Journal of Economic Behavior and Organization 244 (2026) 107494

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

journal homepage: www.elsevier.com/locate/jebo

Research Paper

The cultural foundations of cooperation enforcement

Max Otto

University of Mannheim; A5, 6, 68159, Mannheim, Germany

ARTICLE INFO

Keywords:
Cooperation
Morality
Enforcement mechanisms
Reputation
Revenge
Cross-cultural analysis

ABSTRACT

The human ability to maintain large-scale cooperation despite individual incentives for defection is as perplexing as it is universal. Although research has identified several mechanisms as potential solutions to this puzzle, empirical evidence robustly shows that societies rely on vastly different enforcement practices to achieve cooperative outcomes. The paper addresses this poorly understood phenomenon by developing a game-theoretical model that links cultural differences to incentives to utilize reputation, reciprocity, and revenge to elicit cooperative behavior. The model shows that reputational enforcement is favored by large temporal discounting, group-bound morality, and high degrees of dyadic reciprocity. In contrast, revenge is supported by honor beliefs and cultural long-term orientations, while being inhibited by a greater reliance on reciprocal relationships. A comprehensive empirical evaluation of these predictions across two global datasets spanning different historical periods provides strong support for the theory. The identified relationships are robust across preindustrial societies and a newly compiled dataset of contemporary sub-national regions, suggesting that the proposed framework captures general mechanisms underlying cross-cultural variation in cooperation enforcement.

1. Introduction

Humans cooperate. Despite ubiquitous opportunities for betrayal and exploitation, individuals in virtually all known societies overcome these temptations on scales that far exceed the narrow circles of kinship that often explain similar behavior in other animals (Henrich and Henrich, 2007). Investigations into this puzzle constitute one of the core concerns of the social sciences and have produced a set of plausible explanations through the identification of stable enforcement mechanisms. However, an abundance of ethnographic evidence has long suggested that the ways in which societies stabilize cooperation among their members, that is, the *specific* enforcement mechanisms they rely on, are subject to substantial variation (Garfield et al., 2023; Molho et al., 2024).

Examples of this diversity are plentiful. Among Pashtun tribes, the moral code of Pukhtunwali demands violent revenge ("badal") as a deterrent against dishonesty (Mahdi, 1986); Chinese clans historically relied on collective exclusion from future interactions rather than individual revenge (Liu, 1959); while traditional Quechuan agricultural communities maintained cooperation through norms of reciprocal obligation ("ayni"; Erasmus, 1956). In contemporary contexts, this finding of enforcement heterogeneity is increasingly supported by more systematic empirical evidence. Cross-national cooperation experiments repeatedly demonstrate not only that participants' behavior is highly variable but also that this variability is predominantly explained by cultural background (Balliet and van Lange, 2013; Gächter and Herrmann, 2009; Henrich et al., 2006; Herrmann et al., 2008; Marlowe et al., 2008). Similarly, surveys on enforcement preferences mirror these systematic differences (Eriksson et al., 2017) and suggest the existence of distinct patterns of mechanism co-occurrences (Eriksson et al., 2021).

E-mail address: max.otto@uni-mannheim.de

https://doi.org/10.1016/j.jebo.2026.107494
Received 16 July 2025; Received in revised form 31 December 2025; Accepted 15 February 2026
Available online 21 February 2026

In combination, this growing body of evidence points to an unambiguous conclusion: Although cooperative behavior itself may be universal in human societies (Curry et al., 2019), the ways in which it is enforced are evidently not. However, the description of this variation in enforcement types naturally raises the more fundamental question about their origins, which has received much less attention (see Molho et al., 2024 for a recent review). Addressing this puzzle is the central aim of this paper.

At its core, cooperation is a problem of strategic interdependence: Incentives for cooperation can exist only if cooperation is anticipated to be reciprocated, and such anticipation is justified only if cooperation is indeed incentivized. The same logic extends naturally to the enforcement practices that produce these incentives in the first place. Any sanction or reward aimed at promoting cooperation is credible only if it is expected to be effective — and effective only if it is perceived as credible. This creates a fundamental challenge. For enforcement practices to successfully elicit cooperative behavior, they must be self-sustaining, which means that abandoning enforcement norms cannot be individually rational even when adhering to them is costly. Via a stylized game-theoretical model, this paper addresses two central questions arising from this problem:

1. Which configurations of enforcement mechanisms can be sustained in cooperative equilibria? Since different enforcement strategies are likely interdependent, understanding cooperation enforcement requires an *equilibrium definition* — a formalization of the functional structure underlying stable patterns of co-occurrence among enforcement types.
2. What factors induce divergent equilibria within this space? The existence of stable enforcement strategies neither guarantees their adoption nor accounts for cross-societal variation in enforcement mechanisms. This requires addressing the problem of *equilibrium selection*.

The paper specifically addresses the origins and interdependencies of *decentralized* ways of enforcing cooperation, focusing on the three most important enforcement mechanisms identified in the literature: reciprocity, reputation, and retaliation. The central argument posits that the interrelation of stable enforcement types is exhaustively defined by the comparative costs of their second-order stability — that is, the relative incentive structure to uphold the respective reward or sanctioning structures. I then argue that cultural beliefs — shared and socially learned information — shape stable expectations about others' behavior, which ultimately give rise to divergent equilibria across distinct enforcement practices.

This approach shows that reliance on reputational control is facilitated by larger degrees of dyadic reciprocity, short-term orientations, and moral beliefs that limit ethical expectations to in-group members. Conversely, retaliatory punishment is more likely to develop as an enforcement device when dyadic reciprocity is relatively unimportant, cultures are more future-oriented, and moral beliefs of honor (rather than dignity) are prevalent.

To assess the generalizability and explanatory power of these theoretical results, their subsequent empirical evaluation draws on multiple data sources spanning two global datasets from different historical periods. First, I analyze ethnographic data from 128 preindustrial small-scale societies. I combine recently coded information on reputational sanctions by Garfield et al. (2023) with originally coded data on revenge practices and textual data on traditional folklore by Michalopoulos and Xue (2021), from which I infer cultural traits. I then turn to an analysis of contemporary survey data. Specifically, I compile a new dataset by aggregating four large multinational survey programs, which I manually harmonize to the subnational level. The resulting analytical sample includes 272 regions within 46 countries across all six inhabited continents. For both the historical and contemporary datasets, a range of identification strategies and model specifications provides strong support for the explanatory structure implied by the theoretical model.

The development of the article's arguments proceeds as follows. I first provide a brief overview of relevant results from (evolutionary) game theory and their connection to broader results in the applied social scientific literature, which will inform the subsequent model building. After presenting the assumption structure of the formal model, I successively present and interpret the results of the mathematical deduction of optimal, culturally specific investment in specific enforcement mechanisms. After a brief summary of the deduced causal structure and an explication of the empirical strategy, I turn to testing these predictions, first for historical and then for contemporary societies.

2. Theoretical and empirical background

Over the past decades, the study of cooperation has perhaps made its biggest strides from an abstract, theoretical point of view. Interdisciplinary attention, not only from the social sciences but also from, e.g., biology and mathematics, combined with the development of sophisticated formal models and simulation exercises, has produced considerable progress in identifying institutional and environmental conditions under which cooperative behavior can emerge. In the context of this paper, this specifically relates to an identification of endogenous, non-centralized mechanisms that can enforce cooperative solutions to social dilemmas in the first place. Scholars studying human cooperation typically refer to these solution concepts as the 'three Rs': reciprocation, reputation, and retribution (Boyd and Richerson, 2009; Cuesta et al., 2015; Haley and Fessler, 2005; Henrich et al., 2006; Nowak and Sigmund, 2005). These mechanisms have in common that, within a context of repeated interactions, all three stabilize cooperation through credible threats of direct or indirect utility losses for cheaters. Reciprocity (or positive/direct reciprocity) does so by promising ongoing gains within dyadic relations under the threat of relationship termination. Reputation (or indirect reciprocity) can enforce cooperation by threatening to exclude cheaters from trust or exchange in third-party relationships if information about their past behavior circulates. Retribution (or negative reciprocity) deters cheaters through credible prospects of direct revenge-related utility losses.

While differing in form, all three mechanisms have in common that they are costly to maintain. Individuals must expend effort or resources to reward cooperation, broadcast reputational information, or carry out punitive acts. The rationale for incurring these costs is most straightforward in the case of reciprocity: As long as engaging in a "costly" relationship is preferable in the face of the

alternatives of mutual deceit or non-cooperation, the possibility of enforcement through reciprocal incentives is individually rational.[1] However, this problem is less readily resolved for reputation and revenge. It is not immediately clear why individuals should bear the cost of spreading information about a defector whose reputation they already know, nor why they should incur expenses for revenge rather than simply severing ties with the cheater. From a theoretical point of view, this means that although the "three Rs" may provide stable incentives for cooperation, a complete explanation of their real-world existence also requires understanding their stability in light of this so-called second-order cooperation problem. While previous studies have addressed second-order cooperation problems for specific mechanisms (e.g., Greif, 1994; Henrich and Boyd, 2001; Panchanathan and Boyd, 2004), there has been surprisingly little attempt to integrate these insights into a unified framework. The original theoretical contribution of this paper will later argue that understanding the functional connections between enforcement types necessitates understanding the connections between their respective second-order cooperation problems, that is, the interrelation between investment incentives that ultimately uphold the mechanisms' stability.

However, while of indisputable theoretical interest, resolving the second-order cooperation problem alone is insufficient to explain the observed variation in enforcement strategies across societies. The mere theoretical stability of a potential number of mechanisms (or mechanism combinations) does not imply their adoption, nor why adoption might differ between communities. Elaborations on this equilibrium selection problem remain rare. In many contexts, the question is not treated comparatively, that is, as a selection problem between *different* mechanisms (dos Santos et al., 2011; Rand et al., 2009; Schmid et al., 2021), but more narrowly as a selection problem between variants of the *same* mechanism (Leimar and Hammerstein, 2001; Nowak and Sigmund, 1992; Ohtsuki and Iwasa, 2006). Even fewer publications go beyond general efficiency arguments under broad ceteris paribus assumptions and develop arguments for how empirically existing differences in social and material circumstances would affect a society's enforcement types (see Enke, 2019; Greif and Tabellini, 2017; Nowak et al., 2016 for partial exceptions).

One major reason for this gap lies in the prevailing methodological approaches. A huge and rapidly expanding literature utilizing agent-based computer simulations of evolutionary games provides powerful tools for assessing the viability of abstract strategies under varying conditions, but often fails to link those strategies to real-world cultural or institutional traits. Frequently, strategies are either arbitrarily selected for theoretical interest (Axelrod, 1984) or derived through exhaustive iteration over a given strategy space (e.g., Ohtsuki and Iwasa, 2004). In contrast, this paper adopts a different approach: It begins with empirically observed cultural variation and then explores how these cultural traits map onto different behavioral modes and enforcement practices in the game-theoretical equilibria they induce.

This approach rests on the identification of inter-societal variation in characteristics relevant to the cooperation dilemma in general and the equilibrium selection problem in particular. In this article, I will advance the hypothesis that this variation is cultural, that is, that equilibrium selection is facilitated by shared and socially learned beliefs about the state of the (social) world. Without a definite claim to completeness, I will focus on three plausible cultural dimensions for which allusions to relevance can be inferred from the existing literature.

