $\hat{\rho}$ estimated from the staircase measure using the procedure in Section 2.3. Empirical specifications for the three columns can be found in Section 2.8. Regressions are at the subject region level. All specifications control for gender, language dummies, lifetime mean precipitation, and hemisphere×interview-month seasonality fixed effects. Standard errors are clustered by region only (single-way). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 22: Lifetime Mean Temperature and Risk Aversion (Clustering: Country Only)

	(1)	(2)	(3)
Panel A: Staircase Risk Aversion			
Mean Lifetime Temp	-1.883***	-0.196	-0.722*
	(0.602)	(0.304)	(0.404)
R^2	0.153	0.159	0.212
Panel B: CRRA $\hat{\rho}$			
Mean Lifetime Temp	-0.622***	-0.318***	-0.177**
	(0.102)	(0.052)	(0.082)
R^2	0.124	0.126	0.177
N	78,221	78,221	77,828
Region FE	X	X	X
Cohort FE		X	
Country × Cohort FE			X

Notes: The dependent variable in panel A is the GPS staircase lottery measure (1–32 scale) normalized so higher numbers mean more risk aversion. The dependent variable in Panel B is the CRRA $\hat{\rho}$ estimated from the staircase measure using the procedure in Section 2.3. Empirical specifications for the three columns can be found in Section 2.8. Regressions are at the subject region level. All specifications control for gender, language dummies, lifetime mean precipitation, and hemisphere×interview-month seasonality fixed effects. Standard errors are clustered by country only (single-way). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 23: Lifetime Mean Temperature and Risk Aversion (Clustering: Cohort Only)

	(1)	(2)	(3)
Panel A: Staircase Risk Aversion			
Mean Lifetime Temp	-1.883***	-0.196	-0.722**
	(0.193)	(0.197)	(0.344)
R^2	0.153	0.159	0.212
Panel B: CRRA $\hat{\rho}$			
Mean Lifetime Temp	-0.622***	-0.318***	-0.177
	(0.071)	(0.081)	(0.116)
R^2	0.124	0.126	0.177
N	78,221	78,221	77,828
Region FE	X	X	X
Cohort FE		X	
Country × Cohort FE			X

Notes: The dependent variable in panel A is the GPS staircase lottery measure (1–32 scale) normalized so higher numbers mean more risk aversion. The dependent variable in Panel B is the CRRA $\hat{\rho}$ estimated from the staircase measure using the procedure in Section 2.3. Empirical specifications for the three columns can be found in Section 2.8. Regressions are at the subject region level. All specifications control for gender, language dummies, lifetime mean precipitation, and hemisphere×interview-month seasonality fixed effects. Standard errors are clustered by birth year cohort only (single-way). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.11 National panel risk aversion instruments and histograms

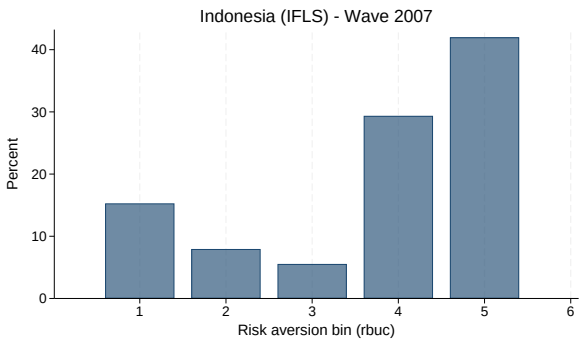
A.11.1 Indonesia

Q	Question text (IFLS4, IFLS5)	Responses & routing
SI01	Suppose you are offered two ways to earn some money. *Option 1*: You are guaranteed Rp 800,000 per month. *Option 2*: You have an equal chance of either the same income, Rp 800,000 per month, or, if you are lucky, Rp 1,600,000 per month. Which option would you choose?	1. Rp 800,000 → **SI02** 2. Rp 1,600,000 or Rp 800,000 → **SI03** 8. Don't know
SI02	Are you sure? In Option 2 you will get at least Rp 800,000 and you may get Rp 1,600,000 per month. In Option 1 you will always get Rp 800,000 per month.	1. Still picks Option 1 → **SI11** 2. Switches to Option 2 8. Don't know
SI03	Now, in Option 2 you have an equal chance of receiving either Rp 1,600,000 or Rp 400,000 per month, depending on how lucky you are. Option 1 guarantees an income of Rp 800,000 per month. Which option would you choose?	1. Rp 800,000 → **SI04** 2. Rp 1,600,000 or Rp 400,000 → **SI05** 8. Don't know
SI04	Now, in Option 2 you have an equal chance of receiving either Rp 1,600,000 or Rp 600,000 per month, depending on how lucky you are. Option 1 guarantees an income of Rp 800,000 per month. Which option would you choose?	1. Rp 800,000 → **SI11** 2. Rp 1,600,000 or Rp 600,000 → **SI11** 8. Don't know
SI05	Now, in Option 2 you have an equal chance of receiving either Rp 1,600,000 or Rp 200,000 per month, depending on how lucky you are. Option 1 guarantees an income of Rp 800,000 per	1. Rp 800,000 → **SI11** 2. Rp 1,600,000 or Rp 200,000 → **SI11** 8. Don't know

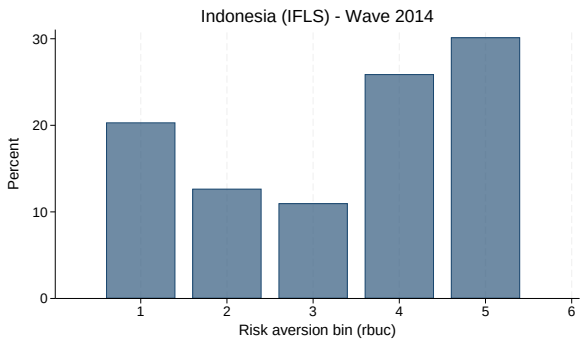
The branching logic produces five terminal nodes, which define the `rbuc` bins (higher = more risk averse).

`rbuc`	Path
1	SI01→risky, SI03→risky, SI05→risky
2	SI01→risky, SI03→risky, SI05→safe
3	SI01→risky, SI03→safe, SI04→risky
4	SI01→risky, SI03→safe, SI04→safe
5	SI01→safe (confirmed in SI02)

Figure 9: Distribution of `rbuc` by wave, Indonesia (IFLS)



(a) IFLS4



(b) IFLS5

A.11.2 Mexico

Q	Question text (MxFLS2)	Responses & routing
RG01	Before we continue, what color chip do you have the highest probability of getting? *(Comprehension check; interviewer corrects if wrong.)*	1. Blue 2. Yellow 3. Same probability 8. DK
RG02	Choose between two bags: *Bag 1*: Blue or yellow chip, you receive $1,000. *Bag 2*: Blue chip $500, yellow chip $2,000. Which bag do you choose?	1. Bag 1 → **RG05** 2. Bag 2 8. DK → **RG05**
RG03	Choose between two bags: *Bag 1*: Blue chip $500, yellow chip $2,000. *Bag 2*: Blue chip $300, yellow chip $3,000. Which bag do you choose?	1 → **RG05** 2 8 → **RG05**
RG04	Choose between two bags: *Bag 1*: Blue chip $100, yellow chip $4,000. *Bag 2*: Blue chip $100, yellow chip $7,000. Which bag do you choose?	1 → **RG08** 2 → **RG08** 8
RG05	Choose between two bags: *Bag 1*: Blue chip $1,000, yellow chip $1,000. *Bag 2*: Blue chip $800, yellow chip $2,000. Which bag do you choose?	1 2 → **RG08** 8 → **RG08**
RG06	Choose between two bags: *Bag 1*: Blue chip $1,000, yellow chip $1,000. *Bag 2*: Blue chip $800, yellow chip $4,000. Which bag do you choose?	1 2 → **RG08** 8
RG07	Choose between two bags: *Bag 1*: Blue chip $1,000, yellow chip $1,000. *Bag 2*: Blue chip $800, yellow chip $8,000. Which bag do you choose?	1 2 8

The branching logic produces 5 terminal bins in MxFLS2.

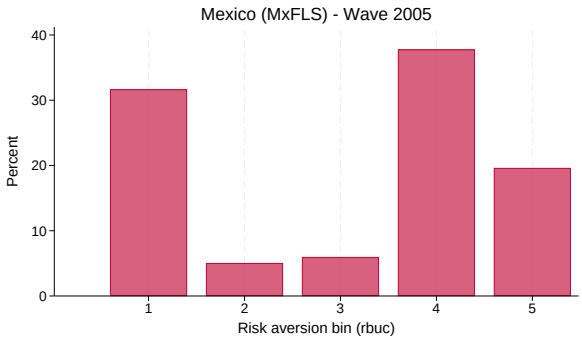
Q	Question text (MxFLS3)	Responses & routing
RG01	Choose between two bags: *Bag 1*: One ball worth $2,500. *Bag 2*: Two balls, one worth $2,500 (same as Bag 1) and one worth $5,000. Which bag do you choose?	1. $2,500 or $2,500 2. $2,500 or $5,000 → **RG03** 8. DK
RG02	Are you sure? If you choose Bag 1 you will win $2,500. If you choose Bag 2, you will win at least $2,500 and probably $5,000, depending on your luck. Which bag do you choose?	1. Still Bag 1 → **RG07** 2. Changes to Bag 2 8. DK
RG03	Choose between two bags: *Bag 1*: Guarantees $2,500. *Bag 2*: One ball worth $2,000, one worth $5,000. Which bag do you choose?	1. $2,500 → **RG08** 2. $5,000 or $2,000 8. DK
RG04	Choose between two bags: *Bag 1*: Guarantees $2,500. *Bag 2*: One ball worth $1,500, one worth $5,000. Which bag do you choose?	1. $2,500 → **RG08** 2. $5,000 or $1,500 8. DK
RG05	Choose between two bags: *Bag 1*: Guarantees $2,500. *Bag 2*: One ball worth $1,000, one worth $5,000. Which bag do you choose?	1. $2,500 → **RG08** 2. $5,000 or $1,000 8. DK
RG06	Choose between two bags: *Bag 1*: Guarantees $2,500. *Bag 2*: One ball worth $500, one worth $5,000. Which bag do you choose?	1. $2,500 2. $5,000 or $500 → **RG08** 8. DK → **RG08**
RG07	Finally, choose between two bags: *Bag 1*: Guarantees $2,000. *Bag 2*: One ball worth $2,500, one worth $5,000. Which bag do you choose?	1. $2,000 2. $5,000 or $2,500 8. DK

The MxFLS3 branching tree produces up to 7 terminal bins. In the analysis, bins 6 and 7 are collapsed into bin 5 to harmonize with the MxFLS2 five-bin scale.

