Ancestral origins of environmental (in)attention

César Barilla
University of Oxford

ⓡ

Palaash Bhargava
University of Chicago

Abstract

How does the climatic experience of past generations affect today's assessment of the importance of environmental concerns? To answer, we combine empirical evidence spanning multiple contemporary surveys and group-level cultural records, with a new theory of how cultural transmission shapes costly environmental attention. We show that the intensity of ancestral climate anomalies has a persistent effect on the perceived stakes of environmental considerations in decision-making. The relationship is U-shaped: descendants of groups who faced more stable or more volatile climates attribute higher importance to environmental concerns, with a dip at intermediate levels. Environmental content in folklore and other cultural narratives displays the same U-shape, consistent with a cultural-transmission channel linking ancestral experience to contemporary perceptions. We propose a general model in which environmental attention is a costly choice made before climate conditions are realized, and perceptions of its stakes are shaped by realized gains and losses through an evolutionary process. Because attention is chosen ex ante, selection pressure is coarse: it only disciplines perceptions through average payoffs under the specific climate distribution a group experiences, generating heterogeneous bias across ethnic groups. When environmental attention serves two functions—using typical conditions effectively and protecting against extreme events—the model rationalizes the U-shaped dependence of perceived stakes on ancestral climate anomalies.

Keywords: Environment; Cultural Transmission; Economic History; Inattention; Learning; Individual Preferences; Ecological Anthropology. **JEL:** Q50, Q54, D83, D91, Z13, N00

This version: February 27, 2026. ⓡ stands for random order author naming convention.

We thank Michael Best, Shivangi Bishnoi, Sandra Black, Laura Boudreau, Yeon-Koo Che, Mark Dean, Léa Dousset, Ruben Durante, Alex Eble, Duarte Gonçalves, Florian Grosset, Navin Kartik, Tushar Kundu, Elliot Lipnowski, Nicolas Longuet-Marx, Nathan Nunn, Cristian Pop-Eleches, Bernard Salanié, Matthias Schündeln, Matthias Sutter, Miguel Urquiola, Akanksha Vardani, and Jack Willis for helpful comments and discussions. We are grateful to seminar participants at Columbia University, Historical Danish Political Economy, MPWZ Text as Data Workshop, Max Planck Alumni Workshop and IPWSD for feedback. Uday

1 Introduction

Sustained interactions between groups and their surroundings create generational memories (Nunn, 2009) which in turn shape the beliefs of descendants and influence their behavior and choices. Classical examples include religious practices (McCleary and Barro, 2006), behavioral traits (Lowes et al., 2017), social norms (Buggle and Durante, 2021; Alesina, Giuliano, and Nunn, 2013), and human capital decisions (S. O. Becker et al., 2020). In this paper, we focus on attitudes towards environmental issues.

Anthropological studies document enduring relationships between communities and their environments, and draw links to descendants' heightened attention to environmental justice and preservation. Many case studies trace such relationships back to experienced environmental conditions and implicitly point to predictability as an essential dimension. For instance, the development of environmental stewardship is sometime attributed to the ritualization of ecological knowledge (for ethnic groups such as the Toraja, Manobo and Balinese) enabled by *stable* typical conditions (Lansing and Kremer, 1993; Berkes, 2017; Atran et al., 2002). By contrast, others point to a high prevalence of anomalies in ancestral climates (for e.g. the Sámi, Quechua and Māori) increasing the need for protection against extreme events (Berkes and Folke, 1998; Gómez-Baggethun and Reyes-García, 2013; Bwambale, Muhumuza, and Nyeko, 2018). In either case, resulting mechanisms of adaptation and coordination are reflected in norms and identities that persist through cultural transmission, independently of the descendant's location.

To understand the effect of the climate experienced by an individual's ethnic ancestors on the importance they attach to environmental issues, we combine quantitative analysis at a global scale with a general theoretical foundation rooted in costly attention and evolutionary dynamics. Our theoretical and empirical exercises aim to respectively formalize and quantify the following intuition. At the individual level, taking environmental considerations and consequences into account when making decisions ("environmental attention") requires costly effort with uncertain returns. The actual value of such effort (the "stakes" of environmental attention) is a complex object, which is not only heterogeneous across situations but also a function of realized climate conditions. Drawing inference from previously documented accounts, we assume that stakes are driven by anomalies, i.e. intensity of deviations from "typical" conditions. We propose that the subjective perception of the stakes of environmental attention is formed via a cultural transmission process through which more successful heuristics are noisily selected. This mechanism drives the link between *realized climate anomalies* in ancestral generations and *perceived value* of descendants.

Bangabash, Pari Bishen, Kevin Li Grant, Gina Nemade and Smriti Raj Sharma provided excellent research assistance. We acknowledge the support of Columbia's Program for Economic Research.

We construct multiple variables to quantify our key objects of interest. The lead outcome variable is a direct proxy of the overall value of "caring about the environment" from the 5th and 6th wave of World Value Surveys (WVS), carried out globally between 2005 and 2014. To corroborate the results, we construct alternative outcome variables by creating composite indices from either environmental actions and opinions related questions in Integrated Value Surveys, or questions from the Environmental Module in the European Social Survey. We follow Giuliano and Nunn (2021) to match our outcome variables to ancestral experiences: both individual and group level outcomes of interest are matched to ethnic groups in Giuliano and Nunn's (2018) Ancestral Characteristics of Modern Population; and the location of ethnic groups is linked to data on historical temperature anomalies from Mann et al. (2009). Our main measure of ancestral climate anomalies is the within-generation average intensity of deviations from (generation-specific) typical conditions, averaged across generations of ethnic ancestors over the period 1600-1920 C.E.

Our main empirical result is that the perceived value of attention to environmental issues is U-shaped in the average intensity of anomalies experienced by ancestors. Consistently with anthropological evidence, descendants of populations that faced highly volatile or stable climate attach more value to environmental issues, whereas intermediate variability levels lead to lower perceived value. This result is robust to alternative specifications for the measure of climatic anomalies, inclusion of higher moments of the temperature distribution, different inputs used to construct pro-environment indices and the choice of sample from Integrated Value Surveys.

We propose a general model of cultural transmission about the value of attention. Individuals in successive generations face many situations, each modeled as a canonical rational inattention problem, where they have to decide how much attention to pay to environmental considerations. Ex post gains and losses depend on climate conditions through a (situation-specific) stakes function, but attention is chosen before realized conditions are observed. We then embed the rational inattention problem into evolutionary *replicator-mutator dynamics*. Each generation is composed of a population of agents characterized by their subjective perception of the average stakes function ("cultural types"), who each make decisions accordingly. The distribution of types evolves via selection of successful types and random mutations mean-reverting to an arbitrary default. Our first main theoretical result (Theorem 1) is a general characterization of the outcome of the cultural transmission process. If the weight of the default is sufficiently small, the average cultural type converges to the closest function to the default which induces the correct *mean value* under ancestral conditions. Crucially, the interplay of attention choices being made *ex ante* with evolutionary dynamics selecting for *ex post* losses creates a wedge, which causes persistent dependency on ancestral conditions that would not otherwise occur. We harness the general characterization of Theorem 1 to

provide a general decomposition of the effect of ancestral conditions into a *transmission effect* and a *climate dissimilarity effect*. We then use this result to establish (Theorem 2) that if the true value of attention is driven by a (sufficiently strong) dual motive of *exploitation of typical conditions* and *protection against extreme events*, then a descendant's perceived value of caring about environmental issues should be (approximately) *U-shaped* in the average intensity of anomalies among their ancestors.

Since our mechanism rests on cultural transmission mediating the influence that successive generations exert over their descendant's perception, a greater perceived value of attention at the group level should be reflected in cultural discourse and retained in the collective memory, thereby adding to the stock of existing environmentally themed cultural narratives and collective memory. To capture this empirically, we use the prevalence of environment-related folklore at the ethnic group level, extracted from Michalopoulos and Xue (2021) using text analysis, as well as complementary environmental indices created using NLP classification of textual records associated with ethnic groups obtained from Wikipedia and the Joshua Project. Accordingly, the effect of ancestral climate variability on narratives should exhibit the same U-shape. This is indeed what we find, and the result is equally robust.

Related Literature Our work builds on a rich anthropological and ecological literature that documents how ancestral experiences with the environment shape enduring norms and cultural identities. Ethnographic and historical accounts describe how descendants of communities across the world developed ritualized ecological knowledge in stable climates or adaptive practices and protective taboos in unstable climates, resulting in environmental stewardship that persists through cultural transmission (Berkes, 2017; Lansing and Kremer, 1993; Atran et al., 2002; Gómez-Baggethun and Reyes-García, 2013). Folklore, rituals, and oral histories in many societies encode such ecological memory, consistent with broader findings on the intergenerational persistence of beliefs and norms (S. O. Becker et al., 2020; Giuliano and Nunn, 2021; Fernández, 2025). Our paper formalizes the mechanisms underlying these instances and shows that they scale globally: using comprehensive datasets containing information from individuals globally, we demonstrate that the anomalies in ancestral climate conditions predict contemporary attention to environmental issues and that this relationship follows a U-shaped pattern. In doing so, we bridge the anthropological literature with the economics of cultural transmission and preference formation, complementing work on the historical origins of norms related to prosociality, cooperation, gender, and economic preferences (Alesina, Giuliano, and Nunn, 2013; Lowes, 2018; Buggle and Durante, 2021; A. Becker, Enke, and Falk, 2020). Studies most closely aligned with our approach are Giuliano and Nunn (2021) and Buggle and Durante (2021), who analyze the effect of dissimilarity of climate experiences *across* generations and inter-annual variation in weather conditions on traditional values and cooperative behavior, respectively. Complementing their findings, we extend this line of inquiry by exploring how the

prevalence of climate anomalies *within* successive generations shapes attention to environmental issues. Closely related to our work, Dewitte (2024) shows that historical exposure to fossil fuel extraction leads to higher levels of climate change denial at the community level in the US, and that this is explained by the development of economic identities which influence belief formation. In a similar vein, Bhattacharya et al. (2025) highlight that interventions imparting environmental education can influence household behavior through intra-family transmission of environmental awareness and actions both from parents to children and also from children to parents.

Our model contributes to theoretical discussions on cultural transmission and preference formation, both expanding on canonical tools and drawing new connections. Our cultural transmission model is anchored in a large literature which adapts tools from evolutionary biology to model culture (see Cavalli-Sforza and Feldman, 1981; Boyd and Richerson, 1988, for seminal contributions). In economics, this ties in to two main strands of literature. A large body of work following the seminal contribution of Bisin and Verdier (2001) focuses on environments which sparsely capture the role of endogenous socialization incentives or institutions (see the review by Bisin and Verdier, 2023, and references therein). Another long standing approach, surveyed in Young (2015), builds on the tools of evolutionary game theory to study dynamics of social norms. Our approach features relatively standard replicator-mutator dynamics (see e.g. Hofbauer and Sigmund, 1998, for a textbook reference), which abstracts away from fine details of the transmission process, but enables us to consider a very rich environment with infinite dimensional cultural types. One contribution of our approach is that this richness of the type space provides scope for new effects that could not be captured in models with finite types, and which can be studied using tools from functional analysis. A second novelty lies in that our modeling of individual level decisions draws on the canonical rational inattention framework – which was pioneered by Sims (2003) and has since found broad applications (for a review, see Maćkowiak, Matějka, and Wiederholt, 2023). The key articulation we propose is to model cultural transmission as a mechanism that shapes *prior perception of value* in the attention problem. Although some previous work highlighted the importance of information is cultural transmission mechanisms (e.g. Adriani and Sonderegger, 2009; Adriani and Sonderegger, 2018; Adriani, Matheson, and Sonderegger, 2018), to the best of our knowledge the formal connection between rational inattention and theories of cultural transmission is novel.

Our study also connects to the literature on narrative economics, which highlights how collective memories and stories shape preferences and decision-making (Shiller, 2017; Akerlof and Snower, 2016). Narratives embedded in folklore and cultural traditions preserve social memory and can serve as vehicles for transmitting values both behavioral (Michalopoulos and Xue, 2021; Harari, 2014) and ecological (Osemeobo, 1994; Amlor and Alidza, 2016; Sanganyado, Teta, and Masiri, 2018). We extend this work by

showing that the presence of environmental themes in ethnic folklore is influenced by ancestral climate anomalies, linking narratives and cultural memory to environmental concern at scale.

Additionally, we contribute to research on climate change perceptions by identifying a novel historical determinant of environmental preferences. Existing studies emphasize factors such as imagery and emotion-based learning (Leiserowitz, 2006), personal experiences and morality (E. Weber, 1997; Hansen, Marx, and E. U. Weber, 2004), behavioral traits, misperceived norms (Andre et al., 2024), political ideology, and economic preferences (Luo and Zhao, 2019; Shi et al., 2016). We show that ancestral climate experiences have a persistent influence on environmental attention, even after accounting for other contemporary drivers.

The paper is structured as follows. Section 2 describes the data and introduces the objects of interest. Section 3 gives an overview of the theoretical framework, introduces the model and describes the main results. Section 4 presents the empirical strategy and the results. Section 5 provides various robustness checks, while Section 6 dives into heterogeneity analysis. Section 7 concludes.

2 Data and definition of relevant variables

This section introduces the key variables we study and the eight different data sources used to construct them (see Figure 1 for an overview). Our main outcome of interest is the *subjective value of attention to environmental issues* at the individual level – for which we provide multiple proxies relying on the Integrated Value Surveys (IVS, 2020) and European Social Survey (Poortinga et al., 2018; ESS-ERIC, 2018). Our second key outcome is *the prevalence of environmental themes in cultural material*, capturing group level beliefs, practices and traditions, and used to provide additional evidence for the cultural transmission channel. To construct proxies for this outcome. we use folklore data from Michalopoulos and Xue (2021), as well as data on written records from The Joshua Project and Wikipedia. For each outcome, we highlight one main variable that will serve as our baseline to illustrate the results, which are then corroborated across multiple alternative measures. Our central explanatory variable is *the average intensity of ancestral climate anomalies*, which we introduce and motivate below. This and additional variables capturing ancestral climatic conditions are constructed using historical temperature data from Mann et al. (2009). Linking individuals to their ancestors from corresponding ethnic groups relies on Giuliano and Nunn (2018)'s Ancestral Characteristics of Modern Populations.

Individual-level data – perceived value of environmental attention Environmental aspects are relevant across many situations where individuals or groups have to make decisions. Internalizing

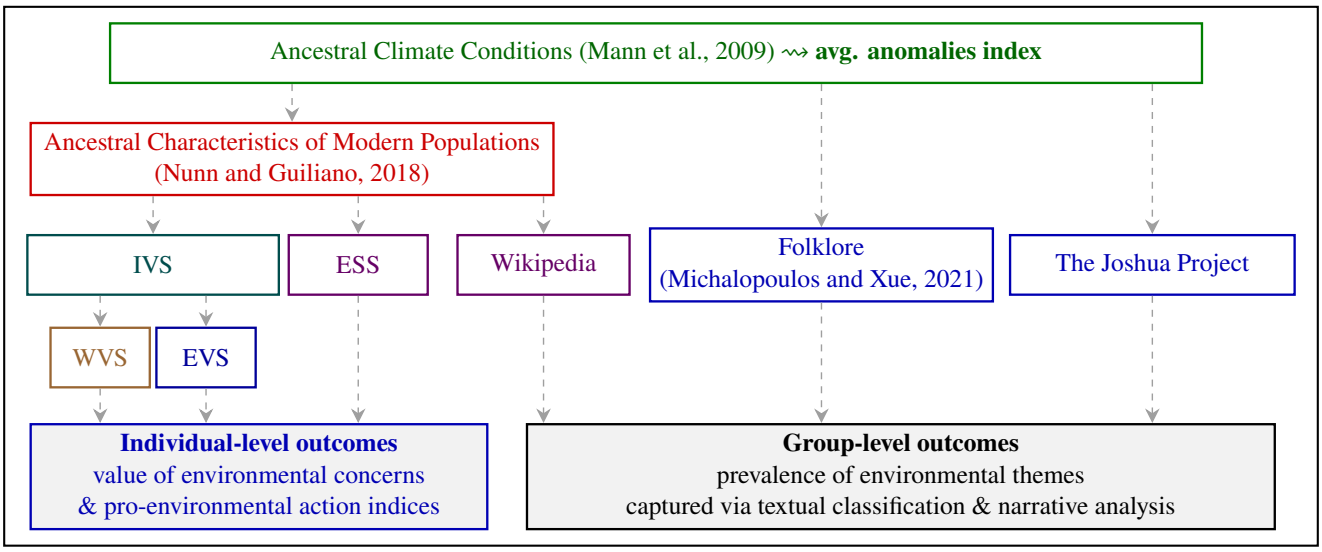


FIGURE 1: Datasets and merging strategy

environmental consequences or understanding how to factor environmental effects into specific choices requires costly effort, which itself is endogenously chosen. Such choices are guided by individuals' underlying perception or subjective belief about *the overall value of environmental concerns* – for which we will give a precise theoretical definition in the next section. This is our essential outcome of interest.

Our main dependent variable captures the level of importance attributed to environmental issues via individual responses to the question: *"How important is it to you to take care of the environment?"* from wave 5 and 6 of **World Value Surveys** section of the Integrated Value Surveys. [1,2]. The broad, agnostic and independent framing of the question makes it a suitable proxy for an individual's assessment of the value of making efforts to accommodate environmental concerns (and figure out how to optimally do so), across situations. Additionally, the Integrated Value Survey provides an array of demographic information on the respondents, including language spoken at home. This is used to link an individual to their ethnic ancestors, via the language or dialect associated to the ethnic groups available in the ancestral characteristics database (see details below).

[1]Responses to this question are captured on a scale of 1 to 6 where 1 corresponds to not important at all and 6 corresponds to extremely important to me.

[2]The Integrated Values Surveys (IVS) is a harmonized global dataset that integrates and standardizes responses from the World Values Survey (WVS) and the European Values Study (EVS), and is designed to facilitate cross-country and within-country analyses over time of human values, beliefs, and preferences using consistent measures. WVS covers nationally representative samples from over 100 countries between 1981 and 2022. Individuals are asked questions along many thematic categories (perceptions of life, environment, work, family, politics, society, religion and national identity). For our lead question of interest, we have a total of 157,142 respondents from 78 countries in waves 5 and 6 (2005-2014).

We also construct alternative proxies for the value of environmental attention. For our second set of proxies, we use other questions from the World Values Survey (WVS) and the European Values Study (EVS) that capture self-reported engagement in pro-environmental actions and support for environmental causes. Since responses to these questions may reflect both demand- and supply-side considerations and often require individuals to make explicit trade-offs, we recover a latent measure of intrinsic environmental attention by conducting a principal component analysis (PCA), and use the first principal component as our variable.[3] Lastly, we use the climate module of the 8th round of the European Social Survey (ESS) to construct a third set of proxy variables.[4] As before, responses may reflect both intrinsic concern and external constraints or policy environments, so we again employ principal component analysis (PCA) to isolate a latent measure of intrinsic attention to environmental issues, and use the first principal component as our index. We construct four alternative measures of environmental attention based on separate PCA exercises. The first three gradually expand the set of included questions (simultaneously reducing the subsample), where questions concern various environmental attitudes and tradeoffs; the fourth PCA is based solely on climate-change related beliefs. [5] As with the IVS, the ESS provides rich demographic information on respondents, which we similarly use to link individuals to their ethnic ancestry (via language) and provide controls.

For all constructions across different datasets, individual survey responses and the first principal component are normalized to lie between 0 and 1, with higher numbers indicating greater subjective value.

Group level data – environmental themes in culture Our lead group-level variable interest is a measure of the *relative proportion of environmentally-related motifs in folklore*, constructed using the data from Michalopoulos and Xue (2021). This folklore database builds on Berezkin's (2015) catalog of motifs and links the linguistic groups therein to the ethnic groups in Murdock's (1957) Ethnographic Atlas, providing a distribution of oral traditions and folktales associated to each ethnic group (see Online Appendix E.3 for details). Following the approach from Michalopoulos and Xue (2021), we use *ConceptNet*[6] (Speer, Chin, and Havasi, 2017) to classify motifs associated to each ethnic group as an environment related or non-environment related motif. We use the seed words: weather, climate, temperature, environment and natural disaster. We then retrieve all associated words and phrases for each seed that have high relevance

[3] See Online Appendix E.1 for further details about the questions we use for our PCA, data construction, sample size and additional demographic characteristics we use as control variables in our regressions related to varibles constructed from IVS.

[4] Similar to IVS, the ESS is a cross-national repeated social survey that measures public attitudes, beliefs, values, and behaviors, primarily across European countries. The 8th round contains a dedicated climate module that elicits respondents' views and behaviors related to climate change, energy use, and environmental policy. See Online Appendix E.2 for more information on the climate module, additional demographic variables used as controls and our associated data construction strategy.

[5] Information on the constructed proxies is available for approximately 24,000 to 33,000 respondents across 14 to 23 countries.

[6] ConceptNet is a knowledge representation project, designed to reflect the general human knowledge and its expression in natural language.

score and manually verify the relevance of each item. This yields a list of 228 words which we employ in a dictionary-based approach to classify motifs as environment vs non-environment related. Using this information, we construct our main measure of environment focus in folklore as:

$$\mathfrak{F}_e := \log\left(1 + \frac{\text{No. of environment related motifs}}{\text{Total no. of motifs}}\right)$$

for each ethnic group in the Ethnographic Atlas sample.

For robustness, we first construct alternative versions of our main outcome using (i) the raw share of environmentally related motifs among all motifs associated with an ethnic group, (ii) an inverse hyperbolic sine transformation of this share, and (iii) a log-link transformation of the share (for log-Poisson regressions). Second, we implement a classification procedure based on syntactic structure rather than full-text thematic scoring. For each motif, we extract subject–verb–object (SVO) triplets using a dependency parser and retain the unique subjects and objects implied by these relational structures, thereby reducing the vocabulary to a compact and interpretable set of actors and recipients. We manually classify these terms as environmental or non-environmental, and label a motif as environmentally related if at least one subject or object refers to a natural entity. This procedure yields an alternative measure of environmental focus in folklore that relies on explicit narrative roles rather than contextual embeddings. Online Appendix E.3 provides full details on parsing, term classification, and construction of the resulting index.

To complement the folklore-based measure, we construct a second group-level index of environmental attention using ethnographic descriptions from the Joshua Project. We collect standardized narrative texts for all ethnic groups with available content, excluding sections explicitly framed for missionary purposes, and use a manual review of a random subset to identify conceptually grounded markers of environmental salience. Guided by this exercise, we implement a BART-based zero-shot classification procedure (fine-tuned on MultiNLI (Williams, Nangia, and Bowman, 2018)) to score each group's text along predefined environmental dimensions, including ecological references, climate exposure, subsistence–environment linkages, resistance to environmental degradation, and nature-related spirituality. These scores are aggregated into a normalized composite index. Geographic identifiers in the database allow us to link groups to historical temperature anomaly data and additional controls. Online Appendix E.4 provides full details on data sources, text processing, classification design, and sample construction.

As a third measure, we construct an index of environmental attention using textual descriptions from English Wikipedia pages on ethnic groups. We manually compile the relevant set of pages corresponding to groups in the ancestral characteristics database and scrape their full narrative content, relying on closely related entries (e.g., language or associated political entities) when a dedicated page is unavailable. A

manual review of a random subset of pages informs the identification of conceptually grounded themes capturing environment-centered cultural narratives. Guided by this exercise, we apply a BART-based zero-shot classification procedure (again, fine-tuned on MultiNLI) to score each group's text along predefined dimensions, including the centrality of nature to identity and livelihood, nature-linked spiritual systems, traditional ecological knowledge, and narratives of resistance to environmentally threatening external actors. These scores are aggregated into a normalized composite index between 0 and 1. We additionally collect information on the number of Wikipedia language editions covering each group as a proxy for documentation breadth, and match the resulting measures to the ancestral characteristics database and temperature anomaly data. Online Appendix E.5 provides full details on page selection, preprocessing, thematic design, classification methodology, and sample construction.

