Cultural transmission and historical origins of attention to environmental issues

César Barilla[†] ⓡ Palaash Bhargava[‡]

October 14, 2024

Abstract

How does the climatic experience of previous generations affect today's attention to environmental questions? Using self-reported beliefs and environmental themes in folklore, we show empirically that the realized intensity of deviations from typical climatic conditions in ancestral generations influences how much descendants care about the environment. The effect exhibits a U-shape where more stable and more unstable climates lead to higher attention, with a dip for intermediate realizations. We propose a theoretical framework where the value of costly attention to environmental conditions depends on the perceived stability of the environment; prior beliefs about which are shaped through cultural transmission by the experience of ethnic ancestors. The U-shape results from a dual purpose of learning about the environment: optimal utilization of typical conditions and protection against extreme events.

Keywords: Environment; Climate; Cultural Transmission; Inattention; Learning; Individual Preferences; Economic History

JEL Classifications: Q50, Q54, D83, D91, Z10, N00

[†]Columbia University

[‡]Columbia University

We would like to thank Sandra Black, Miguel Urquiola, Cristian Pop-Eleches, Jack Willis, Laura Boudreau, Michael Best, Navin Kartik, Yeon-Koo Che, Matthias Schündeln, Elliot Lipnowski, Suresh Naidu, Duarte Gonçalves, Florian Grosset, Akanksha Vardani, Tushar Kundu and Shivangi Bishnoi for helpful comments and discussions. We also thank attendees of the Applied Microeconomics, Development and Theory colloquia at Columbia University for helpful feedback. We thank Pari Bishen, Kevin Li Grant and Gina Nemade for their research assistance. We acknowledge the support of Columbia's Program for Economic Research. ⓡ stands for random order author naming convention.

1 Introduction

The consequences of climate change are immense. Global temperatures are expected to rise by at least 2.7° F by 2100 (IPCC [2013]). This will generate severe economic, social and health implications ranging from increased mortality and violence to reduced human productivity and economic growth (IPCC [2014], Auffhammer [2018]). Yet, public opinion remains divided on fundamental questions such as whether climate change is real, what needs to be done to protect the environment and how much attention do environmental issues warrant. A lack of consensus in beliefs about climate issues in turn stalls any attempts at policy change. In a recent survey carried out by the Yale Program on Climate Change Communications (Howe et al. [2015]), 72% Americans agreed that global warming is happening but only 47% Americans thought that it would affect them personally (Figure 6). As a result, only half of the surveyed population thought that addressing climate change should be a high priority for the next Congress and the President. Global surveys show similar statistics (Bell et al. [2021]).

How does one explain heterogeneity of beliefs about the importance of environmental issues? Previous literature has highlighted the role of social factors such as education levels, political ideology, age and moral values (Bell et al. [2021], Luo and Zhao [2019], Andre et al. [2021]) in shaping "climate preferences". Instead, this paper focuses on the role of ancestral climatic experiences in determining an individual's assessment of the importance of taking care of the environment. We hypothesize and test empirically that the realized variability of the climate in ancestral generations – understood as the intensity of deviations from appropriately defined "normal" conditions – influences their descendant's perception of the value of learning and caring about the environment, and hence their choice of attention. The underlying logic is that ancestors' experiences through cultural transmission and socialization directly or indirectly mould the beliefs of later generations.

Using tools from the literature on rational inattention, we provide a simple and flexible theoretical framework to formalize this effect. The individual's assessment of the importance of environmental issues is modelled as a costly attention problem. Taking care of the environment requires acquiring information about how best to adapt to climate, protect it, use resources, protect oneself against extreme events, etc. Information about adaptation to relevant feature(s) of the environment is available but is costly to acquire precisely.[1] The value of acquiring information (or, equivalently, the losses from imprecise knowledge) in turn depends on the overall *variability* of climate conditions. A given piece of information can either be relevant to optimally adjust to "normal" conditions or protect against extreme events. In the former case, stakes are decreasing in variability (if deviations from normal conditions are small, optimal utilization is more beneficial); in the latter, stakes are increasing in variability (protection against extreme event becomes more relevant when the intensity of deviations from typical conditions is higher). Since the true variability of the climate is unknown, prior beliefs about the *scale* of the variability of the climate that one might face feeds into the problem to determine at-

[1] Costs may be interpreted as coming either from literal acquisition costs, cognitive costs, or opportunity costs of attention not allocated to other issues.

tention. We derive the following theoretical comparative statics result: because of the dual role of learning, attention choice is single-troughed in the scale parameter of prior beliefs, meaning that ex ante perceptions of least and most amount of variability lead to higher attention levels, with a dip for intermediate levels.

To tie in the effect of cultural transmission in the model, we assume that the empirical distribution of variability in ancestor's generations determines an individual's prior beliefs – hence this feeds directly into the attention problem and the comparative statics result. This can be interpreted either literally or as a sparse modelling device. Under the former interpretation, an individual's ancestors mechanically endow them with a prior, i.e. before learning anything about the specific environmental conditions that one faces, perceived stability of the climate is shaped by their ancestor's experiences. Under the latter interpretation, this captures the effect on the perception of the value of attention (e.g. capturing in reduced form some underlying evolutionary process by which successive generations adapt their attention strategy from empirical payoffs). Putting everything together, the model predicts that the choice of attention should be U-shaped in the empirical scale of climatic variability experienced by previous generations.

We test this prediction empirically by regressing two different variables that capture the level of attention (either at an individual or a group level) on the average climatic variability experienced by ethnic ancestral generations. The first variable is what we interpret as a proxy measure of attention to environmental issues from the World Value Surveys. The second variable is the stock of environment related ethnic folklore which we interpret as a direct outcome of the level of attention paid to environment, not only by the current but also across several previous generations. We closely follow Giuliano and Nunn [2021] to construct the relevant data for our empirical analysis. Our measure of ancestral climatic variability is constructed using temperature anomalies faced by an individual's ethnic ancestors between the period 1600-1920 A.D.[2] We combine data from four different sources. We link the data on historical temperature anomalies from Mann et al. [2009] to the location of ethnic groups available in Giuliano and Nunn [2018]'s Ancestral Characteristics of Modern Population.[3] Our primary data source to obtain information on attention to climate issues is the World Value Surveys (referred to as WVS henceforth) carried out globally between 1981 and 2022. Apart from individual level information on questions such as how much one cares about the environment, WVS also provides us with information on the language spoken at home. This helps us match survey respondents from WVS to the ethnic groups present in Ancestral Characteristics of Modern Population. Finally, in order to understand the impact of temperature anomalies on existence of environmental folklore,

[2]Our measure of ancestral climate variability should not be confused with Giuliano and Nunn [2021]'s measure of climate instability. While they use across generation variation in average temperatures, we rely on the within generation climatic variability averaged over generations. We provide more clarification in Subsection 4.1.

[3]This dataset combines information on pre-industrial characteristics of ethnic groups around the world from Murdock's *Ethnographic Atlas*, People of Easternmost Europe (Bondarenko et al. [2005]) and World Ethnographic Sample (Murdock [1957]). Additionally, using version 16 of Ethnologue: Languages of the world (Gordon [2009]), Giuliano and Nunn [2018] link the ethnic groups to over 7000 different dialects and languages spoken around the world today.

we match the temperature data to the ethnic groups available in the Folklore database provided by Michalopoulos and Xue [2021].

Our main empirical result corroborates the predictions of the model: attention is indeed U-shaped – higher and lower experienced variability by ancestor generation lead to more attention in descendants, whereas more "intermediate" levels lead to lower attention. Further, this result is robust to alternative specifications for the measure of variability, inclusion of higher moments of the temperature distribution, as well as the choice of sample from WVS. We provide additional evidence in the support of this hypothesis by harnessing data on folklore associated to a particular ethnic group to measure the prevalence of environmental themes. Interpreting folklore as a cumulative stock of knowledge and collective memory, we argue that a higher attention paid to environmental issues by successive generations should be reflected in a higher presence of environmental themes in folklore at the group level. That is, the effect of experienced climatic variability amongst ancestors on folklore should exhibit the same U-shape. This is indeed what we find empirically. Finally, we provide suggestive evidence of muted response to ancestral climate variability for individuals from more developed ethnic groups. That is, in the case of historical institutional and economic protection at the ethnic group level, variability in climate seems to have a lesser bite. This in turn feeds into the cultural transmission by ancestors and mutes the slope of the level of attention paid by present day individuals.

The rest of the paper is structured as follows. Section 2 presents related literature. Section 3 introduces the theoretical framework and its results. Section 4 describes the data. Section 5 presents the empirical strategy and the corresponding results. Section 6 concludes.

2 Related Literature

This paper contributes to four broad strands of literature. First, our project adds to previous works on perceptions of climate change and the determinants of environmental preferences. We focus on a novel historical factor – how ancestral experiences impact beliefs about climate risks. While this is not directly related to climate change perception, it speaks to the formation of climate preferences through deeper historical mechanisms. Hence, this complements previously studied channels in the active debate about what drives opinion about climate change – see Leiserowitz [2006] for imagery and emotion based learning; Weber [2010], Weber [1997]; Hansen et al. [2004] for personal experiences and morality; Andre et al. [2021] for behavioral traits, moral values and misperceived norms; Luo and Zhao [2019] and Shi et al. [2016] for knowledge internalization, economic preferences and political ideology.

Second, it contributes to the fast growing literature on historical origins of preferences and norms. A large body of work has attempted to study the impact of traditional practices, historical institutions and experiences on current behavioral traits such as prosociality (Le Rossignol et al. [2022]), cooperation (Lowes [2018], Buggle and Durante [2021]), evolution of culture and institutions (Lowes et al.

[2017], Giuliano and Nunn [2021]), gender norms (Becker [2019], Alesina et al. [2013]) and differences in economic preferences (Becker et al. [2020]). The two papers within this context that are close to our exercise are Giuliano and Nunn [2021] and Buggle and Durante [2021]. Both papers use historical climate variability faced by an individual's ethnic or national ancestors. The first paper studies its impact on erosion of traditional values and the second paper studies its impact on current preferences for cooperation. Our paper is the first one to explore the impact of historical climate shocks on current environmental preferences. It also speaks to the growing theoretical literature on cultural transmission and the determinants of preferences (Bisin and Verdier [2001]). Our approach, however, draws from a different toolkit of modeling devices.

Third, our paper contributes to the growing literature on narrative economics. Shiller [2017] has shown both the impact of institutions and experiences on the formation of narratives and the impact of narratives on individual preferences and decisions (Shiller [2017], Akerlof and Snower [2016], Anthony [2021]). Individuals transmit information over generations in the form of stories (Harari [2014]) and these stories can help us uncover characteristics of societies of the past (Michalopoulos and Xue [2021]). In the context of environment, folklore, songs and narratives have been extensively used in Africa to persuade individuals to adopt sustainable practices and worry about environmental degradation (Osemeobo [1994], Amlor and Alidza [2016], Sanganyado et al. [2018]). Historically, the prevalence of institutional changes, environmental surroundings and traditional work practices have made their way into ethnic folklore (Michalopoulos and Xue [2021], Anthony [2021]). Within our paper, we extend this literature by showing another historical determinant: ancestral experiences with respect to climatic variability, which crucially impacts the content of ethnic folklore. Further, we document that climatic variability impacts ethnic folklore in a non-monotonic way since the stakes of optimally interacting with one's environment matter more at the tails.

Fourth, our paper creates a novel bridge (to the best of our knowledge) between the theoretical literature on (in)attention and the applied and theoretical literature on cultural transmission. The simple theoretical framework that we use to approach "caring about the environment" as an attention problem draws directly from the literature on rational inattention – in which the seminal contribution by Sims [2003] has been followed by a large literature with many distinct avenues of investigation, from behavioral content (Caplin and Dean [2015], Caplin et al. [2022]), macroeconomics (Maćkowiak and Wiederholt [2009], discrete choice (Matějka and McKay [2015]), voting (Matějka and Tabellini [2017]), bargaining and trading (Ravid [2020]). For a broad survey of the literature, refer to Maćkowiak et al. [2023].[4] This also broadly connects with the larger active literature in economic theory on costly information acquisition. Our contribution is mainly a methodological and conceptual novelty: by modelling cultural transmission as determining an agent's prior about some parameter of the attention problem, we harness the tools of the information acquisition literature to study the effect of cultural transmission as a comparative statics problem. This provides a new application for the rational inattention approach. As a result, this also contributes to the literature on cultural transmission (Bisin

[4] See also Veldkamp [2011].

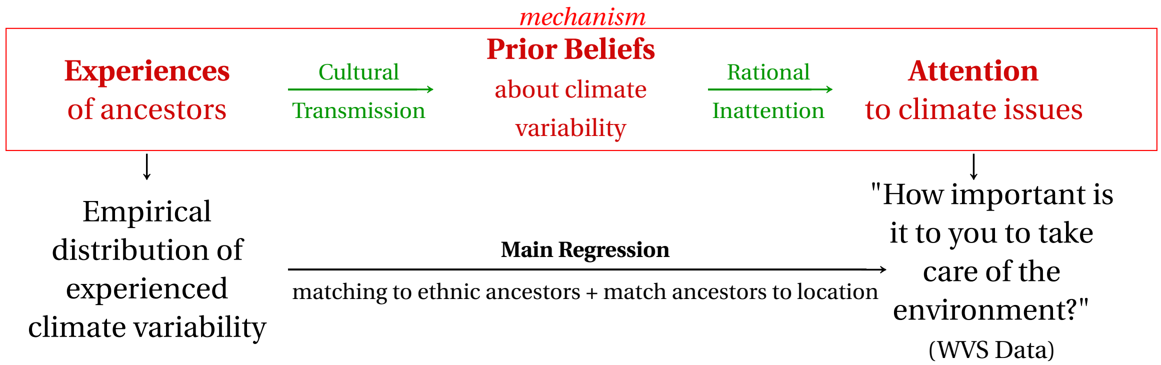


FIGURE 1: Conceptual Framework and Empirical Strategy

and Verdier [2023], Bisin and Verdier [2011]). Our approach differs substantially from the literature since our focus is not on studying explicit transmission mechanisms (see e.g. Bisin and Verdier [2001], Tabellini [2008], Panebianco and Verdier [2017]). Instead, we derive results under a flexible reduced form assumption for the channel of cultural transmission. Some contributions in this literature explicitly consider, like we do, a framework where cultural transmission influences beliefs in the context of uncertainty (see for instance Adriani and Sonderegger [2009, 2018], Adriani et al. [2018]) but their focus is different from ours (and complementary): they emphasize the mechanisms of informational transmission through social learning from earlier to younger generations, whereas we focus on the attention / information acquisition problem of younger generations as our outcome variable, in which behavior is driven by prior shaped through cultural transmission.