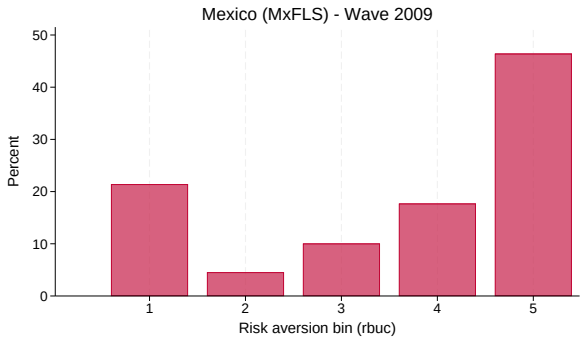
`rbuc`	Path (MxFLS2)
1	RG02→Bag 2, RG03→Bag 2, RG04→Bag 2
2	RG02→Bag 2, RG03→Bag 2, RG04→Bag 1
3	RG02→Bag 2, RG03→Bag 1
4	RG02→Bag 1, RG05→Bag 2
5	RG02→Bag 1, RG05→Bag 1 (safe in all)

`rbuc`	Path (MxFLS3)
1	RG01→Bag 2, RG03→Bag 2, RG04→Bag 2, RG05→Bag 2, RG06→Bag 2
2	RG01→Bag 2, RG03→Bag 2, RG04→Bag 2, RG05→Bag 2, RG06→Bag 1
3	RG01→Bag 2, RG03→Bag 2, RG04→Bag 2, RG05→Bag 1
4	RG01→Bag 2, RG03→Bag 2, RG04→Bag 1
5	RG01→Bag 2, RG03→Bag 1
6	RG01→Bag 1 (switches in RG02)
7	RG01→Bag 1 (confirmed safe in RG02) → RG07

Figure 10: Distribution of `rbuc` by wave, Mexico (MxFLS)



(a) MxFLS2



(b) MxFLS3

A.11.3 Chile

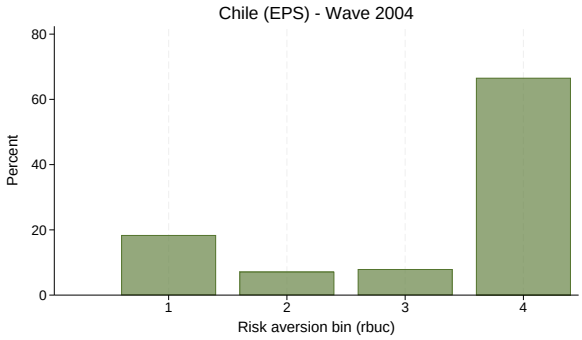
The Chilean Social Protection Survey (EPS) elicits risk preferences via a hypothetical job-choice scenario (Section J1). The respondent is told: "Suppose that your only source of household income is a choice between the following two jobs." Each situation presents Alternative A (a safe job with fixed, stable income for life) versus Alternative B (a risky job with an equal chance of doubling income or losing a specified fraction). The module uses a branching staircase with 3 situations.

Situation	Question text (EPS 2004, 2006, 2009)	Responses & routing
1st	*Alternative A*: A job with fixed, stable income for life. *Alternative B*: A job with an equal chance of doubling your income or keeping only **1/4** of your income, for life. Which do you choose?	1. Alternative A 2. Alternative B → **J2** 4. NS/NR
2nd	*Alternative A*: A job with fixed, stable income for life. *Alternative B*: A job with an equal chance of doubling your income or keeping only **1/2** of your income, for life. Which do you choose?	1. Alternative A 2. Alternative B → **J2** 4. NS/NR
3rd	*Alternative A*: A job with fixed, stable income for life. *Alternative B*: A job with an equal chance of doubling your income or keeping only **3/4** of your income, for life. Which do you choose?	1. Alternative A 2. Alternative B 4. NS/NR

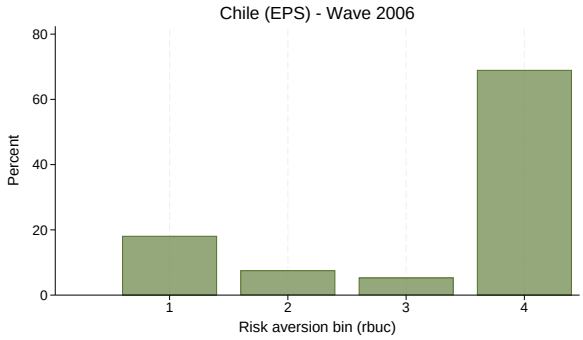
`rbuc`	Path
1	Accepts 1st situation (downside: keep 1/4)
2	Rejects 1st, accepts 2nd (downside: keep 1/2)
3	Rejects 1st and 2nd, accepts 3rd (downside: keep 3/4)
4	Rejects all three situations

The 2002 wave uses a different elicitation with 5 binary questions producing 6 bins; it is excluded from the main specification due to incompatibility with the 4-bin scale.

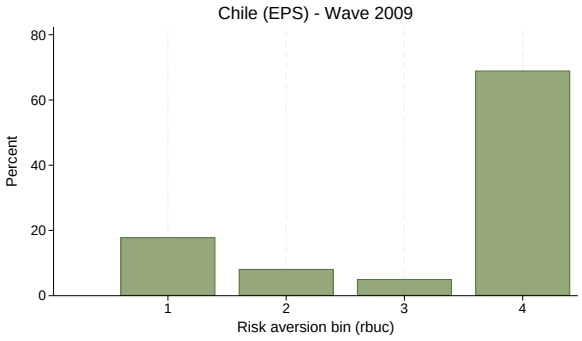
Figure 11: Distribution of `rbuc` by wave, Chile (EPS)



(a) EPS 2004



(b) EPS 2006



(c) EPS 2009

A.11.4 Japan

Suppose that there is a "speed lottery" with a 50% chance of winning ¥100,000. If you win, you get the prize right away. If you lose, you get nothing. How much would you spend to buy

a ticket for this lottery? Choose Option "A" if you would buy it at that price, and choose Option "B" if you would not buy the ticket at that price. (X ONE Box For EACH Row)

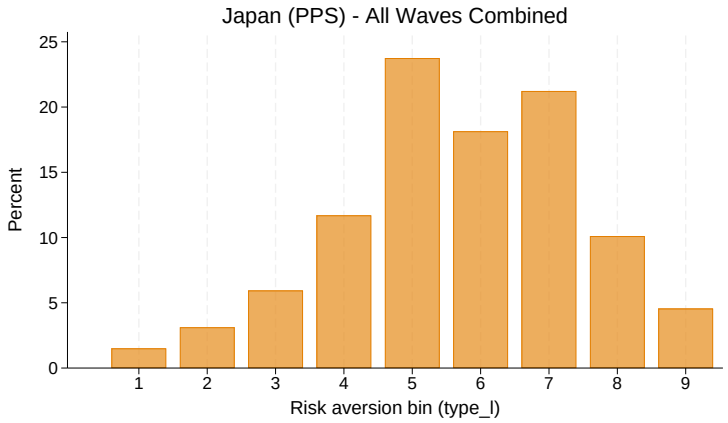
Price of ticket	Option A (buy)	Option B (do not buy)
¥10	☐	☐
¥2,000	☐	☐
¥4,000	☐	☐
¥8,000	☐	☐
¥15,000	☐	☐
¥25,000	☐	☐
¥35,000	☐	☐
¥50,000	☐	☐

Respondents with consistent (monotone) switching patterns are assigned to one of 9 bins:

`type_1`	Response pattern (rows 1–8)	Interpretation
1	(A, A, A, A, A, A, A, A)	Buys at all prices
2	(A, A, A, A, A, A, A, B)	WTP between ¥35,000 and ¥50,000
3	(A, A, A, A, A, A, B, B)	WTP between ¥25,000 and ¥35,000
4	(A, A, A, A, A, B, B, B)	WTP between ¥15,000 and ¥25,000
5	(A, A, A, A, B, B, B, B)	WTP between ¥8,000 and ¥15,000
6	(A, A, A, B, B, B, B, B)	WTP between ¥4,000 and ¥8,000
7	(A, A, B, B, B, B, B, B)	WTP between ¥2,000 and ¥4,000
8	(A, B, B, B, B, B, B, B)	WTP between ¥10 and ¥2,000
9	(B, B, B, B, B, B, B, B)	Never buys

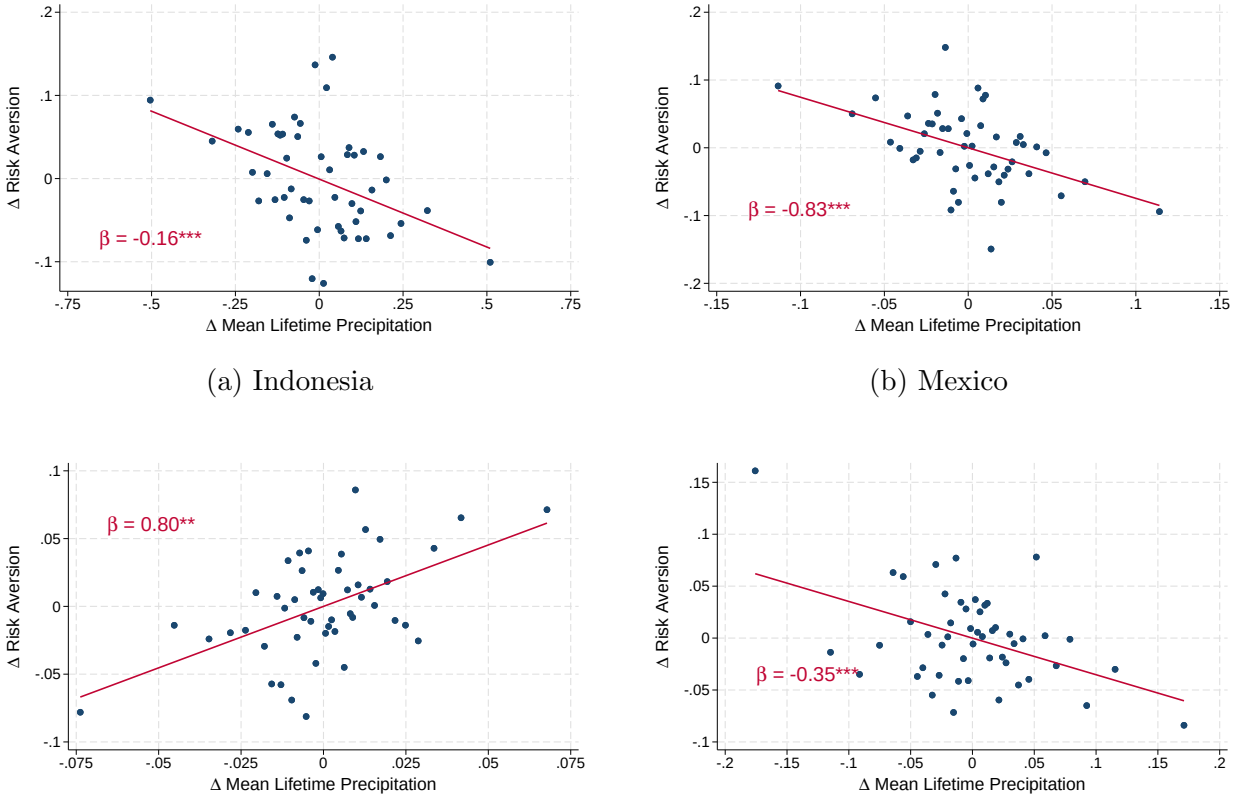
Respondents with non-monotone patterns (multiple switching) are coded as missing. The module is available in waves 2011–2024.