Ancestral ethnic group data We rely on Giuliano and Nunn (2018)'s *Ancestral Characteristics of Modern Populations* database to obtain information on the pre-industrial traits of individuals' ancestors. The dataset links ethnographic information on ethnic groups—drawn from Murdock's *Ethnographic Atlas*, Bondarenko et al.'s (2005)'s dataset on the Peoples of Easternmost Europe, and Murdock's (1957)'s *World Ethnographic Sample*—to over 7,000 contemporary languages and dialects listed in Gordon's (2009)'s *Ethnologue: Languages of the World*, which we use to map individuals in the IVS and ESS to their corresponding ethnic ancestors.[7] The geographic coordinates provided for ethnic groups in the ethnographic sources allow us to match ancestral locations to grid-level climate data from Mann et al. (2009), enabling the construction of measures of ancestral climatic exposure. Additional details about the dataset and the corresponding variables are provided in Online Appendix E.6.

Historical temperature data and anomaly measure Mann et al. (2009) provide information on global patterns of detrended surface temperature for a long historical period using a climate field reconstruction approach.[8] The dataset reports the average annual temperature anomalies (deviations from the 1961-1995 reference-period average measured in degree celsius) at the 5degree-by-5degree ($\approx$ 555km $\times$ 555km) grid cell level since 500 AD. We restrict attention to the temperature data between 1600 AD and 1920 AD, to improve reliability of the data (by restricting to years for which at least 2 types proxies are available) and ensure that our measure is independent from contemporaneous confounds (by retaining a gap of at least 75 years with survey responses).

[7] For our analysis, we also use key pre-industrial characteristics of ethnic groups as controls. Specifically, we consider measures capturing primary subsistence patterns; indicators of settlement and economic complexity; agricultural intensity; community size; and local and supra-local jurisdictional hierarchy.

[8] See Online Appendix E.7 for more details

The main explanatory variable we construct is the *average intensity of deviations from typical conditions* faced by ethnic ancestors across several generations. This is motivated by the idea that predictability of the climate drives the true value of attention to environmental issues – which we will formalize in the next section. Intuitively: the more weather conditions concentrate on their (locally) "typical" range, the greater the potential *gains* from adapting to those conditions. By contrast, the more unpredictable extreme events occur, the higher the potential *losses* from not taking measures to protect oneself against such realisations.

In our baseline measure, we assume that each generation lives for 20 years and they do not overlap.[9] We also assume that all the historical generations associated to a particular ethnic group (details below) live within the same grid cell. [10] Therefore, for our primary specification, each ethnic group has a total of 16 generations associated to them. Fixing a grid cell / ethnic group, use g to denote its g^{th} generation; the associated avg. anomaly measure is constructed as follows:

1. Let $\tau_{g,t}$ the temperature faced by generation g in year t of their lifetime $\mathbb{T}_g$; denote $\hat{F}_g$ the empirical CDF of $\{\tau_{g,t}\}_{t\in\mathbb{T}_g}$. Recast temperature realizations in terms of *deviations from the "typical" range*, defined as the range between the α^{th} and the $(1-\alpha)^{\text{th}}$ percentile of the empirical distribution of temperature values within a generation's lifetime ($\alpha = .2$ in our main specification):

$$z_{t,g} := \underbrace{\mathbb{1}_{\tau_{g,t}<\hat{F}_g^{-1}(\alpha)}}_{\text{if}\tau_{g,t}<\alpha\text{-qtile}} \underbrace{\left|\tau_{g,t} - \hat{F}_g^{-1}(\alpha)\right|}_{\text{dist. to }\alpha\text{-qtile}} + \underbrace{\mathbb{1}_{\tau_{g,t}>\hat{F}_g^{-1}(1-\alpha)}}_{\text{if}\tau_{g,t}>(1-\alpha)\text{-qtile}} \underbrace{\left|\tau_{g,t} - \hat{F}_g^{-1}(1-\alpha)\right|}_{\text{dist. to }(1-\alpha)\text{-qtile}}.$$

2. For generation g, let $\hat{\theta}_g$ the average intensity of deviations from the typical range of conditions:

$$\hat{\theta}_g = \frac{1}{K_g}\sum_{t\in\mathbb{T}_g} z_{t,g},$$

where $K_g := \#\left\{\tau_{g,t} \;\middle|\; \tau_{g,t} < \hat{F}_g^{-1}(\alpha) \text{ or } \tau_{g,t} > \hat{F}_g^{-1}(1-\alpha)\right\}$ is just the normalizing constant that counts the number of excursions outside the $(\alpha, 1-\alpha)$ percentile range.

[9] Our measure of ancestral climate volatility/anomalies should not be confused with Giuliano and Nunn (2021)'s measure of climate instability. They use across generation changes in average temperatures to capture the difference in life experiences of different generations, whereas we are interested in within-generation experiences of climate variability, i.e. we are interested in short-term weather changes and the intensity of those changes from "normal" range of expected weather conditions within a generation.

[10] The assumption that all historical generations of a particular ethnic group live in the same grid is not completely unreasonable since the 5degree-by-5degree grid effectively covers an area of approximately 308,025km^2. If the centroid of the ethnic group is close enough to the centroid of the grid then minor movements historically across generations will be restricted within the grid level at which information is available.

3. Define the average ancestral climatic anomalies for an ethnic group as Avg. anomalies $= \frac{1}{G}\sum_{g=1}^{G}\hat{\theta}_g$, where G is the number of generations (16 is our baseline specification).

Within each generation g, $\hat{\theta}_g$ captures the intensity of deviations from typical climate conditions in the generation's lifetime, i.e how unpredictable and how extreme a climate generation g faced.[11] Averaging over generations captures the overall *scale* of deviations from typical conditions experienced by one's ancestors.[12] In Section 5, we also show the robustness of our results to the choice of the time span of generations, thresholds for extreme events, and alternative measures of volatility/anomalies.

3 Theory

Having defined our objects of interest, we build them into a general theoretical framework to analyze the impact of ancestral experiences on descendant's perceptions via cultural transmission. Before introducing the model and main results, we provide an intuitive overview of the theory in Section 3.1 .

3.1 Overview

We develop a model of the cultural transmission of the perceived value of caring about environmental issues. Individuals repeatedly face situations in which incorporating environmental considerations into their decisions requires costly effort (attention). How much they "care" in a given situation depends on their perceived stakes: what they believe they stand to gain from taking environmental consequences into account, or to lose by ignoring them. A central feature of the framework is that, in any given decision, the true stakes are a function of realized climate conditions. In some realizations, environmental attention matters little; in others, it is decisive. Crucially, however, attention must be chosen before climate conditions are observed. What matters ex ante is therefore the *expected* value of attention, taken under the specific climate distribution individuals face. Our main object of interest is the underlying perceived stakes *function*, mapping climate realizations into the value of attention. The resulting overall perceived importance of environmental care (across situations and under the climate conditions individuals anticipate facing) is the object we measure in the data.

To explain how such perceptions form, we combine two ingredients. First, the individual level effort choice is modeled as a canonical rational inattention problem – this provides a flexible approach that can

[11] This approach is similar to the exercise carried out by climatologists in extreme weather attribution and anomaly analysis (Smith et al., 2008; National Academies of Sciences, Engineering, and Medicine and others, 2016). We vary the definition of the typical range and our results are robust to different specifications.

[12] Variation in our variable of interest at the grid level is reported in Online Appendix Figure 7.

capture "care" across a rich variety of situations while remaining abstract about details of the specific problems. Expected stakes translate into attention effort: given any perceived stakes function, individuals optimally choose how much costly information to acquire. This gives a precise foundation to the "value of caring". Second, we embed this attention problem into population-level evolutionary dynamics that capture cultural transmission. Within each generation, individuals differ in their perceived stakes functions – this defines cultural types. Individuals behave optimally given their perceptions, realize gains and losses under the actual climate conditions they face, and more successful types transmit more effectively to the next generation, subject to noise.

If attention was chosen after climate realizations are observed, evolutionary pressure would select the correct stakes function at every point. In that case, ancestral climate conditions would have no effect on descendants' perceptions. But because attention is chosen ex ante, evolutionary selection is necessarily coarse. Fitness depends only on average realized losses, and therefore only on the expected value of the stakes function under experienced conditions. Evolution attempts to discipline a function but can identify only a moment of that function. The remaining degrees of freedom are shaped by mutation and historical contingency. This generates our central theoretical prediction: descendants' perceived value of environmental attention reflects the true average value of attention under the climate distribution experienced by their ancestors. Figure 2 gives a visual representation of the mechanism.

We further provide a simple micro-foundation for the primitive shape of the true value of attention. Environmental considerations combine two broad motives. The first is exploitation: attention helps individuals make better use of typical climate conditions, so its value declines as conditions move away from the normal range. The second is protection: attention reduces losses from extreme events, so its value rises with the severity of deviations. Aggregating across heterogeneous situations yields a U-shaped relationship between realized climate intensity and the value of attention. Because evolutionary dynamics select only on averages, this primitive U-shape transfers to average perceived value: ancestral populations exposed either to very stable or to highly variable climates converge to higher perceived value of environmental attention, while intermediate variability induces lower perceived value. This delivers the U-shaped pattern in our empirical analysis.

3.2 Model structure

To describe the model, take the point of view of a single lineage of successive ancestral generations and their descendants. We first introduce the generic *environmental attention problem* faced by all individuals in each generation; and then embed it into a evolutionary dynamics framework which drives the *perceived stakes*.

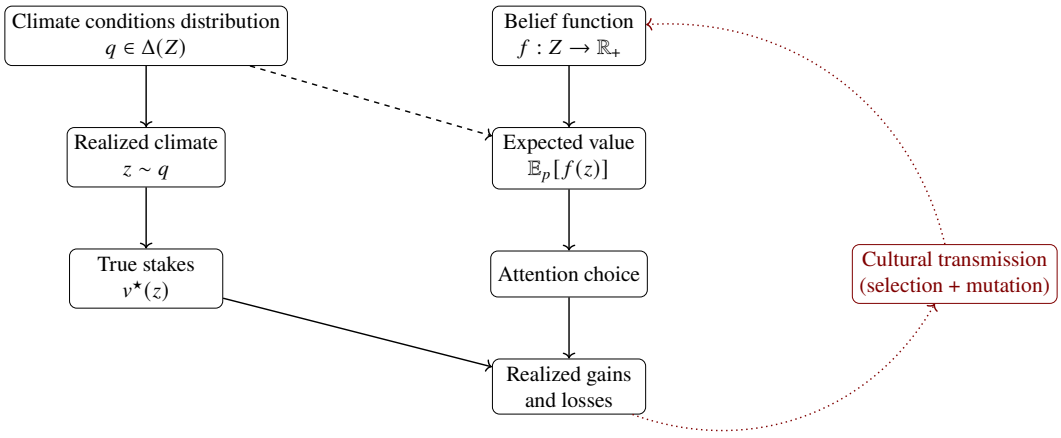


FIGURE 2: The cultural transmission mechanism – selection depends only on the average

ENVIRONMENTAL CONCERNS ACROSS SITUATIONS Each individual faces many situations over the course of their lifetime in which they must decide on some course of action. This requires weighing many complex aspects: acquiring and processing information, thinking about all relevant tradeoffs, evaluating consequences of one's action, etc. This decision is costly, and itself endogenously chosen. To model this sparsely, represent each problem via the *loss from imprecise attention to environmental consequences*. Let s be a particular situation. There is an unknown state of the world $\omega_s \in \mathbb{R}$, and the agent's chosen "action" $a_s \in \mathbb{R}$ induces losses $v_s(z)(\omega_s - a_s)^2$, where $v_s(z) \geq 0$ is a situation-specific function of *realized climate conditions* z. The function $v_s(z)$ is the key object we will micro-found; for now take it as given (the individual's perception) to introduce the problem. Assume that the agent's prior is $\omega_s \sim \mathcal{N}(0, \sigma_s^2)$.

This is a flexible modeling approach. One natural interpretation is to view ω_s as the action that would be optimal under perfect assessment of environmental consequences, normalized so that action zero is chosen when environmental considerations are ignored, and losses scale with the squared distance from this benchmark. Alternatively, this formulation can be seen as a reduced-form representation of losses arising from uncertainty; in that case ω_s represents an abstract state of knowledge, the agent's action corresponds to their best estimate of that state, and losses reflect the expected mean-squared error in estimating ω_s.

CLIMATE CONDITIONS We represent climate conditions by *intensity of deviation* from a well-understood "normal" range, so that $z \in Z \subset \mathbb{R}_+$ measures how far realized conditions depart from typical levels in a given period. What matters for decision-making is precisely this distance from the norm. The climate-sensitive stakes parameter $v_s(z)$ captures both heterogeneity of situations – environmental aspects may matter more in some problems than others (dependence on s) – and variation in realized losses with climate conditions (dependence on z). Each generation knows the distribution of climate conditions they

expect to encounter. As a baseline, assume that ancestral generations, living in a fixed geographical area, face a common distribution $z \sim q \in \Delta(Z)$,[13] with descendants potentially facing different conditions $\tilde{q} \in \Delta(Z)$ (due to migration, climate change, etc.).

Optimal attention In each situation, the agent chooses how much costly attention to devote to environmental consequences. We model this as a standard rational inattention problem: the agent selects an information structure about ω_s, paying a cost proportional to Shannon mutual information. Given quadratic losses and Gaussian priors, it is without loss to summarize the problem as a choice of posterior variance $\xi \leq \sigma_s^2$:

$$\min_{\xi \leq \sigma_s^2} \ \mathbb{E}_p[\hat{v}_s] \cdot \xi \ + \ \kappa \log\left(\frac{\sigma_s^2}{\xi}\right),$$

where $\mathbb{E}_p[\hat{v}_s]$ denotes the expected perceived stakes at the time attention is chosen. The derivation of this form of the problem is standard in the rational inattention literature – see Online Appendix G.1 for details and Maćkowiak, Matějka, and Wiederholt (2023) for a review.

The solution is standard:

$$\xi^* = \min\left\{\sigma_s^2, \ \frac{\kappa}{\mathbb{E}_p[\hat{v}_s]}\right\},$$

so posterior variance is decreasing in expected stakes. Equivalently, higher perceived value of attention induces strictly greater uncertainty reduction and information acquisition. Thus, $\mathbb{E}_p[\hat{v}_s]$ is the sufficient statistic governing attention effort and realized losses. The perceived stakes is precisely the object whose formation we seek to micro-found through cultural transmission. For expositional simplicity, assume from now on that prior uncertainty is symmetric across problems $\sigma_s^2 \equiv \sigma^2$, and solutions are always interior.

Generational structure and true value We embed the individual-level rational inattention block into a population model spanning successive generations. Each ancestral generation, indexed by $g \in G$, contains a population of individuals and lives for T periods indexed $t = 1, ..., T$. In each period t of generation g's lifetime, climate condition $z_{g,t} \sim q$ is realized (drawn iid). In each period, there are $n = 1, ..., N$ attention problems ("situations") of the form described above. Assume that the *true value* of attention $v_{g,t,n}$ in problem (g, t, n) is a random function drawn iid from a given distribution over a suitable class of functions (details below) with mean $v^\star$. The function $v^\star : Z \to \mathbb{R}_+^*$ is the underlying **average stakes function** (across situations).

[13] This is a natural assumption: a fixed geographical area faces a relatively stable *distribution* of conditions (though random *realizations* in any given generation's lifespan). The results, however, extend to assuming heterogeneous climate distributions.

Cultural types and realized losses A **cultural type** is described by a *perceived average stakes function* $f : \mathbb{R}_+ \to \mathbb{R}_+^*$. Let $\mathcal{F}$ a suitable functional class of admissible cultural types, containing the true average $v^\star$ and any situation-specific value v_s.[14] An individual with cultural type f chooses their optimal attention in problem (g, t, n) according to the RI problem using perceived stakes $f(z)$, with $z \sim q$. Define optimal posterior variance as a function of cultural type:[15]

$$\xi_q(f) := \frac{\kappa}{\mathbb{E}_q[f]},$$

which implies that *realised losses* by an agent of cultural type f in problem (t, g, n) are given by:

$$\ell_{g,t,n}(f) := v_{g,t,n}(z_{g,t}) \times \xi_q(f) + \kappa \log\left(\frac{\sigma^2}{\xi_q(f)}\right)$$

Observe that realised losses depend on the attention choice of the agent (hence their cultural type), and also on the realised climate condition $z_{t,g}$ which enters the *true* stakes function. Define the average realized losses of a type f in generation g:

$$L_g(f) := \frac{1}{T}\frac{1}{N}\sum_{t=1}^{T}\sum_{n=1}^{N} \ell_{g,t,n}(f)$$

Evolutionary dynamics: transmission mechanism We now link the distribution of cultural types across generations, using standard ideas from evolutionary theory. In each generation, a plurality of types coexist and each behaves optimally according to their belief. The next generation's population of types is generated from the previous one by *selection* – types which fare empirically better are transmitted more successfully – and *mutation* – capturing a random change component.

Formally, let $P_g \in \Delta(\mathcal{F})$ be the population distribution over cultural types f in generation g. Define the standard *replicator-mutator update* as (up to a normalizing constant):

$$P_{g+1}(df') \propto \int_{\mathcal{F}} \underbrace{M(df'|f)}_{\text{mutation}}\underbrace{\Phi_g(f)}_{\text{selection}} P_g(df)$$

[14] Formally, we place ourselves in the Sobolev space $H^1(\mathbb{R}_+)$ and take $\mathcal{F}$ to be a subset of positive-valued functions with bounded norm, pointwise bounded below by some arbitrarily small constant. See Appendix C.1 for details.

[15] This assumes for simplicity that agents have no problem-specific information and choose attention based on aggregate perceived value; a natural extension can be written where agents have some information about the problem specific-noise $\varepsilon_{t,g,n}$.

where $M(\cdot|f) \in \Delta(\mathcal{F})$ is a mutation kernel on $\mathcal{F}$. This is the usual evolutionary structure: the fitness function Φ_g captures the relationship between a type's "success" or "fit" to the environment they face and reproduction. A better relative fit means more reproduction (overweighting) of that type in the next generation, with cultural types randomly mutating across generations (captured by M).

We make two generic assumptions about the structure of selection and mutations to discipline dynamics. First, we assume that fitness Φ_g is a decreasing function of realized losses. For expositional simplicity, we adopt *exponential selection*; this is a standard choice which gives convenient representations, but results extend to alternative functional forms.

Assumption 1 (Exponential selection). Fitness is a decreasing exponential function of average losses L: $\Phi_g(f) := \varphi\big(L_g(f)\big)$, where $\varphi(L) =: \exp(-\tau L)$ and $\tau > 0$ captures the strength of evolutionary pressure.

Second, we assume that mutation is *locally mean-reverting towards a default* and "direction neutral".

Assumption 2. There exists a function $f_0 \in \mathcal{F}$, constants $\lambda \in [0, 1)$, $\rho > 0$, such that:

$$\mathbb{E}_{f' \sim M(\cdot|f)}[f'] = (1 - \rho\lambda) f + \rho\lambda f_0 \quad \text{and} \quad \mathrm{Cov}_{M(\cdot|f)} = \rho C$$

where C is a self-adjoint trace-class operator with $C \propto I$ on its support.

The intuition for mean-reversion is that, absent the force of selection, mutation on its own tends to revert to a "default" value f_0. A natural choice is to take $f_0 \equiv 0$, so that if evolution does not actively select individuals who do well by paying attention to climate issues, the population reverts to a perception that there is *no value*. The assumption on $\mathrm{Cov}_{M(\cdot|f)}$ (isotropy) amounts to saying that the noise component of mutations is not biased towards one directions – this is essentially the analogue of "spherical noise" in a functional space. Parameter $\rho > 0$ captures the strength of mutation, and we will typically assume it to be small. Parameter λ captures the relative strength of the pull towards the default versus persistence, conditioning on the level of mutation. In the limit case $\lambda = 0$, the default plays no role, so pure inertia and noise dominate.

Lastly, for tractability, we make a standard assumption to capture that the selection and mutation forces are relatively balanced and "not too strong". The main role of this assumption is to enable well-defined local approximation of the population distribution around the average cultural type.

Assumption 3 (Weak selection/small mutation). $\tau = O(\rho)$ ("weak selection") and $\rho \ll 1$ ("small mutations")

3.3 Main results

Our model delivers two main results. First, we characterize the outcome of the evolutionary process, which yields both a general result, establishing the exact form of path-dependency, and a toolbox to study variation across settings. Second, we harness this machinery to show that a simple micro-foundation of the value of learning predicts that descendant's perceived value is U-shaped in ancestral climate volatility.

Characterizing evolutionary outcomes Before stating the first theorem, make a simple observation in order to introduce conveniently condensed notation, and foreshadow the intuition behind the result. Average realized losses $L_g(f)$ for any given type f can be rewritten as a simple function of *ex-ante expected stakes* $\mathbb{E}_q[f]$ and *empirical average stakes* $\hat{\bar{v}}_g$. Indeed expanding for losses $\ell_{t,g,n}(f)$ and optimal attention $\xi_q(f)$ yields:

$$L_g(f) = \mathcal{L}\big(\mathbb{E}_q[f] \,\big|\, \hat{\bar{v}}_g\big); \quad \text{where } \mathcal{L}(\mu|v) := v \times \frac{\kappa}{2\mu} + \kappa \log\left(\frac{2}{\kappa}\sigma^2\mu\right) \quad \text{(loss function)},$$

$$\text{and } \hat{\bar{v}}_g := \frac{1}{T}\frac{1}{N}\sum_{t=1}^{T}\sum_{n=1}^{N} v_{g,t,n}(z_{g,t}) \quad \text{(empirical gen. } g \text{ mean value)}.$$

The "primitive" loss function $\mathcal{L}$ condenses all relevant information about evolutionary fit, and appears prominently in the result – which we now state before expanding on the intuition and implications.

Theorem 1. *Fix an ancestral distribution of conditions q. The average cultural type f_g converges to $\hat{v}$ which is the unique minimizer of the penalized fit objective:*

$$J(f) := \tau\mathcal{L}\big(\mathbb{E}_q[f] \,\big|\, \mathbb{E}_q[v^\star]\big) + \frac{\lambda}{2}\|f - f_0\|^2$$

As $\lambda \downarrow 0$ the solution approaches the closest function to f_0 which correctly identifies average stakes $\mathbb{E}_q[v^\star]$:

$$\hat{v} \xrightarrow[\lambda\to 0]{} \underset{f \,|\, \mathbb{E}_q[f]=\mathbb{E}_q[v^\star]}{\arg\max} \|f - f_0\|$$

One essential observation underpins the intuition for the main result. Since attention choice optimally depends on the *expected value* $\mathbb{E}_q[f]$, because specific climate conditions are not known ex ante, realized losses and hence fitness too only depends on the expected value induced by a cultural type. This means that, in a given generation, evolutionary dynamics can only reward fitness *to the mean* and do not distinguish further properties of the function f. Indeed, observe that two cultural types f_1 and f_2 which induce the same mean value under given conditions i.e. $\int f_1 dp_g = \int f_2 dp_g$ induce the exact same attention choice

$v(f_1) = v(f_2)$, losses $\ell_{t,g,n}(f_1) = \ell_{t,g,n}(f_2)$, and fitness $\Phi_g(f_1) = \Phi_g(f_2)$. In that sense, two cultural types which induce the same mean in a given generation are completely *evolutionary equivalent*: they will have the same reproductive advantage in the next generation.