3 Theoretical Framework

3.1 Intuitive overview of the conceptual framework

We propose a simple stylized theoretical framework for how the climatic experience of previous generations impacts descendants' choice of attention to climate issues. Before introducing and analyzing the model, we provide an intuitive overview of its structure and our results. This overview foreshadows the theoretical predictions that will be tested in our empirical specification, emphasizes the key objects and conceptual ingredients, and outlines the essential mechanisms. Figure 1 provides a visual representation of the conceptual framework and its relationship to the empirical analysis.

The structure of the model revolves around two separate conceptual blocks. First, a simple attention model in which an agent has the possibility to learn (at a cost) about how to best adapt to some feature of the environment (which crops to favor for which climates, learn about prevention value, what materials to choose for construction, how to use, save and protect specific natural resources,...). The value of learning about adaptation to environmental characteristics is not fixed and is driven by two distinct motives: exploitation[5] of "typical" conditions and protection against extreme events. The true value of learning about specific features under both of those motives is determined by the underlying variability of the climate that the individual will face. We make the following intuitive assumption: the value of accurate information is decreasing in variability for features that are relevant in exploiting typical conditions and increasing in variability for features that pertain to protection from extreme events. This true environmental variability parameter, however, is not known by the agent (who holds a prior over it).

The second conceptual ingredient is the assumption that this prior belief on the variability parameter is determined by the climatic experiences of ancestors. We interpret this as a sparse modeling device to capture the effect of complex cultural transmission and socialization mechanisms in shaping an individual's prior beliefs about the environment they might face. In practice, this entails that the climatic variability experienced by previous generation directly influences the *expected stakes* in the descendant's learning problem, hence their attention choice. We derive a general comparative statics result: the attention choice is *single-troughed* in a *scale parameter* of the prior beliefs – i.e attention is highest when beliefs are either most concentrated around a low variability or very spread out towards possibly high variability.[6]

This comparative statics result combined with our cultural transmission assumption gives a direct testable hypothesis: the attention to environmental issues of descendants should be non-monotonic and "U-shaped" in the empirical (realized) scale of climate variability experienced by ancestral generations. This has an intuitive and flexible interpretation: *the perceived importance of attention to climate issues is highest for populations who have historically faced either more consistently stable climates (hence have learnt to rely on exploitation of typical conditions) or more consistently unstable climates (hence have learnt to value protection against unpredictable extreme events)*. This is consistent with both a literal or heuristic interpretation of the model.

The theoretical results we derive hold in a broad class of attention models. We begin by providing one canonical foundation that builds on the standard Gaussian-quadratic rational inattention model but later generalize it to show its flexibility.

As shown in Figure 1, our model structure directly maps to our empirical strategy. Our main regression tests the hypothesis of the existence of a transmission channel from ancestor experience to attention to climate issues. It shows a statistically significant U-shaped relation between an individual's / group's measure of attention and an estimate of the empirical scale of climate variability experienced by the corresponding ethnic ancestors.

[5]By this we mean optimal utilization. Throughout the course of this paper we will use these two phrases interchangeably.

[6]Intuitively, the scale parameter controls how "stretched out" towards high intensities of climate variability the beliefs are – a higher scale will entail both a higher expected level of variability and a thicker tail.

The following subsections rigorously introduce the model and its results, first focusing on the attention problem in isolation, then introducing cultural transmission through priors and deriving the main comparative statics result.

3.2 The attention problem

We begin by providing an intuitive derivation of the structure of the attention problem and its solution. Although we illustrate the derivation by making specific assumptions that are rooted in a modification of the canonical rational inattention model with entropy-based costs, we want to emphasize that this leads to a generalized problem with a minimal structure that underpins our results. The substantial assumption is that the attention problem can be reduced as one where payoffs are (approximately) linear in attention level, costs are convex in attention, and the marginal benefits to higher attention follow the "dual purpose" exploitation/protection structure. Any alternative specification which maps to this class would be equivalent.

The agent's environment and payoffs. Consider the problem of a single agent deciding how much attention to allocate to the problem of how to adapt to some feature of the environment. Denote by $x \in \mathbb{R}$ the true "state of the world" for that adaptation problem, which can be interpreted as the best theoretical course of action. We further assume that knowledge about x can have one of two possible characteristics: either is relevant for exploitation of "typical" climate conditions or for protection against extreme climate events. Denote by $\iota \in \{\mathscr{E}, \mathscr{P}\}$ the variable that captures whether knowledge about x is relevant for exploitation ($\iota = \mathscr{E}$) or protection ($\iota = \mathscr{P}$).

The agent's environment is additionally described by another unknown state $\eta \in \mathbb{R}_+$ which captures the variability of the climate that the agent will face. Intuitively, a lower η captures a very stable climate, in which typical conditions dominate and unpredictable events that deviate from those conditions are rare and mild; a higher η on the other hand corresponds to a highly volatile climate where extreme deviations and unexpected realizations are frequent.

The agent faces a decision problem and chooses some action $a \in A = \mathbb{R}$.[7] Given a choice of action a and states (η, x, ι), the agent's payoffs take the following form :

$$u(a; x, \iota, \eta) = -\gamma(\eta, \iota)(a - x)^2$$

The quadratic loss component $(a-x)^2$ is a standard specification in the literature which captures the idea that losses are proportional to mean-square prediction error, i.e posterior uncertainty about the true optimal course of action. As is standard in such specifications, $\gamma > 0$ is a "stakes" parameter that scales the payoff consequences of imprecise knowledge of the payoff-relevant state x. A novel

[7]Note that the "action" need not be interpreted literally and can simply serve to capture the value of acquiring information as in an estimation problem.

component is that we assume that those stakes are not fixed but instead depend on both the nature of the problem ι and the variability of the agent's environment η. We further assume the following structure :

$$\gamma(\eta,\iota) = \begin{cases} w(\eta) & \text{if } \iota = \mathscr{E} \\ l(\eta) & \text{if } \iota = \mathscr{P} \end{cases}$$

where:

- $w : \mathbb{R}_+ \to \mathbb{R}_+$ is a strictly decreasing and strictly convex $\mathscr{C}^2$ function such that $w' \underset{\infty}{\longrightarrow} 0$, i.e. the benefits of acquiring information about exploitation-type features are higher in more stable environments (which have more exploitation opportunities) and decrease at an increasing rate when moving towards more unstable environmental conditions.
- $l : \mathbb{R}_+ \to \mathbb{R}_+$ is a strictly increasing and strictly convex $\mathscr{C}^2$ function such that $l'(0) = 0$, i.e. the benefits of acquiring information about protection-type features are higher in more unstable environments (which feature more deviations that require protection) and increase at an increasing rate when moving towards more unstable environmental conditions.

Naturally, this decision-making framework need not be taken as literally representing the problem of an agent taking a one-shot action on some environmentally related issue. Instead, it aims to capture sparsely and at a higher level of abstraction the forces that lead to acquiring information about how to adapt to one's environment: information is acquired about *specific independent problems and climate related issues* (e.g. which kind of crops is it best to cultivate in a given climate, how can constructions be made resistant to floods or hurricanes and to what extent should they be, how and how much water should be conserved, to what extent is it optimal to exploit forests for wood,...); each particular problem pertains either to optimally exploiting normal conditions or self-protection against extreme events; in each kind of problem the underlying stability of the environment determines the extent to which information is important and useful: there is more scope for exploitation in stable conditions (higher gains) and more scope for protection in unstable conditions (higher reduction of losses).

Information acquisition. We assume that the agent problem has the following timing :

0. The agent chooses some signal structure s (at a cost to be specified) which is informative about the unobserved state x
1. The agent learns whether the problem is relevant for exploitation ($\iota = \mathscr{E}$) or protection ($\iota = \mathscr{P}$)[8]
2. The signal $s|x$ realizes and is observed by the agent, who then takes some action a

[8]The assumption that the nature of the problem is not known before attention to x is chosen captures the idea that learning opportunities arise exogenously and allocation of attention to environmental issues is decided *before* knowing if the information pertains to exploitation or protection, but that fact becomes transparent once some learning occurs.

Importantly, the agent cannot directly learn about η in the relevant timespan where learning about x is decided – this reflects the fact that η is a significantly more complex hyperparameter that controls the distribution of conditions that the agent faces.

We further assume that the agent holds mutually *independent* prior beliefs about η, ι and x.[9] Denote $p_0 \in \Delta(\mathbb{R}_+)$ the prior over η, $q := \mathbb{P}(\iota = \mathscr{E})$ the prior probability of facing an exploitation-type problem. We only make a substantial assumption here about the prior on x, which is assumed to be Gaussian : $x \sim \mathcal{N}(0, \sigma^2)$ (the normalization of the mean to zero is without loss).

To complete the description of the model, we only need to specify the costs of information acquisition. Following the large literature on rational inattention (see the recent survey Maćkowiak et al. [2023]), we posit that costs of acquiring signal structure s (i.e distribution of signals conditional on x) are linear in the Shannon mutual information between s and x:

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

where p_s and p_x denote the respective distributions of s and x and H denotes the (relative) entropy (for details, see the classical reference Cover and Thomas [2006]). 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 et al. [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."*

Formally, we can then write the agent's choice problem over signal structure s :

$$\max_{p_s} \mathbb{E}\Big[\max_a \mathbb{E}\big[u(a; x, \iota, \eta) \mid s\big]\Big] - \kappa I(p_s, p_x)$$

Solving for optimal attention. We can substantially simplify the problem above, and appeal to classical techniques in the rational inattention literature to obtain a closed form solution. By 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 the following alternative representation of the problem:

[9] This is a simplifying assumptions which is partly made for tractability but can be justified and interpreted as such: independence between x and ι means that the nature of the problem doesn't affect the ex ante belief about the content of its solution; independence between x, ι and η is justified by the fact that x, ι capture local unconditional knowledge about some particular problem and η only impacts payoffs by determining the relevance of that knowledge in context.

can be treated as a constant stakes parameter from the perspective of the attention problem. Now our problem takes the form of a standard rational inattention problem with stakes parameter r. We can appeal to standard results to claim that Gaussian signals are optimal under a Gaussian prior and quadratic loss (see e.g. Maćkowiak and Wiederholt [2009]) which thanks to simplifications of the Shannon mutual information for Gaussians allows us to recast both payoffs and costs simply in terms of the induced posterior variance $\sigma^2_{x|s}$. A last relabeling recasts the problem in terms of the *attention level* $\xi := 1 - \frac{\sigma^2_{x|s}}{\sigma^2} \in [0,1]$ (where recall that σ^2 is the prior variance) and gives the following program :

$$\max_{\xi \in [0,1]} r\xi - \frac{\kappa}{2} \log\left(\frac{1}{1-\xi}\right) \tag{$*$}$$

Which has the familiar solution :

$$\xi^* = \max\left\{0, 1 - \frac{\kappa}{2r}\right\}$$

Importantly, this variable ξ will constitute the main outcome variable of interest in our empirical analysis: ξ captures the agent's level of attention to climate issues, which we equate to a measure of how much one cares about environmental questions. In the model, ξ has a very intuitive interpretation: it is a number between 0 and 1 which capture the *level of attention* ($\xi = 0$ represents no information being acquired and $\xi = 1$ represents acquisition of perfect information, which is naturally only theoretically possible but precluded by unbounded information costs).

A generalization of the attention model. We now introduce a generalization of problem $*$ which conveniently subsumes the Gaussian-quadratic-entropy setup as a particular case. There are two main reasons for this: (1) the language of the general model will make for a more convenient derivation of our main results and (2) this serves to highlight the driving assumptions as well as the flexibility of our framework.[10] From now on, we consider the following attention problem:

$$\max_{\xi} \xi \int_{\mathbb{R}_+} (W(\eta) + L(\eta))\, dp_0(\eta) - c(\xi) \tag{$\bigstar$}$$

Where we make the following assumptions:

(i) The cost of attention c is a non-negative increasing, convex, smooth function on $[0,1)$ with $c(0) = 0$ and $c(\xi) \xrightarrow[\xi \to 1]{} \infty$;

(ii) The adaptation component of the marginal benefit of attention W is a smooth non-negative strictly decreasing and strictly convex function of climate variability η such that $W' \underset{\infty}{\longrightarrow} 0$;

[10] The expositional choice to first present the Gaussian-quadratic-entropy setup was made primarily to provide an intuitive way that the general model form could be derived from primitive ingredients.

(iii) The protection component of the marginal benefit of attention L is a smooth non-negative strictly increasing and strictly convex function of climate variability η with $L'(0) = 0$.