Perhaps the most obvious indicator of different approaches to cooperation dilemmas are the moral beliefs a society holds (Curry et al., 2019; Graham et al., 2013). By defining what is deemed "right" or "wrong," "just" or "unjust," morality serve as fundamental reference points for cooperation rules in different cultures. On the most abstract level, this pertains to convictions regarding to whom moral directives apply (i.e., who is required to behave morally) — both in terms of the social distance between interacting parties (the scope of morality) and the role individuals occupy within these dilemmas (the locus of morality).

Beliefs about the scope of morality define the range of individuals for whom moral behavior is prescribed and expected. The corresponding distinction between limited and generalized morality — sometimes also referred to as particularist versus universalist morality — captures whether moral obligations relate exclusively to one's own in-group or indiscriminately towards generalized others. The consequences of this conceptual divide have received growing attention in recent years (see Enke, 2024 for a review). In particular, isolated empirical studies have begun connecting similar cultural traits to differences in the formation of reputational control (Hajikhameneh, 2024; Horak et al., 2021; Clercq et al., 2023).

The locus of morality relates to beliefs about the targets of moral directives conditional on individuals' role within a cooperation dilemma. While moral systems may emphasize the responsibilities of potential cheaters (e.g., demand universal honesty or particularistic loyalty), they may also assign moral obligations to the cheated, prescribing how victims should respond to experienced offenses (Sommers, 2009). In applied contexts, associated moral conceptions are often referred to as honor beliefs (see Leung and Cohen, 2011; Uskul and Cross, 2020 for reviews), which capture the peculiarity of individuals being able to lose moral status through the actions of others which then need to be regained through an appropriate reaction. I will denote inverse moral beliefs as dignity moralities, meaning that moral status is inalienable in the sense of exclusively being subject to an individual's *own* actions. Since the first conceptual work on this cultural dimension by Cohen and Nisbett (1994), an increasing number of empirical results indicate that honor may be related to a greater utilization of retributive punishment (Cao et al., 2021; Schläpfer, 2024).

A third and more general (non-moral) dimension relevant to all enforcement types concerns intertemporal preferences — specifically, how individuals value future versus present payoffs. Because reciprocity, reputation, and revenge each involve delayed benefits or costs, they rely on some degree of future-oriented thinking. From a theoretical standpoint, this has long been recognized in models of reciprocity (Axelrod and Hamilton, 1981), and experimental studies confirm that the perceived value of future interaction increases the effectiveness of reciprocity-based cooperation (Rand and Nowak, 2013). However, the influence of time preferences on reputation

[1] Unsurprisingly, reciprocity is often understood as a universal, constitutive element in human exchange relations across diverse theoretical schools of thought (see, for example, Fiske, 1992; Gouldner, 1960; Homans, 1974).

and revenge remains underexplored (although see Espín et al., 2012 for a rare empirical exception), despite evidence of considerable cross-cultural variation in patience and long-term orientations (Falk et al., 2018; Wang et al., 2016).

Finally, it is important to emphasize that, from a broader perspective, cultural beliefs rarely operate in isolation. Together with informal enforcement practices and broader institutional arrangements, they are best understood as elements of co-evolutionary systems characterized by mutual dependence and feedback (Richerson and Boyd, 2008; Henrich, 2018). Ecological and economic conditions often select for multiple cultural traits simultaneously, while existing beliefs and institutions, in turn, shape the adoption and persistence of new ones. As a consequence, cultural attributes tend to cluster into coherent bundles. Market integration, for example, has been shown to foster both generalized morality (Enke, 2023) and longer temporal horizons (Salali and Migliano, 2015), whereas group-bounded moral frameworks appear to stabilize moralized beliefs about personal honor (Kashima et al., 2025).

These cultural bundles are closely intertwined with patterns of informal and formal enforcement and are frequently reinforced by the very institutional and socio-structural arrangements they help sustain. One prominent illustration of this broader clustering logic — of particular relevance for the present analysis — is kinship intensity, that is, the organization, salience, and density of family-based social ties. A growing interdisciplinary literature documents systematic associations between kinship structures and moral reasoning (Curtin et al., 2020), time preferences (Henrich, 2020), cooperative behavior (Schulz et al., 2019), and enforcement practices (Enke, 2019). In kinship-intensive environments, specific psychological dispositions, cultural beliefs, and enforcement logics tend to cohere and plausibly reinforce one another, giving rise to persistent social configurations. For instance, Enke (2019) argues that tight kinship societies generate interaction patterns and monitoring intensities that favor the effectiveness of revenge-based punishment over more centralized enforcement mechanisms, alongside cultural features such as limited moral scope and moralized emotions of shame.

At the same time, while some explanations for such interdependencies among specific elements within patterns of co-occurrence have been advanced, the concrete mechanisms underlying these relationships often remain opaque and weakly formalized (Henrich, 2015). In particular, the micro-level incentive structures through which specific cultural beliefs translate into stable enforcement regimes — and through which mutually reinforcing bundles arise more generally — are frequently left implicit or assumed a priori. Because much of the existing research has remained anchored in the empirical analysis of macro-level patterns of co-occurrence, calls for explicit accounts of the mechanisms shaping these interdependencies have grown louder in recent years (see Currie et al., 2021). As a first step in this direction, it is therefore useful to abstract from the full complexity of cultural co-evolution and examine how the three specific cultural beliefs discussed above — taken as given — shape individual incentives to adopt, sustain, and comply with informal enforcement mechanisms, before the empirical analysis subsequently revisits the related identification challenges.

3. A Formal Model of cultural beliefs in social dilemmas

To structure the synthesis of these ideas and extend them towards a constructive theory of culturally induced modes of cooperation, I develop a stylized formal model of a trust game (that is, a prisoner's dilemma in extensive form). The model expands upon elements of previous work by Greif (1994, 2010) and Dixit (2003). Unlike related approaches in the same tradition (e.g., Enke, 2019; Tabellini, 2008), the model deliberately avoids positing any proximate variation in material payoffs associated with cultural traits such as moral beliefs. This minimal assumption approach achieves a higher generalizability insofar as it remains open to the subsequent incorporation of perceptual differences (provided ample empirical justification) without rendering the model's core results dependent on them. For brevity, the exposition focuses on assumptions and equilibrium intuition; full formal derivations of all results are provided in Section A of the Supplementary Material.

Consider a large society in which all individuals belong to a shared social network such that they can potentially interact. This society is divided into $n \geq 2$ groups of equal size, where the membership of an individual is captured by a unique visible marker (or "tag") that is culturally determined and can represent, e.g., descent, ethnicity, dialect, or any other combination of characteristics deemed relevant for group association (Bowles and Gintis, 2004; Cohen, 2012).

Independently of group membership, individuals are randomly matched to play one-shot trust games with a discounted future, where time preferences are captured by a discount factor $\beta \in (0, 1)$. Each population contains Normal players (type N) and a small stable fraction $\epsilon > 0$ of Cheater types (type C), conceptualized as amoral actors. Type Cs always cheat as trustees and may shirk any enforcement norms as trustors even if these norms are behaviorally stable among Type N trustors. The presence of Type Cs pins down off-path beliefs, which ensures that defections cannot be credibly attributed to error and thereby stabilizes expectations outside cooperative equilibria.

Once matched, players observe their respective markers of group membership but not their moral type. The trustor then decides whether to trust by offering a reward $r > 0$ for honesty to the trustee or refrain from trusting, opting instead to wait for a new partner in the next period. In the latter case, both parties receive a payoff normalized to zero. If the trustor chooses to trust, the trustee must decide whether to behave honestly (resulting in a payoff of r for themselves and $1 - r$ for the trustor) or cheat (resulting in a payoff of $1 > r$ for themselves and -1 for the trustor). If cheated, the trustor can invest $p > 0$ in punishment, reducing the trustee's payoff by $P(p) > 0$, where P is strictly increasing and continuous in p. After each round, individuals spread information about their counterpart to a proportion $g \in [0, 1]$ of other players (gossip), modeled as a linear function of individual effort. The information transmission need not be perfectly accurate: all qualitative results are robust to moderate noise in transmission, such as honest players being mistakenly labeled as cheaters or cheaters as honest (see Supplementary Material Section A.5). Throughout, it is assumed that $r < 1 - P(p)$: trustors prefer not cooperating to offering rewards that exceed cheating incentives, and punishment threats alone cannot resolve the dilemma. Fig. 1 summarizes the structure of the game.

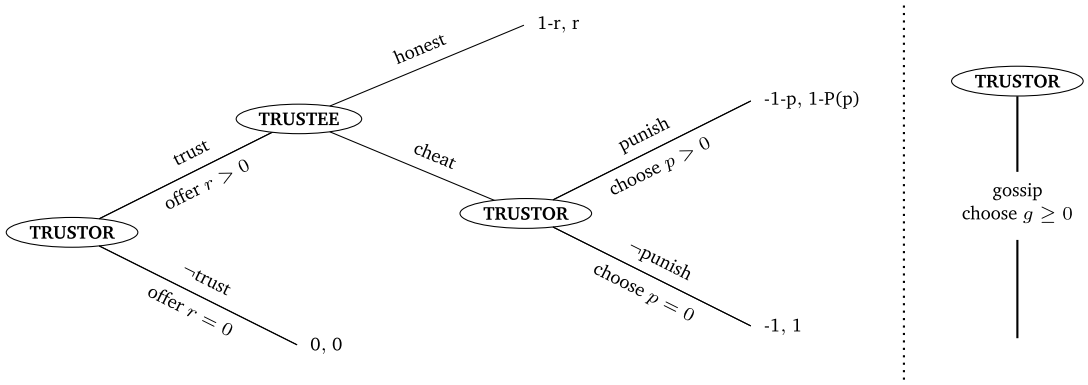

Fig. 1. The one-shot trust-game with punishment and gossiping phase. Offers of $r = 0$ are conceptualized as equivalent to not trusting and investments of $p = 0$ as equivalent to not punishing.

Finally, each society is characterized by a set of moral beliefs which define shared cultural expectations about appropriate ("good" and "bad") behavior.

- Let $\mu \in 0, 1$ denote the *scope of morality*. When $\mu = 1$, moral obligations apply universally (generalized morality). When $\mu = 0$, moral concern is restricted to in-group members (limited morality).
- Let $\pi \in \{0, 1\}$ denote the *locus of morality*. Whenever $\pi = 1$ individuals adhere to honor-based beliefs: moral status can be lost by being wronged and must be restored by retaliation. When $\pi = 0$, individuals hold dignity-based beliefs: Moral status depends solely on one's own actions and cannot be diminished by others' behavior alone.

The cooperative equilibrium of the game is characterized by the minimal reward r that trustors must offer to induce honest behavior. Crucially, this offer differs systematically across moral systems. Under generalized morality ($\mu = 1$), trustors must propose a reward sufficient to elicit honesty from all trustees. Under limited morality ($\mu = 0$), trustors anticipate that out-group trustees face no moral obligation to behave honestly and therefore optimally refrain from trusting them. Correspondingly, trustees expect that cheating out-group members does not damage their moral standing within their own in-group, stabilizing such defection off the equilibrium path. In effect, both belief systems support equilibrium cooperation, but imply distinct interaction structures: universally applied trust under generalized morality and trust confined to in-group interactions under limited morality. Supplementary Material A.1. details the underlying calculations building on this argument, yielding the optimal honesty-evoking offer r^* as follows:

Lemma 1: The unique cooperation-evoking offer r^* in the game's equilibria is given by

$$r^* = \begin{cases} (1 - P(p))\frac{n(1-\beta)}{n(1-\beta)+\beta g}, & \mu = 0 \\ (1 - P(p))\frac{1-\beta}{1-\beta(1-g)}, & \mu = 1 \end{cases} \tag{1}$$

This equilibrium offer admits a natural interpretation as the degree of reciprocity required to sustain cooperation. From the trustee's perspective, r^* represents the minimal immediate value of the relationship that makes honest behavior incentive compatible — the smallest "quid" that warrants the "quo" of honest conduct.