Crucially, this wedge is introduced by the interplay of attention choices being made *ex ante* with evolutionary dynamics selecting for *ex post* losses. As a benchmark, consider a simple alternative model where agents observe the current z *before* choosing attention. In that case, losses would penalize for the fit to the *entire function* $v^\star$; as the weight of the default λ vanishes, we would then obtain that the average cultural type converges to $v^\star$ itself. This hints at a general phenomenon caused by the combination of decision under uncertainty (or information acquisition) across rich situations and evolutionary selection, which to the best of our knowledge is a novel theoretical observation.

A pivotal intermediate result translates this observation into full dynamics: we can rewrite the updates on the average cultural type as successive steps of a kernel stochastic gradient descent, under empirical losses.

Proposition 1. *Under weak selection and small mutation, the population mean evolves according to:*

$$f_{g+1} = f_g - \rho \nabla J_g(f_g) + o(\rho)$$

where J_g is the empirically realized penalized fit objective in generation g:

$$J_g(f) := \tau \mathcal{L}\big(\mathbb{E}_g[f] \,\big|\, \hat{\bar{v}}_g\big) + \frac{\lambda}{2}\|f - f_0\|^2$$

This is both a useful technical step but also provides further intuition into the inner mechanics. Indeed, it allows the use of tools from the statistics and machine-learning literature to study convergence and behavior over time. The complete proof of Theorem 1, which can be found in Appendix C.2 uses Proposition 1 to establish that evolutionary dynamics converge to a minimizer of the *theoretical* penalized fit objective J. We then provide a slightly more general result in Appendix Lemma 2, which gives an explicit solution of this problem – and, in particular, delivers the result of Theorem 1.

Ancestral heterogeneity Let us now recontextualize the implications of Theorem 1 by taking a step back and considering the effect of *effect of ancestral heterogeneity*. This is cleanly captured by considering the effect of varying the distribution q faced by ancestors. Define the ex ante perceived value of attention

for a descendant now facing climate $\tilde{q}$ when their ancestor faced climate q is:

$$B(q|\tilde{q}) := \mathbb{E}_{z\sim\tilde{q}}[\hat{v}_q(z)] = \int \hat{v}_q(z)d\tilde{q}(z),$$

where $\hat{v}_q$ denotes the limit average cultural type (stakes function), characterized in Theorem 1, as a function of ancestral conditions q – from now on, focus on the $\lambda \downarrow 0$ solution and take $f_0 \equiv 0$ for simplicity. $B(q|\tilde{q})$ is the key determinant of attention choice, and is exactly the object we measure in our data. The expression for $B(q|\tilde{q})$ highlights the bias introduced by q. *If* evolutionary dynamics induced effective pointwise learning of $v^\star$, *then we should* see *no effect* of descending from a particular group of ancestors: we would have $\hat{v} \approx v^\star$ and $B(q|\tilde{q}) \approx \mathbb{E}_{\tilde{q}}[v^\star]$. But since evolutionary dynamics only fit to the mean, the climate faced by one's ancestor q distorts today's perceived value. $B(q|\tilde{q})$ is generically biased away from $\mathbb{E}_{\tilde{q}}[v^\star]$. Further, under small λ, $\hat{v}_q$ is the closest function to the default f_0 which verifies $\mathbb{E}_{z\sim q}[\hat{v}_q(z)] = \mathbb{E}_{z\sim q}[v^\star(z)]$. This pushes $\hat{v}_q$ towards overweighting the relative value of conditions that are frequent under q, causing potentially severe mismatch of shape and magnitude between $\hat{v}_q$ and $v^\star$. More generally, Theorem 1 provides a toolbox which reduces studying the effect of ancestral experiences on descendants' attention to a comparative statics exercise (of B, in q). In particular, it implies the following decomposition result.

Proposition 2. *The contemporaneous average belief B can be decomposed into:*

$$B(q|\tilde{q}) = V(q) \times S(q|\tilde{q})$$

where $V(q) := \mathbb{E}_q[v^\star]$ *is average ancestral value in conditions q and S is a climate similarity measure,* $S(q|\tilde{q}) = \frac{\mathbb{E}_{q\otimes\tilde{q}}[G]}{\mathbb{E}_{q\otimes q}[G]}$*, using the similarity kernel* $G(x,y) = \frac{1}{2}\left(e^{-|x-y|} + e^{-|x+y|}\right)$.

This formalizes the intuition that the average value among descendants is essentially the average value among ancestors, modulated by something which depends on how different a climate they face. Notice in particular that $S(q|q) = 1$, by construction. To anchor the comparison and tie it back to our empirical exercise, assume that the possible distributions of climate anomalies faced by different ancestral ethnic groups are parameterized by a scale family, $\{q_\theta\}$:

$$Z \sim q_\theta \iff Z \sim \theta\bar{Z}, \text{ where } \bar{Z} \sim \bar{q},$$

for $\theta \in [\underline{\theta}, 1]$, and where we have chosen without loss to normalize in terms of the maximal scale distribution $\bar{q}$. The scale parameter θ controls "how stretched out to the right" the distribution is, so it parametrizes the average intensity of deviations from the typical range. Proposition 2 can be written with all quantities as function as $\theta, \tilde{\theta}$ for any two distributions $q_\theta, q_{\tilde{\theta}}$; taking log derivatives in θ for fixed $\tilde{\theta}$

implies that the marginal effect of ancestral climates decomposes neatly into:

$$\frac{\frac{\partial}{\partial\theta}B}{B}(q_\theta|q_{\tilde{\theta}}) = \underbrace{\frac{\frac{\partial}{\partial\theta}V}{V}(q_\theta)}_{\text{transmission effect}} + \underbrace{\frac{\frac{\partial}{\partial\theta}S}{S}(q_\theta|q_{\tilde{\theta}})}_{\text{dissimilarity effect}}.$$

If the dissimilarity effect is "not too strong" relative to the transmission effect, then the descendant average value B, as a function of ancestral conditions, will retain the structural properties of ancestral average value V – although it will always alter the strength of the effect. To make this concrete and give specific conditions on the primitives, we turn to a simple micro-foundation of the value of environmental attention.

Rationalizing the U-shape: exploitation and protection Assume that in any given situation, the true value of attention derive from a possibly problem specific mix of two motives. The first is an *exploitation motive*, which induces losses that are decreasing in deviations from normal conditions (there is more to gain from attention when conditions are typical). The second is a *protection motive*, which induces losses that are increasing in deviations from normal conditions (there is more to lose from inattention when conditions are unusual). Formally, in any situation s:

$$v_s(z) = \alpha_s w(z) + \beta_s r(z),$$

where $w(z)$ represents the gains from adaption to typical conditions and $r(z)$ represents the reduction of losses from extreme events. For convenience, assume from assumed convex decreasing and with vanishing slope at the right boundary of Z, , assumed increasing with vanishing slope at the left boundary; $\alpha_s, \beta_s \geq 0$ are problem specific weights. Assume we can average across the range of possible situations to give the representation as before:

$$\mathbb{E}[v_s] = v^\star, \text{ where: } v^\star(z) = \alpha\ w(z) + \beta\ r(z)$$

where $\alpha, \beta > 0$ represent the fundamental average weights of the exploitation and protection motives across the situations that one might encounter. This assumption delivers our second main theoretical result.

Theorem 2. *i.* (Core U-Shape) *There exists a unique $\theta^\dagger$ such that V is strictly decreasing if $\theta < \theta^\dagger$ and strictly increasing if $\theta > \theta^\dagger$.*

ii. (Approximate U-Shape Preservation) *Assume that* $\theta^\dagger \in (\underline{\theta}, \overline{\theta})$ *and that* $v^\star$ *is strongly convex i.e. there exists* $m > 0$ *such that* $(v^\star)'' \geq m > 0$. *Let* $V_{\max} := \max\{V(\underline{\theta}), V(\overline{\theta})\}$ *(expected value upper bound); and:* $\mu_1 := \mathbb{E}_{z\sim\bar{q}}[z]$, $\mu_2 := \mathbb{E}_{z\sim\bar{q}}[z^2]$, *and* $\bar{\ell} := \mathbb{E}_{z\sim\bar{q}}[e^{-\overline{\theta}z}]$ *(distributional constants).*

For any $\delta > 0$, *if:*

$$\frac{m}{V_{\max}}\mu_2\delta > \frac{3\mu_1}{\bar{\ell}^2}$$

then the map $\theta \mapsto B(\theta \mid \tilde{\theta})$ *is strictly decreasing on* $[\underline{\theta}, \theta^\dagger - \delta]$ *and strictly increasing on* $[\theta^\dagger + \delta, \overline{\theta}]$. *In particular, every global minimizer of* $B(\cdot \mid \tilde{\theta})$ *lies in* $[\theta^\dagger - \delta, \theta^\dagger + \delta]$, *uniformly in* $\tilde{\theta}$.

A first step towards the result is the elementary observation that under our exploitation-protection assumption, the primitive $v^\star$ is U-shaped *in climate conditions* z. However, this is misleading: it establishes a U-shaped property, but not the one we observe, since for now this is only about the true value in terms of realized climate conditions. How does this transfer to descendants' perception? To get the intuition behind Theorem 2, first consider an individual facing the same anomaly distribution q_θ as their ancestors. Then average realized $z_{t,g}$ (which is exactly our explanatory variable in the data) is an unbiased estimator of the scale parameter θ, so for enough generations:

$$\sum_g \frac{1}{G} \sum_t \frac{1}{T} z_{t,g} \approx \theta$$

The perceived value of attention is approximately equal to the mean value of attention under ancestral conditions via evolutionary convergence:

$$\mathbb{E}_{q_\theta}[\hat{v}_{q_\theta}] \approx \int v^\star(z) dq_\theta(z) = \alpha W(\theta) + \beta R(\theta) \text{ where } W(\theta) = \mathbb{E}_{q_\theta}[w], R(\theta) = \mathbb{E}_{q_\theta}[R]$$

which the first point in the theorem establishes is U-shaped in θ. However, this approximate equality masks the modulating effect of the descendants' climate distribution $q_{\tilde{\theta}}$. While Proposition 2 gives a general intuition why the U-shape will be preserved provided that climate dissimilarity effects are not too strong, the effects remain very complex. Indeed, there are many possible ways to produce a precise "U-shape preservation" statement, that each require taking a stand on fine properties of $v^\star$ and $\bar{q}$. In Theorem 2, we focus on a simple idea: if $v^\star$ is convex enough (the protection and exploitation marginal effects are sensitive enough to variation in conditions), then the U-shape obtains. How convex it needs to be depends on the distributional properties of $\bar{q}$ (through its moments and Laplace transform) and how closely we want the approximation to hold. See Appendix C.3 for the complete proof of the Theorem and further details.

Together, Theorem 1 and Theorem 2 give a precise and general description of how descendant's average perception of the value of environmental attention is shaped by the distribution of climate conditions their ancestors faced. We have focused on the U-shape, which is our central prediction. Many more interesting comparative statics exercises can be carried out, using this theoretical machinery, to analyze the effect of various parameters – we discuss some of these in Section 6. The theoretical framework is substantially more general than our application, and we hope this may serve as a building block for further research.

4 Empirical results

4.1 Individual level analysis – value of environmental attention

We begin by examining the impact of average ancestral climatic anomalies on an individual's perceived value of attention to environmental issues. We run a simple OLS specification, controlling for individual demographic characteristics, historical ethnic group's social and political characteristics, historical ethnic group's geographical characteristics and respondent's country-by-year fixed effects. Our empirical strategy is equivalent to a continuous treatment framework where the treatment is the average intensity of anomalies in ancestral climate conditions and treatment is assigned by lottery of birth, i.e. an individual doesn't choose which ethnicity they are born in. Our regression equation reads as:

$$y_{iect} = \beta_1 \text{Avg. Anomalies}_e + \beta_2 (\text{Avg. Anomalies}_e)^2 + \mathbf{X_{ict}}\mathbf{\Gamma} + \mathbf{X_e}\mathbf{\Omega} + \alpha_{ct} + \epsilon_{iect}, \tag{1}$$

where i indexes the individual, e indexes the historical ethnic group an individual belongs to, c indexes the country of the individual and t indexes year. Specifically:

- y_{iect} refers to the dependent variable of interest, i.e. an individual's perception of the value of attention, captured by the multiple proxies created from WVS, IVS, and ESS (see Section 2);
- Avg. Anomalies is the index of anomaly intensity in ancestral generations constructed in Section 2;
- $\mathbf{X_{it}}$ is the vector of individual level demographic characteristics which include Age, Gender, Income, Educational level and dummies for Occupation categories;
- $\mathbf{X_e}$ corresponds to the vector of historical ethnic group controls (see Online Appendix E.6 for details about coding) which include primary mode of subsistence, economic development indicators (complexity of settlement, size of local community and intensity of agriculture and level of local and global jurisdictional development), as well as ethnicity level geographical variables such as distance

to closest coast and equator and Köppen climate classification of the spatial grid the ethnic group was historically located in;

- α_{ct} corresponds to country-year fixed effects which captures any contemporaneous variables such as level of economic development of the country (of residence), general population's level of schooling, awareness and perception of climate issues[16] and contemporary climate shocks faced by the country's population.

We cluster the standard errors at the ethnic group level. Our coefficients of interest are β_1 and β_2. Our theory suggests that the level of attention should initially decrease and then increase with the intensity of anomalies faced by ancestral generations. As a result, the expected signs on β_1 and β_2 should be negative and positive respectively, capturing that the perceived value is (i) decreasing over the lower range of values of ancestral climatic variability because the diminution of benefits from exploitation dominates evolutionary dynamics and (ii) increasing over higher values because the increase of the loss-protection motive for attention starts to dominate.

Table 1 reports the results from Equation 1 for our lead variable of interest, i.e. normalized responses to the question: *"How important is it to you to take care of the environment?"*. Column 1 reports the results we obtain from regressing the variable on historical measures of ancestral climatic volatility controlling for country-year fixed effects. Subsequent columns add individual level and historical controls. Across all specifications, our results corroborate the hypothesis. The results we find are all statistically significant at 1% level and consistent across different choice of sample or method of construction of our ancestral climate anomalies index.

For the purpose of a stylized illustration, assume that, on our index of climate anomalies, the most stable ancestral climates scored 0.01 and the most unstable ancestral climates scored 0.09.[17] In that case, the coefficients β_1 and β_2 suggest that a increase of 0.01 ($\sim$ 10%) in the average deviations from the typical climate conditions in the most stable ancestral environments would have led to a 8.5% *decrease* in attention to environment today relative to the sample mean and a increase of 0.01 in the average deviations in the most unstable ancestral environments would have led to a 14.3% percentage point *increase* in attention to environment today.

[16] This can capture all sorts of general equilibrium factors and spillovers within the current generation that can emanate from the country of residence

[17] For the sample in consideration, the average ancestral climate anomalies captured by our constructed index ranges from 0.015 to 0.093. We restrict the sample to remove a small mass of outliers lying outside the 99th percentile of the distribution of our values of average anomalies.

TABLE 1: Coefficients for the impact of average anomalies in ancestral climatic conditions on individual's self-reported attention to the environment

Dependent variable: Self reported measure of attention to environment				
	(1)	(2)	(3)	(4)
Avg. anomalies	-6.291***	-7.095***	-9.194***	-9.877***
	(1.587)	(1.902)	(2.087)	(2.003)
Avg. anomalies sq.	71.038***	78.984***	100.939***	111.767***
	(18.330)	(21.756)	(22.608)	(21.346)
Income		-0.001*	-0.001*	-0.001*
		(0.001)	(0.001)	(0.001)
Male		-0.013***	-0.013***	-0.013***
		(0.003)	(0.003)	(0.003)
Age		0.002***	0.002***	0.002***
		(0.000)	(0.000)	(0.000)
Education level		0.006***	0.006***	0.006***
		(0.001)	(0.001)	(0.001)
Occupation category Controls	N	Y	Y	Y
Historical ethnic group characteristics	N	N	Y	Y
Historical topographic characteristics	N	N	N	Y
Country-year Fixed effects	Y	Y	Y	Y
Mean of dep var	0.702	0.704	0.704	0.704
St. Dev. of dep var	0.252	0.250	0.250	0.250
Min value Avg. anomalies	0.015	0.015	0.015	0.015
Max value Avg. anomalies	0.093	0.093	0.093	0.093
R-sq	0.097	0.113	0.113	0.113
Adj. R-sq	0.097	0.112	0.112	0.113
N	157142	138067	138067	138067
*** p<0.01, **p<0.05, * p<0.1, + p<0.15				

Note: The unit of observation is an individual. The dependent variable is the individual's level of attention paid to environment. The dependent variable ranges between 0 and 1 and increases with the reported level of attention. The variable is constructed by rescaling the answer to the prompt: On a scale of 1 to 6, how important is it for this individual to take care of the environment. Average anomalies refers to the average intensity of deviations from the typical climate conditions (*specific to each ancestral generation*) across generations. Average anomalies sq. refers to the square of the average anomalies term. Average anomalies range between 0.015 and 0.093 within the sample. Historical ethnic group characteristics include measure of development such as agricultural intensity, complexity of settlement, level of political heirarchies, size of the local community and main source of subsistence. Historical topographic characteristics include controls for distance to the equator, distance to the closest coast and geographical Köppen climate classification for the location of the ethnic group obtained from Ethnographic Atlas. Standard errors are clustered at the ethnicity level.

We provide additional evidence on this relationship, using our second proxy, constructed from data on self-reported engagement in pro-environmental actions and support for national environmental causes from the Integrated Values Surveys. We present results from regressing the environmental attention index on historical measures of ancestral climatic volatility, controlling for individual-level characteristics, historical ethnic-group-level controls, and country–year fixed effects in Table 3. Across Columns 1 to 4, we vary the

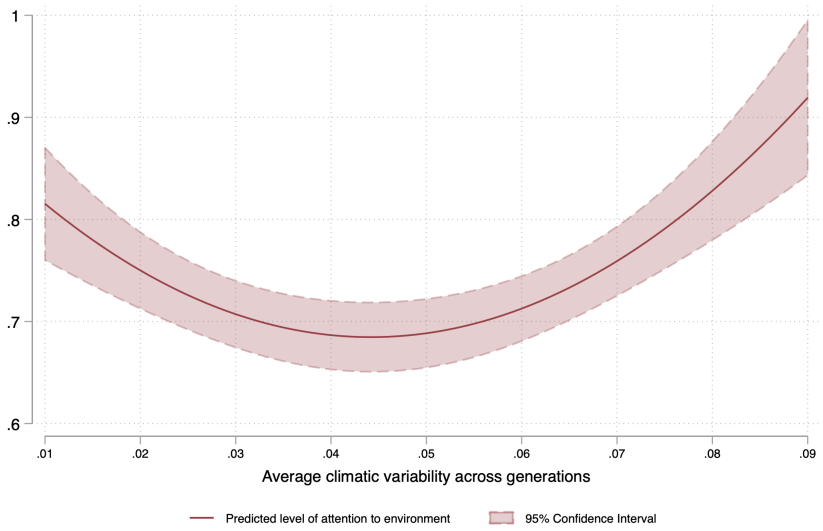


FIGURE 3: Margins plot - Individual's level of attention as a function of average ancestral climatic anomalies

Note: N = 138067. Margins plot corresponding to Equation 1. Average value of climatic anomalies are allowed to range from 0.01 to 0.09. The figure provides concrete evidence of an decreasing level of attention for increases in ancestral climate anomalies (empirical estimate of the scale parameter governing the prior) for values at the lower end (through reduced value of learning due to a lower possibility of exploitating the environment) and an increasing level of attention for increases in ancestral climate variability for values at the higher end (through increased value of learning due to an higher possibility of protection against extreme events). For the margins plot, continuous controls are fixed at their mean values. Discrete / categorical controls are fixed at Male = 1, Historically agrarian society = 1 and Köppen climate classification = D.

choice of sample and variables used to construct the PCA whose normalized first component we use as our index (see Appendix A.1 and Online Appendix E.1 for more details). Across all specifications, we observe a statistically significant U-shaped relationship between ancestral climatic anomalies and environmental attention, with the estimated coefficients satisfying $\beta_1 < 0$ and $\beta_2 > 0$.

We finally replicate this analysis with our third proxy constructed from the climate and energy module included in the 8th round of the European Social Survey (ESS). We estimate Equation 1 with the pro-environment indices as the dependent variables, controlling for the same set of individual-level characteristics and historical ethnic-group-level controls as in the previous specifications. Because all ESS observations are drawn from a single survey year (2016), we include country fixed effects rather than country–year fixed effects. Table 4 reports the results. Across Columns 1 to 4, we vary our choice of variables to construct the PCA whose normalized first component we use as our index (see Appendix A.2 and Online Appendix E.2 for more details). Across all specifications, we again observe a statistically

significant U-shaped relationship between ancestral climatic anomalies and pro-environmental attention, with estimated coefficients satisfying $\beta_1 < 0$ and $\beta_2 > 0$.

4.2 Group level analysis – environmental themes in culture

Beyond individual level analysis, the impact of ancestors' experiences on descendants' perceptions can be captured at the group level through the stock of knowledge and collective memory related to environmental themes. The stories, legends, songs, etc. which constitute folklore and narratives are transmitted within language groups across times and locations. Hence this measure serves two possible purposes. First, if successive generations of a particular ethnic group face climate conditions which increase each generation's level of attention paid to the environment, then increased attention should result into a higher occurrence of environmental themes in folklore and associated narratives. Second, this corroborates the idea that the (evolutionary) transmission across generation passes through cultural material.

We test this theory by first regressing the level of environmental focus in ethnic folklore on our variable of average ancestral climatic volatility, using the data decribed in Section 2. We run the following specification at the ethnicity level:

$$\mathfrak{F}_{ec} = \beta_0 + \beta_1 \text{Avg. anomalies}_{ec} + \beta_2 (\text{Avg. anomalies})^2_{ec} + \mathbf{X_{ec}\Omega} + \alpha_c + \epsilon_{ec} \tag{2}$$

where e indexes an ethnic group, c indexes the country whose current geography would have historically included the ethnic group and $\mathfrak{F}_{ec}$ denotes the previously constructed outcome variable for prevalence of environmental motifs in folklore, i.e:

$$\mathfrak{F}_{ec} := \log\left(1 + \frac{\text{No. of related motifs}_{ec}}{\text{Total no. of motifs}_{ec}}\right)$$

$\mathbf{X_{ec}}$ captures historical ethnic group controls as before such as primary mode of subsistence, economic development indicators (complexity of settlement, size of local community and intensity of agriculture) and level of jurisdictional development (both locally and globally). It also includes ethnicity level geographical variables such as distance to closest coast and equator and Köppen climate classification of the spatial grid the ethnic group was historically located in. In line with Michalopoulos and Xue (2021), we additionally control for the first year of publication and the total number of publications, authors, publishers and languages of the text associated to each ethnic group. Country fixed effects are captured by α_c. We cluster the standard errors at the language group level as specified by Berezkin (2015). Results from this specification are reported in Table 2.