Figure 12: Distribution of `rbuc` pooled across all waves, Japan (PPS)



A.12 Precipitation scatterplots: National panels

Figure 13: ΔLifetime precipitation and Δrisk aversion across four national panels.



(a) Indonesia

(b) Mexico

(c) Chile

(d) Japan

Notes: Each panel plots means of risk aversion against lifetime mean precipitation, both residualized on individual fixed effects, wave fixed effects, and lifetime mean temperature (FWL). The slope of the fitted line equals the coefficient on lifetime precipitation from the full panel regression. 50 quantile bins. Standard errors clustered at birth region × birth year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.13 National panels regression results

Table 24: Panel FE regressions: lifetime temperature and precipitation on risk aversion

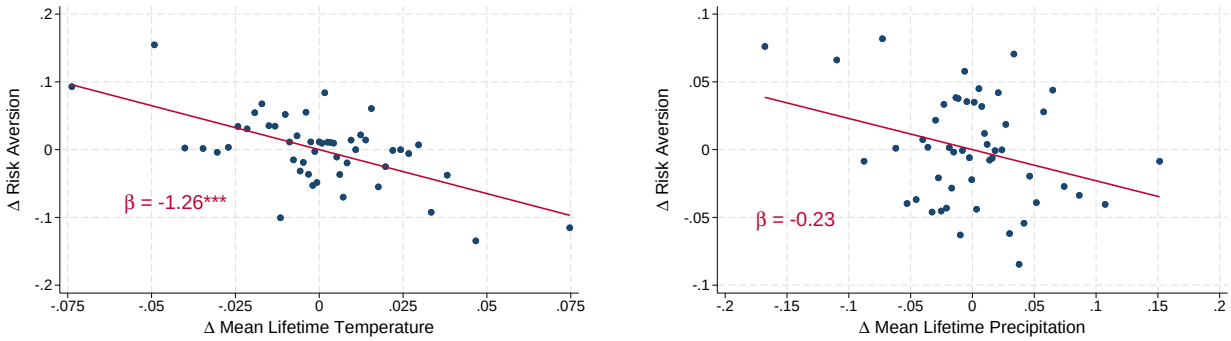
	(1) Indonesia	(2) Mexico	(3) Chile	(4) Japan
Lifetime mean temperature	-3.193***	-2.085***	-0.974*	-1.019***
	(0.433)	(0.727)	(0.519)	(0.305)
Lifetime mean precipitation	-0.159***	-0.832***	0.796**	-0.348***
	(0.054)	(0.299)	(0.323)	(0.124)
Individual FE	✓	✓	✓	✓
Wave/Year FE	✓	✓	✓	✓
Survey waves	2007, 2014	2006, 2009	2004, 2006, 2009	2011–2024
Observations	42,114	24,588	40,672	29,949
Unique subjects	21,057	12,294	15,496	4,980
Within R^2	0.0038	0.0016	0.0006	0.0008

Notes: Standard errors clustered at birth region × birth year in parentheses. Climate location: birth province (Indonesia), birth state (Mexico), birth region (Chile), prefecture at age 15 (Japan). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.14 Robustness: Japan excluding 2011-2012

Hanaoka et al. (2018) show that the 2011 Great East Japan Earthquake shifted risk preferences between the 2011 and 2012 PPS waves. As a robustness check, we re-estimate the Japan specification excluding these two waves. Figure 14 presents binscatter plots for both temperature and precipitation. The temperature coefficient strengthens from -1.02 to -1.26 ($p < 0.001$, $N = 20{,}663$), and the precipitation coefficient becomes insignificant. The results confirm that the baseline Japan estimates are not driven by earthquake-related preference shifts.

Figure 14: Japan regressions excluding 2011–2012 waves



(a) Temperature

(b) Precipitation

Notes: Sample restricted to waves 2013–2024. Both panels residualized on individual FE, year FE, and the other climate variable. 50 quantile bins. Standard errors clustered at prefecture × birth year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B Data Appendix

B.1 Decoding the GPS risk preference sub-scales

The GPS public-release data provide a single composite risk-taking index for each respondent, constructed as a weighted sum of z-scored sub-scales, but do not include the individual sub-scale values. Of the 80,337 GPS respondents, 634 are missing the composite risk-taking variable due to incomplete responses on one or both sub-scales; the decoding procedure described below applies to the remaining 79,703. Because our analysis examines the staircase and Likert measures separately, and because the structural estimation of CRRA parameters (Section 2.3) requires the discrete staircase response, we developed a procedure to recover the underlying sub-scale values from the composite index alone.

B.1.1 Construction of the Composite Index

According to the GPS methodology (Falk et al. (2018) online appendix), the composite risk-taking measure is constructed as:

$$\texttt{risktaking}_i = w_1 \cdot z(S_i) + w_2 \cdot z(L_i) \tag{7}$$

where $S_i \in \{1, \ldots, 32\}$ is the staircase risk-taking response, $L_i \in \{0, \ldots, 10\}$ is the Likert willingness-to-take-risks response, $z(\cdot)$ denotes z-score standardization, and (w_1, w_2) are OLS-derived weights that sum to one (0.4729985 for the staircase and and 0.5270015 for the Likert). For respondents with complete data on both items, the composite is a deterministic function of two integer inputs, yielding at most $32 \times 11 = 352$ distinct values arranged on a regular two-dimensional lattice in composite space. However, the GPS data contain 735 unique values of the composite, because the GPS imputation procedure generated fractional intermediate values for respondents with incomplete item responses. Crucially, the magnitude of the z-scores for each sub-scale is not recorded anywhere in the published literature. The

reconstruction of the sub-scale values from the published combined measure is impossible without a procedure that recovers the z-scores.

B.1.2 Identifying the Lattice Step Sizes

The key insight enabling the decoding is that z-score standardization is a linear transformation that preserves equal spacing. A one-unit increase in the staircase response changes the composite by a constant amount $B = w_1/\mathrm{SD}(S)$, regardless of the starting value; similarly, a one-unit increase in the Likert response changes the composite by $A = w_2/\mathrm{SD}(L)$. The 352 lattice points corresponding to complete-data respondents are therefore separated by exact integer multiples of these two step sizes.

To discover A and B without knowing their values *a priori*, we exploited the fact that the most frequently occurring composite values are overwhelmingly likely to correspond to exact integer subscale combinations, since each lattice point pools all respondents with that (S, L) pair, while imputed values are more idiosyncratic. For each of the 32 most frequent composite values, we computed the distance to every other value in the support and identified distances that appeared *symmetrically*—i.e., there existed values at both $+d$ and $-d$ from the reference point. This symmetric-distance criterion is highly discriminating: imputed values, which are perturbed by regression prediction and floating-point arithmetic, are extremely unlikely to sit at exactly equal distances on both sides of a reference point.

This analysis revealed two step sizes with dramatically higher frequency than any others: $A = 0.251324$ and $B = 0.058957$, together with their integer multiples. The ratio $A/B \approx 4.26$ is consistent with the expected structure: the Likert scale spans 11 levels while the staircase spans 32, so the staircase step should be (in expectation) considerably smaller.

B.1.3 Lattice Identification and Coordinate Assignment

With the step sizes in hand, we identified which of the 735 unique composite values belong to the 352-point lattice using iterative expansion from seed points. Starting from the

two endpoint values (the minimum and maximum of the composite, which correspond to (S=1, L=0) and (S=32, L=10)), we iteratively searched for other values at integer multiples of A or B away (within floating-point tolerance of 10^{-6}). Newly identified lattice points became seeds for the next iteration. The search converged in 17 iterations to a connected component of exactly 352 values—precisely 32×11—confirming the lattice structure. As a robustness check, starting from only a single endpoint required 34 iterations to reach the same 352-point lattice, containing the other endpoint.

We then assigned (S, L) coordinates by propagating labels outward from the anchor point (S=1, L=0) at the minimum composite value. For each labeled point (S_0, L_0), unlabeled lattice points at distance $k \cdot A$ were assigned $(S_0, L_0 + k)$ (a Likert move), and those at distance $k \cdot B$ were assigned $(S_0 + k, L_0)$ (a staircase move), subject to the bounds $S \in [1, 32]$ and $L \in [0, 10]$. This produced a perfect 32×11 cross-tabulation with exactly one composite value per (S, L) cell.

For the 383 non-lattice composite values (corresponding to respondents with imputed subscale data), we assigned the (S, L) coordinates of the nearest lattice point in composite space.

B.1.4 Verification

Table 25 reports verification statistics for the decoding procedure. The procedure recovers exactly 352 lattice points spanning all 32 staircase levels and all 11 Likert levels, with zero duplicate (S, L) assignments. These 352 values account for 89.6% of all respondents. The maximum reconstruction error on lattice points is 9.1×10^{-6}—within floating-point precision—confirming that the composite is indeed a linear function of the subscales with the identified step sizes.