At the same time, this characterization is not intended to constitute an exhaustive theory of reciprocity. Given the level of stylization, r^* is best understood as a model endogenous representation of dyadic reciprocal demands in equilibrium.[2] This modeling choice also reflects a substantive rationale introduced earlier: Reciprocity represents a baseline mechanism that is nearly universal across societies and social dilemmas. Precisely because of its ubiquity, it provides a useful benchmark for analyzing the emergence and comparative advantages of more institutionalized enforcement mechanisms.

3.1. Reputation

One of these enforcement mechanisms is reputation, that is, a system in which trustors condition their trust allocation on circulating information about trustees' past behavior. The first-order stability of this mechanism is immediately palpable. If a trustor learns that a trustee has cheated despite having been offered r^*, their consistent belief must necessarily be that this trustee is of type C, i.e., an amoral individual who will *never* play honestly. For trustees this dynamic creates a natural deterrent for cheating: as more third parties learn of dishonesty, future exchange opportunities shrink. Anticipating reduced partner pools, type N trustees cooperate for smaller direct offers r.

[2] In formal terms, r^* as presented in Proposition 1 is entirely determined by cultural beliefs and other enforcement mechanisms. However, this implies a missing degree of freedom when consequences of changes in reciprocity that are subject to these omitted factors are to be deduced. Parts of the proofs presented in the Supplementary Material solve this problem by applying a real scalar $R \in (0, 1]$ to r^* that serves as a rough approximation of these model-exogenous factors which are assumed to exist empirically.

However, the institutional basis for reputational control comes at a positive cost. For trustors to benefit from the efficiency-enhancing effects of reputational control, they must invest in spreading information (they must "gossip"). In this sense, the question of reputational control can be considered as equivalent to an optimization problem that balances the cost of continuously spreading reputational information against the deterrence-based benefits this information spread produces. Proposition 1, formally derived in Section A.2 of the Supplementary Material, explicates the result arising from this argument.

Proposition 1. *For plausible parameter calibrations, the optimal level of gossip is defined as*

$$g^* = \begin{cases} n - \frac{1}{\beta}(n - \sqrt{n\beta(1-\beta)(1-P(p))}), & \mu = 0 \\ 1 - \frac{1}{\beta}(1 - \sqrt{\beta(1-\beta)(1-P(p))}), & \mu = 1 \end{cases} \tag{2}$$

Such investments are self-sustaining, because engaging is gossip is necessarily a public act: Members of a society will be aware if any trustor shirks on engaging in such information exchange which will (off-equilbrium path) produce the assumption of Type C status and thus come with future losses of gains from cooperation (see Supplementary Material Section A.4. for details). With this in mind, the optimal levels of gossip deduced in Proposition 1 carries several concrete implications.

- **Generalized morality:** Under plausible conditions (sufficiently pronounced long-term orientations), limited-morality societies sustain higher optimal reputational control than generalized-morality societies. Ultimately, this result arises from the different patterns of social interaction these moral frameworks sustain. Because under limited morality trustees can expect cooperation exclusively from in-group members, trustors only need to target in-group members with information on trustee's past behavior to effectively deter cheating. Under generalized morality, on the other hand, trustees may continue interacting with out-group members even after their reputation is damaged within their own group. From an investment standpoint, this implies that each additional effort to spread reputational information has a greater marginal return in limited-morality contexts, where it more effectively constrains the pool of potential exchange partners for known cheaters.
- **Long-term orientation:** For plausible calibrations (i.e., again a sufficiently large β), increases in patience *decrease* optimal levels of reputational control in a society. The substantive reason for this result lies in the fact that under pronounced future-oriented attitudes, even relatively small reputational damages will be perceived to suffer large utility losses. This heightened sensitivity to future consequences means that trustees already face strong incentives to behave honestly, even when only a few others might find out about their misconduct. For trustors, this reduces the marginal utility of further investing in gossip or information sharing.
- **Reciprocity:** Higher levels of dyadic reciprocity *increase* the optimal level of gossip. The rationale behind this result is surprisingly intuitive, when recalling the logic through which reputation and reciprocity enforce honesty. When deliberating whether they should cheat or not, the relevant question for a trustee is "Is this specific *relationship* worth maintaining?" under reciprocal enforcement, whereas the equivalent query under reputational enforcement becomes "Is my standing in this *society* worth protecting?". From trustors' point of view, investments in spreading information are thus essentially a shift in the locus of enforcement: Vis a vis reciprocity, reputational control moves the enforcement burden from the dyadic to the collective level. Hence, whenever the costs of maintaining cooperation in isolated dyadic relationships are comparatively high, shifting enforcement to the collective level, at a given cost of gossiping, becomes more attractive.
- **Revenge:** Larger threats of retaliation *decrease* optimal level of reputational control. The threat of immediate punishment reduces trustees' cheating incentives on the dyadic level. Similar to the intuition outlined above, this means that whenever sizable revenge norms are established, enforcing cooperation solely on the relational level gains comparative efficiency against the alternative of shifting enforcement to the societal level through gossip.

3.2. Revenge

Beyond reciprocity and reputation, enforcement may finally also operate through direct revenge against dishonest trustees. As previously introduced, I model retaliation as an investment decision by trustors: by incurring a cost $p > 0$, a trustor can impose direct harm $P(p) > 0$ on a trustee who has cheated. While the deterrent effect of such threats is immediate — expected punishment naturally lowers trustees' incentives to cheat — the more fundamental question concerns their strategic sustainability. Because retaliation is individually costly and yields no direct material benefit, it is not obvious why rational trustors would be willing to punish in equilibrium. The central issue, therefore, is whether costly revenge can be sustained as part of an equilibrium strategy, and, if so, what bounds the costs trustors are willing to bear for such retaliation.

Proposition 2, derived formally in Supplementary Material Section A.3, addresses this question by characterizing an upper bound on feasible punishment investments in equilibrium.

Proposition 2. *Under plausible calibrations of the model parameters, the constraints on the level of optimal investments in the punishment of cheaters are defined as:*

$$p^* \leq \begin{cases} 0, & \pi = 0 \\ \frac{\beta\left((1-r^*_{\mu=j})(1-P(p))-r^*_{\mu=j}\right)}{(1-\beta)(1-P(p)n^{j-1})}, & \pi = 1,\ j = 0, 1 \end{cases} \tag{3}$$

Once again, the right-hand side of the above inequality defines an implicit function that maps cultural beliefs and existing enforcement mechanisms onto the upper bound of feasible investments in revenge. Analyzing this function yields several testable predictions:[3]

- **Honor:** Honor beliefs ($\pi = 1$) support equilibrium strategies involving positive investments in direct punishment of cheaters, whereas dignity beliefs ($\pi = 0$) do not. This divergence stems from the equilibrium beliefs each cultural system sustains that directly relate to the second-order cooperation problem. Under honor beliefs, retaliation is moralized, which means the expectation that moral and amoral individuals will respond differently to being cheated. Whenever a trustee encounters a trustor known not to have retaliated in the past, they will infer the trustor is amoral (Type C). Given this belief, cheating becomes rational for any offer r that implies an existing punishment threat because the trustee must assume that the trustor will *never* punish. Anticipating this inference, trustors will develop positive incentives to uphold their honor as to sustain the credibility of (implied) punishment threats. Because these retaliation costs are strategically incurred, moral and amoral trustors will indeed exhibit behaviorally distinct punishment patterns, reinforcing the expectations that sustain the moral system in the first place.[4] Under dignity beliefs ($\pi = 0$), this signaling logic breaks down entirely. When retaliation is not moralized, trustees will not interpret non-retaliation as a signal of amorality. Information about a trustor's punishment history no longer provides cues about moral type and thus does not affect the credibility of sanctioning mechanisms that sustain cooperation. Without this signaling function, Type N trustors lack incentives to pay the costs of retaliation that would distinguish them from Type Cs — positive investments in punishment cannot be sustained as part of an equilibrium strategy. As a result there is no difference between Type N trustors and Type C trustors in punishment behavior, and thus no moralization of revenge as such (which again stabilizes the initial belief).
- **Long-term orientation:** Greater patience (β) *raises* the upper bound on feasible retaliation under honor beliefs. Retaliation serves to protect access to future cooperative opportunities by maintaining a credible honor status. In practical terms, maintaining an intact honor status thus equates to incurring some cost for revenge in the present to uphold a credible threat of retaliation in the future. As the value of future cooperative interactions increases, so does the willingness to incur costs to ensure continued access to them.
- **Reciprocity:** Higher levels of dyadic reciprocity *reduce* feasible investments in retaliation. When cooperation hinges on reciprocal incentives, productive relationships are highly contingent on the dyadic level as trustors must incur larger costs to make these relationships incentive-compatible. These costs reduce the net returns from cooperation, lowering the value of preserving an honorable status through retaliation which helps sustain them in the future. Conversely, when cooperation is less dependent on dyadic incentives, the immediate profitability of dyadic exchange is higher, increasing the value of honor and, with it, the willingness to defend it through punishment.
- **Reputation:** Higher levels of reputational control *increase* the upper bound on feasible retaliation. By shifting enforcement from the dyadic to the collective level, reputation reduces the need for costly bilateral incentives and raises the proximate returns from cooperative exchanges. This increases the value of maintaining an honorable reputation and strengthens incentives to defend it through retaliation.[5]

3.3. Summary and empirical strategy

The theoretical model's main predictions are summarized graphically in Fig. 2. Empirically, the analysis will focus on evaluating the predicted effects of secondary variables on the enforcement mechanisms of reputation and revenge. This focus directly reflects the model's treatment of reciprocity as a conceptual reference category, rather than as an explanatory target in itself. The directional expectations for most predictor-outcome relationships implied by this causal structure are straightforward in most cases. One exception is the interplay between reputational control and revenge where the bidirectional dynamic does not generate a clear-cut expectation for the net empirical relationship between the two mechanisms. Whether the reinforcing or attenuating causal direction dominates is thus left theoretically ambiguous and is treated as a purely empirical question.

In conceptual terms, the empirical implementation adapts the theoretical model in two key ways. First, while the model assumes discrete moral types for analytical tractability, the empirical specifications treat these traits as continuous variables to better capture their underlying variation in real-world contexts. Second, given that the model predicts monotonic relationships across all

[3] The formal bound for $\pi = 1$ differs between generalized and limited morality. I do not elaborate on this distinction here, as doing so would require additional assumptions about the functional form of $P(p)$ and the distribution of realized punishment costs between moral types. Accordingly, the following empirical analysis restricts itself to controlling for the degree of generalized morality when predicting punishment investments, without linking this to specific theoretical expectations.

[4] This logic also explains why Proposition 3 defines only an upper bound rather than a precise optimal level of punishment. In honor-based systems, retaliation is undertaken primarily to preserve moral status; the exact investment is endogenous to prevailing beliefs about "sufficient" punishment. Nonetheless, the bound retains explanatory power. Higher limits of p^* imply that even cultural systems demanding high punishment threats can support stable enforcement norms and hence enhance exchange efficiency through deterrence. Alternatively, higher limits may be interpreted as a broader stability criterion, whenever punishment costs vary between individuals. Higher bounds then imply that more individuals can afford to meet the cultural expectation of revenge, which matters because empirical evidence suggests not only that punishment's effectiveness depends on the proportion of unpunished cheating (Egas and Riedl, 2008), but also that widespread non-adherence can erode norm stability itself (Diekmann et al., 2015; Keuschnigg and Wolbring, 2015).