TABLE 2: Coefficients for the impact of ancestral climatic anomalies on occurrences of environment related ethnic folklore

Dependent variable: ln(1 + No. of environment related motifs/Total no. of motifs)				
	(1)	(2)	(3)	(4)
Avg. anomalies	-3.233***	-3.172***	-2.493***	-1.548**
	(0.944)	(0.735)	(0.784)	(0.764)
Avg. anomalies sq.	33.930***	31.539***	25.341***	17.237**
	(9.526)	(7.684)	(8.047)	(7.112)
First year of publication		0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)
Number of authors		0.001	0.002**	0.002+
		(0.001)	(0.001)	(0.001)
Number of languages		-0.014***	-0.011***	-0.006
		(0.003)	(0.004)	(0.004)
Number of publishers		-0.001	-0.000	0.003
		(0.002)	(0.003)	(0.003)
Number of publications		-0.000	-0.002	-0.005**
		(0.002)	(0.002)	(0.002)
Historical ethnic group characteristics	N	N	Y	Y
Historical topographic characteristics	N	N	Y	Y
Country Fixed effects	N	N	N	Y
Mean of dep var	0.276	0.276	0.276	0.277
St. Dev. of dep var	0.097	0.097	0.097	0.096
Min value Avg. anomalies	0.010	0.010	0.010	0.010
Max value Avg. anomalies	0.097	0.097	0.097	0.097
R-sq	0.020	0.116	0.198	0.439
Adj. R-sq	0.018	0.110	0.180	0.366
N	1037	1037	1037	982
*** p<0.01, **p<0.05, * p<0.1, + p<0.15				

Note: The unit of observation is an ethnic group. Environment related motifs are classified using ConceptNet. Average anomalies refers to the average intensity of deviations from the typical climate conditions (*specific to each ancestral generation*) across generations. Average anomalies sq. refers to the square of the average variability term. Average anomalies range between 0.01 and 0.097 within the sample. Historical ethnic group characteristics include measure of development such as agricultural intensity, complexity of settlement, level of political heirarchies and main source of subsistence. Historical topographic characteristics include controls for distance to the equator, distance to the closest coast and geographical Köppen climate classification for the location of the ethnic group obtained from Ethnographic Atlas. Standard errors are clustered at language group level.

As before, the coefficient on Avg. anomalies is negative and statistically significant at 5% level and the coefficient on Avg. anomalies squared is positive and statistically significant at 5% level, thereby providing evidence that the climatic experiences of ancestors feed into the attention parameter in the same way and increase the stock of environmental folkore for either extremely stable or extremely unstable ancestral climates, with a dip in the intermediate range. An increase of 0.01 ($\sim$ 10 %) in average deviations from the typical climate conditions in the most stable environments leads to a $\sim$1% *decrease* in the proportion

of environmentally related folklore and an increase of 0.01 in average deviations in the most unstable environments lead to a ∼2% *increase* in the proportion of environmentally related folklore.

To provide further evidence on this relationship, we additionally evaluate the effects of ancestral climatic anomalies on the prevalence of environmental themes in descriptive narratives drawn from historical records of ethnic groups available on English Wikipedia and the Joshua Project. We first use the composite index capturing environmental themes constructed from English Wikipedia text as our main dependent variable,[18] and re-estimate Equation 2. As before, we control for historical ethnic-group-level characteristics and include country fixed effects. To account for variation in the quality and granularity of available information, we additionally control for the length of the associated English Wikipedia page and the number of distinct language editions in which the ethnic group has a corresponding page. Table r 5eports the results.[19] Across all specifications, we observe a statistically significant U-shaped relationship between environmental content in written narratives and the degree of ancestral climatic anomalies.

We replicate this exercise using the index constructed using written narratives of ethnic groups available in the Joshua Project. We estimate the same specification as in Equation , 2 where $\mathfrak{F}_{ec}$ now denotes the normalized score derived from Joshua Project text capturing the level of environmental attention reflected in written records for each ethnic group. The vector $\mathbf{X}_{ec}$ includes ethnic-group-level characteristics available in the Joshua Project dataset. Specifically, we control for the group's specified religion, whether it is classified as least reached or frontier (i.e., variables capturing the difficulty of reaching these groups as reported by Joshua Project contributors), and whether the group is classified as indigenous. In addition, we include geographic characteristics at the ethnic-group level, such as distance to the nearest coast, distance to the equator, and the Köppen climate classification of the spatial grid corresponding to the ethnic group's location. Country fixed effects are captured by α_c, and standard errors are clustered at the people group level as specified in the Joshua Project. Results from this specification are reported in Table 6. Across all specifications, from Columns 1 through 4, where we sequentially add ethnic-group-level controls and country fixed effects, we again find a statistically significant U-shaped relationship between environmental content in written narratives and the degree of ancestral climatic anomalies.

[18] That is, $\mathfrak{F}_{ec}$ now denotes the normalized score derived from Wikipedia text capturing the level of environmental attention reflected in written records for each ethnic group.

[19] Column 1 presents estimates from specifications that include only pages directly describing the ethnic group, excluding pages related to the geographic or political entity in which the group is located or the language it speaks. Column 2 expands the sample to include ethnic groups for which only pages related to geographic or political entities are available, while Column 3 further includes groups for which only language-based pages are available.

5 Robustness

We verify the robustness of the U-shaped relationship between average ancestral climate anomalies and the individually perceived value of environmental attention. We run multiple alternate specifications to establish that our findings are not driven by the idiosyncrasies of the sample, our choice of lifespan, range of typical conditions or misspecification of the functional form in the regression equation. We further establish that the core theoretical results hold more generally under various natural extensions of the model.

Choice of language sample Matching individuals to their corresponding ethnic ancestors through the use of spoken language could introduce systematic noise. First, even though our approach is agnostic on the channel of association apart from existence of a link between individuals and some of their ancestors through shared language, the mapping of descendants to their ancestors is inherently more imprecise for languages with a larger group of speakers (e.g. English, Spanish, Arabic). Further, if a large part of our sample is composed of individuals belonging to a particular language group, it is natural to test the sensitivity of our results to the inclusion of such groups. To tackle these issues, we consecutively remove three of the largest language groups from our sample and re run Equation 1 for our main variable of interest, i.e. responses to "How important is ti for the individual to take care of the environment?". We initially exclude English and Spanish speakers from our sample (Table 10, column 2) and run Equation 1 again. In the next specification, in addition to English and Spanish speakers, we also remove Arabic speakers (Table 10, column 3). Another way of tackling the bias emanating from large ethnic groups present in our sample is to run a weighted OLS regression where the weights are inverse of the frequency with which individuals of these groups occur in the sample. Column 4 reports results obtained from the weighted regressions instead. Our results and their significance stay consistent across all sample restrictions and alternate regression specifications. Second, for approximately 7% of the sample in consideration, we had proxied for the language spoken at home with the language of the interview. Column 5 in Table 10 reports results from the specification where individuals whose language at home was proxied by the language of the interview are removed. The coefficients on average variability and average variability squared still retain their original signs and significance levels.

Contemporaneous region confounds Individuals belonging to a specific ethnic group within our sample of interest may have decided to migrate to specific regions whose characteristics might drive their attention to environment. Therefore, the coefficients on the average ancestral climatic anomalies might be picking up correlated contemporaneous effects of current location instead of ancestral experience. Firstly, our original empirical specification tackles this through the use of respondent's current country-year fixed

effects. The inclusion of these fixed effects make the identification of coefficients β_1 and β_2 in our sample independent of any variation that would have jointly explained the responses of two individuals from the same ethnic group residing in the same country at the time of survey. One may still worry about the validity of our results if individuals from specific ethnic groups cluster in specific regions within the country and different regions face differential contemporaneous climatic shocks. To further robustly address this concern, we run a specification (Table 10, column 6) where we replace country-year fixed effects with region-year fixed effects.[20] That is, we increase the granularity of the contemporaneous spatial variable capturing the current location of the individual. In this case, the identification of coefficients β_1 and β_2 is independent of any variation that would have jointly explained the responses of two individuals from the same ethnic group residing in the same region of a country at the time of survey. The signs and the statistical significance of the coefficients of interest under this specification stay the same as before.

Construction of ancestral climatic anomalies We also test robustness of the results to our choice of measure of average climatic anomalies across ancestral generations. We vary our choice of lifespan for each generation from 20 to 50 years and we vary our definition of the range of typical conditions (pertinent to each ancestral generation) between $10^{th} - 90^{th}$ percentiles of temperatures within the lifetime, $20^{th} - 80^{th}$ percentiles of temperatures within the lifetime, $30^{th} - 70^{th}$ percentiles of temperatures within the lifetime. Our original specification assumed the lifespan of a generation to be 20 years (in line with Giuliano and Nunn (2021)) and the range of typical temperatures to be the $20^{th} - 80^{th}$ percentiles of temperatures within the lifetime. The first three blocks in Table 11 report the coefficients associated to average anomalies and average anomalies squared from the regression of our main variable of interest, attention to environment, on these variables (with the choice of lifespan dictated by the column and the choice of typical range dictated by the row). We carry out two additional specification checks. In the first one, instead of taking the average deviations from the typical range cut points to construct the within generation climate variability index, we take the average of deviations squared. Across generation average structure still stays the same. In the second one, instead of taking deviations from typical range, we instead assume climate variability within the generation is just indexed by the standard deviation of temperatures faced by the generation. Across all cases, the signs and the significance of the coefficients seem to show that our findings are robust to alternate choices of variable construction. [21]

Other characteristics of temperature distributions and model fit We now test the sensitivity of our results to the inclusion of additional variables that capture either broader features of the historical

[20] Where the region is a sub-territory within the country. See IVS documentation for further details.

[21] A similar exercise with our group level variable of attention to environment (environmental themes in ethnic folklore) shows broadly the same results.

temperature distribution associated with an ethnic group or variation across ancestral generations within a group. We also evaluate model fit using the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). Results from the different specifications, described in detail below, are reported in Table 12.

We re-estimate Equation 1 under several modifications. First, we augment the baseline specification by including higher-order terms of ancestral climatic anomalies, such as the cubic term (Column 3). While this introduces a risk of overfitting, it allows us to assess whether the signs and statistical significance of the linear and quadratic terms are preserved. Second, we include a measure capturing the standard deviation of average temperatures faced across ancestral generations within each ethnic group (Column 4). This addresses concerns that our baseline measure may proxy for dissimilarities in ancestral experiences across generations, consistent with issues raised in the original approach of Giuliano and Nunn (2021). Third, we include higher-order moments, up to the fourth order (mean, standard deviation, skewness, and kurtosis), of the overall temperature distribution experienced by each ethnic group over a 320-year period (Column 5). This specification helps address concerns that our measure of climatic anomalies may capture idiosyncratic geographic characteristics of group locations beyond the topographic and climate-region controls already included in the baseline specification. In addition, we estimate a linear specification that includes only the average level of ancestral climatic anomalies, excluding the quadratic term (Column 1), which we benchmark against our baseline quadratic specification (Column 2).

Across all specifications, both the signs and statistical significance of our coefficients of interest remain robust. Comparing AIC and BIC values across models, we find that the linear specification not only fails to capture the non-linear relationship but also performs substantially worse in terms of model fit. In particular, the quadratic specification yields AIC and BIC values that are more than 200 points lower (i.e., better) than those of the linear model. By contrast, the inclusion of additional higher-order terms or distributional moments does not meaningfully improve model fit.

Alternate functional forms for folklore We also test the sensitivity of our group-level results based on folklore data to the functional form used in constructing the dependent variable. To do so, we estimate four alternative specifications and benchmark them against the baseline results reported in Table 2. First, instead of relying on ConceptNet to classify environmentally related folklore, we use our manual classification based on subject–verb–object (SVO) triplets extracted from motifs using spaCy. Under this approach, a motif is classified as environmentally related if either the subject or the object in any associated SVO triplet is identified as an environment-related term. We then re-estimate the baseline specification using the same log transformation of the resulting folklore shares to assess sensitivity to the classification method. Results from this specification are reported in Column 1 of Table 13. Next,

we revert to the original ConceptNet-based classification and vary the functional form applied to the folklore shares. Specifically, we re-estimate Equation 2 using the untransformed raw shares (Column 3) and the inverse hyperbolic sine transformation of the raw shares (Column 4), addressing potential concerns associated with log transformations. Finally, we estimate a log-Poisson regression using the raw shares as the outcome to further assess robustness to alternative modeling choices. Across Columns 1 and 3–5, we consistently find a statistically significant U-shaped relationship between the prevalence of environmentally related folklore at the group level and average ancestral climatic anomalies.

Adjustments for Spatial Non-Independence Finally, to ensure that our results are not biased by spatial correlation in the standard errors, we re-estimate all specifications across the World Values Survey, Integrated Values Survey, European Social Survey, Ethnic Folklore, Wikipedia, and Joshua Project samples. We allow for spatial dependence in the error structure within radii of 200 km and 500 km. Standard errors for the main coefficients on average climatic variability and its square are reported in curly brackets for the 200 km radius and in square brackets for the 500 km radius in Table 14. Across all specifications and samples, the statistical significance of our estimates largely remains intact.

Extensions and variants of the model The core predictions of the model are robust to many extensions, and indeed the technical results provide a more general toolbox. Some are relatively direct generalizations which have been omitted for clarity of exposition. For instance, we could consider that agent observe a noisy-signal of the problem specific stakes instead of deciding on attention choice based purely on average stakes – this complicates derivations but results are largely unchanged under mild assumptions. We could also relax some tractability assumptions like interior solutions in the attention problem, or isotropic noise; the machinery is straightforward and the forces are similar, but this may allow additional richness in the specific forms of our objects of interest. A more meaningful generalization would be to consider heterogeneous climate distributions among ancestors; in that case, the main results are still largely preserved if more nuanced: instead of fitting to *one* moment, the evolutionary process will fit to multiple – but as long as the diversity of experienced climates (or the sample of experiences in each generation) remain small, bias will be induced. An example of a more substantial relaxation that can be accommodated is to remove the mean-reversion to the default in the mutation structure. This creates a technical difficulty as the evolutionary process is now dominated by inertia and noise, and no longer guaranteed to converge to a unique point; however, the core evolutionary bias remains and this gives additional scope for even stronger forms of path dependency.

6 Heterogeneity Analysis

As a last exercise, we probe our claim of cultural transmission impacting an individual's level of attention to environment by carrying out a series of heterogeneity analyses. To do so, we split our World Value Survey sample in different ways either by the level of historical development of the ethnic group, the current climate conditions faced by the respondents and their characteristics such as migration status.

Historical economic development Recall that the climate experiences of the ancestors in our specifications is modelled by the average temperature deviations from normal conditions across generations. However, the temperature deviations can have differential impact on groups based on their level of development. That is, if an ethnic group was institutionally and economically developed then there is reason to believe that the variability in climate would have had a lesser bite, which in turn should affect the process of prior formation at the individual level. Theoretically, since the prior of the individual is empirically estimated through the expectation of variability distribution, the level of protection endowed by the historical economic and institutional development should reduce the value of this constant and mute the transmission.

We provide suggestive evidence of a muted effect by rerunning Equation 1 separately on sub-samples split by level of economic and institutional development at the ethnic group level. The sample is divided along four characteristics, depending on whether the ethnic ancestors of an individual (1) were hunter gatherers or non-hunter gatherers, (2) had low / high level of economic complexity, (3) had low / high level of local jurisdiction and (4) had low / high level of global jurisdiction.[22]

Table 7 reports the results from our original specification for each sub-sample. Note that, irrespective of how the sample is spliced, the U-shaped relationship is maintained. However, the main purpose of this exercise is to compare the slopes across the two samples, not the levels. Indeed, historical economic development may independently impact other factors such as the level of education or awareness towards climate issues in general. In this exercise, we analyze how transmission through the climatic variability at the ancestral level gets muted or amplified in the presence or absence of historical economic protection. If ancestors had a higher stock of economic and institutional protection then the actual temperature variability would result in a less than one to one transition into climate related ancestral experiences. Therefore, to facilitate comparison, we normalize the margin plots we obtain from our separate regressions, so that the

[22]Economic complexity is captured by the size of local communities (similar to Giuliano and Nunn (2021)). Communities of size above 400 are considered as those with high economic complexity. If a group had a local jurisdiction of at least 4 levels or global jurisdiction of at least one level, they are considered as communities with high local (global) jurisdiction. See IVS (2020) for more details.

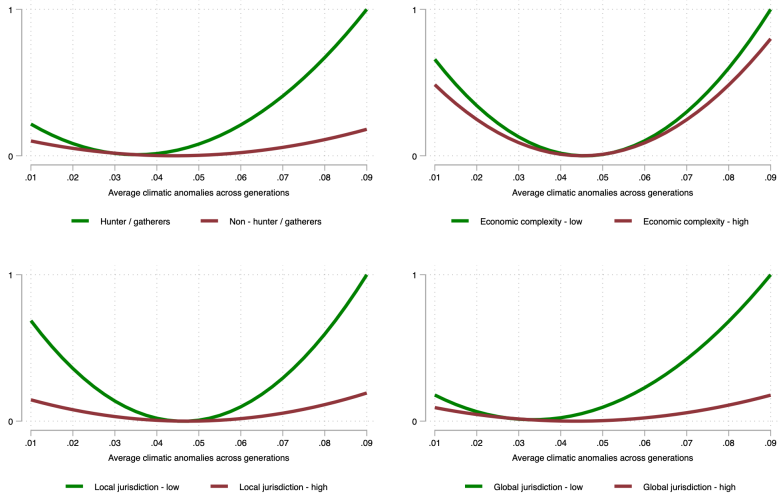


FIGURE 4: Margins plots - Relative slopes of individual level of attention function categorized by the development level of the ethnic group

> **Note:** Margins (corresponding to Equation 1) plotted for samples split by the level of ethnic group's historical development. All values are normalized to facilitate comparison of slopes, i.e. the minimum point of margin plots across the sample split are made equal to each other and all other margin values are normalized to lie between 0 and 1. Controls are fixed at their mean values from the original sample. Top left sub-figure corresponds to margins obtained from hunter / gatherers vs non-hunter / gatherers. Top right (Bottom left / Bottom right) figure corresponds to margins obtained from low vs high level of economic complexity (local / global jurisdiction) respectively.

minimum point of the predicted level of attention function from both the sub-samples are equalized and values are normalized to be in $[0, 1]$.

Figure 4 plots predicted attention functions for each sub-sample under the four different types of sample splits. In all cases, we see that the U-shape of the attention function is more pronounced when the ancestral level of economic and institutional development is low, i.e. for individuals coming from more economically developed ethnic groups, the U-shaped curves are pointwise flatter. This suggests that the experience of ancestors (and not the historical temperature distribution itself) is indeed the key variable which feeds into cultural transmission to shape an individual's prior today.

CONTEMPORANEOUS ENVIRONMENTAL CONDITIONS What determines individuals' attention to environmental issues? So far, we have shown attention is shaped by priors formed through ancestral experiences. These ancestral experiences operate as a supply-side factor in the attention problem. We now examine when and how the relationship between ancestral experiences and current environmental attention becomes

more or less pronounced by studying their interaction with contemporaneous, demand-side factors arising from an individual's current environment.

In periods or locations where climate change is more salient, through higher exposure to natural disasters, the marginal impact of inherited priors on environmental attention may be amplified. Separately, when current climatic conditions more closely resemble ancestral environments, inherited beliefs and heuristics may be more relevant for decision-making, increasing the losses from inattention and strengthening the role of ancestral experiences. As before, our objective is not to assess whether attention levels differ across environments, but to understand how the marginal impact of ancestral climatic variability varies with contemporaneous conditions.

To formally investigate these mechanisms, we split the World Values Survey sample along two dimensions. First, we split the sample by the median level of recent natural disaster impacts at the country level.[23] Second, we divide respondents by the median similarity between current climatic variability and ancestral climatic variability.[24]

Table 8 reports estimates of our baseline specification for each subsample. Irrespective of how the sample is split, the U-shaped relationship between ancestral climatic variability and environmental attention is preserved. However, as before, our primary interest lies in comparing slopes rather than levels across subsamples. To facilitate this comparison, we normalize the predicted margins from each subsample so that the minimum of the predicted attention function is equalized across groups, and all values are scaled to lie in the interval $[0, 1]$.

Figure 5 plots the resulting predicted attention functions. Greater exposure to recent natural disasters heighten the marginal impact of ancestral climatic variability on current environmental attention. Similarly, when current climatic conditions more closely resemble ancestral environments, the influence of ancestral experiences is amplified. In contrast, the U-shaped curves corresponding to individuals facing fewer recent disaster impacts or greater dissimilarity between current and ancestral environments are pointwise flatter.

Migration and language as a channel of transmission A potential concern is that our estimates may be biased if descendants migrate far from their ancestral climate zones or if ancestral climatic conditions are correlated with other long-run historical traits—such as institutions or colonial legacies—that also

[23] We rely on disaster-impact data (EM-DAT, CRED / UCLouvain, 2024) to compute country-level averages over the past 20 years of (i) the share of GDP affected by natural disasters, (ii) the proportion of the population that died due to natural disasters, and (iii) the proportion of the population affected. To reduce dimensionality, we perform a principal component analysis over these variables and use the first component as our measure of disaster exposure.

[24] This measure is constructed as the absolute difference between ancestral climatic anomalies and the current climatic anomalies experienced in the respondent's country of residence.

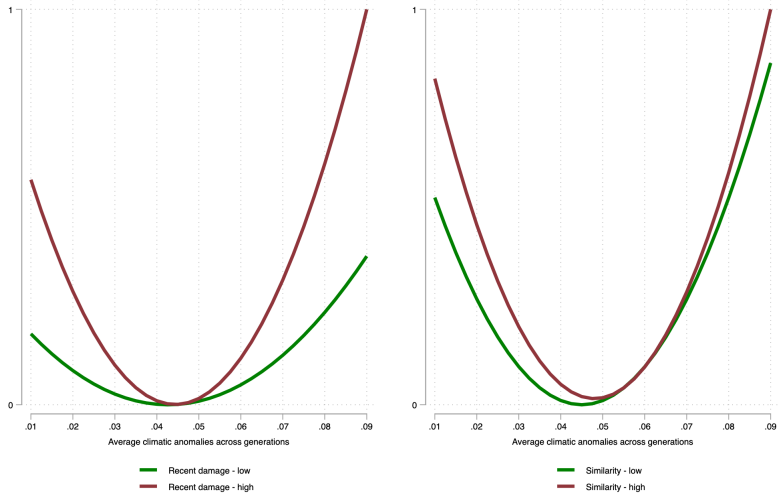


FIGURE 5: Margins plots - Relative slopes of individual level of attention function categorized by current climatic conditions faced by the respondents.

> **Note:** Margins (corresponding to Equation 1) plotted for samples split by the current level of instability, damages and similarity in country's weather and ancestral weather. All values are normalized to facilitate comparison of slopes, i.e. the minimum point of margin plots across the sample split are made equal to each other and all other margin values are normalized to lie between 0 and 1. Controls are fixed at their mean values from the original sample. Left sub-figure corresponds to margins obtained from low vs high level of recent disaster related damages. Right sub-figure corresponds to margins obtained from low vs high level of similarity between current environment and ancestral environment.

shape contemporary environmental attitudes. While we already control for a rich set of historical ethnic-group-level characteristics to mitigate these concerns, this critique is further attenuated in our setting because individuals are linked to ancestral climatic experiences through language spoken, rather than through country of origin or current location. This mapping implies that the relevant channel of transmission operates through linguistic and cultural narratives, such as family values, stories, and cultural practices, that are transmitted intergenerationally and can persist irrespective of individuals' current place of residence. This interpretation is consistent with our emphasis on folklore as a cumulative stock of knowledge and beliefs embodied in language.

Nevertheless, to directly assess whether historical migration weakens the relationship between ancestral climate and contemporary environmental attention, we split the sample based on whether at least one parent of the respondent was an immigrant to the country of residence.[25] Columns 7 and 8 of Table 7 report the results for children of immigrants and non-immigrants, respectively. We find that the estimated

[25]Directly identifying first-generation immigrants is not feasible across all waves of the World Values Survey, as this information is not consistently collected, which would lead to substantial sample loss.

relationships are strikingly similar across the two subsamples: the U-shaped pattern is preserved, and both the magnitudes and confidence intervals of the coefficients almost completely overlap. This evidence suggests that the transmission of ancestral climatic experiences operates primarily through language-based and cultural channels, rather than through country-specific institutions or persistent exposure to ancestral locations.