From the empirical standard deviations of the on-lattice sub-scales ($\mathrm{SD}(S) = 10.74$, $\mathrm{SD}(L) = 2.77$) and the identified step sizes, we can recover the construction weights as $w_1 = B \times \mathrm{SD}(S) = 0.633$ and $w_2 = A \times \mathrm{SD}(L) = 0.695$, yielding weight shares of 0.477 (staircase)

Table 25: Verification of GPS Subscale Decoding

Verification Check	Value
Panel A: Lattice Structure	
Unique `risktaking` values in GPS data	735
Values on decoded lattice	352
Staircase levels recovered	32 (of 32)
Likert levels recovered	11 (of 11)
Share of respondents on lattice (%)	89.6
Max reconstruction error (lattice)	9.10e-06
Panel B: Subscale–Composite Correlations	
Pearson r(risktaking, staircase)	0.6977
Pearson r(risktaking, likert)	0.7767
Pearson r(staircase, likert)	0.0907
Panel C: Z-Score Consistency	
Mean(staircase), on-lattice	10.76
SD(staircase), on-lattice	10.74
Mean(likert), on-lattice	5.21
SD(likert), on-lattice	2.77
Implied staircase weight share	0.4769
Implied likert weight share	0.5231
Pearson r(risktaking, $\hat{r}$ from z-scores)	1.000000
Panel D: Full-Sample Reconstruction	
Pearson r(risktaking, reconstruction)	0.999998
Mean |error|	0.000384
Share within 0.001 (%)	91.7
Share within 0.01 (%)	99.4
Share within 0.1 (%)	100.0

Notes: Panel A confirms the lattice structure: exactly 352 values form a complete 32×11 grid, covering 89.6% of respondents, with reconstruction error below 10^{-5}. Panel B reports correlations between the decoded subscales and the composite index; the cross-subscale correlation of 0.091 confirms that the staircase and Likert measure largely distinct dimensions of risk attitudes. Panel C verifies that the implied construction weights (0.477 staircase, 0.523 Likert) closely match the published GPS weights ($\sim$0.473 and $\sim$0.527). Panel D shows that even including nearest-neighbor imputed values, the reconstruction correlates 0.999998 with the original composite, with 99.4% of observations reconstructed within 0.01.

and 0.523 (Likert). These closely match the published GPS weights of approximately 0.473 and 0.527, providing independent confirmation that the decoding is correct.

Reconstructing the composite from the decoded z-scores using the implied weights produces a correlation of $r = 1.000000$ with the original composite for on-lattice observations. For the full sample (including nearest-neighbor imputed values for non-lattice respondents), the correlation is $r = 0.999998$, with 99.4% of all observations reconstructed within 0.01 of their true composite value.

We repeated the lattice discovery algorithm starting from off-lattice composite values rather than the identified lattice endpoints. If the lattice structure were spurious – arising, for example, from artifacts of the imputation procedure or from numerical coincidence – we would expect to find comparably large connected components originating from off-lattice seeds. In practice, no off-lattice starting point yielded a connected component of more than a handful of points, confirming that the 352-point lattice we identify is a unique and genuine structural feature of the data rather than a statistical artifact.

B.2 Structural Estimation of CRRA Relative Risk Aversion $\hat{\rho}$

This section documents the full procedure for structurally estimating the coefficient of relative risk aversion $\hat{\rho}$ from the GPS staircase lottery data. We describe: (i) construction of proxies for country-specific lottery prizes; (ii) construction of the regional GNI per capita income proxy assigned to individuals; (iii) the CRRA indifference equation and its normalization for numerical stability; (iv) the bracketing assumptions and income integration horizons; (v) the RPM+Tremble estimation model; and (vi) the output, quality checks, and specification choice.

B.2.1 Proxies for Country-Specific Lottery Prizes

The GPS staircase risk module presents each respondent with a sequence of five binary choices between a sure payment and a 50/50 lottery paying X or nothing. The German baseline uses $X = 300$ EUR, with 31 safe payment levels $s_b = 10b$ EUR for $b = 1, \ldots, 31$. For deployment in other countries, all monetary amounts were scaled by a country-specific factors k_c equal to the ratio of national median incomes, and the lottery prize was rounded to a convenient number. The safe payments were then adjusted proportionally (Falk et al., 2018, Online Appendix §AE.2). Unfortunately, the country-specific factors were never published, and the Gallup country-level median household income measures that were likely used to construct them are only available behind the Gallup paywall.

We proxy for the k_c factors using national median per-capita income or consumption from the World Bank Poverty and Inequality Platform (PIP), queried for 2012 with gap-filling. PIP reports incomes in international dollars, converted from local currency using ICP purchasing power parities. We consider two PPP vintages: the 2011 ICP round, which is near-contemporaneous to the GPS survey year, and the 2017 ICP round. Each ICP round conducts a fresh cross-country price comparison, and PIP extrapolates the resulting benchmark PPPs to the query year (2012) using domestic CPI. The two PIP vintages therefore differ in both the PPP conversion factors applied and the underlying survey data and methodology, as PIP

revises its income estimates with each ICP round. As a third alternative, we use publicly available Gallup data listing median household income globally, averaged across 2006–2012. PIP data is available for all 76 GPS countries; the Gallup data covers 71.[15]

For each measure $j \in \{\text{PIP 11}, \text{PIP 17}, \text{Gallup}\}$, we define the calculated scaling factor as

$$k_c^{\text{calc},j} = \frac{\text{median income}_c^j}{\text{median income}_{\text{Germany}}^j}, \tag{8}$$

so that Germany is normalized to $k_c = 1$ by construction. Because each PIP vintage applies a different set of PPP conversion factors to (potentially revised) local-currency survey data, the ratio in Equation (8) is a dimensionless quantity whose value depends on the choice of j.

We select among these measures by comparing against the true k_c for five countries where implementation errors in the QJE Online Appendix (§AG.1) inadvertently reveal the actual local-currency lottery amounts: Indonesia, Pakistan, Ukraine, Vietnam, and Malawi. For each of these countries we extract the implied local-currency-to-EUR exchange rate e_c (LCU per EUR) from the errata. The ground-truth scaling factors under PPP vintage j are

$$k_c^{\text{true},j} = e_c \times \frac{\text{PPP}_{\text{Germany}}^j}{\text{PPP}_c^j}, \tag{9}$$

where PPP_c^j is the ICP conversion factor (LCU per international dollar) for country c under vintage j, extrapolated to 2012 via domestic CPI. This converts e_c from local currency units per EUR into a dimensionless ratio comparable to $k_c^{\text{calc},j}$.

Crucially, both $k_c^{\text{calc},j}$ and $k_c^{\text{true},j}$ must use the *same* PPP vintage for an apples-to-apples comparison: the PIP 11 calculated factors are compared against ground-truth factors computed with 2011 PPP, and the PIP 17 factors against ground-truth factors computed with 2017 PPP. For Gallup, whose PPP vintage is undocumented, we compare against $k_c^{\text{true},11}$.

[15] PIP data is available at https://pip.worldbank.org. ICP PPP conversion factors were downloaded via the World Bank API. The Gallup data, to our knowledge the only public data from Gallup on cross-country income differences in the period, was downloaded from https://news.gallup.com/poll/166211/worldwide-median-household-income-000.aspx.

Table 26: Calculated k_c vs. Ground-Truth k_c: Alternative Data Sources

		Ground Truth		$k_c^{\text{calc},j}$			$k_c^{\text{calc},j}/k_c^{\text{true},j}$		
Country	e_c	$k_c^{\text{true},11}$	$k_c^{\text{true},17}$	PIP '11	PIP '17	Gallup	PIP '11	PIP '17	Gallup
Indonesia	400.0	0.0850	0.0619	0.0869	0.0800	0.0660	1.023	1.294	0.776
Malawi	2.5	0.0263	0.0081	0.0287	0.0309	—	1.093	3.799	—
Pakistan	4.0	0.1298	0.0947	0.0773	0.0739	0.1218	0.596	0.780	0.939
Ukraine	1.0	0.2599	0.1124	0.2636	0.2322	0.3322	1.014	2.066	1.278
Vietnam	2000.0	0.2197	0.2016	0.1456	0.1617	0.1435	0.663	0.802	0.653
Mean $\lvert\log r\rvert$							0.211	0.557	0.247

Notes: e_c is the implied LCU-per-EUR exchange rate from the AG.1 errata. $k_c^{\text{true},11}$ and $k_c^{\text{true},17}$ are the ground-truth scaling factors computed from e_c via Equation (9) using 2011 and 2017 ICP PPP multiples for 2012, respectively. The $k_c^{\text{calc},j}$ columns show the PIP- or Gallup-derived scaling factor (Equation 8) without errata corrections: PIP '11 uses 2011 ICP PPP, PIP '17 uses 2017 ICP PPP, and Gallup uses an undocumented PPP. The right panel shows each calculated factor's ratio to the vintage-matched ground truth; PIP '11 and PIP '17 are each compared to their own ground-truth column, while Gallup is compared to $k_c^{\text{true},11}$. Gallup data is unavailable for Malawi. Mean $\lvert\log r\rvert$: lower is better.

Table 26 reports both ground-truth columns alongside the ratio $r = k_c^{\text{calc},j}/k_c^{\text{true},j}$ for each measure, where $r = 1$ indicates exact agreement. PIP 11 achieves the lowest mean absolute log deviation ($\overline{|\log r|} = 0.211$), compared with 0.557 for PIP 17 and 0.247 for Gallup. We therefore assign $k_c^{\text{calc},11}$ to GPS countries. Finally, for the five errata countries, we substitute the known $k_c^{\text{true},11}$ for the estimated value to reduce error in those settings.

B.2.2 Individual Income Proxies

For the broad bracketing specifications (see below) we need measures of subject background income. Unfortunately, these do not exist in the GPS public data. We develop a proxy for this data for each respondent using the GNI per capita of their sub-national region. Our data source is Chrisendo et al. (2025), who provide sub-national GNI per capita for approximately 151 countries at the first administrative level, covering 1990–2023. The broad and disaggregated geographic coverage of these data are a major advantage for our purposes, as it allows us to construct proxies for subject income with few modeling assumptions, and with variation across regions for the whole sample.