[5] Outside the model's calibration, reputational systems may also amplify the costs of failing to retaliate by facilitating information transmission about dishonor, further reinforcing this relationship.

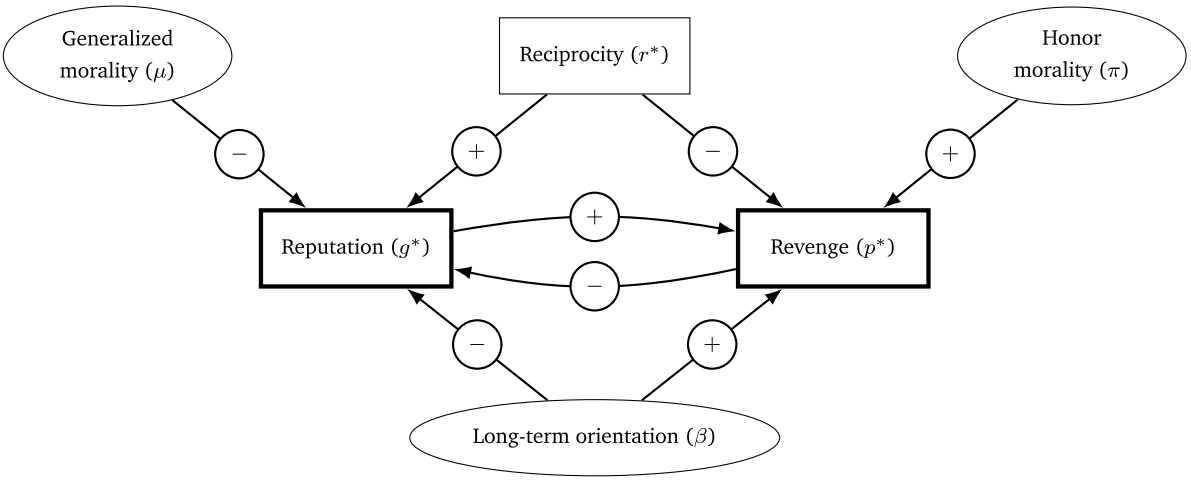

Fig. 2. Schematic depiction of relationships between cultural traits (ovals) and enforcement mechanisms (boxes) as deduced through the game-theoretical model. The main explanda of the following empirical analysis are set in bold.

predictor-outcome relationships, I employ linear approximations as a robust, parsimonious method to test the core of these directional hypotheses.

A related consideration concerns the empirical counterparts to the model's target variables. While the theory derives predictions for enforcement strategies in a stylized trust game, direct global measures of narrowly defined trust-game enforcement are largely unavailable. The empirical analysis therefore relies primarily, though not exclusively, on broader measures of social norm enforcement. This approach retains a highly relevant signal for the theoretical constructs of interest, insofar as the trust game canonically captures the mixed-motive logic addressed by social norms more generally (Bicchieri, 2005), and enforcement preferences exhibit substantial domain independence across contexts (Eriksson et al., 2021).

In the concrete specifications, all models include a measure proxying the number of interacting groups (parameter n in the formal model). Regarding potential confounders beyond the scope of the theory, I adopt a parsimonious control strategy that centers on outcome-relevant confounding (see Pearl, 2011; Steiner and Kim, 2016).[6] Given that even among such plausible confounders bidirectional causal dependencies with cultural traits and enforcement practices are likely, the primary goal of the analysis is best understood as estimating conditional relationships within a given interaction environment rather than deep causal primitives. Put differently, these specifications aim to align the empirical analysis with the model's comparative-static focus, which examines how enforcement equilibria vary with cultural parameters while taking the broader social environment as given.

Within this framework, I control for a small set of relevant background factors. First, and perhaps most importantly, I account for the presence of *centralized* enforcement institutions, which are expected to reduce reliance on decentralized enforcement practices (Ostrom, 2005). Second, I control for heterogeneity in subsistence strategies (e.g., own-use agriculture versus wage labor), as economic organization shapes the most important types of social dilemmas individuals face and, consequently, the demand for enforcement in the first place. Both factors have also been shown to influence cultural traits (e.g., Hackman and Kramer, 2021; Lightner and Hagen, 2021; Rothstein and Stolle, 2008). Third, I account for kinship structure, as this dimension of social organization fundamentally shapes interaction patterns (Henrich, 2020) and has been connected to the feasibility and efficacy of informal enforcement strategies (Enke, 2019).

While these core specifications are designed to hold central features of the immediate interaction environment constant in which informal enforcement operates, they do not fully rule out the possibility that cultural traits and enforcement mechanisms jointly co-evolved with broader ecological and institutional conditions over longer time horizons. To assess the robustness of the results to such co-evolutionary identification threats, I subsequently consider more extensive specifications that incorporate additional (socio-)ecological controls with plausibly exogenous, outcome-relevant causal channels. These controls are introduced in conceptually distinct blocks that capture slow-moving background factors — such as ecological fundamentals and historical-institutional trajectories —

[6] Several covariates commonly included in cross-cultural analyses are intentionally omitted from the main specifications. This choice reflects two considerations. First, many such variables are more plausibly downstream outcomes of the cultural traits or enforcement mechanisms under study rather than their causal antecedents (see, for example, Sunde et al., 2022 on long-term orientation, Gorodnichenko and Roland, 2017 on group boundedness, and Ahlerup et al., 2009 on enforcement institutions and economic development). Second, even when certain ecological variables could be construed as confounders in a narrow sense, they are often more directly related to cultural traits than to enforcement mechanisms. Conditioning on such variables under an inevitably incomplete control strategy can increase bias in the estimated effects (Pearl, 2011). For transparency, the Supplementary Material reports specifications that include an extensive set of additional covariates, including some that plausibly constitute "bad controls." While some coefficients expectedly attenuate under these specifications, the qualitative patterns underlying the paper's main conclusions remain unchanged.

that may have shaped both cultural beliefs and enforcement practices without being directly reflected in more proximate covariates. Importantly, these extended specifications do not alter the qualitative interpretation of the estimated effects as a test of the underlying theory, but serve to demonstrate that the core relationships are not driven by broad, persistent features of the social environment.

4. Historical societies

The empirical evaluation of these predictions begins with an analysis of historical societies. This initial step not only illustrates the generalizability of the theoretical framework but also offers the advantage of excluding the effects of globalization – and thus potential cultural convergence (Barrett et al., 2016; Herkenrath et al., 2005) – while minimizing the influence of formal judicial systems that are widespread in contemporary societies. Unless otherwise noted, the data for this part of the analysis derive from systematic codings of original ethnographic materials collected by anthropologists primarily in the first half of the twentieth century. Specifically, the analysis draws on largely preindustrial societies included in the Standard Cross-Cultural Sample (SCCS; Murdock and White, 2024), which systematizes a broad range of ecological information.

4.1. Data & operationalization

Although the SCCS contains information on cultural traits – albeit often only for subsets of all originally included 186 societies – systematic information on enforcement mechanisms is notably absent. To address this gap, I lean on two complementary sources of data. First, I rely on a recent coding effort by Garfield et al. (2023) which classifies the predominant forms of punishment in different societies. Their dataset includes a binary indicator for whether norm violations are sanctioned through reputational damage, which I use to operationalize reliance on reputational enforcement.[7] Second, because none of the other sanction categories in Garfield et al. adequately capture the logic of individually costly retaliation — existing codes focus on physical and material sanctions or executions — I construct a novel indicator for revenge-based enforcement. Following Garfield et al.'s general approach, I reexamine original ethnographic materials from the electronic Human Relations Area Files (eHRAF), a systematized database of ethnographic texts in which each paragraph is indexed using the Outline of Cultural Materials (OCM; Murdock et al., 2008). Using eHRAF's Advanced Search, I retrieved all paragraphs associated with sanction-related OCM codes containing retaliation- or revenge-related keywords. All resulting passages were read in full for the 131 societies covered by both eHRAF and the SCCS, yielding 2552 paragraphs in total. From this material, I derived a binary indicator that takes a value of one if revenge is described as (i) costly to the administering party (e.g., involving substantial effort or risk of physical harm), (ii) socially acceptable or morally appropriate, and (iii) invoked in response to a broad range of offenses, indicating that it functions as a general mechanism of social control rather than a narrowly targeted response (such as to homicide or adultery alone). Detailed information on the coding procedure — including search strategy, comprehensive coding criteria, inter-rater reliability statistics, bias checks, and illustrative ethnographic examples — is provided in Supplementary Material Section B.

Naturally, both of these binary indicators do not directly capture the continuous investment levels derived from the formal model. However, since the data originate from ethnographic records, these codings can be interpreted as signaling the realization of a salience threshold for a given mechanism. Because societies that invested more heavily in a particular type of enforcement are more likely to have this mechanism documented by ethnographers, the corresponding binary outcomes should serve as a meaningful proxy for its prevalence. Fig. 3 illustrates the spatial distribution of these enforcement practices across the analytical sample.

To capture cultural traits and dyadic reciprocity, I draw on data from Michalopoulos and Xue (2021), who assembled a harmonized corpus of textual data from traditional folklore intended to reflect a society's intergenerationally transmitted beliefs, values, and narratives. The data consists of various economic, psychological, and cultural concepts that characterize the original folklore motifs. For reciprocity and future-oriented attitudes, I mirror the conceptual approach employed by Cao et al. (2021). This means representing the respective trait through a list of related keywords, where the empirical indicator of its cultural prevalence is the number of keywords present in a society's folklore concepts as a proportion of all relevant keywords. For the two moral traits, honor and generalized morality, I slightly modify this approach in recognition of the fact that moral orientations often reflect trade-offs between competing ethical considerations (see similarly Enke, 2023).

The prevalence of honor beliefs is captured as the (standardized) difference between the proportion of terms relating to revenge, aggression, and anger and the proportion of terms relating to forgiveness and mercy. While the former captures norms and motivating emotions associated with the need to defend one's moral status (Cohen et al., 1996), the tendency to forgive has more recently garnered attention as being largely absent from honor cultures while viewed as explicitly prosocial in dignity cultures (Ceylan-Batur et al., 2023; Lin et al., 2022; Uskul et al., 2023).[8]

[7] The reputation indicator developed by Garfield et al. is based on a search strategy restricted to specific norm violations (adultery, religion, (food-) theft, rape, and cowardice in warfare). Although these domains do not map perfectly onto the trust-game framework underlying the theoretical model, reputational enforcement is plausibly not domain-specific. This is because the existence of reputational sanctions implies the presence of a (costly) information-sharing structure that, once established, operates as a general mechanism of social control. Section C.2 of the Supplementary Material provides empirical support for this interpretation, showing that the indicator predicts both topic-specific "idle talk" and domain-independent references to reputational sanctions scraped from eHRAF metadata.