Obviously, the previous setup maps into this formulation by setting $W(\eta) = \sigma^2 q w(\eta)$, $L(\eta) = \sigma^2(1-q)l(\eta)$ and $c(\xi) = \frac{\kappa}{2}\log(1/1-\xi)$. We emphasize that the problem ★ represents the distilled form of our specific model and highlight the minimal structure and assumptions that we will need to derive our main result. Specifically, in the next subsection, we introduce an approach that allows us to frame the effect of cultural transmission as a simple comparative statics exercise in p_0 (the prior on η) within problem ★.

This will lead us to our main stylized fact in Proposition 1 – the single-troughed behavior of attention. Naturally, more specific assumptions within the class of problem ★ would lead to finer predictions. Nevertheless, our aim here is mainly to provide a general theoretical framework and to focus on its robust predictions under minimal assumptions. Given that we interpret this model as a stylized and high-level heuristic representation, overly specific assumptions and predictions should be treated carefully – hence we adopt a cautiously minimal stance which reinforces the generality of the "U-shape" prediction. In that regards, any alternative specification that can be mapped to ★ would be equivalent.

3.3 Cultural transmission via priors and comparative statics

To articulate the influence of ancestral experience in an individual's attention to environmental issues, we hypothesize the following channel: the agent's prior belief p_0 about the climate variability parameter η (which controls the stakes in the learning problem) is shaped, through complex mechanisms of socialization and cultural transmission, by the experience of their ethnic ancestors, passed on through generations. The prior belief p_0, which feeds into the attention problem as a *parameter* is determined (through some process to be specified) by the *empirical realization* of the climate variability experienced by ancestors – hence analyzing the effect of ancestral generations' experience on current generation's attention choices boils down to a comparative static exercise in p_0.

This assumption can be interpreted in one of two ways. On the one hand, we can take the mechanism as literally capturing the agent's perception (conscious or not) of the variability of the climate they could face; in this interpretation cultural transmission is taken as some (non-Bayesian) mechanism by which a literal prior is formed before (Bayesian) learning occurs. Alternatively, we can interpret cultural transmission as acting *directly* on the perceived scale parameter – in which case the "prior" is just a modeling device that captures the structure with which past generations' experiences aggregate into perceived value of learning. In this latter interpretation, it is tempting to frame the mechanism under a broad evolutionary logic: populations who benefit most from whatever learning occurs about environmental conditions are those who face most consistently stable or consistently unstable climates. They are in turn the ones who transmit a higher perceived value (stakes) of information acquisition about environmental adaptation. Taking an explicit stance on the underlying intergenerational transmission mechanisms is beyond the scope of our approach and, on the contrary, we view the interpretative flexibility as an advantage: since either interpretation (or a convex combination of both) is valid, this modeling device provides a sparse and flexible reduced form approach to instrumentally capture the influence of ancestral history.

In order to make the assumption explicit and provide a simple parametric mapping to our quantitative analysis, we assume that the prior belongs to some arbitrary *scale family* of distributions. Formally, this means that the prior belief p_0 is fully described by a one-dimensional scale parameter $\theta > 0$ and a reference probability distribution $\overline{p} \in \Delta(\mathbb{R}_+)$ (assumed absolutely continuous with full support and identified with its density) such that :

$$p_0(\eta) = \frac{1}{\theta}\overline{p}\Big(\frac{\eta}{\theta}\Big)$$

To gain intuition, observe that for a scale family defined over nonnegative reals, increasing the scale intuitively corresponds to stretching the distribution upwards – i.e a belief characterized by a higher likelihood of more unstable climates and a thicker tail (more extreme tail realizations of unstable climates). Observe further that the scale directly determines both the expectation and variance of the distribution; indeed denote by $\mu := \int \eta \overline{p}(\eta)d\eta$, $\nu^2 := \int (\eta-\mu)^2\overline{p}(\eta)d\eta$ the expectation and variance of the reference distribution, then a straightforward change of variable shows:

$$\mathbb{E}_{p_0}[\eta] = \mu\theta \text{ and } \mathbb{V}_{p_0}(\eta) = \nu^2\theta^2$$

This parametrization provides an intuitive approach to doing comparative statics in the prior by simply varying θ. Because it has this stretching effect on the distribution, it intuitively captures the (generally more complex) effect of inducing a uniformly higher climate variability and inducing a corresponding stretch of the expectation of the stakes functions $W(\eta)$ and $L(\eta)$. Indeed, a key step in the proof of Proposition 1 (our main comparative statics result) below will be to argue that the properties of W and L in terms of η transfer to their expectation in terms of θ under the scale parametrization. The scale parametrization for p_0 is an intuitive approach given our monotonicity assumptions and that η is an intensity variable. Scale intuitively captures the dependence of stakes on the belief "how high" (stretched to the right) beliefs about climate variability are. The result of Proposition 1 also provides further ex post justification for considering scale families: they generate a natural structure and robustly produce the U-shape of attention within the generalized model structure ★.

Making the transmission mechanism explicit within the scale parametrization will allow us to directly connect our theoretical results to our empirical analysis. In its general form, transmission can be formalized as follows. Denote the current period agent/generation as generation 0; let $\{\hat{\eta}_g\}_{g=1,\ldots,G}$ a sample of (estimated) *experienced* climate variability for the previous G generations of direct ancestors (indexed backwards). Then there exists some function Φ that captures the mechanisms of compound intergenerational transmission, such that the scale parameter in the prior of the current generation is given by $\theta = \Phi(\hat{\eta}_1,\ldots,\hat{\eta}_G)$. A minimal assumption here would be that Φ is increasing in each of its components. In the next section we will test a even simpler specification where Φ is the sample average:

$$\theta = \frac{1}{G}\sum_{g=1}^{G}\hat{\eta}_g \tag{CT}$$

This captures a direct intuition: that the empirical distribution of experiences of one's ancestors acts as a frame of reference for an individual's prior belief about their environment; furthermore, the sample average can be, up to a multiplicative constant, identified with a scale parameter of the *empirical distribution* of $\hat{\eta}$ in previous generations. Our main hypothesis therefore amounts to the following: the typical intensity of climate variability experienced by ancestral generations, through representations and (non modeled) mechanisms of cultural transmission, determines the (implicit or explicit) expected intensity of climate variability that controls the stakes to learning about one's environment as previously specified.

With those ingredients at hand, we can already anticipate our next section and spell out what our empirical strategy will be. First, we will recover measure of individual attention to environmental issues ξ_i (indexed by individual i). Then, matching each individual to their ancestor generations and each ancestral population to their climatic experience, we will construct measures of experienced climate variability $\eta_{i,g}$ for each ancestor generation g of each individual i. Finally, we will aggregate those $\hat{\eta}_{i,g}$ according to CT into an individual specific measure of the empirical scale of their ancestor's experiences $\theta_i = \frac{1}{G}\sum_{g=1}^{G}\hat{\eta}_{i,g}$. This will allow us to run the following regression specification fitting a quadratic approximation of ξ as a function of θ:

$$\xi_i = \beta_1\theta_i + \beta_2\theta_i^2 + \text{ controls}_i + \varepsilon_i$$

Naturally, the next sections will elaborate in details on how the relevant variables are constructed and the empirical strategy behind the specific regressions we run. The aim of this foreshadowing is merely to tie together the theoretical ingredients with empirical objects.

Before turning to our empirical exercise, there remains one major theoretical component to elucidate. What results does our model predict in the regression specification laid out below, or conversely what rationale can our model provide for observed behavior? Indeed, our model has, so far, outlined and analyzed the channel by which θ (from now on interpreted as condensing the experience of previous generations into a scale parameter) affects attention choice ξ, but what specific predictions does it make? The main general (testable) prediction of the model is laid out in the following proposition, the proof of which is relegated in the appendix.

Proposition 1. *Denote $\xi^*(\theta)$ the optimal level of attention in $\bigstar$ for a given prior scale parameter θ. There exists two thresholds $\underline{\theta},\overline{\theta}$ (not necessarily distinct) such that :*

- *ξ^* is strictly decreasing on* $(0,\underline{\theta}]$
- $\xi^*(\theta) = 0$ *in* $(\underline{\theta},\overline{\theta})$

- *ξ^* is strictly increasing on* $[\overline{\theta}, \infty)$

In words, this states that the model predicts that attention should be single-troughed in the aggregated climate variability experienced by ancestral generations. This results potentially allows for a central "censored" region where $\xi^* = 0$, although this region can be empty.[11] Under some reasonable assumptions (e.g. that total stakes are never too low), we should expect a roughly "U-Shaped" pattern for ξ^*.

Anticipating that our empirical results will indeed robustly confirm this U-shaped pattern of attention, we can interpret the model as giving us a theoretical rationale for the following intuition: because information about adaptation to environmental conditions (*caring about the environment*) is more valuable when climate is either more stable or more unstable and because perceptions of climate instability and/or the value of information are determined by the experiences of successive generations through cultural transmissions, individuals whose ancestors have faced more consistently stable/moderate or unstable/extreme climates are led to pay more attention to climate issues, whereas attention dips for intermediate level of experienced ancestral climate variability.

In Figure 2 below, we provide a simple illustration of the behavior of attention in a particular example. To construct this example, we assume that the exploitation-motive weighting function W takes an inverse form $W(\eta) = \frac{1}{a+\eta}$ with $a > 0$ a constant and the protection-motive weighting function L is quadratic $L(\eta) = b\eta^2$ for a constant $b > 0$. We further assume that the prior p_0 about η belongs to a log-normal family parametrized by scale $\theta > 0$.[12] This allows us to compute expected weights as a function of scale $\mathscr{W}(\theta) = \mathbb{E}[W(\eta)]$ and $\mathscr{L}(\theta) = \mathbb{E}[L(\eta)]$ (also plotted in Figure 2). Lastly, we assume that the attention cost is obtained from the underlying Shannon cost specification previously introduced i.e $c(\xi) = \frac{\kappa}{2}\log(\frac{1}{1-\xi})$. The right panel of Figure 2 plots the resulting optimal attention strategy and illustrates the key "U-Shape" prediction.

4 Data and Definition of relevant variables

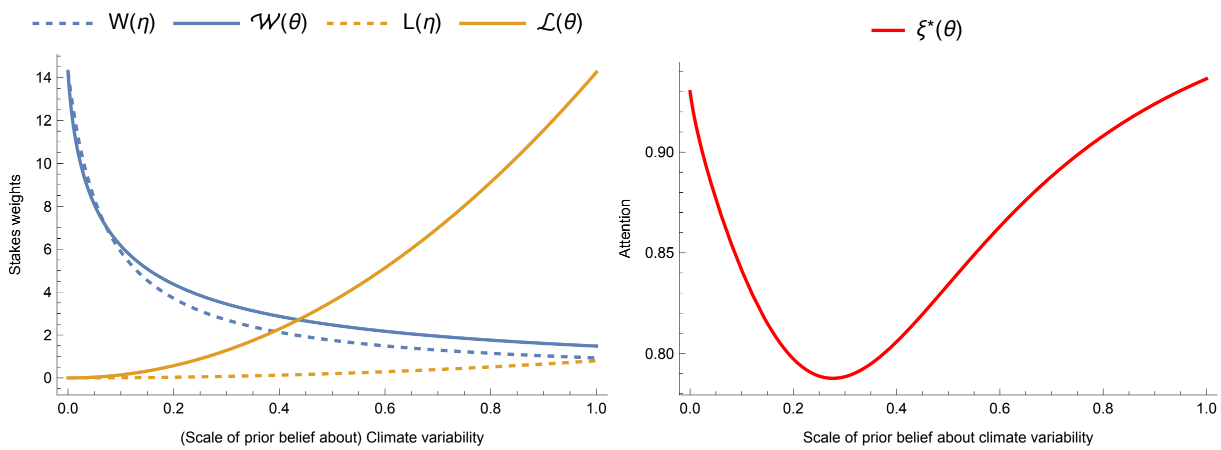
To empirically test the predictions laid down in our theoretical model, we need to construct the following components:

1. The measure of attention paid to the environment at the individual or the group level (ξ_i).
2. The link between individuals and their ancestors from corresponding ethnic groups (whose experiences will feed into beliefs to determine attention).

[11]There is a non-enmpty censored region if and only if the combined expected learning marginal benefit $\mathbb{E}[W(\eta) + L(\eta)]$ dips "low enough" (below $c'(0)$).

[12]Which can be written in the usual (log-expectation/log-standard) parametrization of the lognormal distribution as: $\eta \sim \text{Lognormal}(\log(\theta), \sigma)$ for some $\sigma > 0$.

FIGURE 2: Illustration of payoff structure and optimal attention strategy



Note: numerical simulations for the example of the optimal attention strategy predicted by the model (right panel) for given W and L functions (left panel) and under a log-normal prior assumption for p_0 inducing $\mathscr{W} = \mathbb{E}_{p_0}[W]$, $\mathscr{L} = \mathbb{E}_{p_0}[L]$.

3. The link between ancestors' historical location and climate conditions to obtain the relevant empirical distribution of ancestral climate variability $\hat{\eta}_{i,g}$ and aggregate it into the belief scale parameter θ_i.

We carry out this exercise by combining and constructing data from four different sources. We use Mann et al. [2009] for the historical temperature data and construct our measure of ancestral climate variability (captured by the mean of deviations from "typical" conditions across generations). We rely on Giuliano and Nunn [2018]'s Ancestral Characteristics of Modern Populations for data on ethnographic groups (including historical location and main language of the group), *World Value Survey* (WVS [2020]) for the variable capturing attention paid to environment (*level of care for the environment*) and language spoken at home by the individual and Michalopoulos and Xue [2021] for data on ethnic environmental folklore (our measure of attention paid to environment at the group level).