7 Conclusion

The starting point of this paper is a simple intuition: attention to climate-related issues is influenced by ancestors' climatic experiences through cultural transmission. We formalize this idea into an empirically testable hypothesis and, using several data sources, examine how ancestral climate variability affects individual and group attention to environmental issues. Our results show a significant impact: attention follows a robust U-shaped pattern, with individuals whose ancestors faced consistently stable or volatile climates showing the highest concern for environmental issues, and a dip for intermediate conditions.

We propose a flexible theoretical framework to explain this pattern, rooted in a dual purpose of learning –exploiting typical conditions and protecting against extremes– which is supported by anthropological evidence. Our model captures the effect of cultural transmission by assuming that perceived stakes are influenced by prior beliefs, which aggregate ancestral experiences via a scale parameter. The framework predicts the U-shaped pattern of attention under mild assumptions, underscoring that the dual purpose of learning drives the result. To further support the cultural transmission mechanism, we analyze environmental themes in folklore, which displays the same U-shape, suggesting that folklore retains attention given to environmental issues over generations. Additional heterogeneity analysis reveals a differential sensitivity to ancestral experiences based on characteristics like economic development.

Beliefs about environmental issues are of critical importance in the age of climate change. By highlighting the role of culture and socialization, our work contributes to the broader study of cultural transmission and its impact on climate preferences. Prior work in ecological anthropology had descriptively shown how communities embed ecological knowledge in norms, rituals, and social practices to navigate both stable and volatile environments. Our analysis generalizes these insights to a global scale. With the help of this quantitative analysis along with a systematic conceptual framework, we provide a bridge between anthropological accounts and an economic approach to understanding the role of ancestral experiences in the formation of climate preferences. We hope that this project provides clear evidence and a simple analytical framework, laying groundwork for further research on the cultural and ancestral roots of environmental concerns.

References

Adriani, F., J. A. Matheson, and S. Sonderegger (2018). "Teaching by example and induced beliefs in a model of cultural transmission". *Journal of Economic Behavior & Organization* 145, pp. 511–529 (cit. on p. 4).

Adriani, F. and S. Sonderegger (2009). "Why do parents socialize their children to behave pro-socially? An information-based theory". *Journal of Public Economics* 93.11-12, pp. 1119–1124 (cit. on p. 4).

– (2018). "Signaling about norms: Socialization under strategic uncertainty". *The Scandinavian Journal of Economics* 120.3, pp. 685–716 (cit. on p. 4).

Akerlof, G. A. and D. J. Snower (2016). "Bread and bullets". *Journal of Economic Behavior & Organization* 126, pp. 58–71 (cit. on p. 4).

Alesina, A., P. Giuliano, and N. Nunn (2013). "On the origins of gender roles: Women and the plough". *The quarterly journal of economics* 128.2, pp. 469–530 (cit. on pp. 1, 3).

Amlor, M. Q. and M. Q. Alidza (2016). "Indigenous Education in Environmental Management and Conservation in Ghana: The Role of Folklore". *Journal of Environment and Ecology* 7.1 (cit. on p. 4).

Andre, P. et al. (2024). "Peter Andre and Teodora Boneva and Felix Chopra and Armin Falk". *Review of Economics and Statistics*, pp. 1–46 (cit. on p. 5).

Atran, S. et al. (2002). "Folkecology, cultural epidemiology, and the spirit of the commons: A garden experiment in the Maya lowlands, 19912001". *Current anthropology* 43.3, pp. 421–450 (cit. on pp. 1, 3).

Bahrami-Rad, D., A. Becker, and J. Henrich (2021). "Tabulated nonsense? Testing the validity of the Ethnographic Atlas". *Economics Letters* 204, p. 109880 (cit. on p. 62).

Becker, A., B. Enke, and A. Falk (2020). "Ancient origins of the global variation in economic preferences". *AEA Papers and Proceedings*. Vol. 110, pp. 319–23 (cit. on p. 3).

Becker, S. O. et al. (2020). "Forced migration and human capital: Evidence from post-WWII population transfers". *American Economic Review* 110.5, pp. 1430–1463 (cit. on pp. 1, 3).

Berezkin, Y. (2015). "Folklore and Mythology Catalogue: Its Lay-out and Potential for Research". *The Retrospective Methods Network Newsletter* 10, pp. 58–70 (cit. on pp. 7, 26).

Berkes, F. (2017). *Sacred ecology*. Routledge (cit. on pp. 1, 3).

Berkes, F. and C. Folke (1998). "Linking social and ecological systems for resilience and sustainability". *Linking social and ecological systems: management practices and social mechanisms for building resilience* 1.4, p. 4 (cit. on p. 1).

Bhattacharya, S. et al. (2025). "Do Kids Teach Parents: Impact of Environmental Education on Household Behaviors". *Working Paper* (cit. on p. 4).

Bisin, A. and T. Verdier (2001). "The economics of cultural transmission and the dynamics of preferences". *Journal of Economic theory* 97.2, pp. 298–319 (cit. on p. 4).

– (2023). "Advances in the economic theory of cultural transmission". *Annual Review of Economics* 15 (cit. on p. 4).

Bondarenko, D. et al. (2005). "Ethnographic Atlas XXXI: Peoples of Easternmost Europe". *Ethnology* 44.3, pp. 261–289 (cit. on pp. 9, 62).

Boyd, R. and P. J. Richerson (1988). *Culture and the evolutionary process*. University of Chicago press (cit. on p. 4).

Buggle, J. C. and R. Durante (2021). "Climate risk, cooperation and the co-evolution of culture and institutions". *The Economic Journal* 131.637, pp. 1947–1987 (cit. on pp. 1, 3).

Bwambale, B., M. Muhumuza, and M. Nyeko (2018). "Traditional ecological knowledge and flood risk management: A preliminary case study of the Rwenzori". *Jàmbá: Journal of Disaster Risk Studies* 10.1, pp. 1–10 (cit. on p. 1).

Cavalli-Sforza, L. L. and M. W. Feldman (1981). *Cultural transmission and evolution: A quantitative approach*. 16. Princeton University Press (cit. on p. 4).

EM-DAT, CRED / UCLouvain (2024). *Number of deaths from disasters*. Processed by Our World in Data. EM-DAT, Centre for Research on the Epidemiology of Disasters (CRED) (cit. on p. 35).

Dewitte, E. (2024). "Economic Identities and the Historical Roots of Climate Change Attitudes in the U.S." *Working Paper* (cit. on p. 4).

ESS-ERIC (2018). *ESS3 - Integrated File, Edition 3.7*. Tech. rep. (cit. on p. 5).

Fernández, R. (2025). *Understanding Cultural Change*. Tech. rep. National Bureau of Economic Research (cit. on p. 3).

Giuliano, P. and N. Nunn (Jan. 2018). "Ancestral Characteristics of Modern Populations". *Economic History of Developing Regions* 33.1, pp. 1–17 (cit. on pp. 2, 5, 9, 55, 62).

– (2021). "Understanding cultural persistence and change". *The Review of Economic Studies* 88.4, pp. 1541–1581 (cit. on pp. 2, 3, 10, 30, 31, 33).

Gómez-Baggethun, E. and V. Reyes-García (2013). "Reinterpreting change in traditional ecological knowledge". *Human Ecology* 41.4, pp. 643–647 (cit. on pp. 1, 3).

Gordon, R. (2009). "Ethnologue: Languages of the world, 16th ed." *SIL International, Dallas* (cit. on pp. 9, 62).

Hansen, J. W., S. M. Marx, and E. U. Weber (2004). *The role of climate perceptions, expectations, and forecasts in farmer decision making: the Argentine Pampas and South Florida: Final Report of an IRI Seed Grant Project*. Tech. rep. (cit. on p. 5).

Harari, Y. N. (2014). “Sapiens: A brief history of humankind”. *Publish in agreement with The Deborah Harris Agency and the Grayhawk Agency* (cit. on p. 4).

Hofbauer, J. and K. Sigmund (1998). *Evolutionary games and population dynamics*. Cambridge university press (cit. on p. 4).

IVS (2020). *European Values Study and World Values Survey: Joint EVS/WVS 2017-2021 Dataset (Joint EVS/WVS)*. Tech. rep. (cit. on pp. 5, 33).

Lansing, J. S. and J. N. Kremer (1993). “Emergent properties of Balinese water temple networks: Coadaptation on a rugged fitness landscape”. *American Anthropologist* 95.1, pp. 97–114 (cit. on pp. 1, 3).

Laouenan, M. et al. (2022). “A cross-verified database of notable people, 3500BC-2018AD”. *Scientific Data* 9.1, p. 290 (cit. on p. 62).

Leiserowitz, A. (2006). “Climate Change Risk Perception and Policy Preferences: The Role of Affect, Imagery, and Values”. *Climate Change* 77, pp. 45–72 (cit. on p. 5).

Lowes, S. (2018). *Matrilineal Kinship and Spousal Cooperation: Evidence from the Matrilineal Belt*. Tech. rep. UCSD (cit. on p. 3).

Lowes, S. et al. (2017). “The evolution of culture and institutions: Evidence from the Kuba Kingdom”. *Econometrica* 85.4, pp. 1065–1091 (cit. on p. 1).

Luo, Y. and J. Zhao (2019). “Motivated Attention in Climate Change Perception and Action”. *Frontiers in Psychology* 10 (cit. on p. 5).

Maćkowiak, B., F. Matějka, and M. Wiederholt (2023). “Rational inattention: A review”. *Journal of Economic Literature* 61.1, pp. 226–273 (cit. on pp. 4, 14, 70).

Maćkowiak, B. and M. Wiederholt (2009). “Optimal sticky prices under rational inattention”. *American Economic Review* 99.3, pp. 769–803 (cit. on p. 70).

Mann, M. E. et al. (2008). “Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia”. *Proceedings of the National Academy of Sciences* 105.36, pp. 13252–13257 (cit. on p. 63).

Mann, M. E. et al. (2009). “Global Signatures and Dynamical Origins of the Little Ice Age and Medieval Climate Anomaly”. *Science* 326.5957, pp. 1256–1260 (cit. on pp. 2, 5, 9, 60, 62–64).

McCleary, R. M. and R. J. Barro (2006). “Religion and economy”. *Journal of Economic perspectives* 20.2, pp. 49–72 (cit. on p. 1).

Michalopoulos, S. and M. M. Xue (Jan. 2021). “Folklore*”. *The Quarterly Journal of Economics* 136.4, pp. 1993–2046 (cit. on pp. 3–5, 7, 26, 58).

Murdock, G. P. (1957). “World ethnographic sample”. *American Anthropologist* 59, pp. 664–687 (cit. on pp. 7, 9, 62).

National Academies of Sciences, Engineering, and Medicine and others (2016). "Committee on Extreme Weather Events and Climate Change Attribution, Board on Atmospheric Sciences and Climate, Division on Earth and Life Studies (2016) Attribution of extreme weather events in the context of climate change". *Attribution of extreme weather events in the context of climate change* (cit. on p. 11).

Nunn, N. (2009). "The importance of history for economic development". *Annu. Rev. Econ.* 1.1, pp. 65–92 (cit. on p. 1).

Osemeobo, G. J. (1994). "The role of folklore in environmental conservation: evidence from Edo state, Nigeria". *The International Journal of Sustainable Development & World Ecology* 1.1, pp. 48–55 (cit. on p. 4).

Poortinga, W. et al. (2018). "European attitudes to climate change and energy. Topline results from Round 8 of the European Social Survey" (cit. on p. 5).

Sanganyado, E., C. Teta, and B. Masiri (2018). "Impact of African traditional worldviews on climate change adaptation". *Integrated Environmental Assessment and Management* 14.2, pp. 189–193 (cit. on p. 4).

Shi, J. et al. (2016). "Knowledge as a driver of public perceptions about climate change reassessed". *Nature Climate Change* 6.8, pp. 759–762 (cit. on p. 5).

Shiller, R. J. (2017). "Narrative economics". *American economic review* 107.4, pp. 967–1004 (cit. on p. 4).

Sims, C. A. (2003). "Implications of rational inattention". *Journal of monetary Economics* 50.3, pp. 665–690 (cit. on p. 4).

Smith, T. M. et al. (2008). "Improvements to NOAA's historical merged land–ocean surface temperature analysis (1880–2006)". *Journal of climate* 21.10, pp. 2283–2296 (cit. on p. 11).

Speer, R., J. Chin, and C. Havasi (2017). "Conceptnet 5.5: An open multilingual graph of general knowledge". *Thirty-first AAAI conference on artificial intelligence* (cit. on pp. 7, 58).

Weber, E. (1997). *Perception and expectation of climate change: Precondition for economic and technological adaptation In Bazerman M., Messick D., Tenbrunsel A., and Wade-Benzoni K.(Eds.) Psychological perspectives to environmental and ethical issues in management (314–341)* (cit. on p. 5).

Williams, A., N. Nangia, and S. R. Bowman (2018). "A Broad-Coverage Challenge Corpus for Sentence Understanding through Inference". *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT)* (cit. on pp. 8, 60).

Young, H. P. (2015). "The evolution of social norms". *Annual Review of Economics* 7.1, pp. 359–387 (cit. on p. 4).

Appendix

A Additional Tables for Individual Level Analysis

A.1 Integrated Value Surveys - Pro-Environment Index

TABLE 3: Coefficients for the impact of average ancestral climatic anomalies on individual's self-reported attention to the environment

Dependent variable: Pro-environment index (IVS)				
	(1)	(2)	(3)	(4)
Avg. Anomalies	-2.975**	-2.980**	-3.119*	-3.087**
	(1.317)	(1.408)	(1.756)	(1.374)
Avg. Anomalies sq.	35.442**	35.555**	37.140*	34.611**
	(14.899)	(15.734)	(19.527)	(14.808)
Income	0.004***	0.004***	0.005***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)
Male	-0.010***	-0.006**	-0.002	-0.007**
	(0.003)	(0.003)	(0.003)	(0.003)
Age	-0.000	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Education level	0.010***	0.010***	0.015***	0.008***
	(0.001)	(0.001)	(0.001)	(0.001)
Occupation category Controls	Y	Y	Y	Y
Historical ethnic group characteristics	Y	Y	Y	Y
Historical topographic characteristics	Y	Y	Y	Y
Country-year Fixed effects	Y	Y	Y	Y
WVS Sample only	N	Y	N	N
Exc. Env. Confidence from PCA	N	N	Y	N
Exc. Env. vs Growth from PCA	N	N	N	Y
Mean of dep var	0.443	0.450	0.334	0.365
St. Dev. of dep var	0.230	0.234	0.310	0.237
Min value Avg. Anomalies	0.015	0.015	0.015	0.015
Max value Avg. Anomalies	0.093	0.093	0.093	0.093
R-sq	0.124	0.121	0.118	0.146
Adj. R-sq	0.123	0.120	0.117	0.145
N	219900	180293	238141	295507
*** p<0.01, **p<0.05, * p<0.1, + p<0.15				

Notes on Table 3: The unit of observation is an individual. The dependent variable is the individual's level of attention paid to environment. The dependent variable ranges between 0 and 1 and increases with the reported level environmental engagement and action. The variable is constructed by rescaling the first PCA from answers to pro-environment actions and support for environmental causes (in particular, their voluntary membership in environmental groups, confidence in and approval of environmental movements and willingness to choose environmental protection over growth). WVS Sample only excludes respondents from European Value Survey. Column 3 and 4 recalculate the PCA by removing individual responses to either confidence in and approval of environmental movement or choice between environment vs growth. Average anomalies refers to the average intensity of deviations from the typical climate conditions (*specific to each ancestral generation*) across generations.

TABLE 4: Coefficients for the impact of average ancestral climatic anomalies on individual's self-reported attention to the environment

Dependent variable: Pro-environment index (ESS)				
	(1)	(2)	(3)	(4)
Avg. Anomalies	-2.391***	-4.369***	-4.700***	-2.018***
	(0.841)	(1.030)	(0.963)	(0.773)
Avg. Anomalies sq.	17.951***	27.973***	29.219***	14.816***
	(6.511)	(6.252)	(4.952)	(4.496)
Income	0.002***	0.002**	0.002**	0.001*
	(0.001)	(0.001)	(0.001)	(0.000)
Male	-0.013***	-0.019***	-0.021***	-0.012***
	(0.004)	(0.004)	(0.004)	(0.002)
Age	-0.000***	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Educ	0.012***	0.008***	0.008***	0.013***
	(0.001)	(0.001)	(0.001)	(0.001)
Occupation category Controls	Y	Y	Y	Y
Historical ethnic group characteristics	Y	Y	Y	Y
Historical topographic characteristics	Y	Y	Y	Y
Country Fixed effects	Y	Y	Y	Y
PCA inc. renewable resources	N	Y	Y	N
PCA inc. climate change impact	N	N	Y	N
PCA w. climate change opinions only	N	N	N	Y
Mean of dep var	0.580	0.552	0.557	0.589
St. Dev. of dep var	0.136	0.148	0.143	0.172
Min value Avg. Anomalies	0.022	0.022	0.022	0.019
Max value Avg. Anomalies	0.119	0.119	0.119	0.119
R-sq	0.162	0.158	0.161	0.124
Adj. R-sq	0.160	0.155	0.156	0.123
N	23787	11696	9241	33055
*** p<0.01, **p<0.05, * p<0.1, + p<0.15				

Average anomalies sq. refers to the square of the average anomalies term. Average anomalies ranges between 0.015 and 0.093 within the sample. Historical ethnic group characteristics include measure of development such as agricultural intensity, complexity of settlement, level of political heirarchies, size of the local community and main source of subsistence. Historical topographic characteristics include controls for distance to the equator, distance to the closest coast and geographical Köppen climate classification for the location of the ethnic group obtained from Ethnographic Atlas. Standard errors are clustered at the ethnicity level.

A.2 European Social Survey - Pro-Environment Index

Notes on Table 4: The unit of observation is an individual. The dependent variable is the individual's level of attention paid to environment captured by the first component of the PCA carried over relevant actions detailed above. The dependent variable ranges between 0 and 1 and increases with the reported level environmental engagement and action. The variable is constructed

by rescaling the first PCA from answers to pro-environment actions and support for environmental causes. Average anomalies refers to the average intensity of deviations from the typical climate conditions (*specific to each ancestral generation*) across generations. Average anomalies sq. refers to the square of the average anomalies term. Average anomalies ranges between 0.019 and 0.119 within the sample. Historical ethnic group characteristics include measure of development such as agricultural intensity, complexity of settlement, level of political heirarchies, size of the local community and main source of subsistence. Historical topographic characteristics include controls for distance to the equator, distance to the closest coast and geographical Köppen climate classification for the location of the ethnic group obtained from Ethnographic Atlas. Standard errors are clustered at the ethnicity level.

B Tables for Group Level Analysis

B.1 Environmental Themes in Wikipedia Description of Ethnic Groups

TABLE 5: Coefficients for the impact of climatic shocks on occurrences of environment related themes in description of ethnic groups on Wikipedia

Dependent variable: Environment score for description on Wikipedia			
	(1)	(2)	(3)
Avg. Anomalies	-4.088**	-3.623*	-3.295*
	(2.013)	(1.955)	(1.913)
Avg. Anomalies sq.	42.798**	40.107**	38.738**
	(19.281)	(18.744)	(18.501)
Page Length	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)
Number of editions	0.001*	0.001*	0.001*
	(0.000)	(0.000)	(0.000)
Historical ethnic group characteristics	Y	Y	Y
Historical topographic characteristics	Y	Y	Y
Country Fixed effects	Y	Y	Y
Language pages included	N	N	Y
Geographical / Political Entity pages included	N	Y	Y
Mean of dep var	0.508	0.502	0.494
St. Dev. of dep var	0.212	0.214	0.218
Min value Avg. Anomalies	0.010	0.010	0.010
Max value Avg. Anomalies	0.097	0.097	0.097
R-sq	0.325	0.336	0.336
Adj. R-sq	0.161	0.180	0.188
N	1031	1081	1137
*** p<0.01, **p<0.05, * p<0.1, + p<0.15			

Notes on Table 5: The unit of observation is an ethnic group. Average anomalies refers to the average intensity of deviations from the typical climate conditions (*specific to each ancestral generation*) across generations. Average anomalies sq. refers to the square of the average variability term. Average anomalies ranges between 0.01 and 0.097 within the sample. Topographic

characteristics include controls for distance to the equator, distance to the closest coast and geographical Köppen climate classification for the location of the ethnic group. Page length is captured by the number of words in the Wikipedia biography for the ethnic group. Number of editions refers to the number of Wikipedia language editions the ethnic group has a biography in. Language pages refer to biographies that primarily captures information on the language spoken by the ethnic group instead of the group itself. Geographical / Political Entity pages refer to biographies that primarily capture information on the geographical or political entity that the ethnic group resided in instead of the group itself.

B.2 Environmental Themes in Descriptions of Ethnic Groups from Joshua Project

TABLE 6: Coefficients for the impact of climatic shocks on occurrences of environment related themes in description of ethnic groups in Joshua Project

Dependent variable: Environment score for description in Joshua Project				
	(1)	(2)	(3)	(4)
Avg. Anomalies	-3.734***	-3.365***	-2.871***	-1.390**
	(0.455)	(0.452)	(0.469)	(0.569)
Avg. Anomalies sq.	33.568***	30.165***	26.856***	13.163**
	(4.525)	(4.465)	(4.565)	(5.539)
Frontier - Y		-0.056***	-0.050***	-0.020***
		(0.005)	(0.005)	(0.006)
Indegenous Group - Y		0.058***	0.038***	0.049***
		(0.005)	(0.005)	(0.006)
Least Reached - Y		0.008	0.015*	0.014+
		(0.009)	(0.009)	(0.009)
Religion Fixed effects	N	Y	Y	Y
Topographic characteristics	N	N	Y	Y
Country Fixed effects	N	N	N	Y
Mean of dep var	0.236	0.238	0.238	0.238
St. Dev. of dep var	0.154	0.154	0.154	0.154
Min value Avg. Anomalies	0.010	0.010	0.010	0.010
Max value Avg. Anomalies	0.097	0.097	0.097	0.097
R-sq	0.008	0.065	0.080	0.149
Adj. R-sq	0.007	0.064	0.079	0.131
N	12898	11873	11873	11866
*** p<0.01, **p<0.05, * p<0.1, + p<0.15				

Notes on Table 6: The unit of observation is an ethnic group. Average anomalies refers to the average intensity of deviations from the typical climate conditions (*specific to each ancestral generation*) across generations. Average anomalies sq. refers to the square of the average anomalies term. Average anomalies ranges between 0.01 and 0.097 within the sample. Topographic characteristics include controls for distance to the equator, distance to the closest coast and geographical Köppen climate classification for the location of the ethnic group obtained from the Joshua Project website. Least Reached and Frontier categories capture the difficulty highlighted by the Joshua Project in terms of reaching the group in consideration and the success they had in converting individuals to Christianity. Standard errors are clustered at people group level.

C Model: additional details and omitted proofs

C.1 Preliminaries

Let $H = H^1(\mathbb{R}_+)$ the standard Sobolev space, which is a separable Hilbert space equipped with the scalar product:

$$\langle f, g \rangle = \int_{\mathbb{R}_+} fg + f'g'.$$

All derivatives from now on are understood in the Sobolev sense. Denote $\|\cdot\|$ the associated Sobolev norm. Recall that there is a continuous Sobolev embedding $H \hookrightarrow C_b^0(\mathbb{R}_+)$ ie there exists a constant α such that for all $f \in H$, $\|f\|_\infty \leq \alpha \|f\|$. Define the admissible set as:

$$\mathcal{F} = F_{\varepsilon,K} := \left\{ f \in H \;\middle|\; f(x) \geq \varepsilon \; \forall x, \|f\| \leq K \right\}.$$

for some $\varepsilon > 0$, $K > 0$. Although there are many possible ways to specify a consistent functional class, using a Sobolev space is a clean general way to make pointwise statements rigorous and stable under the Hilbert norm,[26] while assuming minimal regularity. The pointwise lower bound is a technical requirement to make derivations simpler, by guaranteeing that there is always strictly positive value of attention, but we can take ε arbitrarily small without loss of generality.