[8] An intuition of why honor moralities are structurally incompatible with pronounced notions of forgiveness can easily be retraced over the formal argumentation of the model. For a trustor who has been cheated, forgiving entails refraining from punishing the dishonest trustee. Within a dignity-based moral framework, such forgiveness is, in principle, viable. If the trustee can credibly signal that they are not of Type C — that is, not

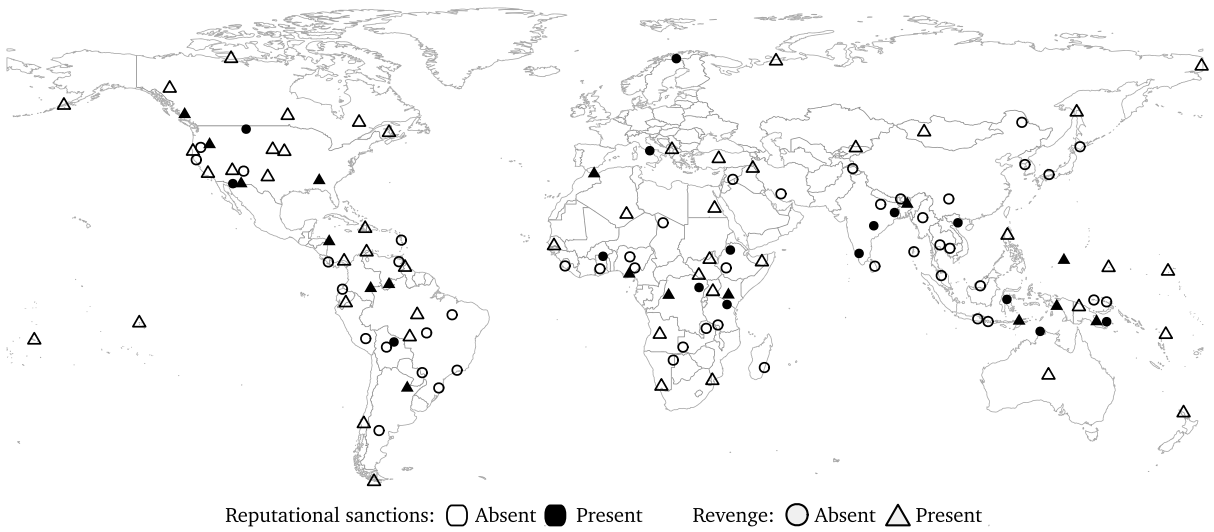

Fig. 3. Geographical distribution of historical societies included in the analytical sample and their respective reliance on reputation and revenge as enforcement practices.

For generalized morality, I similarly (and perhaps more intuitively) use the standardized difference between the prevalence of terms relating to universal, nonrelational notions of justice and the proportion of terms relating to particularist in-group orientations or out-group demarcations. Sections C.1 and C.2. of the Supplementary Material provide a detailed account of the construction of all cultural indicators, including complete lists of keywords, along with bivariate correlations between these indices and conceptually related traits from the SCCS as a test of convergent validity.

All models include a set of control variables to account for potential confounders and measurement concerns. First, because the likelihood of detecting any given motif mechanically increases with corpus size, I control for the (logged) number of motifs recorded in the folklore dataset. To approximate the number of interacting groups (parameter n in the theoretical model), I use population density estimates from Pryor (1985). The models also control for the presence of centralized enforcement institutions, proxied by the existence of specialized judiciaries within a society (Tuden and Marshall, 1972). In addition, I include a measure of subsistence heterogeneity, given evidence that hunter-gatherer societies differ systematically in social structure and dominant social dilemmas compared to those with more complex economic systems (Gowdy and Krall, 2016; Hill et al., 2011). To address the possibility that kinship structures simultaneously shape enforcement strategies and cultural traits, I control for the Kinship Intensity Index (KII; Schulz et al., 2019) which aggregates five dimensions of kin-based institutional organization (cousin marriage preference, polygamy prevalence, extended family co-residence, lineage organization, and community organization).

To further mitigate identification concerns related to co-evolutionary dynamics and broader ecological confounding, additional specifications include controls for a range of geographic and environmental characteristics. These include terrain ruggedness (from Nunn and Puga, 2012), caloric suitability for agriculture (from Özak, 2015), irrigation potential (from FAO and IIASA, 2025), and interannual variation in temperature and percipitation (from Fick and Hijmans, 2017). For these spatial variables, values are averaged within a 100 km buffer around each society's location to approximate its resource catchment area while limiting overlap with neighboring groups; results are robust to using 50 km and 150 km buffers. Finally, in extended specifications I control for distance to the nearest urban center in 1600 CE (based on data by Klein Goldewijk et al., 2011), social complexity as captured through a scale developed by Murdock and Provost (1973), and a dummy indicating the society's integration into a larger nation-state. Details on variable construction and a discussion of their relevance for identification are provided in Supplementary Material Section C.1.

To address potential non-independence of observations due to shared ancestry or cultural diffusion, all standard errors are clustered at the language subfamily level, based on the linguistic taxonomy developed by Burton (1999).

4.2. Results

Table 1 presents the results of a series of linear probability models predicting the presence of reputational punishment in 128 historical societies, which constitute the analytical sample of the following analysis. Overall, the findings align closely with the theoretical expectations derived from the game-theoretical framework. Both generalized morality and long-term orientation are consistently and significantly negatively associated with the presence of reputational sanctions across all model specifications. Substantively, this

an amoral individual — the trustor may rationally choose to extend trust again in the future or withhold negative gossip. In contrast, honor-based moral systems necessitate punishment as a public signal that the victim is not dishonorable themselves. Even if the trustor privately believes that the trustee is not inherently amoral, they remain compelled to punish to protect their own moral standing within the community.

Table 1
Linear probability models predicting the presence of reputational punishment in historical societies. Robust standard errors clustered at the language-subfamily level are reported in parentheses. Ecology includes controls for terrain ruggedness, caloric suitability for agriculture, irrigation potential, and interannual variation in temperature and precipitation. Socio-institutional development includes controls for distance to the nearest urban center in 1600 CE, nation-state integration, and social complexity. All continuous independent variables are standardized.

	Dependent variable: Presence of reputational punishment				
	(1)	(2)	(3)	(4)	(5)
Generalized morality	−0.177***	−0.173***	−0.163***	−0.167***	−0.168***
	(0.036)	(0.034)	(0.032)	(0.034)	(0.033)
Long-term orientation	−0.190**	−0.182**	−0.161*	−0.172*	−0.175*
	(0.087)	(0.086)	(0.093)	(0.088)	(0.099)
Reciprocity	0.210***	0.237***	0.250***	0.238***	0.260***
	(0.071)	(0.077)	(0.079)	(0.077)	(0.080)
Presence of revenge	0.092	0.110	0.147*	0.132	0.166*
	(0.082)	(0.083)	(0.084)	(0.085)	(0.087)
Folklore motifs (log)	Yes	Yes	Yes	Yes	Yes
Population density	Yes	Yes	Yes	Yes	Yes
Centralized enforcement		Yes	Yes	Yes	Yes
Hunter/gatherer subsistence		Yes	Yes	Yes	Yes
Kinship intensity		Yes	Yes	Yes	Yes
Ecology			Yes		Yes
Socio-institutional development				Yes	Yes
Observations	128	128	128	128	128
R^2	0.149	0.161	0.201	0.180	0.222
Adjusted R^2	0.107	0.097	0.101	0.095	0.102

Note: *p<0.1; **p<0.05; ***p<0.01

corroborates that societies with more universalistic moral outlook — those extending moral concern beyond in-group boundaries — and those marked by more future-oriented thinking are indeed less likely to rely on reputational sanctions as an enforcement device. Conversely, and equally in line with the model's equilibrium predictions, reciprocity is positively associated with the presence of reputational punishment: societies that place greater emphasis on dyadic obligations are more likely to exhibit systems of reputational control. The estimated effects are substantively large. A one standard deviation increase in any of these predictors corresponds to an 18–25 percentage point change in the average predicted probability of (reported) reputational punishment. Consistent with theoretical expectations of an attenuating feedback between reputational sanctions and revenge, the estimated effect of revenge is generally smaller and does not reach conventional levels of statistical significance in most specifications. Nevertheless, it is consistently positive, with the presence of revenge practices associated with an increase of roughly 10–15 percentage points in the probability that a society also relies on reputational control.

A similar pattern of empirical support for the theoretically derived relationships emerges when analyzing the determinants direct retaliation (see Table 2). Societies with stronger honor-based moralities are indeed more likely to employ revenge as an enforcement mechanism. Similarly, the results indicate that such punishment is more common in cultures with stronger future-oriented attitudes. In contrast, reciprocity — understood as an emphasis on the intrinsic value of dyadic relationships — tends to reduce the likelihood of direct retaliation. These relationships remain statistically robust across specifications — including under expansive controls for socio-ecological characteristics. As in the inverse case discussed above, the effect of reputational punishment on revenge is less pronounced and statistically weaker, however, again consistently positive across models. Notably, the inclusion of institutional controls (social complexity, distance to cities, and nation state integrations; models 4 and 5) in particular leads to an increase in size, again suggesting that, within this sample, the reinforcing pathway from reputational control to revenge may be empirically more pronounced than the attenuating pathway operating in the opposite direction.

The Supplementary Material (Section C.3) presents an extensive set of robustness checks aimed at assessing the credibility of these findings. First, I examine the sensitivity of the results to alternative constructions of the dictionaries underlying the key explanatory variables, including conceptual reductions to three core terms each, embedding-based "no-neighbor" variants, and embedding-centroid-based dictionary expansions. Second, I re-estimate the main specifications using logistic regression models as an alternative to the linear probability models employed in the baseline analysis. Third, I account for potential spatial dependence by implementing (Conley, 1999) standard errors, which adjust for spatial autocorrelation, in place of the language-cluster-clustered standard errors used above. In addition, I systematically exclude geographically proximate societies to ensure that the results are not disproportionately driven by regional clustering. Finally, to address concerns about omitted variable bias, I substantially expand the set of control

Table 2
Linear probability models predicting the presence of revenge-taking in historical societies. Robust standard errors clustered at the language-subfamily level are reported in parentheses. Ecology includes controls for terrain ruggedness, caloric suitability for agriculture, irrigation potential, and interannual variation in temperature and precipitation. Socio-institutional development includes controls for distance to the nearest urban center in 1600 CE, nation-state integration, and social complexity. All continuous independent variables are standardized.

	Dependent variable: Presence of revenge				
	(1)	(2)	(3)	(4)	(5)
Honor morality	0.196***	0.177***	0.179***	0.190***	0.187***
	(0.047)	(0.046)	(0.046)	(0.049)	(0.049)
Long-term orientation	0.222**	0.226**	0.214**	0.196**	0.213**
	(0.098)	(0.092)	(0.088)	(0.099)	(0.097)
Reciprocity	−0.207***	−0.239***	−0.202***	−0.217***	−0.197***
	(0.074)	(0.070)	(0.071)	(0.067)	(0.067)
Reputational punishment	0.138	0.147	0.179*	0.180*	0.201**
	(0.099)	(0.096)	(0.094)	(0.096)	(0.095)
Folklore motifs (log)	Yes	Yes	Yes	Yes	Yes
Population density	Yes	Yes	Yes	Yes	Yes
Generalized morality	Yes	Yes	Yes	Yes	Yes
Centralized enforcement		Yes	Yes	Yes	Yes
Hunter/gatherer subsistence		Yes	Yes	Yes	Yes
Kinship intensity		Yes	Yes	Yes	Yes
Ecology			Yes		Yes
Socio-institutional development				Yes	Yes
Observations	128	128	128	128	128
R^2	0.247	0.281	0.351	0.345	0.397
Adjusted R^2	0.203	0.220	0.265	0.271	0.297

Note: *p<0.1; **p<0.05; ***p<0.01

variables to include additional demographic characteristics, beliefs in moralizing high gods, more detailed measures of subsistence strategies, the frequency of armed conflict, and a broad set of ecological and geographic covariates (including additional climate variables, absolute latitude, distance to navigable waterways, and geographic fixed effects among others). Across all of these alternative specifications, the core results reported in Tables 1 and 2 remain both substantively and statistically robust.