In the following subsections, we describe our methodology to merge all the datasets and construct the relevant variables in further detail.

4.1 Historical Temperature Data

Our primary data source to obtain historical climate variability faced by an individual's ethnic ancestors is Mann et al. [2009]. This dataset provides information on global patterns of detrended surface temperature for a long historical period using a climate field reconstruction approach. The construction uses proxy data with global coverage comprising of 1036 tree-ring series, 32 ice core series, 15 marine coral series, 19 documentary series, 14 speleothon series, 19 lacustrine sediment series and 3 marine sediment series (Mann et al. [2008]). 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 by 555km) grid cell level since 500 AD. For the purpose of our analysis, we restrict attention to the temperature data between 1600 AD and 1920 AD. The choice of the time period is governed by two reasons. First, since the quality of the data becomes less reliable at an annual level when we go further back in the past, we restrict the data to the years for which at least 2 types of the proxy series are available. Second, we want our measure of the empirical estimate of expectation of ancestral climatic volatility to be independent of factors that might also influence an individual's own learning during their lifetime. Therefore, we choose to retain a gap of at least 80 years with the first year for which we have measures of attention paid to the environment by a survey respondent.[13]

The key variable we construct from the historical temperature data is the average climate variability (henceforth referred to as avg. variability) faced by ethnic ancestors across several generations.[14] The avg. variability defines the scale parameter[15] of the distribution from which an individual's or a group's prior is drawn about the variability that they might face and hence determines their choice of attention they want to pay to their climate. Following Giuliano and Nunn [2021], we assume that each generation lives for 20 years and they do not overlap.[16] We also assume that all the historical generations associated to a particular ethnic group (details to follow in the next subsection) live within the same 5degree-by-5degree grid.[17] Therefore, for our primary specification, each ethnic group has a total of 16 generations associated to them. The associated avg. variability measure is constructed as follows:

- Fix a grid / ethnic group.
- For the given ethnic group, use g to denote the g^{th} generation between 1600-1920.
- Let $z_{g,t}$ denote the temperature faced by generation g in year t of their lifetime (denote $\mathbb{T}_g$ the set of those years); denote by $\hat{F}_g$ the empirical CDF of $\{z_{g,t}\}_{t\in\mathbb{T}_g}$.

[13]This also allows us to keep a meaningful timeline in consideration to analyse the impact of climatic variability on environmental occurrences in folklore data, i.e. we do not use information on climatic variability that may have occurred in recent times and do not add to the stock of ethnic folklore we are analysing.

[14]Climate variability as a concept is hard to define (specifically as one goes back in the past). One may require the use of several indicators such as variability in temperature, precipitation, humidity etc. both at a more granular level spatially and temporally. Due to its correlation with other climate manifestations, reliance on historical annual temperature data in the absence of more precise sources is a reasonable proxy.

[15]Note that, given our assumption that the prior belongs to a scale family of distribution, we can also justify this variable by arguing that the avg. variability is proportional to *empirical scale* of the distribution of climate conditions faced by ancestors – hence our transmission mechanism could also be viewed as equivalent to a naive parametric updating from ancestors' experiences.

[16]Our results are robust to changing the lifetime of each generation as shown in Subsection 5.3.

[17]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.

- For each generation g compute $\hat{\eta}_g$ the average intensity of deviations from the "typical" range of temperatures. We define the typical range of temperatures to be between the α^{th} and the $(1-\alpha)^{\text{th}}$ percentile of the empirical distribution of temperature values within a generation's lifetime. Our main specification assumes $\alpha = .2$.[18] Formally, we compute:

$$\hat{\eta}_g = \frac{1}{K_g} \sum_{t \in \mathbb{T}_g} \Big(\underbrace{\mathbb{1}_{z_{g,t} < \hat{F}_g^{-1}(\alpha)}}_{\text{if} z_{g,t} < \alpha\text{-qtile}} \underbrace{\left| z_{g,t} - \hat{F}_g^{-1}(\alpha) \right|}_{\text{dist. to } \alpha\text{-qtile}} + \underbrace{\mathbb{1}_{z_{g,t} > \hat{F}_g^{-1}(1-\alpha)}}_{\text{if} z_{g,t} > (1-\alpha)\text{-qtile}} \underbrace{\left| z_{g,t} - \hat{F}_g^{-1}(1-\alpha) \right|}_{\text{dist. to } (1-\alpha)\text{-qtile}} \Big)$$

Where $\hat{F}_g^{-1}(\alpha), \hat{F}_g^{-1}(1-\alpha)$ denote respectively the α and $1-\alpha$ quantiles of the empirical distribution $\hat{F}_g$ and $K_g := \#\{z_{g,t} \mid z_{g,t} < \hat{F}_g^{-1}(\alpha) \text{ or } z_{g,t} > \hat{F}_g^{-1}(1-\alpha)\}$ is just the normalizing constant that counts the number of excursions outside the $(\alpha, 1-\alpha)$ percentile range. This formula has a very intuitive interpretation: for $y_{g,t}$ that falls outside the $(\alpha, 1-\alpha)$ quantile range we measure the distance to the corresponding quantile (intensity of the deviation) and then average over all such $z_{g,t}$.

- Define the average ancestral climatic variability for an ethnic group as $\text{Avg. variability} = \frac{1}{G}\sum_{g=1}^{G} \hat{\eta}_g$.[19]

Essentially, our measure of ancestral climate variability captures the average magnitude, across generations, of the intensity of deviations (extreme events) from (possibly generation-specific) typical climate conditions in each generation. Intuitively, the $\hat{\eta}_g$'s capture how unpredictable and how extreme was the climate that a given individual's ancestors faced. When we average over generations, this captures the *overall scale* of climate variability experienced by one's ancestors. Variation in our variable of interest at the grid level is reported in Figure 7.

4.2 Ancestral Ethnic Group Data

In order to obtain information on the ancestral characteristics of an individual, we rely on Giuliano and Nunn [2018]'s database on Ancestral Characteristics of Modern Populations. This dataset combines the pre-industrial characteristics (such as mode of subsistence, jurisdictional hierarchy, complexity of settlements, size of community, intensity of agriculture)[20] of ethnic groups available in *Mur-*

[18]We redefine the "typical" range in several different ways for robustness checks. Our results are consistent across our choice of definition (see Subsection 5.3).

[19]Where the denominator G is the number of generations; in our baseline specification $G = 16$ and this number will change in Subsection 5.3 based on our choice of the lifespan of the generation.

[20]For the purpose of our analysis, we restrict our attention to the following pre-industrial characteristics of the ethnic groups: Primary mode of subsistence (Fishing, Hunting and Gathering, Animal Husbandry and Pastoralism and Agriculture), Economic and political development indicators such as settlement / economic complexity (values go from 1 to 8. 1: fully-nomadic, 2: semi-nomadic, 3: semi-sedentary, 4: compact but impermanent settlements, 5: neighborhoods of dispersed family homesteads, 6: separated hamlets forming a single community, 7: compact and relatively permanent, 8: complex settlements), agricultural intensity (ranging from 1 to 6 where 1 corresponds to no agriculture, 2 to casual agriculture, ... , 6 to intensive irrigated agriculture), community size (ranging from 1 to 8 where 1: fewer than 50, 2: 50-99,

dock's Ethnographic Atlas, Bondarenko et al. [2005]'s dataset on People of Easternmost Europe and Murdock's World Ethnographic sample (Murdock [1957]). Information on the pre-industrial characteristics of the 1265 ethnic groups in the Ethnographic Atlas has been coded for the earliest period for which satisfactory ethnographic data is available or can be reconstructed.[21]. The other two samples, Bondarenko et al. [2005] and World Ethnographic Sample Murdock [1957]) which Giuliano and Nunn [2018] use, add 27 more ethnic groups to the 1265 groups available from the Ethnographic Atlas. Finally, Giuliano and Nunn [2018] manually match over 7,000 different dialects and languages from Ethnologue: Languages of the World (Gordon [2009]) to the ethnic groups from the ethnographic data sources. The availability of the latitude and longitude of the ethnic groups location in the ethnographic datasets help us match the ethnographic groups to the surface grids in Mann et al. [2009] temperature anomalies database. This also allows us to assign the Köppen climate classification, distance to the coast and equator to additionally control for the overall climate and geographical conditions clusters of ethnic groups may have faced. To see the distribution of ethnic groups globally from the original dataset and the ones we are able to use for our analysis, refer to Figure 8.

4.3 Attention to environment - World Value Surveys

Our first main dependent variable is the level of attention an individual pays to the environment. As laid down in the theoretical section, an individual's assessment of the importance of environmental issues can be modelled as a costly attention problem where her priors would drive the level of importance she places on or the level of attention she pays to the environment. This level of attention / importance to environment is captured by the individual's response to the question: *"How important is it to you to take care of the environment?"* from wave 5 and 6 of World Value Surveys. 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. We normalize this variable to take a value between 0 and 1 and interpret it as our level of attention parameter.[22]

World Value Surveys is a repeated cross section data comprising of respondents from over 100+ countries over the time period 1981-2022. Individuals from the participating countries are randomly sampled to construct a nationally representative sample and asked questions along the thematic categories related to perceptions of life, environment, work, family, politics, society, religion and national identity. Some questions, such as the one we are interested in are included only in specific waves. As

3: 100-199, 4: 200-399, 5: 400-1,000, 6: 1,000-4,999, 7: 5,000-50,000 and 8: $> 50,000$) and global and local jurisdictional hierarchy (ranging from 0 to 4, denoting the number of levels within a jurisdiction.)

[21] 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 et al. [2021] for the relevance of *Ethnographic Atlas* in capturing actual traditional practices.

[22] The value 1 in the transformed variable corresponds to 6 in the original scale, i.e. if the individual responds with a 6 to the question in consideration, the value our variable would take is 1.

a result, all the responses to our question of interest are from individuals surveyed between 2005 and 2014 in wave 5 and 6.[23]

The survey, additionally, provides an array of demographic information on the respondents which is helpful for our empirical analysis. We use information on age, education level[24], occupational category, income level (ten categories, standardized within the country), gender, language spoken at home and language of the interview (see Table 6 for summary statistics). Information on language spoken at home helps us map the self-reported mother tongue of the individual to the language / dialect associated to the ethnic groups available in the ancestral characteristics database. This provides us the link which we exploit to map an individual to their ethnic ancestors and hence the climatic variability faced by them. In some 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.[25,26] Figure 9 provides a weighted word cloud corresponding to the distribution of languages in our sample of interest.

4.4 Folklore

Our second dependent variable which captures the level of attention to environment at the group level is constructed using the folklore data made available by Michalopoulos and Xue [2021]. This folklore database builds on Berezkin's catalog of motifs and links the linguistic groups found in Yuri Berezkin's catalog to the ethnic groups found in Murdock's *Ethnographic Atlas*. A motif, according to the authors, is an episode or an image found in the set of narratives recorded in an ethnolinguistic community. 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 (Berezkin [2015]). 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. Michalopoulos and Xue [2021] further link Berezkin's groups to the ethnographic groups present in the Ethnographic Atlas, thereby, giving us a distribution of oral traditions and folktales associated to each ethnic group whose history of climatic variability and geographic location have been mapped as discussed in the previous sub-sections. In line with our theo-

[23] We have a total of 157,142 respondents over a period of 2 waves (4 years each) from 78 countries who responded to this question of interest.

[24] Ranging from 1: Inadequately completed elementary education to 8: University with degree/higher education - upper-level tertiary.

[25] Our sample of interest from World Value Surveys contains 202 different language / ethnic groups.

[26] Less than 5% of the individuals in an main sample of consideration have multiple ethnic groups associated to them.

retical framework, if successive generations of an ethnic group face extremely stable or unstable environments (captured by our measure of climate variability) then each generation would pay a higher attention to climate issues[27], thereby adding to the stock of environmentally themed folklore.

To construct the relevant variable to test this, we carry out two different exercises. First, using the relationship between Berezkin's group and the groups from Ethnographic Atlas provided by Michalopoulos and Xue [2021], we link the Folklore data to the climatic variability data available from Mann et al. [2009] at ethnic group level. Second, following suit from Michalopoulos and Xue [2021], we employ the use of *ConceptNet* (Speer et al. [2017]) to classify motifs associated to each ethnic group as an environment related or non-environment related motif.[28] Using this information, we construct our 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. We will treat this measure as a proxy for total stock of attention paid to environmental issues over time at the group level. A relative word count across all descriptions associated to the motifs available in Berezkin's catalog can be found in Fig. 10.