Because Chrisendo et al. (2025) use Global Data Lab (GDL) region codes rather than the GADM codes used in our GPS region-matching procedure (Appendix B.3), we match

GPS survey regions to Chrisendo/Kummu regions via a centroid-in-polygon spatial join: we compute each GPS region's geographic centroid from its GADM boundary polygon and identify the Chrisendo/Kummu region containing that centroid. This yields coverage for 1,143 of our 1,148 GPS survey regions (the five unmatched regions lack geometry in the source GeoPackage). As a cross-validation, we compare the Chrisendo/Kummu GNI values against the DOSE database of subnational economic output (Wenz et al., 2023), which covers 83 countries with officially reported subnational GDP. The Pearson correlation is $r = 0.91$ in log–log space, providing independent confirmation of the income estimates. Regional GNI ranges from $1,336 (rural Malawi) to $210,827 (the Washington, DC area).

B.2.3 The CRRA Indifference Equation

Under CRRA utility $u(c) = c^{1-\rho}/(1-\rho)$ with broad bracketing, a respondent with background income w is indifferent between the 50/50 lottery paying X or nothing and a safe payment of s when

$$\frac{1}{2}\,\frac{(w+X)^{1-\rho}}{1-\rho} + \frac{1}{2}\,\frac{w^{1-\rho}}{1-\rho} = \frac{(w+s)^{1-\rho}}{1-\rho}, \tag{10}$$

where ρ is the coefficient of relative risk aversion. At each of the 31 staircase boundaries, we solve Equation (10) for the indifference ρ_b, where $s_b = k_c \cdot 10b/\text{PPP}_{\text{Germany}}$ and $X = k_c \cdot 300/\text{PPP}_{\text{Germany}}$. This yields a sequence of boundary ρ values that is monotonically decreasing from ρ_{32} (most risk-averse) to ρ_3 (most risk-seeking among solvable boundaries).

Rearranging Equation (10), we define the raw function

$$f(\rho) = \tfrac{1}{2}(w+X)^{1-\rho} + \tfrac{1}{2}\,w^{1-\rho} - (w+s)^{1-\rho}.$$

This function has a critical numerical deficiency: it possesses a universal spurious root at $\rho = 1$. At $\rho = 1$, the exponent $(1-\rho) = 0$, so $x^0 = 1$ for all positive x, giving $f(1) = \frac{1}{2} + \frac{1}{2} - 1 = 0$ regardless of w, X, and s. This root is mathematically exact but economically meaningless, as it arises from the degenerate behavior of power functions at the log-utility exponent, not

from any indifference condition. Moreover, $f(\rho)$ is *non-monotone*: for large ρ, the $w^{1-\rho}$ term (smallest base, largest negative exponent) dominates, sending f back toward positive values. A naive root-finder applied to f would therefore face two roots and could converge to the wrong one.

We resolve both problems by dividing by $(1-\rho)$ to obtain the normalized function

$$g(\rho) = \frac{1}{1-\rho}\Big[\tfrac{1}{2}(w+X)^{1-\rho} + \tfrac{1}{2}\,w^{1-\rho} - (w+s)^{1-\rho}\Big], \tag{11}$$

with the log-utility limit (via L'Hôpital's rule)

$$g(1) = \tfrac{1}{2}\ln(w+X) + \tfrac{1}{2}\ln(w) - \ln(w+s). \tag{12}$$

The key property of $g(\rho)$ is that it is *monotonically decreasing*, so $g = 0$ has at most one root. The division by $(1-\rho)$, which changes sign at $\rho = 1$, eliminates both the spurious root and the non-monotonicity, enabling reliable root-finding.

We evaluate g on a coarse grid $\rho \in [-1{,}000,\ 1{,}000]$ with unit step size (2,001 grid points), vectorized across all N individuals simultaneously. For each individual–boundary pair, we identify the unique positive-to-negative sign change. Within the bracketing interval, vectorized bisection (60 iterations, tolerance 10^{-8}) refines the root. The wide grid range accommodates roots from $\rho \approx +150$ (low safe payments under monthly wealth integration) to $\rho \approx -200$ (high safe payments under daily wealth integration), as well as the more extreme values that arise under yearly horizons.

Boundaries 1 and 2 have $s \geq X$ — the safe payment meets or exceeds the lottery's maximum payout. No finite ρ satisfies indifference at these boundaries: even an infinitely risk-seeking agent under EUT would prefer a sure payment exceeding the lottery's best outcome. These boundaries yield missing values by construction, and respondents in the corresponding staircase bins (the two most risk-seeking bins) require special end-bin treatment.

B.2.4 Bracketing Assumptions and Income Horizons

The most consequential modeling choice in our structural estimation is the *scope of income bracketing*, i.e. how much of each respondent's background income is integrated with the one-time lottery gain when evaluating expected utility. Under *narrow bracketing* ($w = 0$), the respondent is assumed to evaluate the lottery in isolation. The indifference equation admits the closed-form solution $\rho_b = 1 - \ln(2)/\ln(X/s_b)$, which is cross-country invariant (the country scaling factor k_c cancels due to CRRA homogeneity) and maps the 31 boundaries into a compressed range where ρ never exceeds 1. Under *broad bracketing* ($w > 0$), the respondent integrates background income, and the same staircase responses map to ρ values in the range typically encountered in the macroeconomics and finance literatures ($\rho \approx 1$–5 or higher).

The two approaches reflect a genuine theoretical tension. Rabin (2000) demonstrates that small-stakes risk aversion under expected utility implies absurd levels of large-stakes risk aversion unless income integration is accounted for, providing a theoretical rationale for broad bracketing and suggesting that narrow-bracketing ρ values may not have a coherent interpretation outside the specific experimental context. Conversely, Andersen et al. (2018b), using Danish administrative tax records linked to laboratory lottery experiments, find that subjects integrate essentially zero of their total wealth into small-stakes lottery decisions, providing empirical support for narrow bracketing. We adopt broad bracketing as our primary framework because it yields cardinal ρ parameters that are comparable to the risk aversion estimates used throughout the macroeconomics and finance literatures, enabling direct quantitative interpretation of our results. We implement narrow bracketing as a robustness test of our results throughout.

The choice of broad bracketing then necessitates a choice of *bracketing horizon* for subjects, or in other words the share of their income that is integrated into the risky choice, operationalized by their income over a unit of time. The "true" horizon for subjects is unknown. Under broad bracketing, the w/X ratio is the primary determinant of both the level and the dispersion of $\hat{\rho}$. When w is small (e.g. daily income), the lottery represents

a substantial fraction of the relevant income, yielding moderate ρ values with wide cross-sectional dispersion. When w is large (e.g. annual income), the lottery is negligible relative to income, and even modest behavioral risk aversion maps to very large ρ.

We take a data-driven approach to the selection of the bracketing horizon, by implementing five focal horizons and selecting the specification that is most consistent with the choice data. For the first four horizons we divide annual regional GNI per capita by 365 (daily), 52 (weekly), 12 (monthly), and 1 (yearly). For the fifth, we implement an individualized approach that exploits respondents' time preference measures in the GPS. The GPS patience module elicits intertemporal preferences via a separate procedure, yielding a combined measure for each respondent. We rank individuals by patience and map their rank to an income divisor on [1, 365] using a geometric transformation, so that the most patient respondent brackets against annual income and the least patient against daily income.

As we show below, the daily horizon is the only one of the five that is numerically stable and shows consistent results (using a choice monotonicity diagnostic) for the sample. Since this short bracketing horizon is also the most consistent with the empirical evidence from lab studies on narrow bracketing, we adopt it for our primary specification.

B.2.5 Structural Estimation Using the Random Preference Model with Tremble

With the bin boundaries, income proxies, and bracketing horizon candidates in hand, we face two additional issues in the structural estimation. First, for the most risk averse end bin 32, we do not have an upper bound on the risk aversion parameter. Second, a significant fraction of our sample, roughly 8,400 subjects, are in bins 1 and 2, which means that they choose the lottery even when the safe payment exceeded the prize, a first-order stochastic dominance (FOSD) violation. FOSD violations like these are a well-documented feature of experimental measures of risk preferences, and need not reflect irrationality: they may arise from trembling-hand errors, inattention, or a taste for gambling distinct from EUT risk aversion.

To deal with these two issues we implement a probabilistic structural model that combines the Random Preference Model (RPM) of Apesteguia and Ballester (2018) with a tremble parameter (Harless and Camerer, 1994). This model takes into account the information contained in all five choices that subjects make in the staircase procedure, and allows the choice to be the result of a mix between a deterministic decision revealing preferences and a stochastic decision noise driver. For each of the five binary choices at boundaries $b_1, \ldots, b_5$ traversed by a respondent's staircase path is modeled as

$$P(\text{choose lottery at boundary } b) = (1-\omega)\,\Phi\left(\frac{\rho_b - \rho_i}{\sigma}\right) + \omega \cdot \tfrac{1}{2}, \tag{13}$$

where ρ_b is the CRRA indifference value at boundary b (solved from $g(\rho) = 0$ as in Appendix B.2.3), ρ_i is the individual's latent risk aversion parameter, $\sigma > 0$ is a population-level noise parameter in ρ-space, $\omega \in [0, 1]$ is the tremble probability (the chance of a random coin-flip choice), and Φ is the standard normal CDF. The five choices are assumed independent conditional on ρ_i.