5. Contemporary societies

Building on the preceding analysis, a natural next step is to examine whether the theoretical predictions — and the historical patterns they help explain — also hold in contemporary contexts. To this end, I compile a global regional-level dataset from multiple original sources. Regions are defined as major socioeconomic units within countries, typically comprising populations of 5 to 15 million inhabitants, and are aligned with existing administrative divisions wherever possible (see Supplementary Material Section D.1 for details). Compared to coarser country-level analyses, this approach has the obvious advantage of preserving relevant granularity, particularly within large and culturally diverse nations. All individual-level variables were mean aggregated to the regional level between surveys (if available) and waves (if existent). To ensure the reliability of estimates, regions with fewer than 30 observations for any key variable were excluded from the analysis.[9]

5.1. Data & operationalization

I primarily draw on data from the World Values Survey (WVS; Haerpfer et al., 2021), the European Values Survey (EVS; EVS, 2022), the European Social Survey (ESS; ESS, 2024), and the Global Preference Survey (GPS; Falk et al., 2018). The GPS is particularly

[9] The analysis is restricted to survey data collected between 2002 and 2022. Pooling across surveys and years is unavoidable to ensure data availability and sufficiently large samples for reliable estimates at fine-grained geographic levels. However, it also introduces potential biases. These include (1) measurement non-invariance across surveys and time periods, and (2) possible temporal shifts in cultural traits. Although the extent of such biases is difficult to assess directly, prior research suggests that both cultural beliefs in general (Kiley and Vaisey, 2020; Beugelsdijk et al., 2015) and moral convictions in particular (Luttrell and Togans, 2021; Wheeler et al., 2019) tend to be relatively stable over time. Additional reassurance comes from evidence of measurement invariance for at least some relevant variables (Reeskens and Hooghe, 2008). Robustness checks in the Supplementary Material address these concerns indirectly, including replications using varied case inclusion criteria.

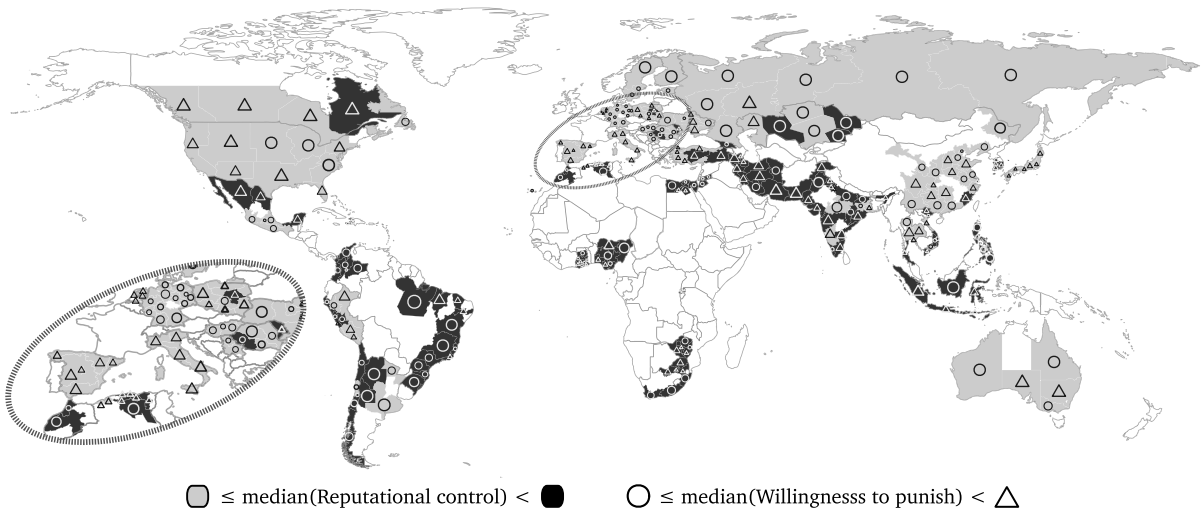

Fig. 4. Geographical distribution of regions included in the contemporary (survey based) sample and their respective reliance on reputation and revenge as enforcement practices.

well suited to the current study, as it directly includes several key variables of interest derived from a globally administered and experimentally validated questionnaire. Specifically, it provides direct measures of long-term orientation, reciprocity, and willingness to engage in costly punishment in response to perceived unfairness (negative reciprocity). Each of these indicators is operationalized through weighted indices based on three survey items, optimized for predictive validity in theoretically relevant behavioral outcomes (see Falk et al., 2023 and Supplementary Material D.1 for details). To capture levels of reputational control, I follow previous research (Danielson et al., 2023; Vonasch et al., 2018) by using the self-rated similarity of respondents to a person described as "To avoid doing anything people would say is wrong, it is important for this person to always behave properly" on a 6-point scale, for which data is drawn from the World Values Survey (waves 5 and 6) and the European Social Survey (waves 1 to 11). Fig. 4 presents an overview of the global distribution of enforcement practices based on these measures.

The two moral dimensions – honor and generalized morality – are constructed over items from the WVS and the EVS. Following the approach outlined in the previous section, honor is operationalized as the scaled difference between an orientation toward aggression and (violent) punishment on the one hand and an orientation toward lenience and forgiveness on the other. Here, the latter is represented by the predominant characteristic of honor cultures of being particularly restricting to (women's) sexuality (Schneider, 1971; Vandello and Cohen, 2003), by being measured as an additive scale of attitudes regarding the justifiability of prostitution, casual sex, premarital sex, and divorce. The aggression component captures regional attitudes toward the justifiability of interpersonal violence and the death penalty. As, to the best of my knowledge, no established measurements of honor moralities over the utilized data exist, this operationalization naturally requires validation. To this end, Section D.2 of the Supplementary Material correlates the honor morality index with several aggregate measures of honor beliefs from the existing literature. To obtain more comprehensive global reference data, I additionally extract the prevalence of "honor-talk" as defined in a dictionary by Gelfand et al. (2015) in countries' current in-force constitutions. Furthermore, I obtain country-level rates of female victim homicides by close relatives and partners from UN data (that is, capturing femicides as a proxy for so-called "honor killings"; Elakkary et al., 2014; Dayan, 2021). All of these variables show meaningful associations with the honor measure utilized in the forthcoming analysis, supporting its convergent validity.

Generalized morality is measured analogously to Section 4 as the scaled difference between universalistic and particularistic moral orientations. Universalism is indicated by self-identification as a world citizen, the valuation of independence in children, and generalized trust. Particularism is inversely captured by identification with one's local community, the emphasis on obedience in children, and trust in personally known individuals. Section D.2 of the Supplementary Material assesses the validity of this operationalization against a number of related attitudinal measures from the existing literature. To further validate the variable against behavioral indicators, I draw on country-level data on voluntary blood donations to nonfamily members by the World Health Organizations and the results of a cross-cultural experimental by Gächter and Schulz (2016) that provides evidence of the willingness to cheat strangers. Again, the original measure used in the forthcoming analysis is well predictive of all of these variables.

To proxy the number of interacting groups, all models include a control for regional degrees of urbanization, extracted from geospatial data based on UN definitions (Schiavina et al., 2023). Alternative specifications using the proportion of WVS/EVS respondents residing in urban areas, or country-level urbanization rates from the World Bank (World Bank, 2024), yield qualitatively equivalent results. To account for the potential confounding role of centralized legal enforcement, I include the World Justice Project's Civil Justice Index,[10] which reflects expert assessments of the effectiveness, fairness, and accessibility of legal institutions in civil disputes.

[10] Data retrieved from https://worldjusticeproject.org/ (19.12.2024). I compute a rolling average from the first available year (2012) to the latest survey wave (2022). However, temporal variation within countries is minimal (average within-country SD = 0.02 on a 0–1 scale).

Robustness checks substituting this measure with regional averages of trust in the legal system from the WVS/EVS yield qualitatively similar results.

I further control for heterogeneity in dominant economic modes across regions. While wage employment naturally predominates across most societies in the sample, some less developed and rural areas continue to rely heavily on own-use agricultural production (Kostov and Lingard, 2004; Sibhatu and Qaim, 2017). This heterogeneity is relevant because the land-tethered and largely self-sufficient nature of subsistence farming potentially amplifies reputational control via reduced geographic mobility while simultaneously reducing the overall need for stringent enforcement in the first place. Given the limited availability of cross-national data on own-use agricultural production, I construct an indirect proxy using the first principal component of three country-level indicators: poverty rates (World Bank, 2025), the population share informally working in agriculture (International Labour Organization, 2025), and average agricultural holding size (as an inverse measure; FAOSTAT, 2025). To increase regional specificity, this index is then weighted by the rural population share. As shown in the Section D.2 of the Supplementary Material, the resulting measure correlates well with (sparsely) available direct indicators, outperforming cruder proxies such as GDP per capita. Finally, I include a measure capturing the strength of family ties, adapting the operationalization proposed by Alesina and Giuliano (2010) using survey data from the WVS, EVS, and ESS. Owing to data availability constraints, their original principal-component index based on three items is reduced to a single component that is consistently observed across waves; nonetheless, the resulting measure exhibits a very high correlation with the original index.

Mirroring the structure of the previous section, subsequent specifications control for deeper historical and ecological confounders, grouped into two conceptually distinct blocks (see Supplementary Material D.1 for details on variable construction and identification strategy). The first block again captures biogeographic characteristics, including terrain ruggedness, caloric suitability, irrigation potential, and interannual climatic variability averaged from geospatial data within GADM-derived regional boundaries. The second block addresses long-run institutional exposure. Specifically, I include the population-weighted distance to the nearest urban center in 1600 CE as a proxy for historical market access and economic conditions (from Klein Goldewijk et al., 2011), alongside colonizer fixed effects (from Becker, 2023) and an ancestry-adjusted measure of premodern jurisdictional integration (derived from Giuliano and Nunn, 2018). The latter two variables capture historical exposure to variants of centralized enforcement institutions, which may have durably shaped cultural orientations and preferences for informal enforcement beyond current institutional quality.

5.2. Results

As an initial empirical assessment of the formal model's predictions using contemporary survey data, Table 3 reports OLS estimates of regional levels of reputational control. In line with theoretical expectations, both generalized morality and long-term orientation are negatively associated with this measure: Regions characterized by a broader scope of moral concern and a stronger future orientation tend to exhibit lower reliance on reputational enforcement. These associations remain substantively meaningful across specifications, although effect sizes attenuate when proximate controls are introduced. For long-term orientation in particular, the inclusion of closely related covariates — most notably the quality of legal institutions and kinship importance — leads to a noticeable reduction in the estimated coefficient in model (2), reflecting the partialling out of shared variation. Once deeper ecological and historical factors that influence these covariates are included (models 3–5), the estimated effect stabilizes and increases again.