5 Empirical Results

5.1 Self reported attention to environment - Individual level analysis

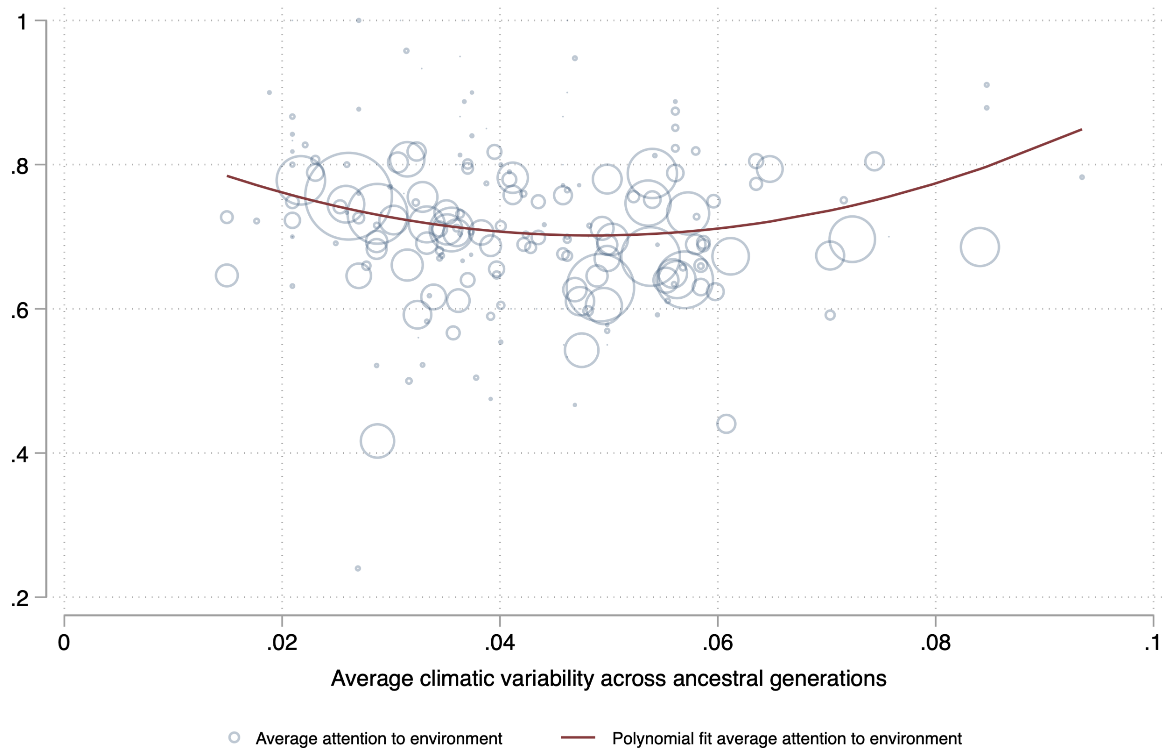
We begin by examining the impact of average climatic variability (our empirical estimate of the scale parameter) on an individual's self reported attention to the environment. That is, we empirically assess the result laid down in Proposition 1 at an individual level. Recall that empirical distribution of an individual's ancestors' climatic experiences inform their prior which then feed into their optimal level of attention they want to pay to their environment. Since the gains from exploitation are decreasing in the variability of the climate and the losses from the lack of protection are increasing in the variability of climate, the value of learning is higher on the tails. Given our assumption that the individual's prior belongs to an arbitrary scale family of distributions, the empirical mean of ancestral experiences is proportional to the scale parameter characterizing the distribution. As a result, the optimal attention level is U-shaped in the average climatic variability experienced by ancestral generations.

Figure 4 presents a preliminary relationship between average level of attention amongst individuals belonging to an ethnic group and the average climatic variability faced by the ancestors. We plot the raw scatter plot between with the two variables and an unrestricted polynomial fit. We do not control for any demographic and ethnic characteristics. The figure points towards two facts: 1. There is con-

[27]And as a consequence talk about it, embed it in lullabies and stories and retain it in their collective memory

[28]We obtain a list of words related to weather, climate, temperature, environment and natural disaster from ConceptNet (Speer et al. [2017]) and then check the occurrence of the words from this list in the description of the motifs.

FIGURE 3: Average level of attention to environment - Ethnic group level



Note: N = 202. The figure plots the preliminary relationship between average level of attention paid to the environment by individuals of a particular ethnic group and the average climatic variability faced by their ancestors of the same ethnic group over a period of 320 years. Our measure of attention to environment is the normalized individual response to the question "On a scale of 1-6, how important is it for you to take care of the environment?" from wave 5 and 6 of World Value Surveys. Our measure of the average climatic variability is the average intensity of deviations from typical climate conditions (specific to each ancestral generation) across 16 ancestral generations each spanning 20 years. Each circle on the graph denotes a particular ethnic group and the size of the circle is proportional to the number of individuals from the ethnic group that are present in our data. The line corresponds to an unrestricted polynomial fit. Further, a regression of average attention on average ancestral variability and average ancestral variability squared gives a coefficient of -7.1 and 73.1 respectively both of which are significant at the 1% level. The plot does not control for any demographic or other ethnic group characteristics.

siderable variation in the average level of attention paid to the environment by individuals belonging to different ethnic groups and 2. The average level of attention paid to the environment is a U-shaped function of the average variability in climate conditions experienced by an ethnic group's ancestors.

We probe the second finding further, by robustly analysing the relationship between attention paid to the environment and the average variability of ancestral climate conditions at a disaggregated individual level. 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 country-by-year fixed effects. Our empirical strategy is equivalent to a continuous treatment framework where the treatment is the average variability of 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, in line with Proposition 1 reads as:

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

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} corresponds to our main dependent variable of interest, i.e. an individual's level of attention to environment captured by their normalized response to the question: *"How important is it to you to take care of the environment?"*
- Avg. Variability captures average climate variability across all ancestral generations (see Section 4.1).
- $\mathbf{X_{it}}$ corresponds to 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 which include primary mode of subsistence[29], economic development indicators (complexity of settlement[30], size of local community[31] and intensity of agriculture[32]) and level of local and global jurisdictional development[33]. 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.
- 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[34] 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.

[29]Fishing, Hunting, Gathering, Animal Husbandry and Pastoralism and Agriculture

[30]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

[31]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: $> 50,000$

[32]Ranging from 1 to 6 where 1 corresponds to no agriculture, 2 to casual agriculture, ... , 6 to intensive irrigated agriculture

[33]Ranging from 1 to 5, denoting the number of levels within a jurisdiction.

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

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

Dependent variable: Self reported measure of attention to environment				
Avg. variability	-6.291***	-7.095***	-9.194***	-9.877***
	(1.587)	(1.902)	(2.087)	(2.003)
Avg. variability 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. Variability	0.015	0.015	0.015	0.015
Max value Avg. Variability	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 variability refers to the average intensity of deviations from the typical climate conditions (*specific to each ancestral generation*) across generations. Average variability sq. refers to the square of the average variability term. Average variability 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 Koppen climate classification for the location of the ethnic group obtained from Ethnographic Atlas. Standard errors are clustered at the ethnicity level.

As laid down in our theoretical framework, the value of learning and correspondingly the level of attention is initially decreasing and then increasing in the prior (characterized by the scale of the distribution). The average climatic variability in our specifications is an empirical estimate of the scale (and hence the prior). As a result, the expected signs on β_1 and β_2 should be – and + respectively. This captures that the level of individual attention is (i) decreasing over the lower range of values of ancestral climatic variability because the diminution of benefits from exploitation dominates and (ii) increasing over higher values because the increase of the loss-protection motive for attention starts to dominate.

Table 1 below reports the results from Equation 1. Column 1 reports the results we obtain from regressing an individual's self reported measure of attention to environment on historical measures of ancestral climatic variability controlling for country-year fixed effects. Subsequent columns add individual level and historical controls. Across all specifications, our results corroborate the findings from the main proposition.

For the purpose of a stylized illustration, assume that the most stable ancestral climates scored 0.01 on our index of climate variability and the most unstable ancestral climates scored 0.09 on our index of climate variability.[35] 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 9% *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 similar percentage point *increase* in attention to environment today.

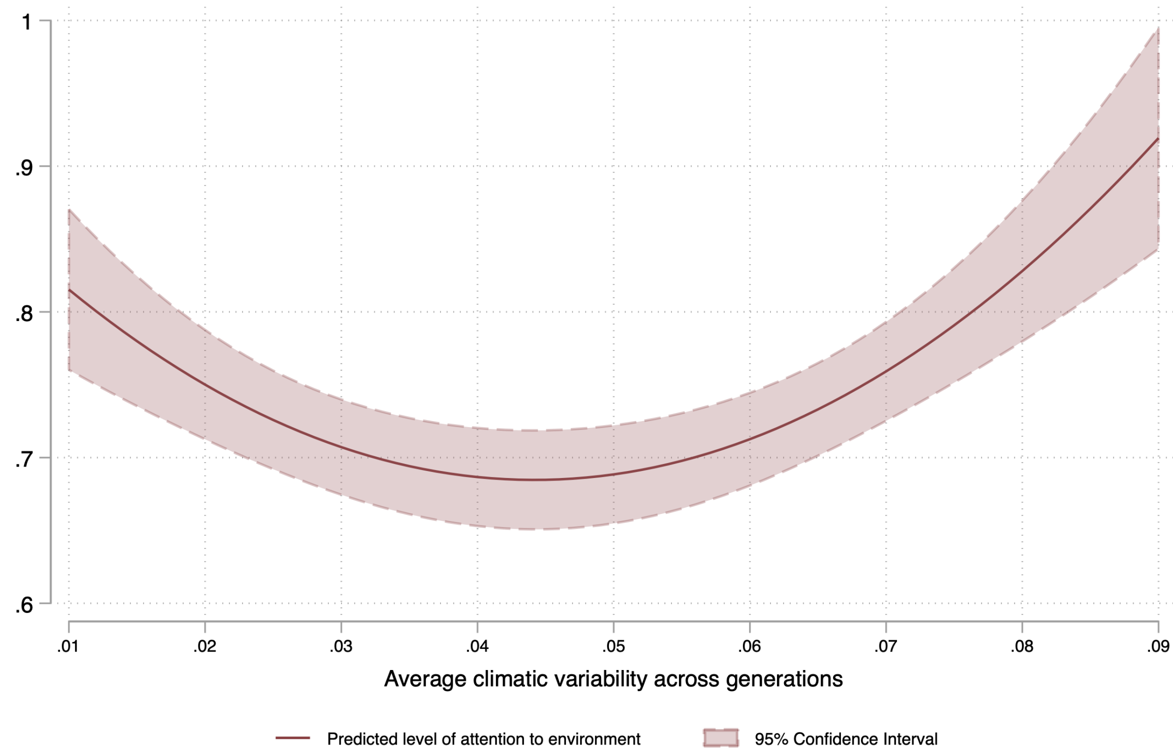
The results we find are all statistically significant at 1% level and consistent across different choice of sample or method of construction of our climate variability index. We provide more details on robustness checks in Subsection 5.3 below.

5.2 Environmental folklore - Group level analysis

Beyond the individual level analysis, the impact of ancestral average climatic variability on the level of attention can also be captured by group level variables such as stock of knowledge and collective memory related to environmental themes. Memories are stored and relayed in stories. A wider and a long lasting dissemination of these stories within communities across time and space often takes the form of folklore. An immediate consequence of our theoretical framework is if the ancestors of successive generations of a particular ethnic group face particular climate conditions that consequently increases each generation's level of attention paid to their environment, then this attention should result into a higher occurrence of environmental themes in the stock of folklore.

[35] For the sample in consideration, the average variability in ancestral climate conditions 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 99$^{\text{th}}$ percentile of the distribution of our values of average variability.

FIGURE 4: Margins plot - Individual's level of attention as a function of average ancestral climatic variability



Note: N = 138067. Margins plot corresponding to Equation 1. Average value of climate variability 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 variability (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.

We test this implication and provide suggestive evidence of impact of variability on group level measures of attention by regressing the level of environmental focus in ethnic folklore on our variable of average ancestral climatic variability. We use a sample of 1053 ethnic groups from the Michalopoulos and Xue [2021] database combined with the historical temperature data from Mann et al. [2009].[36] We run the following specification at the ethnicity level:

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

[36]As before we drop outliers whose Average variability values lie beyond 0.1 (99^{th} percentile of the distribution).

where e indexes an ethnic group, c indexes the country in which the ethnic group was located 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. Finally, 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. α_c corresponds to country of ethnic group fixed effects. We cluster the standard errors at the language group level as specified by Berezkin [2015]. Results from this specification are reported in Table 2.

As before, the coefficient on Avg. variability is negative and statistically significant at 5% level and the coefficient on Avg. variability 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. In vein of the same stylized illustration as before, an increase of 0.01 (~ 10 %) in average deviations from the typical climate conditions in the most stable environments lead to a ~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 ~1.7% increase in the proportion of environmentally related folklore.

5.3 Robustness checks and Heterogeneity Analysis

We now return to our sample of respondents from World Value Surveys and test the robustness of the U-shaped relationship between average ancestral climate variability and the level of attention paid to the environment by an individual. To ascertain that this relationship is not driven by the idiosyncrasies of the sample, our choice of lifespan / range of typical conditions or the misspecification of the functional form in the regression equation, we run a series of alternate specifications. We also shed some light on the role of cultural transmission in the formation of these priors by testing the sensitivity of the *U-shapedness* of attention w.r.t. the historical ethnic group development indicators.

Choice of sample: In practice, matching individuals to their corresponding ethnic ancestors through the use of spoken language may introduce systematic noise and confound our results. First, there might be reason to believe that mapping of big language groups speakers to their corresponding ethnic group is inherently imprecise, for e.g. a considerable population around the world might speak

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

Dependent variable: ln(1 + No. of environment related motifs/Total no. of motifs)				
Avg. variability	-3.233***	-3.172***	-2.493***	-1.548**
	(0.944)	(0.735)	(0.784)	(0.764)
Avg. variability 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. Variability	0.010	0.010	0.010	0.010
Max value Avg. Variability	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. Average variability refers to the average intensity of deviations from the typical climate conditions (*specific to each ancestral generation*) across generations. Average variability sq. refers to the square of the average variability term. Average variability 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. They also include a dummy for whether the primary mode of subsistence is dominated by females of the society. Historical topographic characteristics include controls for distance to the equator, distance to the closest coast and geographical Koppen climate classification for the location of the ethnic group obtained from Ethnographic Atlas. Standard errors are clustered at language group level.