An convenient property of H that we will make ample use of to simplify proofs is that we can write expectations using Riesz representers. For any probability measure $q \in \Delta(\mathbb{R}_+)$, the expectation operator $f \mapsto \mathbb{E}_q[f]$ is a bounded linear functional on H; hence by Riesz's Theorem there is a unique representer $m_q \in H$ such that $\mathbb{E}_q[f] = \langle f, m_q \rangle, \forall f \in H$.[27] Furthermore by construction the representer is non-negative $m_q(z) \geq 0$ for all z,[28] and we can actually write explicitly:

$$m_q(z) = \int_0^\infty G(z,t) dq(t) \text{ where } G(z,t) := \frac{1}{2}\left(e^{-|z-t|} + e^{-|z+t|} \right)$$

In what follows, all functions considered are implicitly assume to belong to $\mathcal{F}$, furthermore all $\mathcal{F}$-valued integrals are understood as Bochner integrals.

[26] Pointwise evaluation is a continuous linear functional in H, unlike in e.g. L^2 where functions are only up to a.e. equality.
[27] H is a reproducing kernel Hilbert space so m_q is effectively a standard kernel embedding. This machinery is convenient for optimization, and is often used for kernel regression in the statistics and machine learning literature – which we will draw an explicit connection to later on.
[28] It is the weak solution of a positivity-preserving elliptic operator

C.2 Proof of Theorem 1

The proof of Theorem 1 proceeds in two steps. First, Proposition 1 derives the representation of dynamics for the average cultural type as a kernel stochastic gradient descent step, using local approximation methods. We then use the result to show that dynamics converge to the unique minimizer of the penalized fit object – which reduces the characterization to solving a convex functional optimization problem.

Proof of Proposition 1 By using Fubini-Tonelli in the replicator-mutator dynamics, we obtain a recursive characterisation of the average:

$$f_{g+1} = \frac{\int_{\mathcal{F}} \mathbb{E}_{f' \sim M_g(\cdot|f)}[f']\, \Phi_g(f)\, P_g(df)}{\int_{\mathcal{F}} \Phi_g(f)\, P_g(df)}.$$

By assumption $\mathbb{E}_{f' \sim M_g(\cdot|f)}[f'] = (1-\rho\lambda)f + \rho\lambda f_0$ so we can replace:

$$f_{g+1} = \frac{\int_{\mathcal{F}} \big((1-\rho\lambda)f + \rho\lambda f_0\big)\, \Phi_g(f)\, P_g(df)}{\int_{\mathcal{F}} \Phi_g(f)\, P_g(df)} = (1-\rho\lambda)\mathbb{E}_{f \sim P_g^{\text{sel}}}[f] + \rho\lambda f_0$$

where P_g^{sel} is the "selection only" (or post-selection, pre-mutation) transformation of P_g, i.e. $P_g^{\text{sel}}(B) := \frac{\int_B \Phi_g(f) P_g(df)}{\mathbb{E}_{P_g}[\Phi_g]}$ for any measurable set B. The "selection shift" $\mathbb{E}_{Q_g^{\text{sel}}}[f] - f_g$ can be rewritten into the so-called "Price decomposition":

$$\mathbb{E}_{P_g^{\text{sel}}}[f] - f_g = \frac{\text{Cov}_{P_g}(f, \Phi_g(f))}{\mathbb{E}_{P_g}[\Phi_g(f)]} = \frac{\mathbb{E}_{P_g}[(f-f_g)\Phi_g(f)]}{\mathbb{E}_{P_g}[\Phi_g(f)]}.$$

This is a canonical result in the evolution literature: see e.g. Page and Nowak (2002) for the equivalence between replicator–mutator and Price decompositions and its uses. Let Σ_g denote the covariance operator of P_g. Take the Fréchet expansion of Φ_g (which is C^2 by construction) around f_g:

$$\Phi_g(f) = \Phi_g(f_g) + \langle \nabla\Phi_g(f_g), f - f_g \rangle + R_g(f),$$

where $R_g(f)$ denotes the remainder. To discipline the remainder and ensure validity of the approximation locally around the mean, we use the following "variance control" lemma.

Lemma 1 (Variance control under weak selection and small mutations)**.** *Let $(Q_g)_{g\geq 0}$ evolve according to replicator–mutator dynamics. Assume the selection intensity satisfies $\tau = O(\rho)$ ("weak selection") and $\rho \ll 1$ ("small mutations"). Make the following regularity assumption: there exists $B < \infty$ such that*

$\sup_g \mathbb{E}_{Q_g}\|f\|^4 \le B$. *Then:*

$$\Sigma_g = \rho C + O(\rho^2)$$

uniformly in g, in trace norm.

This establishes that we can control the spread of the population distribution using the mutation kernel – this is essentially what allows us to discipline local approximations around the mean. The proof of this technical lemma is in Online Appendix G.2. Using the variance control lemma, we have $\mathbb{E}\|f - f_g\| \cdot |R_g(f)| = o(\rho^2)$. Then, using what is essentially an infinite-dimensional analogue of the multivariate delta method:

$$\begin{aligned}\mathrm{Cov}_{Q_g}(f, \Phi_g(f)) &= \mathbb{E}[(f - f_g)\,\Phi_g(f)] = \mathbb{E}[(f - f_g)\,\langle \nabla\Phi_g(f_g), f - f_g\rangle] + \mathbb{E}[(f - f_g)\,R_g(f)] \\ &= \Sigma_g\,\nabla\Phi_g(f_g) + o(\rho^2).\end{aligned}$$

Moreover, under concentration $\mathbb{E}[\Phi_g(f)] = \Phi_g(f_g) + o(1)$, hence dividing by $\mathbb{E}[\Phi_g]$ gives

$$\mathbb{E}_{Q_g^{\mathrm{sel}}}[f] = f_g + \Sigma_g\,\nabla \log \Phi_g(f_g) + o(\rho^2).$$

Next, observe that since $\Phi_g(f_g) = \exp(-\tau\mathcal{L}(\langle f_g, m_q\rangle|\hat{\bar{v}}_g))$:

$$\nabla \log \Phi_g(f_g) = -\tau\mathcal{L}'\big(\mathbb{E}_q[f] \,\big|\, \hat{\bar{v}}_g\big)m_q$$

Where the derivative of $\mathcal{L}$ is taken in its first argument. Plugging this into the law of motion for f_g, we get:

$$\begin{aligned}f_{g+1} &= (1 - \rho\lambda)\Big(f_g - \Sigma_g\tau\mathcal{L}'\big(\mathbb{E}_q[f] \,\big|\, \hat{\bar{v}}_g\big)m_q + o(\rho^2)\Big) + \rho\lambda f_0 \\ &= f_g - \rho\lambda(f_g - f_0) - \tau\mathcal{L}'\big(\mathbb{E}_q[f] \,\big|\, \hat{\bar{v}}_g\big)\Sigma_g m_q + o(\rho^2)\end{aligned}$$

Using the variance control lemma again, replace for $\Sigma_g = \rho I + o(\rho^2)$ and rearrange to get:

$$f_{g+1} = f_g - \rho\lambda(f_g - f_0) - \rho\tau\mathcal{L}'\big(\mathbb{E}_q[f] \,\big|\, \hat{\bar{v}}_g\big)m_q + o(\rho) = f_g - \rho\underbrace{\Big(\lambda(f_g - f_0) + \tau\mathcal{L}'\big(\mathbb{E}_q[f] \,\big|\, \hat{\bar{v}}_g\big)m_q\Big)}_{=\nabla J_g(f_g)} + o(\rho)$$

Minimizing the penalized fit objective The gradient step in Proposition 1 is in terms of the *empirical* penalized fit objective $J_g(f) = \mathcal{L}(\mathbb{E}_q[f]|\hat{\bar{v}}_g)$ – i.e. fit is evaluated against realized average values $\hat{\bar{v}}_g$, taken under the conditions drawn in generation g and with situation-specific noise. But, by construction, each $\hat{\bar{v}}_g$ is an unbiased estimator of $\bar{v}^\star := \mathbb{E}_q[v^\star]$. Each generational update hence amounts

to a stochastic gradient step on the *theoretical* penalized fit objective $J(f) = \mathcal{L}(\mathbb{E}_q[f]|\mathbb{E}_q[v^\star])$, where the loss is evaluated using a noisy but unbiased estimator of the theoretical target $\mathbb{E}_q[v^\star]$. This fits squarely into standard kernel regression theory, since J is strictly convex, standard arguments establish that the stochastic gradient step converges to its unique minimizer.

This reduces the problem of analyzing the outcome of evolutionary dynamics to a tractable functional optimization problem – we only need to characterize the minimizer of J. The following lemma characterizes the solution of this problem explicitly, and completes the proof of the theorem.

Lemma 2. *The penalized fit objective J has a unique minimizer which is of the form:*

$$f^* = f_0 + \frac{c}{\|m_q\|^2} m_q,$$

for some constant c that is the unique explicit solution of a cubic equation. As $\lambda \downarrow 0$, this converges to the element in the set $\{f \in \mathcal{F} | \mathbb{E}_q[f] = \mathbb{E}_q[v^\star]\}$ which is closest to f_0.

The proof of this lemma (which relies on a direct first-order approach), as well as complete details of the explicit solution are provided in Online Appendix G.2.

C.3 Proof of Theorem 2

To prove the first part of the theorem (core U-shape), define:

$$W(\theta) := \int_{\mathbb{R}_+} w(z) dp_\theta(z)$$
$$R(\theta) := \int_{\mathbb{R}_+} r(z) dp_\theta(z)$$

It is straightforward to observe that W and R inherit the properties of w and r almost directly. Taking for instance W, by performing a simple change of variable:

$$W(\theta) := \int_{\mathbb{R}_+} w(z) \frac{1}{\theta} d\overline{p}\left(\frac{z}{\theta}\right) = \int_{\mathbb{R}_+} w(\theta z) d\overline{p}(z)$$

Hence, assuming differentiability only for simplicity of exposition:

$$W'(\theta) = \int_{\mathbb{R}_+} z w'(\theta z) d\overline{p}(z) \leq 0$$
$$W''(\theta) = \int_{\mathbb{R}_+} z^2 w''(\theta z) d\overline{p}(z) \geq 0$$

and $W' \underset{\infty}{\longrightarrow} 0$ follows from dominated convergence. Similarly, we get $R \geq 0$, $R' \geq 0$, $R'' \geq 0$ and $R'(0) = 0$.

The second part combines three lemmas. First establish an elementary lemma which guarantees that B has an approximate U-shape if we can order bound the log slopes of V and S. Second, derive a lower bound on $|(\log V)'|$ from strong convexity of $v^\star$. Lastly, obtain uniform moment/Laplace bound on the similarity function term $|(\log S)'|$, using the explicit form of the kernel.

Lemma 3 (Sign preservation under bounded log-slope perturbations)**.** *Let $V, S : (\underline{\theta}, \overline{\theta}) \to (0, \infty)$ be C^1 and set $B(\theta \mid \tilde{\theta}) = V(\theta)S(\theta \mid \tilde{\theta})$. Fix $\theta^\dagger \in (\underline{\theta}, \overline{\theta})$ and $\delta > 0$ with $[\theta^\dagger - \delta, \theta^\dagger + \delta] \subset (\underline{\theta}, \overline{\theta})$. Suppose there exist constants $\kappa > 0$ and $c \geq 0$ such that for all $\tilde{\theta}$:*

$$\partial_\theta \log V(\theta) \leq -\kappa \quad \forall \theta \in [\underline{\theta}, \theta^\dagger - \delta], \qquad \partial_\theta \log V(\theta) \geq \kappa \quad \forall \theta \in [\theta^\dagger + \delta, \overline{\theta}],$$

and

$$\left|\partial_\theta \log S(\theta \mid \tilde{\theta})\right| \leq c \quad \forall (\theta, \tilde{\theta}) \in [\underline{\theta}, \overline{\theta}]^2.$$

If $\kappa > c$, then for every $\tilde{\theta}$,

$$\partial_\theta B(\theta \mid \tilde{\theta}) < 0 \ \text{ on } [\underline{\theta}, \theta^\dagger - \delta], \qquad \partial_\theta B(\theta \mid \tilde{\theta}) > 0 \ \text{ on } [\theta^\dagger + \delta, \overline{\theta}].$$

Proof. Since $B = VS$, $\partial_\theta \log B = \partial_\theta \log V + \partial_\theta \log S$. On $[\underline{\theta}, \theta^\dagger - \delta]$, $\partial_\theta \log V \leq -\kappa$ and $\partial_\theta \log S \geq -c$, hence $\partial_\theta \log B \leq -(\kappa - c) < 0$, so $\partial_\theta B < 0$. The right side is analogous. □

Lemma 4 (Strong convexity of $v^\star$ yields a quantitative log-slope for V)**.** *Assume $v^\star \in C^2$ satisfies $(v^\star)'' \geq m > 0$, and let $\mu_2 := \mathbb{E}[\bar{Z}^2] \in (0, \infty)$. Then $V(\theta) = \mathbb{E}[v^\star(\theta \bar{Z})]$ is C^2 and*

$$V''(\theta) \geq m\,\mu_2 \qquad \forall \theta \in (\underline{\theta}, \overline{\theta}).$$

If additionally V has a unique minimizer $\theta^\dagger \in (\underline{\theta}, \overline{\theta})$, then for any $\delta > 0$ with $[\theta^\dagger - \delta, \theta^\dagger + \delta] \subset (\underline{\theta}, \overline{\theta})$,

$$\partial_\theta \log V(\theta) \leq -\frac{m\mu_2\,\delta}{V_{\max}} \quad \forall \theta \in [\underline{\theta}, \theta^\dagger - \delta], \qquad \partial_\theta \log V(\theta) \geq \frac{m\mu_2\,\delta}{V_{\max}} \quad \forall \theta \in [\theta^\dagger + \delta, \overline{\theta}],$$

where $V_{\max} := \max\{V(\underline{\theta}), V(\overline{\theta})\}$.

Proof. Differentiate under the expectation:

$$V''(\theta) = \mathbb{E}\big[\bar{Z}^2 (v^\star)''(\theta \bar{Z})\big] \geq m\,\mathbb{E}[\bar{Z}^2] = m\mu_2.$$

Since $V'(\theta^\dagger) = 0$ and V' is increasing with slope at least $m\mu_2$, one has $V'(\theta) \le -m\mu_2\,\delta$ for $\theta \le \theta^\dagger - \delta$ and $V'(\theta) \ge m\mu_2\,\delta$ for $\theta \ge \theta^\dagger + \delta$. Finally, V is convex, hence its maximum over $[\underline{\theta}, \overline{\theta}]$ is attained at an endpoint, so $V(\theta) \le V_{\max}$, and dividing gives the stated bounds on $\partial_\theta \log V = V'/V$. □

Lemma 5 (Uniform log-Lipschitz bound for the similarity term)**.** *Let* $G(x,y) = \frac{1}{2}\big(e^{-|x-y|} + e^{-|x+y|}\big)$ *on* $\mathbb{R}_+^2$, *and define*

$$S(\theta \mid \tilde\theta) = \frac{\mathbb{E}[G(\theta\bar Z, \tilde\theta\bar T)]}{\mathbb{E}[G(\theta\bar Z, \theta\bar T)]}, \qquad \bar Z, \bar T \overset{iid}{\sim} \bar q.$$

Assume $\mu_1 := \mathbb{E}[\bar Z] < \infty$ *and* $L_* := \mathbb{E}[e^{-\overline{\theta}\bar Z}] > 0$. *Then* S *is* C^1 *in* θ *and*

$$\sup_{(\theta,\tilde\theta)\in[\underline{\theta},\overline{\theta}]^2} \big|\partial_\theta \log S(\theta \mid \tilde\theta)\big| \;\le\; \frac{3\,\mu_1}{L_*^2}.$$

Proof. Write $N(\theta,\tilde\theta) := \mathbb{E}[G(\theta\bar Z, \tilde\theta\bar T)]$ and $D(\theta) := \mathbb{E}[G(\theta\bar Z, \theta\bar T)]$, so $S = N/D$. A direct check gives $|\partial_x G(x,y)| \le G(x,y) \le 1$ and $|\partial_y G(x,y)| \le G(x,y) \le 1$ (a.e.). Hence, differentiating under the expectation,

$$|\partial_\theta N(\theta,\tilde\theta)| = \Big|\mathbb{E}\big[\bar Z\,\partial_x G(\theta\bar Z, \tilde\theta\bar T)\big]\Big| \le \mathbb{E}[\bar Z] = \mu_1,$$

and

$$|\partial_\theta D(\theta)| = \Big|\mathbb{E}\big[\bar Z\,\partial_x G(\theta\bar Z, \theta\bar T) + \bar T\,\partial_y G(\theta\bar Z, \theta\bar T)\big]\Big| \le \mathbb{E}[\bar Z + \bar T] = 2\mu_1.$$

Moreover, since $G(x,y) \ge e^{-(x+y)}$ for all $x, y \ge 0$,

$$N(\theta,\tilde\theta) \ge \mathbb{E}\big[e^{-(\theta\bar Z + \tilde\theta\bar T)}\big] = L(\theta)L(\tilde\theta) \ge L_*^2, \qquad D(\theta) \ge \mathbb{E}\big[e^{-(\theta\bar Z + \theta\bar T)}\big] = L(\theta)^2 \ge L_*^2,$$

where $L(\theta) := \mathbb{E}[e^{-\theta\bar Z}]$ and $L_* := L(\overline{\theta})$. Finally,

$$\big|\partial_\theta \log S\big| = \Big|\frac{\partial_\theta N}{N} - \frac{\partial_\theta D}{D}\Big| \le \frac{\mu_1}{L_*^2} + \frac{2\mu_1}{L_*^2} = \frac{3\mu_1}{L_*^2}.$$

□

Putting everything together, by Lemma 4, strong convexity of $v^\star$ implies that for the chosen δ,

$$\partial_\theta \log V(\theta) \le -\kappa \;\text{ on } [\underline{\theta}, \theta^\dagger - \delta], \qquad \partial_\theta \log V(\theta) \ge \kappa \;\text{ on } [\theta^\dagger + \delta, \overline{\theta}],$$

with $\kappa := \frac{m\mu_2\,\delta}{V_{\max}}$. By Lemma 5, one has the uniform bound

$$\left|\partial_\theta \log S(\theta \mid \tilde{\theta})\right| \le c \quad \forall(\theta, \tilde{\theta}) \in [\underline{\theta}, \overline{\theta}]^2, \qquad c := \frac{3\mu_1}{L_*^2}.$$

Applying Lemma 3 yields the claimed monotonicity of $B(\cdot \mid \tilde{\theta})$ outside $[\theta^\dagger - \delta, \theta^\dagger + \delta]$, uniformly in $\tilde{\theta}$.

D Heterogeneity Analysis

D.1 Sample split by historical economic development of ethnic groups

Notes on Table 7: The unit of observation is an individual. The dependent variable is the individual's level of attention paid to environment. The dependent variable ranges between 0 and 1 and increases with the reported level of attention. The variable is constructed by rescaling the answer to the prompt: On a scale of 1 to 6, how important is it for this individual to take care of the environment. Average anomalies refers to the average intensity of deviations from the typical climate conditions (*specific to each ancestral generation*) across generations. Average anomalies sq. refers to the square of the average anomalies term. Average anomalies ranges between 0.015 and 0.093 within the sample. Historical ethnic group characteristics include measure of development such as agricultural intensity, complexity of settlement, level of political heirarchies, size of the local community and main source of subsistence. Historical topographic characteristics include controls for distance to the equator, distance to the closest coast and geographical Köppen climate classification for the location of the ethnic group obtained from Ethnographic Atlas. Standard errors are clustered at the ethnicity level.

D.2 Sample split by current level of climatic instability faced by and migration status of individuals

Notes on Table 8: The unit of observation is an individual. The dependent variable is the individual's level of attention paid to environment. The dependent variable ranges between 0 and 1 and increases with the reported level of attention. The variable is constructed by rescaling the answer to the prompt: On a scale of 1 to 6, how important is it for this individual to take care of the environment. Average anomalies refers to the average intensity of deviations from the typical climate conditions (*specific to each ancestral generation*) across generations. Average anomalies sq. refers to the square of the average anomalies term. Average anomalies ranges between 0.015 and 0.093 within the sample. Historical ethnic group characteristics include measure of development such as agricultural intensity, complexity of settlement, level of political heirarchies, size of the local community and main source of subsistence. Historical topographic characteristics include controls for distance to the equator, distance to the closest coast and geographical Köppen climate classification for the location of the ethnic group obtained from Ethnographic Atlas. Recent Disaster levels are captured by the impact of natural disasters on GDP and population affected.

TABLE 7: Coefficients for the impact of average ancestral climatic anomalies on individual's self-reported attention to the environment

Dependent variable: Self reported measure of attention to environment								
	Hunter-Gatherer		Economic Complexity		Local Jurisdiction		Global Jurisdiction	
	Yes	No	Low	High	Low	High	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Avg. Anomalies	-29.650***	-9.741***	-10.085***	-7.634***	-20.416***	-4.190*	-25.868*	-8.989***
	(0.568)	(2.063)	(3.750)	(2.772)	(4.160)	(2.204)	(13.129)	(2.307)
Avg. Anomalies sq.	421.557***	110.145***	109.993**	84.733***	220.699***	44.355*	387.463**	103.220***
	(7.796)	(22.096)	(41.799)	(29.555)	(39.900)	(23.527)	(180.468)	(25.330)
Income	-0.002	-0.001*	0.001	-0.002***	-0.002***	-0.000	0.001	-0.002**
	(0.006)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Male	-0.009	-0.013***	-0.001	-0.016***	-0.017***	-0.008**	-0.006	-0.013***
	(0.007)	(0.003)	(0.003)	(0.004)	(0.005)	(0.003)	(0.005)	(0.003)
Age	0.001***	0.002***	0.001***	0.002***	0.002***	0.001***	0.001***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Education level	0.007***	0.006***	0.004*	0.007***	0.007***	0.006***	0.006***	0.006***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)
Occupation category Controls	Y	Y	Y	Y	Y	Y	Y	Y
Historical ethnic group characteristics	Y	Y	Y	Y	Y	Y	Y	Y
Historical topographic characteristics	Y	Y	Y	Y	Y	Y	Y	Y
Country-year Fixed effects	Y	Y	Y	Y	Y	Y	Y	Y
Mean of dep var	0.683	0.704	0.703	0.704	0.710	0.695	0.721	0.702
St. Dev. of dep var	0.256	0.250	0.252	0.249	0.243	0.258	0.246	0.250
Min value Avg. Anomalies	0.015	0.015	0.015	0.015	0.018	0.015	0.015	0.015
Max value Avg. Anomalies	0.084	0.093	0.080	0.093	0.084	0.093	0.074	0.093
R-sq	0.061	0.115	0.075	0.127	0.118	0.114	0.097	0.117
Adj. R-sq	0.054	0.115	0.072	0.126	0.117	0.112	0.092	0.116
N	3925	134140	25801	112262	76519	61543	12076	125981
*** p<0.01, **p<0.05, * p<0.1, + p<0.15								

Similarity of Environments captures how similar are the stock of contemporaneous climate anomalies to the average ancestral climatic anomalies faced by the ancestors. Children of immigrants are respondents whose either parent was an immigrant to the country where the individual is currently located. Standard errors are clustered at the ethnicity level.