The empirical relationship between reciprocity and reputational control presents an inverse case. Consistent with the model, the estimated coefficient of reciprocity is positive throughout and increases substantially once controlled for the reliance on subsistence agriculture and the quality of legal institutions. However, the coefficient only obtains conventional levels of statistical significance *after* ecological and historical institutional controls are introduced. On closer inspection, the attenuated effect sizes in models (1) and (2) are largely driven by systematically lower average levels of positive reciprocity in (Sub-Saharan) Africa — a region otherwise characterized by high degrees of reputational control. While it is difficult to pinpoint the exact cause producing this result, one possible (conjectural) explanation could lie in the finding that the concept of reciprocity seems to be culturally instantiated in unusual ways in some African populations (Aju and Beddewela, 2020; Vollan, 2012). This phenomenon could potentially influence regional baseline levels as measured by the underlying survey items — a speculation that is supported by the fact that alternative specifications of model (2), which incorporate an Africa dummy ($\hat{\beta}_{Recip.} = 0.161$), exclude all African regions ($\hat{\beta}_{Recip.} = 0.213$), or consider a subsample of *only* African regions ($\hat{\beta}_{Recip.} = 0.314$), consistently produce robust estimates that align with the initial theoretical predictions. At the same time, the fact that ecological and historical institutional controls substantially reduce this attenuation points to clustered confounding as an equally plausible explanation. In either case, the richer specifications likely mitigate region-level heterogeneity that obscures the underlying relationship between reciprocity and reputational control in more parsimonious models.

Finally, the regionally aggregated willingness to punish unfair behavior does not show a consistent relationship with reputational enforcement. Mirroring the patterns previously found in the historical data, the effects are consistently positive but small and characterized by comparatively large statistical uncertainty.

Next, Table 4 reports OLS estimates of regional willingness to incur personal costs to punish behavior perceived as unfair, providing an empirical assessment of the formal model's predictions regarding contemporary reliance on direct retaliation as an enforcement mechanism. Data availability constraints for the honor morality index restricts this analysis to 31 of the original 46 countries, yielding an effective sample of 184 regions. As in the historical analysis, the estimates broadly support the theoretical relationships. Regions characterized by a stronger prevalence of honor moralities exhibit a significantly higher willingness to punish, even after conditioning on a wide range of proximate institutional, geographic, and historical covariates. The estimated magnitudes are substantial: a one standard deviation increase in honor morality is associated with an increase of more than one quarter of a standard deviation in punishment propensity.

Table 3
OLS estimates predicting regional levels of reputational control. Robust standard errors clustered at the country level are reported in parentheses. Ecology includes controls for terrain ruggedness, caloric suitability for agriculture, irrigation potential, and interannual variation in temperature and precipitation. Historical institutions includes controls for population weighted distance to the nearest urban center in 1600 CE, colonial history (as colonizer fixed effects), and ancestry adjusted historical exposure to jurisdictional centralization. All continuous independent variables are standardized.

	Dependent variable: Reputational control				
	(1)	(2)	(3)	(4)	(5)
Generalized morality	−0.326***	−0.253***	−0.180**	−0.229**	−0.202**
	(0.116)	(0.094)	(0.086)	(0.102)	(0.081)
Long-term orientation	−0.260**	−0.128	−0.161**	−0.190**	−0.159**
	(0.117)	(0.088)	(0.076)	(0.095)	(0.073)
Reciprocity	0.048	0.113	0.129**	0.212**	0.146***
	(0.087)	(0.081)	(0.063)	(0.083)	(0.054)
Willingness to punish	0.079	0.022	0.059	0.109	0.076
	(0.083)	(0.071)	(0.050)	(0.067)	(0.054)
Urbanization	Yes	Yes	Yes	Yes	Yes
Subsistence agriculture		Yes	Yes	Yes	Yes
Civil Justice Index (WJP)		Yes	Yes	Yes	Yes
Kinship importance		Yes	Yes	Yes	Yes
Ecology			Yes		Yes
Historical institutions				Yes	Yes
Observations	272	272	272	272	272
R^2	0.310	0.461	0.607	0.563	0.656
Adjusted R^2	0.297	0.444	0.586	0.539	0.628

Note: *p<0.1; **p<0.05; ***p<0.01

The estimated effects of time preferences and reciprocity are likewise consistent with the model's implications. Greater long-term orientation is associated with a higher willingness to punish, whereas a stronger emphasis on dyadic reciprocity is systematically linked to a lower reliance on punitive behavior across all specifications. The empirical peculiarity regarding positive reciprocity in Sub-Saharan Africa discussed above does not affect the estimates reported in Table 4, as Ghana, Rwanda, and South Africa are excluded from this analysis due to missing data.

Finally, reputational control is positively associated with willingness to punish across all specifications. As for the historical data, the coefficient is relatively small and imprecise in the parsimonious baseline models but increases markedly once controls for geography and long-run institutional legacies are introduced. In particular, conditioning on colonial history substantially amplifies the estimated association, suggesting that colonial-era institutional bundles may have partially obscured the underlying relationship between reputational and punitive enforcement. From a theoretical perspective, these results are consistent with the conjecture that the reinforcing effect of reputational control on retaliatory behavior may, in practice, be stronger than the opposing attenuating effect of revenge threats on reputational control.

The Supplementary Material provides an extensive set of robustness checks that further support these results and their interpretation. Most importantly, Section E presents a series of instrumental-variable estimations that exploit plausibly (conditionally) exogenous variation in historical socio-ecological conditions to address concerns that the estimated effects of cultural traits — generalized morality, honor, and long-term orientation — may be solely driven by reverse causality. The IV results are fully consistent with the baseline estimates and remain robust to reasonable violations of the exclusion restriction.

In addition, Section D.3 reports a range of robustness checks for the OLS analyses. First, I re-estimate the main specifications using datasets with varying minimum regional observation thresholds (set to 1, 50, and 100) and assess the stability of the results by sequentially excluding regions from each continent to address concerns that the findings may be driven by specific geographical clusters. Second, to evaluate sensitivity to the choice of survey items, I replicate the main analyses using two alternative — albeit arguably less precise — operationalizations of both honor and generalized morality respectively. Finally, the results are shown to remain robust to substantially more extensive control strategies, including indicators of political and social stability, cultural and demographic characteristics, economic conditions, geographic factors (such as distance to navigable waterways, climate, absolute latitude, and geographic fixed effects), and historical traits (including state history, time since the Neolithic Revolution, exposure to the slave trade, and ancestry-adjusted kinship intensity among others).

Table 4
OLS estimates predicting the regional willingness to punish at a personal cost. Robust standard errors clustered at the country level are reported in parentheses. Ecology includes controls for terrain ruggedness, caloric suitability for agriculture, irrigation potential, and interannual variation in temperature and precipitation. Historical institutions includes controls for population weighted distance to the nearest urban center in 1600 CE, colonial history (as fixed effects), and ancestry adjusted historical exposure to jurisdictional centralization. All continuous independent variables are standardized.

	Dependent variable: Willingness to punish				
	(1)	(2)	(3)	(4)	(5)
Honor morality	0.261**	0.284*	0.300**	0.241*	0.311**
	(0.127)	(0.146)	(0.152)	(0.137)	(0.150)
Long-term orientation	0.248**	0.212**	0.220*	0.225*	0.240**
	(0.100)	(0.101)	(0.121)	(0.119)	(0.118)
Reciprocity	−0.283**	−0.255**	−0.267***	−0.388***	−0.391***
	(0.121)	(0.119)	(0.104)	(0.090)	(0.080)
Reputational control	0.078	0.037	0.124	0.327**	0.309**
	(0.135)	(0.140)	(0.150)	(0.130)	(0.125)
Generalized morality	Yes	Yes	Yes	Yes	Yes
Urbanization	Yes	Yes	Yes	Yes	Yes
Subsistence agriculture		Yes	Yes	Yes	Yes
Civil Justice Index (WJP)		Yes	Yes	Yes	Yes
Kinship importance		Yes	Yes	Yes	Yes
Ecology			Yes		Yes
Historical institutions				Yes	Yes
Observations	184	184	184	184	184
R^2	0.173	0.188	0.226	0.384	0.428
Adjusted R^2	0.145	0.146	0.157	0.329	0.354

Note: *p<0.1; **p<0.05; ***p<0.01

6. Discussion

This paper has developed and empirically evaluated a game-theoretic framework for analyzing how cultural beliefs shape informal enforcement mechanisms in cooperation dilemmas. Through comparative statics — examining equilibrium behavior under different cultural configurations — the model generates transparent predictions about how generalized morality, honor beliefs, and long-term orientation influence incentives for reputational sanctions and direct retaliation. The empirical analysis provides consistent support for these predictions across a wide range of specifications.

At the same time, the paper deliberately abstracts from the endogenous origins and joint evolution of cultural traits, enforcement practices, and institutions. In reality, these elements are embedded in co-evolutionary systems characterized by feedback and mutual reinforcement (Bowles, 2010; Currie et al., 2021; Nunn, 2021). Cultural beliefs are therefore unlikely to be fully exogenous, as assumed by the model, but instead co-develop with broader socio-ecological conditions and institutional environments.

One aspect of this abstraction concerns interdependencies among cultural traits themselves. Honor beliefs and limited moral scope, for instance, are empirically and conceptually related (Oyserman, 2017), and both have been linked to kinship organization and patterns of social embeddedness that also shape time preferences (Henrich, 2020; Falk et al., 2018). More broadly, enforcement practices may feed back into cultural beliefs. Systems of reputational control, for example, can stabilize group-bounded moralities by increasing the perceived risks of interacting outside established information networks (Greif, 1994, 2006). Cultural traits and informal enforcement regimes are also closely tied to a society's structural and institutional features. Reputational control, limited morality, and honor are likely to favor tighter kinship structures, which in turn reinforce these same traits (Schulz et al., 2019; Enke, 2019), whereas universalistic moral beliefs may facilitate the emergence of formal, centralized enforcement institutions that expand the scope of profitable interaction with outsiders and, once established, further select for moral universalism (Tabellini, 2008).

While fully disentangling these dynamics over observational data is difficult, the empirical analysis takes several measures to mitigate related identification concerns. Supplementary Material Sections C.3 and D.3 show that the main results are robust to specifications in which cultural traits enter individually rather than jointly, suggesting that the estimated effects are not merely artifacts of co-occurring cultural orientations. Instrumental-variable estimates reported in Section E further indicate that the observed relationships are unlikely to be driven exclusively by reverse causality from enforcement practices to cultural beliefs. Finally, the baseline and extended specifications control for a broad set of ecological, institutional, and historical factors that plausibly mediate long-run culture-enforcement co-evolution. The stability of the estimated effects under these conservative specifications is therefore reassuring.

From a theoretical perspective, explicitly modeling the dynamic co-evolution of cultural beliefs, enforcement strategies, and institutions remains an important direction for future research. Such dynamics are often non-trivial (see Bisin and Verdier, 2017) but central to broader questions of institutional development. Nonetheless, the static approach adopted here provides a necessary intermediate step. By isolating the incentive mechanisms through which specific cultural beliefs support particular enforcement regimes, the paper contributes micro-foundational clarity to a literature that has thus far been fragmented and largely descriptive (Molho et al., 2024). The results delineate which culture-enforcement configurations are internally consistent in equilibrium — constraints that any comprehensive dynamic theory must ultimately address as well. The strong empirical alignment with these predictions suggests that the framework offers a productive foundation on which future work on the co-evolution of culture, enforcement, and institutions can build.

7. Conclusion

Reputational control, credible threats of retaliation, and dyadic reciprocity each offer distinct pathways for sustaining the large-scale cooperation ubiquitously observable across human societies. However, while extensive empirical work documents substantial cross-societal variation in reliance on these enforcement mechanisms, coherent theoretical explanations for this variation have remained limited.