English at home but that doesn't mean that they all share an ancestral lineage to the ethnic group originally belonging to the British Isles.[37] Therefore, if our result is driven by the mass of English speakers in the sample then we are capturing effects on attention to environment of other unobservable variables associated to the language rather than the effect of ancestral climatic variability. Independently, if a large part of our sample is composed of individuals belonging to a particular language group, we would also want to test the sensitivity of our results to the inclusion or exclusion of such groups.

To tackle this issue, we consecutively remove three of the largest language groups from our sample. We initially exclude English and Spanish speakers from our sample (Table 3 column (2)) and run Equation 1 again. In the next specification, in addition to English and Spanish speakers, we also remove Arabic speakers (Table 3 column (3)). Our results and their significance stay consistent across all such sample restrictions.

Second, for approximately 7% of the sample in consideration, we had proxied for the language spoken at home with the language of the interview.[38] Table 3 column (4) 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 variability 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 country-year fixed effects. The inclusion of these fixed effects tackles the aforementioned issue through two different sources of variation: identification of the coefficients β_1 and β_2 in our sample is based on 1. the comparison between individuals belonging to the same ethnic group but across different countries and 2. the comparison between individuals belonging to different ethnic groups staying in the same country at the time of survey. However, to robustly assuage this concern, we additionally run a specification (Table 3 column (5)) where we replace country-year fixed effects with region-year fixed effects.[39] That is, we increase the granularity of the contemporaneous spatial variable capturing the current location of the individual. In this case, the comparisons driving our results are between individuals of the same ethnic group but living in different regions within the same country or between

[37]Cultural transmission may happen not only from the experiences shared by ethnic ancestors in the form of bed time lullabies. It may also happen in the form of textbooks taught, stories foretold and in general the language used. If that is indeed the case then even when families adopt a language dissimilar to their original mother tongue, they also adopt the information passed on by the ancestors of that language. Our results primarily are agnostic on the channel of association apart from creating a link between individuals and one of their (potentially many) ancestors though shared language. We nevertheless still test the sensitivity of our results by excluding individuals where the relationship to the ancestors based on the language spoken at home might be loose.

[38]Since the language spoken at home was missing.

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

individuals of different ethnic groups living in the same region within a country. The signs and the statistical significance of the coefficients of interest under this specification stay the same as before.

Construction of ancestral climatic variability: We also test our results against our choice of assumptions to construct the measure of average climatic variability 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 (multi-rows) in Table 4 report the coefficients associated to average variability and average variability squared from the regression of 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.[40]

Inclusion of higher order terms, other moments and characteristics of overall temperature distribution: As a final robustness check, we test the sensitivity of our results to the inclusion of other variables capturing either the overall characteristics of the historical temperature distribution associated to a particular ethnic group or variation across generations within an ethnic group. We rerun Equation 1 with three different specifications: 1. While including higher order terms such as the cube of average ancestral climate variability runs the risk of overfitting the model, we nevertheless additionally include it in our specification to test whether the signs and statistical significance on the linear and the quadratic term is still preserved.[41] 2. Following suit from Giuliano and Nunn [2021], we include a term capturing the standard deviation across average temperature faced by each ethnic ancestral generation.[42] And 3. we include moments up to the 4^{th} order (mean, standard deviation, skewness and kurtosis) of the overall temperature distribution associated to an ethnic group over a span of 320

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

[41]We also try a specification where we include the 4^{th} order term of ancestral climatic variability. Plotting the margins plot from the corresponding regression still exhibit a U-shaped relationship. We do not report the results here since the 3rd order and the 4th order terms show a significant level of collinearity with the linear and quadratic terms.

[42]We do so to assuage any concerns of our variables acting as a proxy of dissimilarities in experiences across ancestral experiences. Since the lifespan of a generation can be redefined in a number of ways, our average ancestral climate variability variables may capture the standard deviation of mean temperatures faced by each generation within their lifetime.

years to assuage any concerns regarding the fact that our variables may capture some idiosyncratic geogrpahical characteristics of the ethnic group's location beyond the topographic and climate region controls already included in the original specification. Results are reported in Table 5.

As before, both the statistical significance and signs on our variable of interest are broadly robust to the inclusion of these additional terms.

Heterogeneous effects by the level of development of ancestral ethnic groups: Can similar temperature deviations still produce different ancestral experiences and thus differentially impact contemporaneous attention to environment? As a final test of our results, we probe our claim of cultural transmission impacting an individual's level of attention to environment by exploiting the heterogeneity in ancestral experiences and characteristics at the ethnic group level. 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 [43] is empirically estimated through the expectation of variability distribution (*up to a multiplicative constant*), 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 this by rerunning Equation 1 separately on sub-samples bifurcated by the level of economic and institutional development at the ethnic group level. We do it in four different ways, i.e. we split the sample by whether the ethnic ancestors of an individual 1. were hunter gatherers or non-hunter gatherers, 2. had high / low level of economic complexity, 3. had high / low level of local jurisdiction and 4. had high / low level of global jurisdiction. Our first sample split relies on the intuition that social groups in their process of economic evolution often transition from hunting and gathering to animal husbandry / agriculture over a period of time. The fact that an individual's ethnic ancestors were hunter-gatherers implies that both due to the primary subsistence activity and the level of development of the ancestors made them more susceptible to climate variations. For our second sample split, we categorize ethnic groups under high level of economic development if they had compact and permanent settlements and under low level of economic development otherwise. For our third sample split, we use the definition of local jurisdictional heirarchy and categorize all ethnic groups as highly institutionally / economically developed if the jurisdiction exhibited a structure beyond family and band ties, i.e. for e.g. the existence of a clan and village. For our final sample split, we categorized ethnic groups as highly developed if any level of global jurisdiction existed beyond the local jurisdiction. All these variables, in one form or another, capture historically the level of protection against / susceptibility towards climate shocks a particular ethnic group may have had.

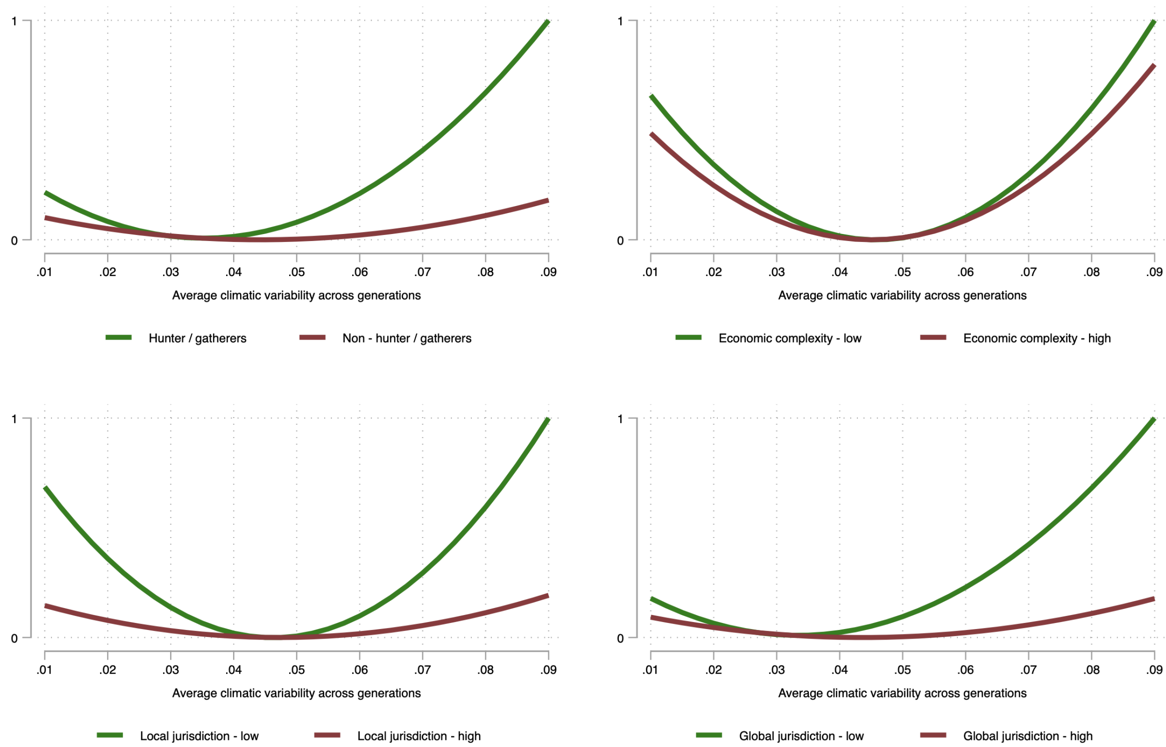
[43] which drives their attention to the climate

The main purpose of this exercise is to compare the slopes across the two samples, not the levels. This is because historical economic development independently can impact other factors such as the level of education or awareness towards climate issues in general. What we want to analyse is 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. We do so in the following ways: we normalize them such that the minimum point of the predicted level of attention function from both the sub-samples are equalized. Further, all values in the predicted attention function are normalized to take a value between 0 and 1. Figure 5 plots these predicted attention functions for each sub-sample under the four different types of sample splits. In all cases, we see that the *U-shapedness* of the attention function is relatively more salient when the level of economic / institutional development is low. This suggests that cultural transmission indeed has a role to play in the formation of an individual's prior today. Put differently, the experiences of the ancestors pertaining to their environment seem to be the key variable, and not the historical temperature distribution itself.

6 Conclusion

The starting point of this paper is a simple intuition: the amount of attention that people pay to climate-related issues is influenced by their ancestors' climatic experiences through cultural transmission mechanisms. We formalize this intuition into an empirically testable hypothesis. By combining four different data sources, we test the effect of ancestral climatic variability on an individual's and a group's level of attention to environment. All our results at the individual level and further robustness checks confirm the existence of a significant impact of ancestral experience on stated attention. Further, it exhibits a robust U-shaped pattern: the marginal effect of ancestral climate variability on attention is first decreasing and then increasing. In other words, individuals whose ancestors have faced most consistently stable or unstable climates care most about environmental issues, with a dip for intermediate realizations. We propose a simple explanation for this fact, which is rooted in a flexible theoretical framework. We assume that the intrinsic value of attention to climate issues is driven by a dual purpose of learning – exploitation of typical conditions and protection against extreme events. To analyze the effect of this structure, we construct a simple model of costly attention where the stakes of learning are contingent on climate variability. The effect of cultural transmission is captured sparsely by assuming that the *perceived stakes* of an individual are driven by prior beliefs which aggregate ancestral experiences, using a general scale parametrization. Our main result robustly predicts the U-shaped property of attention under mild and generic assumptions, thereby highlighting that the result is really driven by the "dual purpose of learning" assumption. Our theoretical framework provides a flexible reduced-form way to study the (potentially intricate) effect of the distribution of ancestral experiences on descendant's attention. We hope that this provides a

FIGURE 5: 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.

novel methodological contribution that can find more applications in future work. To further substantiate and study the effect in our context, as well as indirectly substantiate the narrative of cultural transmission, we perform our main regression with a different outcome variable: the prevalence of environmental themes in folklore. This exhibits the same U-shape, which is consistent with interpreting retained folklore as a stock which is fueled by successive generation's attention to particular issues. We also perform heterogeneity analysis that highlights a differential sensitivity to ancestral experiences according to several intuitive characteristics like the level of economic development.

The formation of beliefs and preferences about environmental issues is a first order concern in the age of climate change. Although it is a complex phenomenon with many causal channels, we hope to shed some light on one specific aspect which is the effect of culture and socialization. Our work ties in more broadly to the active research agenda on cultural transmission by studying a specific application and context, as well as providing a novel approach. Much more work remains to be done in understanding and unpacking the effects of cultural transmission and the formation of climate preferences. We hope that this project can provide grounds for further analysis by providing clear evidence and a simple analytical framework of the effect of experienced ancestral climate variability on attention to climate issues.

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

Dependent variable: Self reported measure of attention to environment					
Avg. variability	-9.877***	-9.243***	-8.903***	-9.897***	-6.649***
	(2.003)	(1.988)	(2.206)	(2.011)	(2.023)
Avg. variability sq.	111.767***	107.503***	103.956***	111.862***	71.929***
	(21.346)	(21.557)	(25.184)	(21.505)	(22.402)
Income	-0.001*	-0.001+	-0.002+	-0.001+	-0.001+
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Male	-0.013***	-0.012***	-0.014***	-0.013***	-0.014***
	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)
Age	0.002***	0.002***	0.002***	0.001***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Education level	0.006***	0.006***	0.007***	0.006***	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
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
Country-year Fixed effects	Y	Y	Y	Y	N
Region-year Fixed effects	N	N	N	N	Y
Sample restrictions					
English speakers excluded	N	Y	Y	N	N
Spanish speakers excluded	N	Y	Y	N	N
Arabic speakers excluded	N	N	Y	N	N
Lang. at home only	N	N	N	Y	N
Mean of dep var	0.704	0.700	0.689	0.703	0.705
St. Dev. of dep var	0.250	0.251	0.249	0.251	0.250
Min value Avg. Variability	0.015	0.015	0.015	0.015	0.015
Max value Avg. Variability	0.093	0.093	0.093	0.093	0.093
R-sq	0.113	0.122	0.118	0.114	0.160
Adj. R-sq	0.113	0.121	0.117	0.113	0.151
N	138067	110484	93184	128532	135228
*** 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 attentiopn. 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 variability refers to the average intensity of deviations from the typical climate conditions (*specific to each ancestral generation*) across generations. Average variability sq. refers to the square of the average variability term. Average variability 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 the closest coast and geographical Koppen 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.