TABLE 8: Coefficients for the impact of average ancestral climatic anomalies on individual's self-reported attention to the environment

Dependent variable: Self reported measure of attention to environment						
	Recent Disaster Levels		Similarity in Environments		Children of Immigrants	
	Low	High	Low	High	Yes	No
	(1)	(2)	(3)	(4)	(5)	(6)
Avg. Anomalies	-4.317**	-12.341***	-10.368***	-14.070***	-9.291***	-8.662***
	(2.163)	(2.566)	(2.360)	(3.119)	(3.492)	(2.077)
Avg. Anomalies sq.	50.572**	139.489***	123.843***	146.427***	99.097***	98.057***
	(24.278)	(27.289)	(25.435)	(31.526)	(36.977)	(22.082)
Income	-0.003***	-0.000	-0.002**	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)
Male	-0.020***	-0.008***	-0.015***	-0.008**	-0.019***	-0.013***
	(0.004)	(0.003)	(0.004)	(0.003)	(0.005)	(0.003)
Age	0.002***	0.001***	0.002***	0.001***	0.002***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Education level	0.007***	0.006***	0.006***	0.006***	0.005***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Occupation category Controls	Y	Y	Y	Y	Y	Y
Historical ethnic group characteristics	Y	Y	Y	Y	Y	Y
Historical topographic characteristics	Y	Y	Y	Y	Y	Y
Country-year Fixed effects	Y	Y	Y	Y	Y	Y
Mean of dep var	0.692	0.712	0.698	0.716	0.708	0.707
St. Dev. of dep var	0.248	0.251	0.250	0.248	0.252	0.251
Min value Avg. Anomalies	0.021	0.015	0.015	0.018	0.015	0.015
Max value Avg. Anomalies	0.093	0.093	0.093	0.093	0.085	0.093
R-sq	0.120	0.110	0.120	0.101	0.117	0.113
Adj. R-sq	0.119	0.109	0.119	0.100	0.107	0.112
N	59581	78486	97432	40632	10966	111600
*** p<0.01, **p<0.05, * p<0.1, + p<0.15						

Online Appendix

E Data: more details on construction and descriptions

E.1 Integrated Value Survey

The Integrated Values Surveys (IVS) is a harmonized global dataset that integrates and standardizes responses from the World Values Survey (WVS) and the European Values Study (EVS), and is designed to facilitate cross-country and within-country analyses over time of human values, beliefs, and preferences using consistent measures. WVS covers nationally representative samples from 104 countries between 1981 and 2022. Similarly, EVS covers nationally representative samples from 49 European nations between 1981 and 2022. Individuals in each survey are asked questions along many thematic categories (perceptions of life, environment, work, family, politics, society, religion and national identity). Each survey capture respondent level information on demographic characteristics such as age, education level, occupational category, income level (ten standardized categories) and gender, which we use as controls. Respondents also report the primary language that they speak at home, which allows us to match them to the corresponding ethnic ancestors from the Ancestral Characteristics database (Giuliano and Nunn, 2018).

For our lead question of interest, *"How important is it to you to take care of the environment?"*, we have a total of 157,142 respondents from 78 countries in waves 5 and 6 (2005-2014). In marginal cases (7% of the total sample), where we do not have information on the language spoken at home, we exploit the language of the interview as a proxy for the language spoken at home. In cases where an individual reports multiple languages spoken at home or the same language is spoken by multiple ethnic groups, we link the individual to all the potential historical ethnic groups and assign equal weight to them. Our sample of interest for main variable of interest from World Value Surveys contains 202 different language / ethnic groups. Less than 5% of the individuals in the main sample of consideration have multiple ethnic groups associated to them.

For the PCA exercises conducted to construct alternate proxies capturing pro-environmental sentiment and actions, we rely on measures capturing (i) voluntary membership in environmentally related organizations (WVS Waves 1–7; EVS Waves 1–5), (ii) approval of and confidence in country-specific environmental and nature movements (WVS Waves 2–7; EVS Waves 2, 4, and 5), and (iii) willingness to prioritize environmental protection over economic growth (WVS Waves 3–7; EVS Waves 3 and 5) – see Table 9. All variables are normalized so that higher numbers indicate greater subjective value, and the first

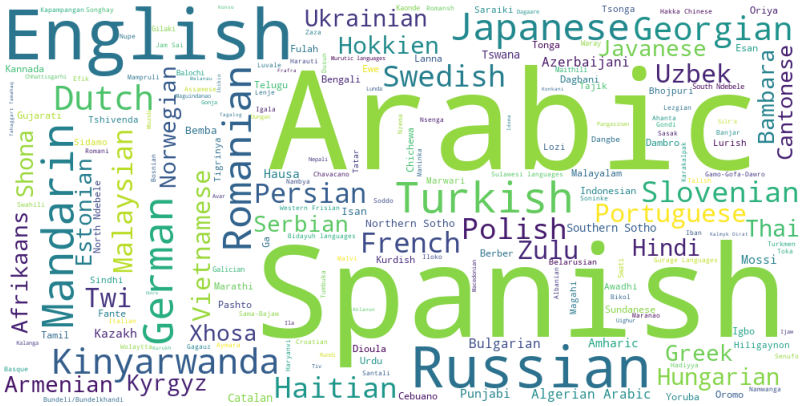
TABLE 9: Variable source (Wave and Study) - stated environmental preferences

Variable	**WVS**							**EVS**				
Wave	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(1)	(2)	(3)	(4)	(5)
Starting year	1981	1989	1994	1999	2005	2010	2017	1981	1990	1999	2008	2017
Membership in Env. Groups	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Voluntary Work for Env. Groups		✓		✓				✓	✓	✓	✓	
Participated in and Provided Money for Env. Actions			✓			✓						
Env. Participation	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Approval: Ecology movement or nature protection		✓							✓			
Confidence: Environmental protection movement			✓	✓	✓	✓	✓				✓	✓
Env. Support		✓	✓	✓	✓	✓	✓		✓		✓	✓
Env. over Growth			✓	✓	✓	✓	✓			✓		✓

Note: Source for each stated environment preference measure described in section 2. Environment is abbreviated as Env.

PCA component is scaled to lie between 0 and 1. Depending on the specification[29], information on the constructed proxy is available for approximately 220,000 to 300,000 respondents across 67 to 93 countries.

FIGURE 6: Self-reported mother tongue (language spoken at home)



Note: Word cloud of self-reported languages spoken at home which allow us to match individuals to the historical ethnic groups. Word size is weighted by the proportion of individuals in the sample of WVS that reported the language as their primary language and answered the environmental question of interest.

E.2 European Social Survey

The European Social Survey is a cross-national repeated social survey that measures public attitudes, beliefs, values, and behaviors across a broad set of countries, primarily in Europe. Each round includes

[29] Our main results use an index constructed as the first principal component from a PCA over individual responses to all the questions described above. To assess the sensitivity of our findings to the specific inputs used in the PCA, we construct alternative indices based on narrower subsets of variables. In particular, we retain voluntary environmental group membership (which is available across most survey waves) and, in turn, combine it with either (i) willingness to prioritize environmental conservation over economic growth or (ii) approval of and confidence in national environmental movements, but not both simultaneously.

rotating modules focused on specific thematic domains. The 8th round contains a dedicated climate module that elicits respondents' views and behaviors related to climate change, energy use, and environmental policy.

Using responses from this module, we conduct four PCA exercises:

1. In the first PCA, we use responses to questions capturing: (i) confidence in one's ability to reduce energy consumption, (ii) frequency of actions taken to reduce energy use, (iii) concern about national dependence on fossil fuels, (iv) prior consideration of climate change, and (v) support for climate-related policies, including increasing taxes on fossil fuels, subsidizing renewable energy, and banning the sale of the least energy-efficient household appliances.[30]

2. In the second PCA exercise, we augment this set with questions, asked to a smaller subsample, regarding respondents' views on the appropriate share of energy that should be generated from renewable sources such as hydroelectric, nuclear, solar, wind, and bioenergy.

3. In the third PCA, we further expand the variable set by including questions, asked to an even smaller subsample, on the likelihood of purchasing highly energy-efficient household appliances and beliefs about whether climate change will have negative global consequences.

4. In the fourth PCA, we solely rely on beliefs related to climate-change. We use responses to questions on prior consideration of climate change, belief in its existence, concern about its impacts, and attribution of climate change primarily to human activity.

To conduct the PCA, as with IVS, we first normalize all variables so that higher numbers indicate greater subjective value, and the first PCA components which we use as our indices is scaled to lie between 0 and 1. ESS also provides rich demographic information on respondents. We use information on the language spoken at home to link individuals to their ethnic ancestry through the language–ethnicity mappings available in the ancestral characteristics database. Additional control variables, which we use on our specifications, include age, gender, education level, income deciles (standardized within country), and occupation categories based on ISCO-08 classifications.

[30]The exact survey questions include: (1) confidence in using less energy; (2) frequency of actions to reduce energy use; (3) worry about national dependence on fossil fuels; (4) extent of prior thought about climate change; (5) belief about whether climate change is driven by human activity (coded higher when attributed mainly to human causes); (6) worry about climate change; (7) support for increasing fossil fuel taxes; (8) support for subsidizing renewable energy; and (9) support for banning the least energy-efficient household appliances.

E.3 Folklore catalog

Berezkin's original catalog consists of motifs related to the mythology, folklore and oral traditions for 958 groups worldwide. After parsing over 6,239 books and journal articles (documenting oral traditions) from 4,041 authors edited by 4,932 publishing houses in 32 different languages, Yuri Berezkin categorized 2,564 motifs and linked them to the ethnolinguistic groups. As per the original catalog, a motif reflected a combination of images, episodes, or structural elements found in two or more texts, including sacred and profane ones. The median group in Berezkin's catalog has 62 motifs. Each motif is accompanied with a title and a short description of an image or an episode in the group's oral tradition.

We used two different methods to classify environmental themes in folklore.

1. **ConceptNet Classification:** ConceptNet (Speer, Chin, and Havasi, 2017) is a large, multilingual semantic knowledge graph that encodes general human common-sense knowledge in the form of structured relationships between concepts. It represents information as nodes (words or short phrases) connected by labeled edges that describe relations such as IsA, PartOf, UsedFor, CapableOf, Causes, and RelatedTo. The knowledge in ConceptNet is aggregated from a variety of sources, including crowdsourced contributions (notably the Open Mind Common Sense project), lexical resources, and structured data repositories. Each connection is weighted to reflect confidence and frequency. Designed to support common-sense reasoning in natural language processing systems, ConceptNet provides a structured representation of everyday knowledge that can be used for semantic similarity, reasoning, and knowledge-augmented modeling. As mentioned before, we construct a ConceptNet list using the seed words: weather, climate, temperature, environment and natural disaster. We then retrieve all associated words and phrases for each seed that have a relevance score of greater than 1 and then manually verify the relevance of each item. We finally end up with a list of 228 words which we employ to classify motifs as environmental vs non-environmental related using a dictionary based approach like Michalopoulos and Xue (2021), i.e. we check the occurrence of the words from this list in the description of the motifs to classify whether a motif is environmentally oriented or not.[31]

2. **Classification procedure based on syntactic structure.** To reduce dimensionality, we extract subject–verb–object (SVO) triplets from each motif using spaCy's dependency parser. Subject–verb–object (SVO) triplets are a structured way of representing the basic relational content of text by identifying who is performing an action, what the action is, and who or what is affected by

[31] For more details on ConceptNet, see (Speer, Chin, and Havasi, 2017).

it. This structure captures the core meaning of a sentence in a compact and interpretable form.[32] Each motif text is first parsed to identify its grammatical structure. For every verb within the motif, we then identify the corresponding subject and object based on their syntactic relationships to the verb, expanding these arguments to full noun phrases when relevant. We retain the set of unique subjects and objects obtained from the SVO triplets, which yields a substantially smaller and more interpretable vocabulary. We manually classify these terms as environmental or non-environmental (e.g., whether they refer to plants, animals, or other natural entities). Using this classification, we label a motif as environmentally related if it contains at least one subject or object classified as environmental. We use this alternative procedure to construct an additional measure of environmental focus in folklore, which serves as a robustness check for our main classification approach.

E.4 The Joshua Project

We construct an alternative group-level measure of environmental attention using textual descriptions from the Joshua Project. Similar in spirit to the folklore-based measure, this approach treats written ethnographic descriptions as reflecting the cumulative stock of experiences, beliefs, and knowledge embodied within an ethnic group over time.

The Joshua Project dataset aggregates information from a wide range of secondary sources - including global, regional, and national researchers, mission organizations, census materials, and other published records - which are integrated to provide a consistent global picture of ethnic groups, including their geographic distribution, population estimates, and descriptive narratives of social organization and cultural practices. While the project originated in a religious and missionary context, its dataset draws on diverse ethnographic, linguistic, and historical sources and has been used in academic research as a practical resource for harmonized ethnic group identifiers and descriptive cultural information, particularly in settings where official census data on ethnicity are sparse or inconsistent.

We begin by downloading the full Joshua Project dataset and identifying all ethnic groups for which descriptive information is available. For each group, we scrape textual content from the project's official website, which provides narrative descriptions of ethnic groups in a relatively standardized format. These

[32] By reducing unstructured text to SVO triplets, complex narratives can be transformed into standardized relational units that are comparable across documents, speakers, or contexts. SVO triplets are widely used in text analysis because they preserve directional and role-based information that is lost in bag-of-words or topic-based approaches. Unlike word counts or embeddings alone, SVO representations explicitly encode agency and responsibility by distinguishing actors from recipients of actions, making them particularly useful for analyzing narratives.

descriptions are generally organized into key sections such as a brief introduction, accounts of daily life and subsistence, and discussions of beliefs and worldviews.[33]

To understand how environmental attention is expressed in these texts, we manually processed a random subset of 100 group descriptions with the help of trained research assistants. This exercise was used to identify recurring thematic structures and linguistic patterns through which environmental conditions, ecological constraints, and nature-related beliefs are discussed. In particular, this manual review informed our understanding of how environmental attention appears in narratives related to subsistence activities, exposure to climate and weather shocks, natural resource dependence, environmental degradation, and spirituality linked to nature.

Guided by this analysis, we then apply natural language processing tools to the full corpus of Joshua Project texts. Specifically, we use a BART-based zero-shot text classification approach, fine-tuned over MultiNLI (Williams, Nangia, and Bowman, 2018).[34,35] We classify each group's textual description along several predefined dimensions capturing environmental relevance: (i) references to environment and ecology, (ii) discussion of climate or weather shocks as central features of group life, (iii) embedding of natural resources and ecological constraints in subsistence activities, (iv) resistance to or concern about environmental degradation and threats, and (v) expressions of spirituality or belief systems explicitly linked to nature. We aggregate these classification scores to construct a composite index, which we normalize to lie between 0 and 1 and interpret as our second group-level measure of attention to environmental issues.

The Joshua Project database also provides information on the geographic location of ethnic groups, which allows us to match groups to surface grid cells in the temperature anomaly dataset constructed by Mann et al. (2009). Additional control variables include indicators for whether a group is classified as indigenous,

[33]The website also includes a separate section focused on religious needs, which is explicitly oriented toward Christian missionary activity. Because this content reflects an external normative framing rather than endogenous cultural beliefs, we exclude it from our analysis.

[34]BART is a transformer-based sequence-to-sequence model that can be adapted for classification tasks using a zero-shot framework grounded in natural language inference. Specifically, we use a version of BART fine-tuned on the MultiNLI corpus, which trains the model to determine whether a premise sentence entails, contradicts, or is neutral with respect to a hypothesis. In a zero-shot setting, classification is implemented by converting each candidate thematic category into a short hypothesis statement and scoring the degree to which the input text entails that hypothesis. The model assigns the category with the highest entailment probability, without any task-specific retraining or fine-tuning on our data.

[35]The use of a MultiNLI-trained model is particularly well-suited for descriptive classifications of ethnographic and encyclopedic text—such as content drawn from the Joshua Project website or Wikipedia—because MultiNLI exposes the model to a wide range of genres (including fiction, government documents, spoken dialogue, and written nonfiction) during training. This broad genre coverage improves the model's ability to generalize to diverse narrative and descriptive styles. Importantly, the entailment-based formulation aligns closely with our conceptual objective: we are not asking the model to learn dataset-specific labels, but rather to evaluate whether a given description semantically implies the presence of particular thematic content. This approach allows for flexible, transparent, and reproducible categorization of text into conceptually meaningful themes while maintaining interpretability and eliminating researcher-driven retraining choices.

measures of religious composition (which also proxy for the extent and quality of available descriptive information), and classifications identifying frontier or least-reached groups. Our final sample includes 12,989 ethnic groups distributed globally.

E.5 Wikipedia

To construct the final group-level measure of environmental attention we rely on textual descriptions scraped from English Wikipedia pages on ethnic groups. Similar to the Joshua Project–based measure, this approach treats encyclopedic written records as reflecting a cumulative stock of knowledge, experiences, and cultural narratives associated with a group, while recognizing that such descriptions are curated and externally authored.

We first manually compile a list of English Wikipedia pages for the corresponding ethnic groups in the ancestral characteristics database. We then scrape the full text of the corresponding Wikipedia page which typically provide narrative content summarizing a group's livelihood strategies, history, social organization, and interactions with external actors. In a small number of cases, information on the ethnic group itself is unavailable; in such instances, we rely on pages describing the language spoken by the group or the political or geographic entity with which the group is associated, which often contain brief ethnographic descriptions of the group.

To identify how environmental attention is expressed in these narratives, we first manually reviewed a random subset of 50 Wikipedia pages with the assistance of trained research assistants. This manual review was used to identify recurring themes, narrative structures, and linguistic patterns through which environmental conditions, ecological knowledge, and nature-centered beliefs are discussed. Based on this exercise, we defined four thematic dimensions capturing environment-centered cultural narratives: (i) nature's central to group identity and livelihood; (ii) indigenous religious or spiritual systems that link ancestors or spirits directly to specific natural entities; (iii) detailed traditional ecological knowledge related to desert, marine, forest, or island environments; and (iv) narratives of cultural resistance to external actors, including corporations, missionaries, or states, that are perceived as a threat to environmental factors.

Guided by these themes, we apply a natural language classification approach to the full corpus of Wikipedia texts using a BART-based zero-shot classification framework, fine-tuned over MultiNLI. In this zero-shot setting, the model evaluates each group's textual description against the four predefined themes without requiring task-specific retraining, allowing for flexible and interpretable classification of ethnographic narratives. For each group, we aggregate classification scores across the four dimensions to construct a composite index, which we normalize to lie between 0 and 1 and interpret as a measure of environmental relevance embedded in encyclopedic descriptions of ethnic groups.

In addition, following Laouenan et al. (2022), we collect data on the total number of Wikipedia language editions in which a biography or description of the ethnic group exists, which we use as a proxy for the breadth of documentation available for each group. Finally, we match the resulting environmental attention scores back to the ancestral characteristics database, which allows us to link ethnic groups to surface grid cells in the temperature anomaly dataset constructed by Mann et al. (2009).

E.6 Ancestral Ethnic Groups

Giuliano and Nunn (2018) database on Ancestral Characteristics of Modern Populations combines the pre-industrial characteristics of ethnic groups available in Murdock's Ethnographic Atlas, Bondarenko et al. (2005) dataset on People of Easternmost Europe and Murdock (1957) World Ethnographic sample. with the 1265 groups available from the Ethnographic Atlas. Giuliano and Nunn (2018) manually match over 7,000 different dialects and languages from Gordon's (2009) Ethnologue: Languages of the World to the ethnic groups from the ethnographic data sources.

Information on the pre-industrial characteristics of the ethnic groups in the ethnic atlases has been coded for the earliest period for which satisfactory ethnographic data is available or can be reconstructed. In total, 23 ethnicities are observed during the seventeenth century or earlier, 16 during the eighteenth century, 310 during the nineteenth century, 876 between 1900 and 1950 and 31 after 1950. See Bahrami-Rad, A. Becker, and Henrich (2021) for the relevance of the *Ethnographic Atlas* in capturing traditional practices. The other two samples that Giuliano and Nunn (2018) use, Bondarenko et al. (2005) and Murdock's (1957) World Ethnographic Sample, add 27 more ethnic groups to the 1265 groups available from the Ethnographic Atlas. The availability of the latitude and longitude of the ethnic groups location in the ethnographic datasets allows us to match the ethnographic groups to the surface grids in the Mann et al. (2009) temperature anomalies database.

Historical pre-industrial characteristics of ethnic group controls that we inclue in our regressions include:

- Primary mode of subsistence; Fishing, Hunting, Gathering, Animal Husbandry and Pastoralism and Agriculture.
- complexity of settlement: values go from 1 to 8. 1: fully-nomadic, 2: semi-nomadic, 3: semisedentary, 4: compact but impermenant settlements, 5: neighborhoods of dispersed family homesteads, 6: separated hamlets forming a single community, 7: compact and relatively permanent, 8: complex settlements.
- Size of local community: Ranging from 1 to 8 where 1: fewer than 50, 2: 50-99, 3: 100-199, 4: 200-399, 5: 400-1,000, 6: 1,000-4,999, 7: 5,000-50,000 and 8: $> 50,000$.

- Intensity of agriculture: ranging from 1 to 6 where 1 corresponds to no agriculture, 2 to casual agriculture, ... , 6 to intensive irrigated agriculture.
- Level of local and global jurisdictional development: ranging from 1 to 5, denoting the number of levels within a jurisdiction.
- Ethnicity level geographical variables such as distance to closest coast and equator and Köppen climate classification of the spatial grid the ethnic group was historically located in.

E.7 Climate conditions

We construct measures of ancestral climatic exposure using the climate field reconstruction (CFR) data from Mann et al. (2009). The CFR approach reconstructs past surface temperature fields by combining instrumental temperature records with a large network of climate proxy data, allowing for spatially resolved estimates of temperature variability over long historical horizons. The reconstruction relies on a multiproxy dataset with broad global coverage, comprising 1,036 tree-ring series, 32 ice core series, 15 marine coral series, 19 documentary series, 14 speleothem series, 19 lacustrine (lake) sediment series, and 3 marine sediment series (Mann et al., 2008). These proxy records are statistically calibrated against observed temperature data over the instrumental period and then used to infer historical temperature patterns at the grid-cell level. Details on the spatial distribution of climate conditions are provided in Figure 7.

F Robustness Checks Results

The following section presents detailed results for the robustness checks in Section 4.

F.1 Choice of language sample and Contemporaneous Region Confounds

Notes on Table 10: The unit of observation is an individual. The dependent variable is the individual's level of attention paid to environment. The dependent variable ranges between 0 and 1 and increases with the reported level of attention. The variable is constructed by rescaling the answer to the prompt: On a scale of 1 to 6, how important is it for this individual to take care of the environment. Average anomalies refers to the average intensity of deviations from the typical climate conditions (*specific to each ancestral generation*) across generations. Average anomalies sq. refers to the square of the average anomalies term. Average anomalies ranges between 0.015 and 0.093 within the sample. Region-year fixed effects control for geography specific year fixed effects at a more granular level than the country. Historical ethnic group characteristics include measure of development such as agricultural intensity, complexity of settlement, level of political heirarchies, size of the local community and main source of subsistence. Historical topographic characteristics include controls for distance to the equator, distance to

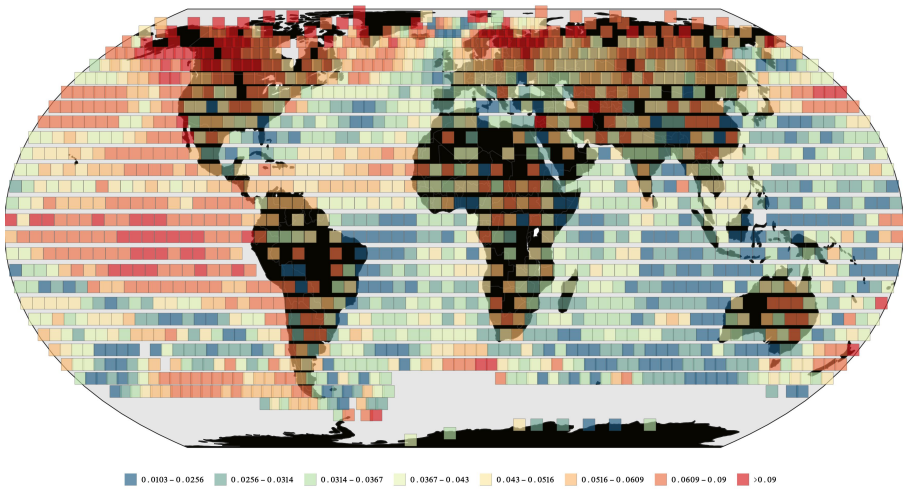


FIGURE 7: Spatial distribution of Average Ancestral Climatic Variability

Note: Distribution of average ancestral climatic variability constructed from Mann et al. (2009). The variable reports the average intensity of deviations from a defined range of normal temperatures specific to a generation. Generation lifespan is assumed to be 20 years and the normal range of temperature are the temperatures within the 20^{th} and 80^{th} percentile of temperatures within the lifespan of a generation. Variable is constructed at the 5 by 5 degree gird level. For more details refer to Section 2.

the closest coast and geographical Köppen climate classification for the location of the ethnic group obtained from Ethnographic Atlas. Lang. at home only refers to the case where we remove all individuals for whom we proxied the language spoken at home by the language of interview. Standard errors are clustered at the ethnicity level.