The critical modeling decision here is whether noise enters in utility space (the Fechner model of Hey and Orme, 1994; Harrison and Rutstrom, 2008) or in preference-parameter space (the RPM). We initially implemented the Fechner specification, $P(\text{lottery}) = \Phi(g(\rho)/\sigma)$, but found that σ is not identified in the GPS staircase data: the forced-monotone tree structure ensures that every observation is perfectly consistent with some ρ, so the likelihood always pushes σ toward zero. With $\sigma \approx 0$, FOSD violators receive anomalous positive $\hat{\rho}$ values (extreme risk aversion compresses all utility differences to near zero, making all choices approximately 50/50). The RPM specification avoids this pathology because noise operates on ρ directly, guaranteeing monotone stochastic choice—the probability of choosing the risky option is monotonically increasing in $\rho_b - \rho_i$ regardless of the utility parameterization (Apesteguia and Ballester, 2018). The tremble component ω is essential for handling FOSD violations. At dominated boundaries (where $s \geq X$, so no finite ρ satisfies indifference and

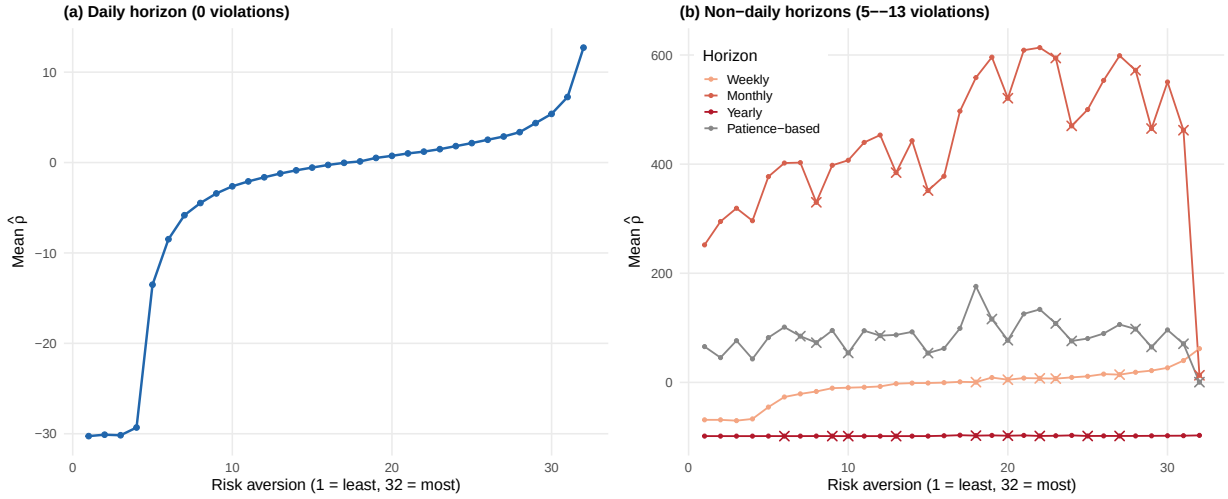
we set $\rho_b = -\infty$), the RPM term assigns probability 1 to choosing the safe option for any finite ρ_i. The tremble probability ω attributes a fraction of dominated lottery choices to random error rather than extreme risk-seeking, producing finite $\hat{\rho}$ for all respondents including gamble seekers.

To estimate the parameter, under each bracketing horizon we reconstruct the binary decision tree to determine which five of the 31 boundaries each staircase bin traverses, then precompute $g(\rho)$ on a grid $\rho \in [-100, 100]$ with step size 0.5, vectorized across all N individuals and all five decision steps. Boundary values ρ_b are identified as the zero crossings of g at each step for each individual. The population-level parameters (σ, ω) are estimated by profile maximum likelihood: for each candidate (σ, ω) on a coarse grid ($\sigma \in [0.005, 20.0]$, $\omega \in [0.001, 0.95]$; 255 points), we compute each individual's log-likelihood maximized over ρ on the evaluation grid. We then refine with a fine grid ($50 \times 50 = 2{,}500$ points) centered on the coarse optimum. At the optimal $(\hat{\sigma}, \hat{\omega})$, each individual's $\hat{\rho}_i$ is extracted as the value maximizing their personal log-likelihood, and is further refined via bounded scalar optimization (tolerance 10^{-6}) over a $\pm 3\hat{\sigma}$ interval around the grid optimum.

We then test the construct validity of each of the resulting five measures using the monotonicity diagnostic recommended by Apesteguia and Ballester (2018). This confirms that the structural mapping from staircase responses to CRRA coefficients under the daily horizon preserves the ordinal ranking implied by the survey instrument. Non-daily horizons fail this diagnostic: the RPM+Tremble estimates exhibit multiple monotonicity violations out of 28 non-FOSD adjacent bin pairs under weekly, monthly, yearly, and patience-based horizons (Table 27 and Fig. 15).

Further, under the daily horizon, the estimated tremble rate is $\hat{\omega} = 0.089$ (8.9%), indicating that fewer than 10% of choices reflect random trembles. The low noise parameter $\hat{\sigma}$ confirms that subjects' choices are highly consistent with a single latent ρ_i. By contrast, non-daily horizons yield dramatically higher tremble rates: $\hat{\omega} \approx 0.11$ for weekly, ≈ 0.78 for monthly, and 0.99 for yearly, indicating that the model attributes the vast majority of choices to

Figure 15: Monotonicity Validation: Mean $\hat{\rho}$ by Risk Aversion Position Across Income Horizons



Notes: Each point shows the mean RPM+Tremble CRRA coefficient $\hat{\rho}$ within each of the 32 staircase positions, ordered by risk aversion (1 = least risk-averse, 32 = most risk-averse). Panel (a) displays the daily income horizon under the regional GNI wealth proxy, which exhibits strict monotonicity—mean $\hat{\rho}$ increases at every adjacent position. Panel (b) overlays the weekly, monthly, yearly, and patience-based horizons, all of which exhibit multiple monotonicity violations (× markers). See Table 27 for the estimation parameters.

Table 27: Estimation Parameters and Monotonicity Validation Across Specifications

Horizon	$\hat{\sigma}$	$\hat{\omega}$	Violations (/28)	Passes
Narrow bracketing	0.003	0.990	1	
Daily	0.040	0.089	0	✓
Weekly	0.003	0.109	5	
Monthly	0.003	0.781	10	
Yearly	0.526	0.990	9	
Patience	0.121	0.226	13	

Notes: Each row reports the RPM+Tremble estimation results for a different income horizon. $\hat{\sigma}$ is the population-level noise parameter in ρ-space; $\hat{\omega}$ is the tremble probability. "Violations" counts the number of non-FOSD adjacent bin pairs (out of 28; bins 1–29) where mean $\hat{\rho}$ fails to increase monotonically from a less to a more risk-averse staircase position. FOSD bins 30–32 are excluded because the safe payment meets or exceeds the lottery prize at those boundaries. Only the daily horizon and the narrow-bracketing specification pass strict monotonicity. For non-daily horizons, $\hat{\omega}$ rises dramatically—from 9% (daily) to 11% (weekly), 78% (monthly), and 99% (yearly)—indicating that the model attributes the vast majority of choices to random noise.

random noise rather than systematic preferences (Table 27). We conclude that the daily horizon is the model best identified and with the best match to the data, and adopt it as our main specification.

B.2.6 Measurement Error Control and Bayesian shrinkage of CRRA Estimates

We implement four approaches to address measurement error in the individual-level CRRA estimates $\hat{\rho}_i$, each embodying distinct assumptions about the structure of noise in staircase-derived risk preference measures. The three Bayesian shrinkage methods all operate in IHS space ($y_i = \text{asinh}(\hat{\rho}_i)$) using profile-likelihood standard errors from the structural RPM+Tremble model, transformed via the delta method.

Our lightest-touch approach truncates the $\hat{\rho}$ distribution at fixed percentiles through **Winsorization** (1/99 and 2.5/97.5). This ensures no single observation exerts undue leverage on regression estimates but makes no distinction between measurement error and genuine preference heterogeneity: all observations beyond a given percentile are treated identically regardless of their individual precision. Winsorization serves as a transparent, assumption-free baseline, which is widely employed in applied economics settings.

The second approach we implement is **Normal Empirical Bayes Shrinkage**. This method uses the classical Gaussian hierarchical model (James and Stein, 1961; Morris, 1983; Efron and Morris, 1975):

$$y_i \mid \theta_i \sim \mathcal{N}(\theta_i, \hat{\sigma}_i^2), \qquad \theta_i \mid c \sim \mathcal{N}(\mu_c, \tau_c^2),$$

where $\hat{\sigma}_i^2$ is the individual measurement error variance from the profile likelihood and τ_c^2 is the between-individual variance of true preferences within country c, estimated via DerSimonian and Laird (1986). Country means are estimated as precision-weighted (GLS) averages, naturally downweighting end-bin respondents whose profile-LL SEs are two orders of magnitude larger than interior-bin respondents'. The posterior mean shrinks each individual

toward their country mean by a factor $B_i = \hat{\sigma}_i^2/(\hat{\sigma}_i^2 + \hat{\tau}_c^2)$, with noisily measured individuals shrunk more. This estimator is standard in the literature for analogous settings with heterogeneous measurement precision, including teacher value-added (Kane and Staiger, 2008; Chetty et al., 2014), neighborhood effects on intergenerational mobility (Chetty and Hendren, 2018), and school effectiveness (Angrist et al., 2017); see Walters (2024) for a recent survey.

Next, we implement ***t*-Distribution Empirical Bayes**. This approach blends classical empirical Bayes with the robust statistics literature (Huber and Ronchetti, 2009), in order to try and account for upstream noise sources not fully captured by the profile-LL SEs: uncertainty in the income proxy k_c, misspecification of the bracketing horizon, or occasional gross errors in survey administration. Here we replace the Gaussian level-1 assumption with a heteroskedastic t-distribution (Lange et al., 1989), providing data-dependent downweighting of observations that are discrepant relative to their precision. Country means are estimated via iteratively reweighted least squares, where outlier weights $w_i = (\nu + 1)/(\nu + d_i^2/\hat{\sigma}_i^2)$ inflate the effective variance of discrepant observations, increasing their shrinkage. We report $\nu = 3$ (heavy tails) and $\nu = 5$ (the lowest value guaranteeing finite kurtosis, following Lange et al., 1989). The MLE over ν selects $\nu^* = 10{,}000$—the normal limit—indicating that with profile-LL SEs, the Gaussian measurement error model is adequate and the additional robustness is unnecessary.

Finally, we implement a Bayesian **Mixture Model** over subject types. This approach models the respondent population as heterogeneous in attentiveness: it assumes that within each country, a fraction $\hat{\pi}_c$ of end-bin respondents are "confused": their staircase responses reflect satisficing, inattention, or non-EUT gambling preferences rather than CRRA risk aversion. The confusion fraction is estimated as the excess end-bin share beyond what a normal distribution fitted to interior-bin respondents predicts. For confused respondents, the measurement error variance is inflated by a factor proportional to $\hat{\pi}_c$, driving their shrinkage weight $B_i \to 1$ and effectively replacing their individual estimate with the country mean. This directly addresses the fact that a large fraction of subjects are in the end bins of the

staircase distribution, and provides the most aggressive shrinkage, which serves as an upper bound on measurement error correction.

The four approaches span a natural set of error control assumptions in our setting. Winsorization makes no assumptions about the error structure, but truncates the tails of the distribution arbitrarily. The Normal EB imposes a parametric structure on errors but lets the data determine each individual's shrinkage through the realized residual from the estimation model. The t-distribution EB takes a similar approach, but allows for a more robust error process that can potentially account for measurement error not directly captured by the estimation model. The mixture model takes the strongest stance by allowing a subset of respondents to be treated as entirely uninformative.