TABLE 4: Coefficient on average variability of ancestral climatic conditions - 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)

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. 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 variability term. Every regression includes controls for demographic characterstics, 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 variability 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 20th and 80th percentile of 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

TABLE 5: Coefficients for the impact of average variability of ancestral climatic conditions on individual's self-reported attention to the environment

Dependent variable: Self reported measure of attention to environment				
Avg. variability	-9.877***	-17.402**	-9.055***	-9.677***
	(2.003)	(7.035)	(2.108)	(2.135)
Avg. variability sq.	111.767***	276.862*	105.537***	109.681***
	(21.346)	(145.560)	(21.700)	(21.829)
Avg. variability cub.		-1130.897		
		(951.849)		
Variability across generations			-0.166	
			(0.123)	
Occupation category Controls	Y	Y	Y	Y
Historical ethnic group characteristics	Y	Y	Y	Y
Historical topographic characteristics	Y	Y	Y	Y
Moments of full temp. distribution	N	N	N	Y
Country-year Fixed effects	Y	Y	Y	Y
Mean of dep var	0.704	0.704	0.704	0.704
St. Dev. of dep var	0.250	0.250	0.250	0.250
Min value Avg. Variability	0.015	0.015	0.015	0.015
Max value Avg. Variability	0.093	0.093	0.093	0.093
R-sq	0.113	0.114	0.114	0.114
Adj. R-sq	0.113	0.113	0.113	0.113
N	138067	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 variability refers to the average intensity of deviations from the typical climate conditions (*specific to each ancestral generation*) across generations. Average variability sq. (cub.) refers to the square (cube) of the average variability 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 temeprature distribution associated to an ethnic group in the 320 years (1600-1920) used to construct the per-generation variability index. Average variability 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 Koppen climate classification for the location of the ethnic group obtained from Ethnographic Atlas. Standard errors are clustered at the ethnicity level.

References

Fabrizio Adriani and Silvia Sonderegger. Why do parents socialize their children to behave prosocially? an information-based theory. *Journal of Public Economics*, 93(11-12):1119–1124, 2009. 6

Fabrizio Adriani and Silvia Sonderegger. Signaling about norms: Socialization under strategic uncertainty. *The Scandinavian Journal of Economics*, 120(3):685–716, 2018. 6

Fabrizio Adriani, Jesse A Matheson, and Silvia Sonderegger. Teaching by example and induced beliefs in a model of cultural transmission. *Journal of Economic Behavior & Organization*, 145:511–529, 2018. 6

George A Akerlof and Dennis J Snower. Bread and bullets. *Journal of Economic Behavior & Organization*, 126:58–71, 2016. 5

Alberto Alesina, Paola Giuliano, and Nathan Nunn. On the origins of gender roles: Women and the plough. *The quarterly journal of economics*, 128(2):469–530, 2013. 5

Martin Q. Amlor and Matthew Q. Alidza. Indigenous Education in Environmental Management and Conservation in Ghana: The Role of Folklore. *Journal of Environment and Ecology*, 7(1), 2016. 5

Peter Andre, Teodora Boneva, Felix Chopra, and Armin Falk. Fighting Climate Change: The Role of Norms, Preferences, and Moral Values. Working Papers 2021-036, Human Capital and Economic Opportunity Working Group, July 2021. URL `https://ideas.repec.org/p/hka/wpaper/2021-036.html`. 2, 4

Patrick Anthony. Labour, folklore, and environmental politics in german mining around 1800. *The Historical Journal*, 64(3):583–605, 2021. 5

Maximilian Auffhammer. Quantifying Economic Damages from Climate Change. *Journal of Economic Perspectives*, 32(4):33–52, Fall 2018. 2

Duman Bahrami-Rad, Anke Becker, and Joseph Henrich. Tabulated nonsense? testing the validity of the ethnographic atlas. *Economics Letters*, 204:109880, 2021. 19

Anke Becker. On the economic origins of restrictions on women's sexuality. 2019. 5

Anke Becker, Benjamin Enke, and Armin Falk. Ancient origins of the global variation in economic preferences. In *AEA Papers and Proceedings*, volume 110, pages 319–23, 2020. 5

James Bell, Jacob Poushter, Moira Fagan, and Christine Huang. In response to climate change, citizens in advanced economies are willing to alter how they live and work. Technical report, September 2021. 2

Yuri Berezkin. Folklore and mythology catalogue: Its lay-out and potential for research. *The Retrospective Methods Network Newsletter*, 10:58–70, 2015. 20, 27

Alberto Bisin and Thierry Verdier. The economics of cultural transmission and the dynamics of preferences. *Journal of Economic theory*, 97(2):298–319, 2001. 5, 6

Alberto Bisin and Thierry Verdier. The economics of cultural transmission and socialization. In *Handbook of social economics*, volume 1, pages 339–416. Elsevier, 2011. 6

Alberto Bisin and Thierry Verdier. Advances in the economic theory of cultural transmission. *Annual Review of Economics*, 15, 2023. 5

Dmitri Bondarenko, Alexander Kazankov, Daria Khaltourina, and Andrey Korotayev. Ethnographic Atlas XXXI: Peoples of Easternmost Europe. *Ethnology*, 44(3):261–289, 2005. 3, 19

Johannes C Buggle and Ruben Durante. Climate risk, cooperation and the co-evolution of culture and institutions. *The Economic Journal*, 131(637):1947–1987, 2021. 4, 5

Andrew Caplin and Mark Dean. Revealed preference, rational inattention, and costly information acquisition. *American Economic Review*, 105(7):2183–2203, 2015. 5

Andrew Caplin, Mark Dean, and John Leahy. Rationally inattentive behavior: Characterizing and generalizing shannon entropy. *Journal of Political Economy*, 130(6):1676–1715, 2022. 5

Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory 2nd Edition (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, July 2006. ISBN 0471241954. 10

Paola Giuliano and Nathan Nunn. Ancestral Characteristics of Modern Populations. *Economic History of Developing Regions*, 33(1):1–17, January 2018. 3, 16, 18, 19, 46

Paola Giuliano and Nathan Nunn. Understanding cultural persistence and change. *The Review of Economic Studies*, 88(4):1541–1581, 2021. 3, 5, 17, 30

RG Gordon. Ethnologue: Languages of the world, 16th ed. *SIL International, Dallas*, 2009. 3, 19

James W Hansen, Sabine M Marx, and Elke U Weber. 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. 2004. 4

Yuval Noah Harari. Sapiens: A brief history of humankind. *Publish in agreement with The Deborah Harris Agency and the Grayhawk Agency*, 2014. 5

Peter D. Howe, Matto Mildenberger, Jennifer R. Marlon, and Anthony Leiserowitz. Geographic variation in opinions on climate change at state and local scales in the USA. *Nature Climate Change*, 5 (6):596–603, June 2015. 2, 44

IPCC. Climate change 2013: The physical science basis. contribution of working group i to the fifth assessment report of the intergovernmental panel on climate change. Technical report, 2013. 2

IPCC. Climate change 2014: Synthesis report. contribution of working groups i, ii and iii to the fifth assessment report of the intergovernmental panel on climate change. Technical report, 2014. 2

Etienne Le Rossignol, Sara Lowes, and Nathan Nunn. Traditional supernatural beliefs and prosocial behavior. Technical report, National Bureau of Economic Research, 2022. 4

Anthony Leiserowitz. Climate Change Risk Perception and Policy Preferences: The Role of Affect, Imagery, and Values. *Climate Change*, 77:45–72, 2006. 4

Sara Lowes. Matrilineal kinship and spousal cooperation: Evidence from the matrilineal belt. Technical report, UCSD, 2018. 4

Sara Lowes, Nathan Nunn, James A Robinson, and Jonathan L Weigel. The evolution of culture and institutions: Evidence from the kuba kingdom. *Econometrica*, 85(4):1065–1091, 2017. 4

Yu Luo and Jiaying Zhao. Motivated attention in climate change perception and action. *Frontiers in Psychology*, 10, 2019. 2, 4

Bartosz Maćkowiak and Mirko Wiederholt. Optimal sticky prices under rational inattention. *American Economic Review*, 99(3):769–803, 2009. 5, 11

Bartosz Maćkowiak, Filip Matějka, and Mirko Wiederholt. Rational inattention: A review. *Journal of Economic Literature*, 61(1):226–273, 2023. 5, 10, 42

Michael E Mann, Zhihua Zhang, Malcolm K Hughes, Raymond S Bradley, Sonya K Miller, Scott Rutherford, and Fenbiao Ni. Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia. *Proceedings of the National Academy of Sciences*, 105(36): 13252–13257, 2008. 16

Michael E. Mann, Zhihua Zhang, Scott Rutherford, Raymond S. Bradley, Malcolm K. Hughes, Drew Shindell, Caspar Ammann, Greg Faluvegi, and Fenbiao Ni. Global signatures and dynamical origins of the little ice age and medieval climate anomaly. *Science*, 326(5957):1256–1260, 2009. 3, 16, 19, 21, 26, 45

Filip Matějka and Alisdair McKay. Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, 105(1):272–298, 2015. 5

Filip Matějka and Guido Tabellini. Electoral competition with rationally inattentive voters. *Available at SSRN*, 2017. 5

Stelios Michalopoulos and Melanie Meng Xue. Folklore*. *The Quarterly Journal of Economics*, 136(4): 1993–2046, 01 2021. 4, 5, 16, 20, 21, 26, 27, 46

George Peter Murdock. World ethnographic sample. *American Anthropologist*, 59:664–687, 1957. 3, 19

Gbadebo J Osemeobo. The role of folklore in environmental conservation: evidence from edo state, nigeria. *The International Journal of Sustainable Development & World Ecology*, 1(1):48–55, 1994. 5

Fabrizio Panebianco and Thierry Verdier. Paternalism, homophily and cultural transmission in random networks. *Games and Economic Behavior*, 105:155–176, 2017. 6

Doron Ravid. Ultimatum bargaining with rational inattention. *American Economic Review*, 110(9): 2948–2963, 2020. 5

Edmond Sanganyado, Charles Teta, and Busani Masiri. Impact of african traditional worldviews on climate change adaptation. *Integrated Environmental Assessment and Management*, 14(2):189–193, 2018. 5

Jing Shi, Vivianne HM Visschers, Michael Siegrist, and Joseph Arvai. Knowledge as a driver of public perceptions about climate change reassessed. *Nature Climate Change*, 6(8):759–762, 2016. 4

Robert J Shiller. Narrative economics. *American economic review*, 107(4):967–1004, 2017. 5

Christopher A Sims. Implications of rational inattention. *Journal of monetary Economics*, 50(3):665–690, 2003. 5

Robyn Speer, Joshua Chin, and Catherine Havasi. Conceptnet 5.5: An open multilingual graph of general knowledge. In *Thirty-first AAAI conference on artificial intelligence*, 2017. 21

Guido Tabellini. The scope of cooperation: Values and incentives. *The Quarterly Journal of Economics*, 123(3):905–950, 2008. 6

Laura L. Veldkamp. *Information Choice in Macroeconomics and Finance*. Princeton University Press, 2011. 5

Elke U. Weber. What shapes perceptions of climate change? *Advanced Review*, 1(3):332–342, 2010. 4

EU Weber. 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), 1997. 4

WVS. European values study and world values survey: Joint evs/wvs 2017-2021 dataset (joint evs/wvs). 2020. 16, 46

Appendix

A Omitted proofs

Most of the derivations of Subsection 3.2 are either straightforward or very usual in the rational inattention literature, therefore we do not provide detailed proofs; the reader can refer to the survey by Maćkowiak et al. [2023] for a review of the literature and techniques used here.

The only result that warrants a detailed proof beyond the arguments mentioned in the main text's analysis of our theoretical framework is Proposition 1. The proof is simple and intuitive. Nevertheless, it is included for completeness. Recall that for the statement of Proposition 1, we are working with the generalized version of the attention of model, which is given by:

$$\max_{\xi\in[0,1]} \int_{\mathbb{R}_+} \xi(W(\eta)+L(\eta))\,dp_0(\eta) - c(\xi) \tag{$\bigstar$}$$

under the assumptions (i)-(iii) on W,L,c detailed in Subsection 3.2. Recall that we assume that the prior p_0 over η belongs to a parametrized scale family of absolutely continuous distributions, i.e there is a reference distribution $p \in \Delta(\mathbb{R}_+)$ which has a density[44] and a scale parameter $\theta \in \mathbb{R}_+$ such that $p_0(\eta) = \frac{1}{\theta} p\left(\frac{\eta}{\theta}\right)$.

This asssumption allows us to rewrite the problem as:

$$\max_{\xi\in[0,1]} (\mathscr{W}(\theta)+\mathscr{L}(\theta))\xi - c(\xi)$$

Where we have defined:

$$\mathscr{W}(\theta) := \int_{\mathbb{R}_+} W(\eta)\,dp_0(\eta)$$
$$\mathscr{L}(\theta) := \int_{\mathbb{R}_+} L(\eta)\,dp_0(\eta)$$

It is straightforward to observe that $\mathscr{W}$ and $\mathscr{L}$ inherit the properties of W and L. Indeed, taking for instance W, observe that by performing a simple change of variable:

$$\mathscr{W}(\theta) := \int_{\mathbb{R}_+} W(\eta)\frac{1}{\theta}\,dp\left(\frac{\eta}{\theta}\right) = \int_{\mathbb{R}_+} W(\theta\eta)\,dp(\eta)$$

Hence $\mathscr{W}$ is $\mathscr{C}^2$ and:

$$\mathscr{W}'(\theta) = \int_{\mathbb{R}_+} \eta w'(\theta\eta)\,dp(\eta) \le 0$$

[44] We abuse notations and identify the measure and its density.

$$\mathscr{W}''(\theta) = \int_{\mathbb{R}_+} \eta^2 w''(\theta\eta) dp(\eta) \geq 0$$

and $\mathscr{W}' \underset{\infty}{\longrightarrow} 0$ follows from dominated convergence. Similarly, we get $\mathscr{L} \geq 0$, $\mathscr{L}$ is $\mathscr{C}^2$ and $\mathscr{L}' \geq 0$, $\mathscr{L}'' \geq 0$ and $\mathscr{L}'(0) = 0$.