F.2 Construction of Ancestral Climatic Anomalies

Notes on Table 11: The unit of observation is an individual. The dependent variable is the individual's level of attention paid to environment. The dependent variable ranges between 0 and 1 and increases with the reported level of attention. The variable is constructed by rescaling the answer to the prompt: On a scale of 1 to 6, how important is it for this individual to take care of the environment. Avg. refers to the average intensity of deviations from the typical climate conditions (*specific to each ancestral generation*) across generations. Avg. sq. refers to the square of the average anomalies term. Every regression includes controls for demographic characteristics, ethnicity level group and topographic characteristics and country-year fixed effects. Each column varies the lifespan of ancestral generation in our method of construction of the average anomalies term. Each row varies the choice of the typical range. Dev. from 20-80 (our main specification) refers to the deviations from the cut points of the range of typical temperatures for each generation, where typical temperatures lie between 20^{th} and 80^{th} percentile of

TABLE 10: Robustness checks: Coefficients for the impact of average ancestral climatic anomalies on individual's self-reported attention to the environment

Dependent variable: Self reported measure of attention to environment						
	(1)	(2)	(3)	(4)	(5)	(6)
Avg. Anomalies	-9.877***	-9.243***	-8.903***	-14.751***	-9.897***	-6.649***
	(2.003)	(1.988)	(2.206)	(2.647)	(2.011)	(2.023)
Avg. Anomalies sq.	111.767***	107.503***	103.956***	163.717***	111.862***	71.929***
	(21.346)	(21.557)	(25.184)	(27.890)	(21.505)	(22.402)
Income	-0.001*	-0.001+	-0.002+	-0.002***	-0.001+	-0.001+
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Male	-0.013***	-0.012***	-0.014***	-0.020***	-0.013***	-0.014***
	(0.003)	(0.004)	(0.004)	(0.005)	(0.003)	(0.003)
Age	0.002***	0.002***	0.002***	0.002***	0.001***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Education level	0.006***	0.006***	0.007***	0.007***	0.006***	0.007***
	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
Occupation category Controls	Y	Y	Y	Y	Y	Y
Historical ethnic group characteristics	Y	Y	Y	Y	Y	Y
Historical topographic characteristics	Y	Y	Y	Y	Y	Y
Country-year Fixed effects	Y	Y	Y	Y	Y	Y
Region-year Fixed effects	N	N	N	N	N	Y
Sample restrictions						
English speakers excluded	N	Y	Y	N	N	N
Spanish speakers excluded	N	Y	Y	N	N	N
Arabic speakers excluded	N	N	Y	N	N	N
Weights at ethnicity level	N	N	N	Y	N	N
Lang. at home only	N	N	N	N	Y	N
Mean of dep var	0.704	0.700	0.689	0.704	0.703	0.705
St. Dev. of dep var	0.250	0.251	0.249	0.250	0.251	0.250
Min value Avg. Anomalies	0.015	0.015	0.015	0.015	0.015	0.015
Max value Avg. Anomalies	0.093	0.093	0.093	0.093	0.093	0.093
R-sq	0.113	0.122	0.118	0.117	0.114	0.160
Adj. R-sq	0.113	0.121	0.117	0.117	0.113	0.151
N	138067	110484	93184	1.89e+09	128532	135228
*** p<0.01, **p<0.05, * p<0.1, + p<0.15						

generation specific temperatures. Dev. from 10-90 and 30-70 are defined analogously. Dev. (sq.) from 20-80 instead squares the deviations from the cut points before taking the average. Std. Dev. instead of defining the deviations from the typical range just takes takes the Standard deviation of temperatures within a generation as a measure of climatic shocks. ***p<0.01, **p<0.05, *p<0.1, + p<0.15

F.3 Other characteristics of temperature distribution and model fit

Notes on Table 12: The unit of observation is an individual. The dependent variable is the individual's level of attention paid to environment. The dependent variable ranges between 0 and 1 and increases with the reported level of attention. The variable is constructed by rescaling the answer to the prompt: On a scale of 1 to 6, how important is it for this individual to take care of

TABLE 11: Coefficient on average ancestral climatic anomalies - Robustness by method of variable construction

Dependent variable: Self reported measure of attention to environment					
Lifespan of generation:		20 years	30 years	40 years	50 years
Dev. from 10-90	Avg.	-14.742***	-13.926***	-10.453***	-5.884***
		(4.739)	(2.433)	(2.067)	(1.770)
	Avg. sq.	316.651***	235.773***	143.257***	64.708***
		(103.216)	(40.585)	(25.865)	(18.135)
Dev. from 20-80	Avg.	-9.877***	-5.064***	-5.618***	-4.423***
		(2.003)	(1.770)	(1.159)	(1.256)
	Avg. sq.	111.767***	44.420***	44.969***	31.565***
		(21.346)	(16.391)	(8.496)	(8.888)
Dev. from 30-70	Avg.	-7.082***	-2.055+	-3.529***	-3.402***
		(1.602)	(1.399)	(1.095)	(1.132)
	Avg. sq.	57.857***	12.856	23.773***	19.033***
		(12.367)	(9.672)	(6.999)	(6.150)
Dev. (sq.) from 20-80	Avg.	-9.064*	-13.880***	-7.059**	-5.105**
		(4.686)	(4.641)	(2.750)	(2.144)
	Avg. sq.	1076.265***	1016.342***	400.162***	238.084***
		(377.298)	(342.086)	(127.251)	(74.105)
Standard Deviation	Avg.	-4.131***	-1.675*	-1.529*	-1.707**
		(0.993)	(0.993)	(0.922)	(0.799)
	Avg. sq.	21.710***	6.709+	6.903+	6.353**
		(5.015)	(4.347)	(4.316)	(3.039)

the environment. Average anomalies refers to the average intensity of deviations from the typical climate conditions (*specific to each ancestral generation*) across generations. Average anomalies sq. (cub.) refers to the square (cube) of the average anomalies term. Variability across generations capture the standard deviations in the average temperatures across generations, i.e. how distinct were the overall climate conditions between generations. Moments of the full temperature distribution include the mean, standard deviation, skewness and kurtosis of the whole temperature distribution associated to an ethnic group in the 320 years (1600-1920) used to construct the per-generation anomalies index. Average anomalies ranges between 0.015 and 0.093 within the sample. Historical ethnic group characteristics include measure of development such as agricultural intensity, complexity of settlement, level of political hierarchies, size of the local community and main source of subsistence. Historical topographic characteristics include controls for distance to the equator, distance to the closest coast and geographical Köppen climate classification for the location of the ethnic group obtained from Ethnographic Atlas. Standard errors are clustered at the ethnicity level.

TABLE 12: Coefficients for the impact of average ancestral climatic anomalies on individual's self-reported attention to the environment

Dependent variable: Self reported measure of attention to environment					
	(1)	(2)	(3)	(4)	(5)
Avg. Anomalies	0.177	-9.877***	-17.402**	-9.055***	-9.677***
	(0.428)	(2.003)	(7.035)	(2.108)	(2.135)
Avg. Anomalies sq.		111.767***	276.862*	105.537***	109.681***
		(21.346)	(145.560)	(21.700)	(21.829)
Avg. Anomalies cub.			-1130.897		
			(951.849)		
Variability across generations				-0.166	
				(0.123)	
Occupation category Controls	Y	Y	Y	Y	Y
Historical ethnic group characteristics	Y	Y	Y	Y	Y
Historical topographic characteristics	Y	Y	Y	Y	Y
Moments of full temp. distribution	N	N	N	N	Y
Country-year Fixed effects	Y	Y	Y	Y	Y
Mean of dep var	0.704	0.704	0.704	0.704	0.704
St. Dev. of dep var	0.250	0.250	0.250	0.250	0.250
Min value Avg. Anomalies	0.015	0.015	0.015	0.015	0.015
Max value Avg. Anomalies	0.093	0.093	0.093	0.093	0.093
R-sq	0.112	0.113	0.114	0.114	0.114
Adj. R-sq	0.111	0.113	0.113	0.113	0.113
N	138067	138067	138067	138067	138067
AIC	-7632.03	-7822.44	-7827.27	-7832.44	-7868.15
BIC	-7346.80	-7527.37	-7522.37	-7527.54	-7533.75
*** p<0.01, **p<0.05, * p<0.1, + p<0.15					

F.4 Alternate functional forms for Folklore

Notes on Table 13: The unit of observation is an ethnic group. Average anomalies refers to the average intensity of deviations from the typical climate conditions (*specific to each ancestral generation*) across generations. Average anomalies sq. refers to the square of the average anomalies term. Average anomalies ranges between 0.01 and 0.097 within the sample. Historical ethnic group characteristics include measure of development such as agricultural intensity, complexity of settlement, level of political heirarchies and main source of subsistence. Historical topographic characteristics include controls for distance to the equator, distance to the closest coast and geographical Köppen climate classification for the location of the ethnic group obtained from Ethnographic Atlas. Standard errors are clustered at language group level.

TABLE 13: Coefficients for the impact of climatic shocks on occurrences of environment related ethnic folklore (Different Specifications)

	SVO Triplets	ConceptNet			
	log(1+Share)	log(1+Share)	Share	Asinh(Share)	Log-Poisson
Avg. Anomalies	-1.524***	-1.548**	-1.943*	-1.884*	-5.527*
	(0.561)	(0.764)	(1.061)	(0.975)	(3.017)
Avg. Anomalies sq.	16.219***	17.237**	22.423**	21.324**	63.612**
	(5.881)	(7.112)	(9.964)	(9.116)	(27.998)
First year of publication	0.000	0.000	0.001	0.000	0.001
	(0.000)	(0.000)	(0.001)	(0.000)	(0.002)
Number of authors	0.001	0.002+	0.003	0.003+	0.009*
	(0.001)	(0.001)	(0.002)	(0.002)	(0.005)
Number of languages	0.001	-0.006	-0.008+	-0.007	-0.026+
	(0.002)	(0.004)	(0.006)	(0.005)	(0.018)
Number of publishers	0.002	0.003	0.004	0.003	0.012
	(0.002)	(0.003)	(0.004)	(0.003)	(0.011)
Number of publications	-0.003	-0.005**	-0.006**	-0.006**	-0.020**
	(0.002)	(0.002)	(0.003)	(0.003)	(0.009)
Historical ethnic group characteristics	Y	Y	Y	Y	Y
Historical topographic characteristics	Y	Y	Y	Y	Y
Country Fixed effects	Y	Y	Y	Y	Y
Mean of dep var	0.434	0.277	0.325	0.317	0.325
St. Dev. of dep var	0.068	0.096	0.133	0.122	0.133
Min value Avg. Anomalies	0.010	0.010	0.010	0.010	0.010
Max value Avg. Anomalies	0.097	0.097	0.097	0.097	0.097
R-sq	0.323	0.439	0.417	0.432	
Adj. R-sq	0.235	0.366	0.341	0.358	
N	982	982	982	982	982
*** p<0.01, **p<0.05, * p<0.1, + p<0.15					

F.5 Conley standard errors to account for spatial correlation

Notes on Table 14: The unit of observation is an individual for columns (1), (2) and (3). The unit of observation is an ethnic group for columns (4), (5) and (6). Average anomalies refers to the average intensity of deviations from the typical climate conditions (*specific to each ancestral generation*) across generations. Average anomalies sq. refers to the square of the average anomalies term. Each regression includes the specified controls used in our primary analysis. Individual level controls include dummies for gender, occupation categories, age, income and education level available for individuals in the respective datasets. Historical ethnic group characteristics include measures such as measure of development such as agricultural intensity, complexity of settlement, level of political hierarchies and main source of subsistence. Historical topographic characteristics include controls for distance to the equator, distance to the closest coast and geographical Köppen climate classification for the location of the ethnic group. Conley Standard Errors which account for spatial correlation within the range of 200 kms and 500 kms are reported in curly and square brackets respectively.

TABLE 14: Coefficients for Climatic Anomalies with Conley Standard Errors

	Individual Level Variables			**Group Level Variables**		
	WVS	**IVS**	**ESS**	**Folklore**	**Wikipedia**	**Joshua**
	(1)	(2)	(3)	(4)	(5)	(6)
Avg. Anomalies	-9.877***	-2.975**	-2.391***	-1.548**	-4.088**	-1.390**
	(2.003)	(1.317)	(0.841)	(0.764)	(2.013)	(0.569)
	{1.929}	{1.279}	{0.830}	{0.852}	{1.975}	{0.660}
	[1.852]	[1.174]	[0.802]	[1.036]	[1.924]	[0.653]
Avg. Anomalies sq.	111.767***	35.442**	17.951***	17.237**	42.798**	13.163**
	(21.346)	(14.899)	(6.511)	(7.112)	(19.281)	(5.539)
	{20.717}	{14.512}	{6.455}	{8.858}	{18.998}	{6.105}
	[19.685]	[12.956]	[6.259]	[9.906]	[18.829]	[6.072]
Relevant individual controls	Y	Y	Y	N	N	N
Relevant ethnic group controls	Y	Y	Y	Y	Y	Y
Country Fixed effects	Y	Y	Y	Y	Y	Y
Mean of dep var	0.704	0.443	0.580	0.277	0.508	0.238
St. Dev. of dep var	0.250	0.230	0.136	0.096	0.212	0.154
Min value Avg. Anomalies	0.015	0.015	0.015	0.010	0.015	0.010
Max value Avg. Anomalies	0.093	0.093	0.093	0.097	0.093	0.097
R-sq	0.113	0.124	0.162	0.439	0.325	0.149
Adj. R-sq	0.113	0.123	0.160	0.366	0.161	0.131
N	138067	219900	23787	982	1031	11866
*** p<0.01, **p<0.05, * p<0.1, + p<0.15						

G Additional Theoretical Results and Proofs

G.1 The Rational Inattention Problem

Most of the derivations for the rational inattention block are either fairly straightforward or standard in the rational inattention literature; the reader can refer to the survey by Maćkowiak, Matějka, and Wiederholt, 2023 for a review of the literature and techniques used.

Take a generic problem, let $\omega \sim \mathcal{N}(0, \sigma^2)$ the unknown state and $u(a; \omega) := -v(\omega - a)^2$ the loss function (where we blackbox the stakes parameters $v > 0$). The cost of information is formally given by:

$$I(x, \omega) := H(x) - \mathbb{E}[H(\omega|x)]$$

where x denotes the signal about the state ω and H denotes the entropy function; for any random variable $y \sim p$, $H(y) = -\int \log(p)dp$. It is usual to justify the rational inattention framework and its canonical entropy cost specification as a heuristic model for situations where attention can be flexibly allocated and there is enough time for attention to adapt and focus on relevant features of the environment. In their recent survey of the literature, Maćkowiak, Matějka, and Wiederholt, 2023 provide the following general intuition which captures well our context of analysis: *"We consider RI to be an "as-if model" or a benchmark that applies well in repeated choice situations, or in choices over the long term. In these cases, the agent thinks about the optimal strategy once, and then applies it many times with little additional effort. Alternatively, it can be a strategy that the agent gradually learned through experience or stumbled upon it due to some evolutionary reasons."*

The agent's rational inattention problem problem is a choice over signal structure x:

$$\max_{x:\Omega \to \Delta(X)} \mathbb{E}\left[\max_a \mathbb{E}\left[u(a; \omega) \;\middle|\; x\right]\right] - \kappa I(x, \omega)$$

We can substantially simplify the problem above, and appeal to classical techniques in the rational inattention literature to obtain the closed form solution in Section 3.2. First, making u explicit, exchanging integrals and observing that quadratic loss entails the optimal action for a given signal structure to be simply the conditional expectation, we get a representation in the form of a canonical Gaussian-Quadratic RI problem. Gaussian signals are optimal under a Gaussian prior and quadratic loss (Maćkowiak and Wiederholt, 2009) which thanks to simplifications of the Shannon mutual information for Gaussian distributions allows us to recast both payoffs and costs first in terms of the induced posterior variance $\sigma^2_{x|s}$, as stated in the main text.

G.2 Technical Lemmas

Proof of Lemma 1 By the law of total covariance,

$$\Sigma_{g+1} = \mathbb{E}_{Q_g}[\mathrm{Cov}_M(f)] + \mathrm{Cov}_{Q_g}(\mathbb{E}[f' \mid f]) + \Delta_g^{\mathrm{sel}},$$

where Δ_g^{sel} collects the distortion induced by fitness-weighted reproduction. By assumption:

$$\mathbb{E}_{Q_g}[\mathrm{Cov}_M(f)] = \rho C, \qquad \mathrm{Cov}_{Q_g}(\mathbb{E}[f' \mid f]) = (1 - \rho\lambda)^2 \Sigma_g.$$

The selection distortion term Δ_g^{sel} is of order $O(\tau\|\Sigma_g\|)$: fitness weighting shifts mass linearly in $f - f_g$, and under the regularity assumption on fourth moments the resulting contribution to second moments is controlled. Since $\tau = O(\rho)$, this implies: $\|\Delta_g^{\mathrm{sel}}\| = O(\rho\|\Sigma_g\|)$.

Combining terms yields

$$\Sigma_{g+1} = \rho C + (1 - \rho\lambda)^2 \Sigma_g + O(\rho\|\Sigma_g\|).$$

Iterating this recursion and using $\rho \ll 1$ shows that $\|\Sigma_g\| = O(\rho)$ uniformly in g, and hence

$$\Sigma_g = \rho C + O(\rho^2),$$

as claimed.

Proof of Lemma 2 To simplify notations in the proof, define the general functional objective:

$$J(f) := \tau \mathcal{L}\big(\langle f, m\rangle \,\big|\, \hat{v}\big) + \frac{\lambda}{2}\|f - f_0\|^2$$

over $\mathcal{F}$, fixing $\hat{v} > 0$ and $m \in H$ such that $m \geq 0$. This is J_g if we take $m = m_p$ and $\hat{v} = \hat{\bar{v}}_g$, and the original J if we let $\hat{v} = \mathbb{E}_q[v^\star]$. First note that J is strictly convex for any $\lambda > 0$, since $\mu \mapsto \mathcal{L}(\mu|v)$ is strictly convex in μ for $\mu > 0$ and $\langle f, m\rangle$ is linear in f. Further since $f_0 \in \mathcal{F}$ and $m \geq 0$ we can look for an unconstrainted (interior) minimizer and verify straightforwardly that the positivity constraint implied by $f \in \mathcal{F}$ does not bind. Hence we can look for a minimizer by solving for a FOC of the Fréchet gradient:

$$\nabla J(f) = \tau \mathcal{L}'\big(\langle f, m\rangle \,\big|\, \hat{v}\big)m + \lambda(f - f_0) = \frac{\kappa}{2\langle f, m\rangle}\left(1 - \frac{\hat{v}}{\langle f, m\rangle}\right)m + \lambda(f - f_0)$$

We look for solutions f of the functional first-order condition $\nabla J(f) = 0$. First, claim that any solution must be such that $f - f_0$ lies in the span of m; indeed assume $\nabla J(f) = 0$, take any g such that $\langle g, m\rangle = 0$

and then the inner product to get:

$$\langle \nabla J(f), g\rangle = \frac{\kappa}{2\langle f, m\rangle}\left(1 - \frac{\hat{v}}{\langle f, m\rangle}\right)\langle m, g\rangle + \lambda\langle f - f_0, g\rangle \quad \Longrightarrow \quad \lambda\langle f - f_0, g\rangle = 0$$

since this holds for all $g \perp m$, by definition $f - f_0 \in \text{span}(m)$ i.e. there exists a constant $c > 0$ such that $f - f_0 = cm$. This further implies that by definition $\langle f, m\rangle = \langle f_0 + cm, m\rangle = \langle f_0, m\rangle + c\|m\|^2$. Now we can plug this back into the FOC and simplify for m to obtain:

$$\frac{\kappa}{2\big(\langle f_0, m\rangle + c\|m\|^2\big)}\left(1 - \frac{\hat{v}}{\langle f_0, m\rangle + c\|m\|^2}\right) + \lambda c = 0.$$

With the change of variable to $a = \langle f_0, m\rangle + c\|m\|^2$, multiply by a^2 to rearrange into a cubic equation:

$$\frac{\lambda}{\|m\|^2}a^3 - \frac{\langle f_0, m\rangle}{\|m\|^2}a^2 + \frac{\kappa}{2}a - \frac{\kappa}{2}v = 0 \iff a^3 - \beta a^2 + \gamma a - \gamma v = 0$$

where $\beta = \langle f_0, m\rangle$ and $\gamma = \frac{\kappa\|m\|^2}{2\lambda}$. This cubic must have a unique positive solution a^* (otherwise we would contradict convexity of J) – this can actually be checked directly and the solutions found explicitly (see below). Then the unique solution f^* of $\nabla J(f^*) = 0$ is:

$$f^* = f_0 + \frac{a^* - \langle f_0, m\rangle}{\|m\|^2}m$$

and one can directly verify that this is indeed the unique minimiser of the regularised loss J. As $\lambda \downarrow 0$, a direct examination of the cubic shows $a^* \to v$, so that we get:

$$f^* \xrightarrow[\lambda\downarrow 0]{} f_0 + \frac{v - \langle f_0, m\rangle}{\|m\|^2}m = \underset{f|\langle f,m\rangle = v}{\arg\min}\ \|f - f_0\|$$

Explicit solution for a^*. In the case where $f_0 \equiv 0$, $\beta = 0$ and this is directly a depressed cubic so we can apply Cardano's formula to get:

$$a^* = \sqrt[3]{\frac{\gamma v}{2} + \sqrt{\frac{\gamma^2 v^2}{4} + \frac{\gamma^3}{27}}} + \sqrt[3]{\frac{\gamma v}{2} - \sqrt{\frac{\gamma^2 v^2}{4} + \frac{\gamma^3}{27}}} \text{ where } \gamma = \frac{\kappa\|m\|^2}{2\lambda}$$

In the general case, first reduce it to a depressed cubic, then apply Cardano's formula to get:

$$a = \frac{\beta}{3} + \sqrt[3]{-\frac{\phi}{2} + \sqrt{\Delta}} + \sqrt[3]{-\frac{\phi}{2} - \sqrt{\Delta}}.$$

where:

$$\phi = -\gamma v + \frac{\gamma\beta}{3} - \frac{2\beta^3}{27} \text{ and } \Delta = \left(\frac{\phi}{2}\right)^2 + \left(\frac{\gamma - \frac{\beta^2}{3}}{3}\right)^3$$