We report all four approaches throughout to demonstrate robustness (Section 4), but use the Normal EB shrinkage as our preferred specification. The key consideration is that the profile-likelihood standard errors from the structural model provide individual-level precision estimates that vary by two orders of magnitude across respondents, and Normal EB is the lightest parametric structure that fully exploits this heterogeneity: each individual's shrinkage is governed by the ratio of their measurement error variance to the between-individual variance, so precisely estimated interior-bin respondents retain most of their individual signal while noisy end-bin respondents are pulled toward the country mean. Winsorization discards this information entirely, treating a precisely estimated $\hat{\rho}$ at the 99th percentile identically to a noisy end-bin estimate at the same percentile. The t-distribution EB, meanwhile, solves a problem that the data indicate does not exist in our setting: the MLE over ν selects $\nu^* = 10{,}000$, the normal limit, confirming that the Gaussian measurement error model is adequate once we condition on profile-likelihood standard errors, and the additional robustness parameters add degrees of freedom the data do not reward. The mixture model takes the strongest stance by treating all excess end-bin respondents as entirely uninformative, but this is difficult to justify when many end-bin respondents have profile-likelihood SEs only modestly larger than interior-bin respondents'; Normal EB already downweights these individuals in

proportion to their precision, without the binary classification into "confused" and "attentive" types that the mixture model imposes. Normal EB also has direct precedent in applied settings with the same core structure (individual-level estimates with known, heterogeneous precision), including teacher value-added (Kane and Staiger, 2008; Chetty et al., 2014), neighborhood effects (Chetty and Hendren, 2018), and school effectiveness (Angrist et al., 2017).

B.3 Region Matching: Global Analysis

This appendix documents the procedure used to match region names in the individual-level GPS (Global Preferences Survey) dataset to administrative boundary polygons from the GADM 4.1.0 (Global Administrative Areas Database). The GPS dataset contains 80,337 observations across 76 countries and 1,148 unique country–region pairs. The GADM 4.1.0 GeoPackage contains 356,508 boundary polygons covering all countries at up to five levels of administrative subdivision, in the WGS 84 coordinate reference system (EPSG:4326).

B.3.1 Automated String Matching

We match survey region names to GADM administrative unit names through a multi-pass procedure that progresses from high-confidence exact matching to fuzzy string matching.

Country name reconciliation. Three country names differ between the survey and GADM: "Bosnia Herzegovina" (survey) maps to "Bosnia and Herzegovina" (GADM), "Czech Republic" to "Czechia," and "Mexico" to "México."

Pass 1: Normalized exact matching. For each survey region, we apply the following normalization: (a) strip leading and trailing whitespace; (b) convert to lowercase; (c) remove diacritical marks using Unicode transliteration (e.g., é→e, ü→u); and (d) collapse multiple whitespace characters to a single space. Normalized survey region names are matched against both the primary GADM admin-1 name (`NAME_1`) and alternate names (`VARNAME_1`, which contains pipe-separated alternate spellings). This pass yields 649 exact matches covering 39,323 observations (48.9% of the data).

Pass 2: Aggressive normalization. For regions unmatched in Pass 1, we apply additional transformations: (a) remove parenthetical content (e.g., "Carinthia (Kärnten)" → "Carinthia"); (b) strip common administrative suffixes including *Voivodeship*, *Province*, *Region*, *District*, *Emirate*, *Oblast*, *County*, *Municipality*, and *City*; and (c) strip the prefix

"County" (for "County X" → "X" patterns appearing in Moldova and Lithuania). The same aggressive normalization is applied to GADM names before re-attempting exact matching. This pass yields 110 additional matches (8,079 observations, 10.1%), resolving cases such as "Dubai Emirate" → "Dubai," "Beijing municipality" → "Beijing," and "Vilniaus county" → "Vilniaus."

Pass 3: Fuzzy string matching. Remaining unmatched regions (for countries not handled by manual crosswalks; see below) are matched using token-sort-ratio fuzzy string matching via the `rapidfuzz` library. For each survey region, three candidate scores are computed: (a) aggressively normalized survey name versus aggressively normalized GADM names; (b) basic-normalized survey name versus basic-normalized GADM names; and (c) extracted parenthetical content versus basic-normalized GADM names. The highest score across (a)–(c) is selected. Matches are classified as *high confidence* (score ≥ 80), *medium confidence* (score 60–79), or *low confidence* (score < 60).

B.3.2 Manual Crosswalks

For many countries, the survey and GADM use fundamentally different administrative classification systems, rendering automated string matching ineffective. We construct manual crosswalks based on official government definitions for each case. In total, manual crosswalks are constructed for 48 countries, covering 405 region mappings.

Dissolve maps (survey broader than GADM). In 15 countries, the survey reports macro-regions that are composed of multiple GADM admin-1 units. For these, we dissolve (union) the constituent GADM polygons to construct the survey's macro-region boundaries.

- **Japan** (11 survey macro-regions → 47 GADM prefectures): The survey uses the traditional Japanese macro-regions (Hokkaido, Tohoku, Kanto, Hokuriku, Koshinetsu, Chukyo, Tokai, Kansai, Chugoku, Shikoku, Kyushu). Prefectures are assigned to macro-regions following the conventional Japanese regional classification.

- **Kenya** (8 former provinces → 47 GADM counties): The survey uses Kenya's former eight-province system, abolished under the 2010 Constitution. Counties are mapped to their former parent province based on the constitutional transition provisions.

- **Uganda** (4 statutory regions → 56 GADM districts): Districts are assigned to the four statutory regions (Central, Eastern, Northern, Western) per the Uganda Bureau of Statistics classification. Two GADM water-body polygons are excluded.

- **Rwanda** (5 survey regions → 5 GADM provinces): A one-to-one mapping between English cardinal directions and Kinyarwanda province names (e.g., North → Amajyaruguru, Kigali → Umujyi wa Kigali).

- **Malawi** (3 macro-regions → 28 GADM districts): Districts are assigned to the Northern, Central, and Southern regions per the Malawi National Statistical Office classification.

- **Hungary** (7 NUTS-2 statistical regions → 20 GADM counties): The survey uses the seven EU NUTS-2 planning regions; GADM reports 20 counties (*megyek*) plus Budapest. Counties are assigned to NUTS-2 regions per the Eurostat classification (e.g., Central Hungary = Budapest + Pest; Northern Great Plains = Hajdú-Bihar + Jász-Nagykun-Szolnok + Szabolcs-Szatmár-Bereg).

- **Sri Lanka** (9 survey provinces → 25 GADM districts): The survey uses Sri Lanka's nine provinces; GADM reports 25 districts. Districts are assigned to provinces per the standard Sri Lankan administrative hierarchy (e.g., Western Province = Colombo + Gampaha + Kalutara).

- **Estonia** (5 survey macro-regions → 16 GADM counties): The survey uses five directional macro-regions (Central, North, North-East, South, West Estonia); GADM reports 16 counties (*maakonnad*).

- **Sweden** (8 survey NUTS-2 regions → 21 GADM counties): The survey uses Sweden's eight NUTS-2 statistical regions; GADM reports 21 counties (*län*). Counties are

assigned to NUTS-2 regions per the Eurostat classification (e.g., East Middle Sweden = Uppsala + Södermanland + Östergötland + Örebro + Västmanland).

- **Switzerland** (7 survey Grossregionen → 26 GADM cantons): The survey uses the seven Swiss "great regions" (*Grossregionen*); GADM reports 26 cantons. Cantons are assigned to Grossregionen per the Swiss Federal Statistical Office classification (e.g., Lake Geneva Region = Genève + Vaud + Valais).

- **Peru** (8 survey macro-regions → 26 GADM departments): The survey uses eight geographic macro-regions (Costa Norte, Costa Centro, Costa Sur, Sierra Norte, Sierra Centro, Sierra Sur, Selva, Lima). Departments are assigned to macro-regions following standard Peruvian geographic classification.

- **Venezuela** (6 survey zones → 25 GADM states): The survey uses six geographic zones (Capital, Central, Centro-Occidental, Oriente, Los Andes, Zuliana). States are assigned to zones per the Venezuelan geographic classification.

- **Serbia** (7 survey macro-regions → 25 GADM districts): The survey uses macro-regions (Belgrade, Central/Eastern/Southern/Western Serbia, Vojvodina North/South); GADM reports 25 districts (*okruzi*).

- **Philippines** (4 survey island-groups → 81 GADM provinces): The survey uses four broad island-group categories (NCR, Balance Luzon, Visayas, Mindanao). Provinces are assigned to island-groups following the standard Philippine geographic classification.

- **Portugal** (7 survey NUTS-2 regions → 20 GADM districts): The survey uses NUTS-2 statistical regions (Norte, Centro, Lisboa, Alentejo, Algarve, Açores, Madeira). Districts are assigned to NUTS-2 regions per the Eurostat classification.

Aggregate-up maps (survey finer than GADM). In four countries, the survey provides greater geographic detail than GADM admin-1. Survey regions are mapped up to their

containing GADM unit.

- **France** (22 survey regions → 13 GADM regions): The survey uses France's pre-2016 system of 22 administrative regions. GADM uses the post-2016 system of 13 regions following the January 2016 territorial reform (Loi No. 2015-29). Survey region names use English translations (e.g., "Brittany" → Bretagne). Multiple survey regions merge into a single GADM region (e.g., Alsace + Champagne-Ardenne + Lorraine → Grand Est), meaning observations from different survey regions share the same GADM polygon.

- **United Kingdom** (12 survey regions → 4 GADM nations): The survey uses nine English sub-regions plus Scotland, Wales, and Northern Ireland. GADM admin-1 contains four nations. All nine English sub-regions map to "England."

- **Greece** (13 survey regions → 8 GADM regions): GADM uses merged administrative regions (e.g., "Thessaly and Central Greece") while the survey uses the finer 13-region classification. Survey regions are mapped up to their containing GADM unit.

- **Bosnia and Herzegovina** (4 survey sub-entity regions → 3 GADM entities): The survey divides Republika Srpska into three sub-regions (RS East, RS West, RS South) and groups the rest as "Rest of Bosnia." These map to GADM's two entities plus Brčko District.

Direct translation and language maps. For several countries, survey region names and GADM names refer to the same administrative units but use different languages or transliteration conventions. These require one-to-one translation mappings:

- **Poland** (16 voivodeships): English to Polish names (e.g., Greater Poland Voivodeship → Wielkopolskie, Silesian Voivodeship → Ślaskie, Subcarpathian Voivodeship → Podkarpackie).