Given our assumptions on c, it is clear that the optimal level of attention as a function of the scale parameter θ is 0 if $W(\theta) + L(\theta) < c'(0)$ and given by a first-order condition otherwise, hence:

$$\xi^*(\theta) = \begin{cases} 0 & \text{if } \mathscr{W}(\theta) + \mathscr{L}(\theta) < c'(0) \\ (c')^{-1}\Big(\mathscr{W}(\theta) + \mathscr{L}(\theta)\Big) & \text{otherwise} \end{cases}$$

Where $(c')^{-1}$ denotes the inverse of the marginal cost function. Recall that by assumption c' is an increasing nonnegative function therefore its inverse is an increasing nonnegative function. Furthermore, the function $F : \theta \mapsto \mathscr{W}(\theta) + \mathscr{L}(\theta)$ is also convex (this is immediate since $\mathscr{W}'', \mathscr{L}'' \geq 0$) and single-throughed. Indeed, observe that $(\mathscr{W}' + \mathscr{L}')(0) = \mathscr{W}'(0) \leq 0$ and since $\mathscr{W}'$ goes to zero at infinity and L' is positive and strictly increasing, there must exist some smallest $\tilde{\theta}$ such that for all $\theta \geq \tilde{\theta}$, $(\mathscr{W}' + \mathscr{L}')(\theta) \geq 0$. By continuity and monotonicity for all $\theta \leq \tilde{\theta}$, $(\mathscr{W}' + \mathscr{L}')(\theta) \leq 0$.

Consider now the level set $\mathscr{S} := \{\theta | F(\theta) < c'(0)\}$. Because of the convexity of F, this set must be an interval (potentially empty): we denote $(\underline{\theta}, \overline{\theta}) = \mathscr{S}$ where $\underline{\theta} < \overline{\theta}$ iff $\mathscr{S} \neq \emptyset$. We extend our notation to the case $\mathscr{S} = \emptyset$ by setting $\underline{\theta} = \overline{\theta} = \tilde{\theta}$ in this case. Furthermore, because of the monotonicity of F and our notational convention for the case $\mathscr{S} = \emptyset$ it is clear that we must have $\underline{\theta} \leq \tilde{\theta} \leq \overline{\theta}$

Now observe the following

- Over $[0, \underline{\theta})$ ξ^* is differentiable and we have:

$$(\xi^*)'(\theta) = \underbrace{(\mathscr{W}'(\theta) + \mathscr{L}'(\theta))}_{\leq 0 \text{ because } \theta < \underline{\theta} \leq \tilde{\theta}} \times \underbrace{((c')^{-1})'(\mathscr{W}(\theta) + \mathscr{L}(\theta))}_{\geq 0}$$

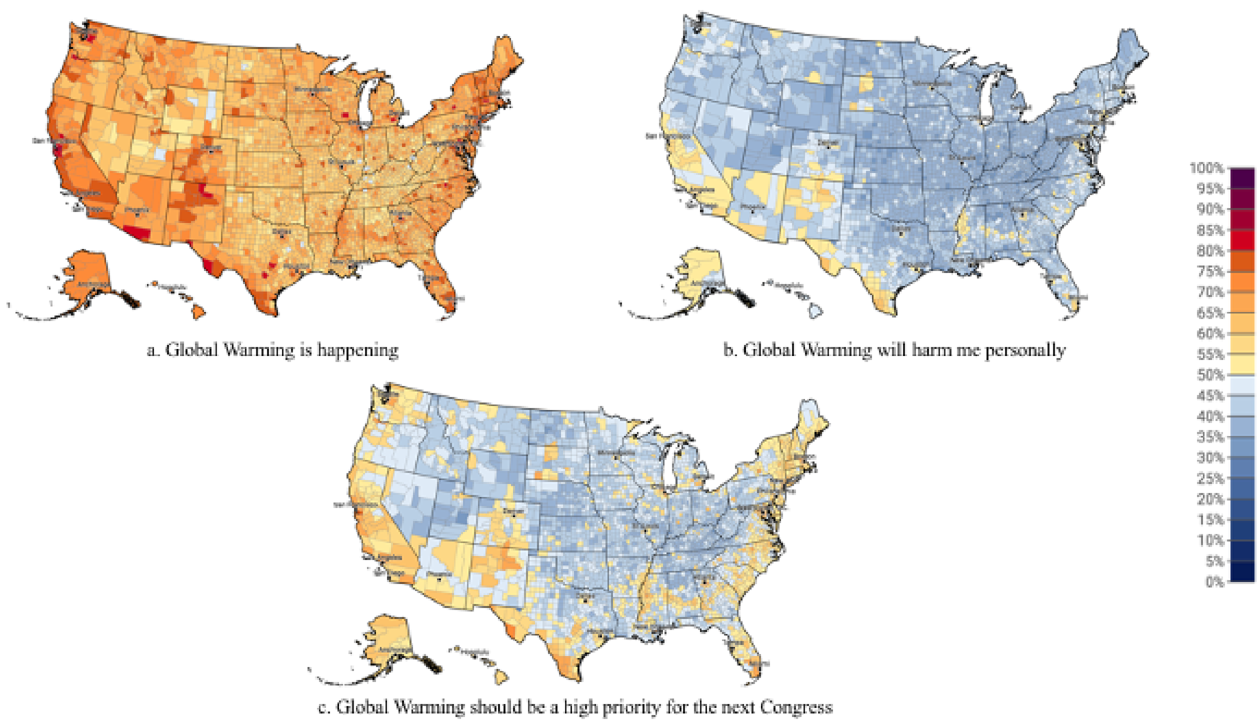
- Over $(\underline{\theta}, \overline{\theta})$, $\xi^* \equiv 0$ is constant.
- Over $(\overline{\theta}, \infty)$ ξ^* is differentiable and we have:

$$(\xi^*)'(\theta) = \underbrace{(\mathscr{W}'(\theta) + \mathscr{L}'(\theta))}_{\geq 0 \text{ because } \theta > \overline{\theta} \geq \tilde{\theta}} \times \underbrace{((c')^{-1})'(\mathscr{W}(\theta) + \mathscr{L}(\theta))}_{\geq 0}$$

Hence putting it all together since ξ^* is clearly continuous, we obtain the result of Proposition 1.

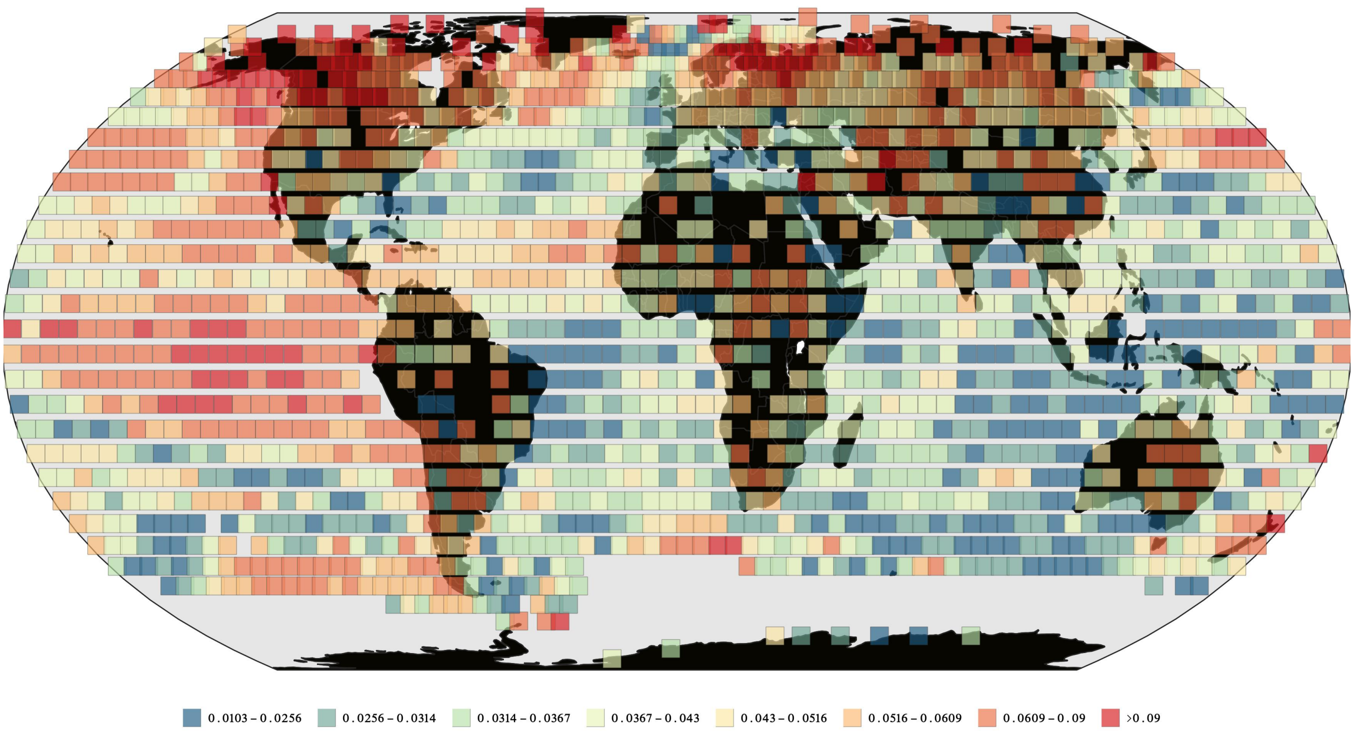
B Additional Figures and Tables

FIGURE 6: County-wise distribution of climate Opinions within the US (2021)

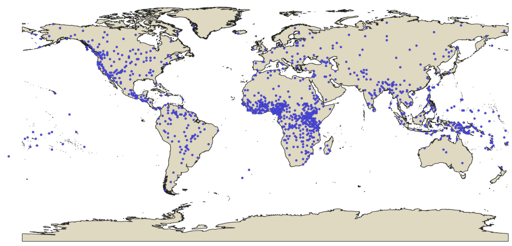


Source: Yale Climate Communication [2021], Howe et al. [2015]. The figure reflects geographical variation within the United States of America at a county level in terms of proportion of residents who believe global warming is happening (top-left figure), if global warming will harm them personally (top-right figure) and if global warming should be a high priority for the next Congress (bottom figure). National average for the 3 questions stands at 72%, 47% and 55% respectively.

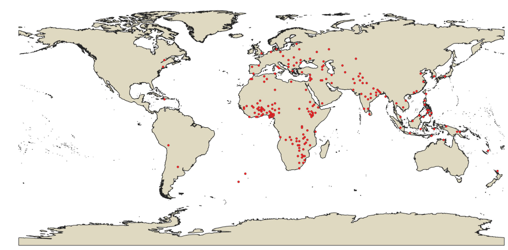
FIGURE 7: Spatial distribution of Average Ancestral Climatic Variability



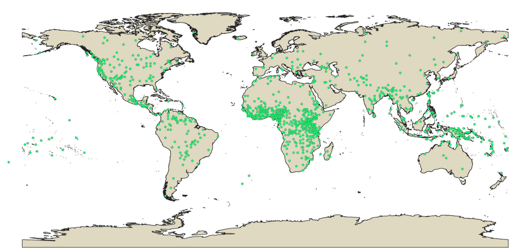
Note: Distribution of average ancestral climatic variability constructed from Mann et al. [2009]. The variable reports the average intensity of deviations from a defined range of normal temperatures specific to a generation. Generation lifespan is assumed to be 20 years and the normal range of temperature are the temperatures within the 20th and 80th 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 Subsection 4.1.



((A)) Ethnic groups available in Giuliano and Nunn [2018]



((B)) Ethnic groups matched to the WVS [2020] question in consideration



((C)) Ethnic groups matched to the Michalopoulos and Xue [2021] database

FIGURE 8: Geographic location of ethnic groups in our sample of analysis

Note: Approximate centroid of the location of ethnic groups in our sample. Subfigure [A] provides the location of all ethnic groups in the Giuliano and Nunn [2018]. Subfigure [B] provides the location of the ethnic groups whose individuals answered the environmental question of interest (*"How important is it for you to take care of the environment?"*) in wave 5 and 6 of World Value Surveys. Subfigure [C] provides the location of ethnic groups that could be matched to Michalopoulos and Xue [2021].

TABLE 6: Summary statistics - demographic characteristics (World Value Surveys)

Variable	Mean	Std. Dev.	Min.	Max.	N
Income Category	4.782	2.176	1	10	148620
Male	0.479	0.5	0	1	157010
Age	41.659	16.52	15	102	156800
Education level	4.815	2.202	1	8	145688
Agriculture	0.037	0.19	0	1	157142
Armed Forces	0.005	0.071	0	1	157142
Employer / Manager	0.025	0.155	0	1	157142
Manual	0.117	0.321	0	1	157142
Non-Manual	0.099	0.298	0	1	157142
Never had a job	0.017	0.13	0	1	157142
Other	0.002	0.05	0	1	157142

FIGURE 9: 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.

FIGURE 10: Non-Tokenized words from the description of motifs available in Berezkin's catalog

Note: Word cloud of non-tokenized words from the descriptions of motifs available in Berezkin's catalog. Russian motifs are translated into English and stop words are removed.