- **Indonesia** (14 provinces): English cardinal directions to Indonesian (e.g., Central Java → Jawa Tengah, East Kalimantan → Kalimantan Timur, South Sulawesi → Sulawesi Selatan).
- **Haiti** (7 departments): English to French (e.g., North → Nord, South-East → Sud-Est).
- **Czech Republic** (5 regions): English to Czech (e.g., Moravian-Silesian Region → Moravskoslezský, Central Bohemia Region → Středočeský).
- **Finland** (4 regions): Finnish to English (e.g., Etelä-Suomi → Southern Finland, Itä-Suomi → Eastern Finland).
- **Russia** (14 regions): Various transliteration standardizations (e.g., Bashkortostan Republic → Bashkortostan, Republic of Tatarstan → Tatarstan).

Individual corrections. An additional 245 region mappings across 42 countries are resolved through individual manual corrections. These address cases of minor spelling differences (e.g., "Sar-e-Pol" → "Sar-e Pol" in Afghanistan), Arabic transliteration variants (e.g., "Ad-Daqhleya" → "Ad Daqahliyah" in Egypt), administrative name changes (e.g., "Orissa" → "Odisha" in India; "NWFP" → "Khyber-Pakhtunkhwa" in Pakistan), and encoding artifacts. The complete set of corrections is documented in the replication code.

Admin-level-2 matching (Morocco). The survey's 43 province and prefecture names correspond to GADM admin level 2 (`NAME_2`), not level 1. Matching is performed against `NAME_2` using the same normalization procedure as Passes 1–2. Thirty-eight provinces match exactly after normalization; five require manual assignment due to spelling differences (e.g., "Agadir Ida ou Tanan" → "Agadir-Ida ou Tanane") or boundary changes (e.g., Nouaceur, created circa 2009, is mapped to Casablanca). Note that Berkane and Taourirt, separate in the survey, both map to the combined GADM unit "Berkane Taourirt."

B.3.3 Metro and Non-Metro Splits

For Australia and Canada, the survey splits major cities from their surrounding state or provincial territory (e.g., "Sydney" versus "New South Wales, excluding Sydney"). GADM contains no built-in metropolitan boundaries, so we approximate metro areas using groupings of GADM admin-level-2 units.

Australia (16 survey regions → GADM Local Government Areas). Metro areas are approximated using groupings of GADM `NAME_2` Local Government Areas (LGAs) based on the Australian Bureau of Statistics Greater Capital City Statistical Area (GCCSA) definitions. Each state is fully partitioned: the metro LGAs plus the remaining non-metro LGAs account for all LGAs in the state. Known approximations include the fact that GADM LGA boundaries reflect a specific vintage that may predate or postdate council amalgamations (e.g., the 2016 NSW council mergers), and that some GCCSA-defined LGAs may not have exact GADM name matches. In total, 559 GADM LGAs are assigned across the 16 survey regions.

Canada (8 survey regions → GADM Census Divisions). Metro Census Metropolitan Areas (CMAs) are approximated using GADM `NAME_2` Census Divisions (CDs) or Regional County Municipalities (RCMs):

- The Toronto CMA is mapped to five Ontario Census Divisions (Toronto, Peel, York, Durham, Halton). The actual CMA boundary extends into small portions of Simcoe County, but the standard five-CD approximation is used.
- The Montreal CMA is mapped to 19 Quebec RCMs centered on Communauté-Urbaine-de-Montréal and Laval. Some peripheral RCMs have only partial CMA overlap but are included in their entirety since CDs cannot be split.
- The Vancouver CMA is mapped to the single "Greater Vancouver" Regional District,

which aligns well with the CMA. Portions in the Fraser Valley Regional District (e.g., Abbotsford) remain in "British Columbia, excluding Vancouver."

Multi-province groupings "Atlantic" (New Brunswick, Nova Scotia, PEI, Newfoundland and Labrador; 47 CDs) and "Prairies" (Alberta, Saskatchewan, Manitoba; 60 CDs) are defined by pooling all CDs within the constituent provinces. Three northern territories (Yukon, Northwest Territories, Nunavut; 6 CDs) are not covered by the survey.

B.3.4 Matching Results

In total, 1,143 regions (98.7%) are matched by our procedure. These comprise exact normalized matches (642 regions), exact aggressively normalized matches (102), manually verified corrections (245), manual dissolve maps (38), manual aggregate-up maps (34), manual direct translation maps (16), Morocco admin-level-2 matches (43), and Australia/Canada metro groupings (23). Four country–region pairs with empty region fields remain formally unmatched (France: 11 observations; Germany: 1; United Kingdom: 4; and Jammu and Kashmir: 30). Croatia (992 observations) also has an empty region field in the GPS data, but because no subnational region information exists for any Croatian respondent, we treat Croatia as a single-region country: we assign all Croatian respondents to a single "Croatia" region and use country-level (population-weighted) climate variables in place of region-level ones.

B.4 Climate Experience Data Construction

This section documents the construction of climate variables used in the global analysis (Section 2.5). We describe the source datasets, the spatial aggregation procedure, and the construction of lifetime experience measures and interview-day temperature anomalies. Climate variables are constructed at the survey-region level; approximately 2,200 respondents whose regions could not be matched to GADM boundaries (see **??**) lack region-level climate data and are excluded from the climate regressions.

B.4.1 Source Data

Our primary climate data are the University of Delaware Air Temperature and Precipitation dataset, version 5.01 (Willmott and Matsuura, 2001). This dataset provides monthly mean temperature (°C) and monthly total precipitation (cm) on a global $0.5° \times 0.5°$ latitude–longitude grid (roughly 55 km at the equator) from January 1900 through December 2017, yielding 1,416 monthly layers for each variable. The data are interpolated from station observations using a climatologically aided interpolation algorithm.

Because UDel v5.01 ends in December 2017 and our counterfactual analyses require climate data through recent years, we extend the monthly record using two additional datasets. For January 2018 through December 2024, we splice in monthly temperature and precipitation from the Climatic Research Unit Time-Series, version 4.09 (CRU TS v4.09; Harris et al., 2020), which provides the same variables on the same 0.5° grid from 1901 to 2024. Because CRU TS reports precipitation in mm/month while UDel uses cm/month, we convert CRU TS precipitation to cm before splicing. We validate the splice by comparing UDel and CRU TS over their 2010–2017 overlap period at the survey-region level, finding high correlation and small mean absolute differences. For January 2025 onward, we aggregate CPC Global Unified daily temperature and daily precipitation (Fan and van den Dool, 2008) from daily to monthly resolution (computing monthly mean temperature from daily $T_{\text{mean}} = (T_{\min} + T_{\max})/2$ and monthly total precipitation from daily accumulations) and append to the extended series.

The final spliced dataset provides monthly temperature and precipitation on a consistent 0.5° grid from January 1900 through the most recent complete month of 2025, yielding 1,500+ monthly layers.

For interview-day temperature anomalies, we use NOAA's CPC Global Unified daily temperature data (Fan and van den Dool, 2008), which provides daily minimum and maximum temperature on a 0.5° grid from 1979 to the present. We derive daily mean temperature as the average of minimum and maximum.

Survey regions are matched to GADM 4.1.0 administrative boundary polygons, as described in Appendix B.3, yielding 1,144 GPS survey regions across 76 countries and 263 country-level polygons.

B.4.2 Spatial Aggregation

We aggregate gridded climate data to survey regions by constructing *area-weighted* measures. We compute fractional-area zonal statistics for the intersection of each polygon with the climate raster grid. For a region r with polygon P_r overlapping grid cells $\{g_1, \ldots, g_J\}$, the area-weighted mean temperature in month t is:

$$T_{r,t}^{\text{area}} = \frac{\sum_{j=1}^{J} w_j^{\text{area}} \cdot T_{g_j,t}}{\sum_{j=1}^{J} w_j^{\text{area}}},$$

where w_j^{area} is the fraction of grid cell g_j's area that falls within P_r. Precipitation is aggregated identically.

B.4.3 Lifetime Climate Experience Variables

For each GPS respondent with survey region r and birth year b (computed as survey year minus age), we construct lifetime climate experience as the average over monthly climate

values from January of year b through December 2011:

$$\overline{T}_{r,b} = \frac{1}{N_b} \sum_{t \in \{\text{Jan}(b),\ldots,\text{Dec}(2011)\}} T_{r,t}^{\text{area}}, \qquad \text{SD}(T)_{r,b} = \left[\frac{1}{N_b - 1} \sum_{t} \left(T_{r,t}^{\text{area}} - \overline{T}_{r,b} \right)^2 \right]^{1/2},$$

where $N_b = 12 \times (2012 - b)$ is the number of months in the experience window. Analogous measures are constructed for precipitation (mean and standard deviation). Because all individuals within the same region and birth year share the same climate history, these variables are collapsed to the region $\times$ birth-year level before merging to individual-level data. Observations with birth years before 1900 (the start of the UDel record) are dropped. We truncate the experience window at December 2011, the last full calendar year before the 2012 GPS survey, rather than extending it to the interview date to avoid conflating lifetime climate trends with short-run weather shocks near the interview, which could introduce reverse causality or salience effects.

B.4.4 Interview-Day Temperature Anomalies

To ensure that our estimated effects of lifetime climate experience are not contaminated by contemporaneous weather effects on the day of the survey or the night before we construct temperature anomaly variables from CPC daily data. We use a climatological baseline for each of the 1,144 survey regions computed from the day-of-year (DOY) weather over the WMO standard 1981–2010 reference period. This gives us the historical mean and standard deviation of daily mean temperature for each of the 366 possible days of the year (February 29 is included), stratified by region. Using this, we can calculate for each respondent's interview date d in region r, the interview-day z-score anomalies:

$$z_{r,d}^{\text{day}} = \frac{T_{r,d} - \bar{T}_{r,\text{doy}(d)}}{\text{SD}(T)_{r,\text{doy}(d)}},$$

where $\bar{T}_{r,\text{doy}(d)}$ and $\text{SD}(T)_{r,\text{doy}(d)}$ are the climatological mean and standard deviation for that day of year in region r. We construct two anomaly measures:

1. $z_{\text{mean}}^{\text{day}}$: interview-day mean temperature anomaly;

2. $z_{\text{min}}^{\text{day}}$: night-before minimum temperature anomaly (relevant for sleep disruption).