This paper addresses this gap by demonstrating that shared cultural beliefs coordinate equilibrium selection among enforcement mechanisms. The model shows how moral and perceptual beliefs shape individual incentives to adhere to costly second-order enforcement norms, which ultimately give rise to divergent but stable enforcement strategies. Specifically, reputational control is favored by group-bounded moralities, shorter time horizons, and high levels of dyadic reciprocity. Credible threats of direct retaliation, conversely, are facilitated by honor moralities, greater patience, and less pronounced dyadic reciprocal commitments. Empirical evidence from two global datasets — covering both contemporary sub-national regions and historical small-scale societies — provides consistent support for these predictions across markedly different contexts, underscoring the generalizability of the proposed mechanisms.

The results contribute to broader debates on cooperation and enforcement in economics and the social sciences in several ways. First, with few exceptions (e.g., Schmid et al., 2021), existing theoretical work has largely analyzed enforcement mechanisms in isolation. By jointly modeling multiple enforcement types and deriving explicit relationships between stable mechanism combinations, this paper shows that enforcement strategies are neither simple substitutes nor independently chosen. Instead, they interact in systematic ways: strong dyadic reciprocity, for example, stabilizes reputational enforcement while simultaneously undermining the credibility of large retaliatory threats.

Second, because enforcement equilibria are sustained by shared cultural beliefs, the analysis speaks directly on the interaction between culture and institutions. While much of this work has focused on the relationship between cultural beliefs and centralized, formal enforcement, the findings suggest that such an emphasis may be incomplete. The emergence, functioning, and effectiveness of centralized legal institutions are likely conditioned by the informal enforcement systems societies already rely on to solve problems of social order. Because many prominent theories of institutional development invoke cultural traits that this paper shows to be consequential for variation in informal enforcement (see Alesina and Giuliano, 2015; Licht et al., 2007; Tabellini, 2008), future work should explicitly account for these interdependencies.

Relatedly, the analysis connects to recent work on kinship systems and family structure. The co-evolution of cultural beliefs and informal enforcement mechanisms — such as reputational control — may generate path dependencies that shape kinship organization, which has been linked to a wide range of prosocial traits and long-run economic outcomes (Schulz et al., 2019; Schulz, 2022). From this perspective, observed associations between family structures, cooperation norms, and development may partly reflect the enforcement regimes that co-evolved with particular cultural bundles.

More broadly, by establishing systematic links between specific cultural dimensions and enforcement equilibria, this paper provides micro-foundations for understanding the observable diversity in solutions to cooperation dilemmas. These results points toward a more integrated framework in which culture, informal enforcement, and eventually formal institutions are analyzed as jointly evolving components of social order, rather than as isolated determinants of cooperative behavior.

Data Availability

Data and replication material are publicly available at OSF: https://osf.io/4d53r/.

Funding

This work was supported by the University of Mannheim's Graduate School of Economic and Social Sciences (GESS).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

I thank two anonymous reviewers for useful suggestions that significantly improved the paper. I also want to thank Nergüzel Aleyna Yavuzer for research assistance as well as Malte Grönemann, Henning Hillmann, and Thomas Gautschi for helpful feedback on earlier versions of the manuscript.

Supplementary material

Supplementary material associated with this article can be found in the online version at 10.1016/j.jebo.2026.107494.

References

Ahlerup, P., Olsson, O., Yanagizawa, D., 2009. Social capital vs institutions in the growth process. Eur. J. Polit. Econ. 25 (1), 1–14. https://doi.org/10.1016/j.ejpoleco.2008.09.008
Aju, O., Beddewela, E., 2020. Afrocentric attitudinal reciprocity and social expectations of employees: the role of employee-centred CSR in africa. J. Bus. Ethics 161 (4), 763–781. https://doi.org/10.1007/s10551-019-04346-x
Alesina, A., Giuliano, P., 2010. The power of the family. J. Econ. Growth 15 (2), 93–125. Publisher: Springer; Springer US. https://doi.org/10.1007/s10887-010-9052-z
Alesina, A., Giuliano, P., 2015. Culture and institutions. J. Econ. Lit. 53 (4), 898–944. https://doi.org/10.1257/jel.53.4.898
Axelrod, R., Hamilton, W.D., 1981. The evolution of cooperation. Science 211 (4489), 1390–1396. https://doi.org/10.1126/science.7466396
Axelrod, R.M., 1984. The evolution of cooperation. Basic Books, New York.
Balliet, D., van Lange, P. A.M., 2013. Trust, punishment, and cooperation across 18 societies: a meta-analysis. Perspec. Psychol. Sci. 8 (4), 363–379. https://doi.org/10.1177/1745691613488533
Barrett, H.C., Bolyanatz, A., Crittenden, A.N., Fessler, D. M.T., Fitzpatrick, S., Gurven, M., Henrich, J., Kanovsky, M., Kushnick, G., Pisor, A., Scelza, B.A., Stich, S., von Rueden, C., Zhao, W., Laurence, S., 2016. Small-scale societies exhibit fundamental variation in the role of intentions in moral judgment. Proc. National Acad. Sci. 113 (17), 4688–4693. https://doi.org/10.1073/pnas.1522070113
Becker, B., 2023. Colonial Dates Dataset (COLDAT). https://doi.org/10.7910/DVN/T9SDEW
Beugelsdijk, S., Maseland, R., van Hoorn, A., 2015. Are scores on hofstede's dimensions of national culture stable over time? a cohort analysis. Global Strategy J. 5 (3), 223–240. https://doi.org/10.1002/gsj.1098
Bicchieri, C., 2005. The Grammar of Society: The Nature and Dynamics of Social Norms. Cambridge University Press. Google-Books-ID: 4N1FDIZvcI8C.
Bisin, A., Verdier, T., 2017. On the Joint Evolution of Culture and Institutions. https://doi.org/10.3386/w23375
Bowles, S., 2010. The coevolution of institutions and preferences: history and theory. In: Institutional and Social Dynamics of Growth and Distribution. Edward Elgar Publishing, Cheltenham, pp. 35–50. Section: Institutional and Social Dynamics of Growth and Distribution.
Bowles, S., Gintis, H., 2004. Persistent parochialism: trust and exclusion in ethnic networks. J. Econ. Behav. Organiz. 55 (1), 1–23. https://doi.org/10.1016/j.jebo.2003.06.005
Boyd, R., Richerson, P.J., 2009. Culture and the evolution of human cooperation. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 364 (1533), 3281–3288. https://doi.org/10.1098/rstb.2009.0134
Burton, M.L., 1999. Language and region codes for the standard cross-cultural sample. Cross-Cultural Res. 33 (1), 63–83. https://doi.org/10.1177/106939719903300105
Cao, Y., Enke, B., Falk, A., Giuliano, P., Nunn, N., 2021. Herding, warfare, and a culture of honor: global evidence. NBER Work. Paper No. 29250 Cambridge, MA; National Bureau of Economic Research. https://doi.org/10.3386/w29250
Ceylan-Batur, S., Uskul, A.K., Gul, P., 2023. Forgive and forget? honor-oriented individuals are less forgiving of transgressing peers. Pers. Individ. Dif. 206, 112147. https://doi.org/10.1016/j.paid.2023.112147
Clercq, D., Haq, I.U., Azeem, M.U., 2023. Gossiping about outsiders: how time-related work stress among collectivistic employees hinders job performance. J. Manage. Organiz. 29 (2), 191–206. https://doi.org/10.1017/jmo.2019.56
Cohen, D., Nisbett, R.E., 1994. Self-protection and the culture of honor: explaining southern violence. Personal. Social Psychol. Bull. 20 (5), 551–567. https://doi.org/10.1177/0146167294205012
Cohen, D., Nisbett, R.E., Bowdle, B.F., Schwarz, N., 1996. Insult, aggression, and the southern culture of honor: an "experimental ethnography". J. Pers. Soc. Psychol. 70 (5), 945–959. https://doi.org/10.1037/0022-3514.70.5.945
Cohen, E., 2012. The evolution of tag-based cooperation in humans. Curr. Anthropol. 53 (5), 588–616. https://doi.org/10.1086/667654
Conley, T.G., 1999. GMM Estimation with cross sectional dependence. J. Econom. 92 (1), 1–45. https://doi.org/10.1016/S0304-4076(98)00084-0
Cuesta, J.A., Gracia-Lázaro, C., Ferrer, A., Moreno, Y., Sánchez, A., 2015. Reputation drives cooperative behaviour and network formation in human groups. Sci. Rep. 5 (1), 7843. https://doi.org/10.1038/srep07843
Currie, T.E., Campenni, M., Flitton, A., Njagi, T., Ontiri, E., Perret, C., Walker, L., 2021. The cultural evolution and ecology of institutions. Philos. Transac. Royal Society B: Biol. Sci. 376 (1828), 20200047. https://doi.org/10.1098/rstb.2020.0047
Curry, O.S., Mullins, D.A., Whitehouse, H., 2019. Is it good to cooperate? testing the theory of morality-as-cooperation in 60 societies. Curr. Anthropol. 60 (1), 47–69. https://doi.org/10.1086/701478
Curtin, C.M., Barrett, H.C., Bolyanatz, A., Crittenden, A.N., Fessler, D. M.T., Fitzpatrick, S., Gurven, M., Kanovsky, M., Kushnick, G., Laurence, S., Pisor, A., Scelza, B., Stich, S., von Rueden, C., Henrich, J., 2020. Kinship intensity and the use of mental states in moral judgment across societies. Evol. Human Behav. 41 (5), 415–429. https://doi.org/10.1016/j.evolhumbehav.2020.07.002
Danielson, S.W., Stewart, E., Vonasch, A., 2023. The morality map: does living in a smaller community cause greater concern for moral reputation? Curr. Res. Ecol. Social Psychol. 4, 100120. https://doi.org/10.1016/j.cresp.2023.100120
Dayan, H., 2021. Female honor killing: the role of low socio-economic status and rapid modernization. J. Interpers. Violence 36 (19–20), 10393–10410. https://doi.org/10.1177/0886260519872984
Diekmann, A., Przepiorka, W., Rauhut, H., 2015. Lifting the veil of ignorance: an experiment on the contagiousness of norm violations. Rational. Soc. 27 (3), 309–333. https://doi.org/10.1177/1043463115593109
Dixit, A., 2003. Trade expansion and contract enforcement. J. Polit. Econ. 111 (6), 1293–1317. https://doi.org/10.1086/378528
Egas, M., Riedl, A., 2008. The economics of altruistic punishment and the maintenance of cooperation. 275 (1637), 871–8. https://doi.org/10.1098/rspb.2007.1558
Elakkary, S., Franke, B., Shokri, D., Hartwig, S., Tsokos, M., Püschel, K., 2014. Honor crimes: review and proposed definition. Forensic. Sci. Med. Pathol. 10 (1), 76–82. https://doi.org/10.1007/s12024-013-9455-1
Enke, B., 2019. Kinship, cooperation, and the evolution of moral systems. Q. J. Econ. 134 (2), 953–1019. https://doi.org/10.1093/qje/qjz001
Enke, B., 2023. Market exposure and human morality. Nat. Hum. Behav. 7 (1), 134–141. https://doi.org/10.1038/s41562-022-01480-x
Enke, B., 2024. Moral boundaries. Annu. Rev. Econom. 16 (1), 133–157. https://doi.org/10.1146/annurev-economics-091223-093730
Erasmus, C.J., 1956. Culture structure and process: the occurrence and disappearance of reciprocal farm labor. Southwest J. Anthropol. 12 (4), 444–469. https://doi.org/10.1086/soutjanth.12.4.3629069
Eriksson, K., Strimling, P., Andersson, P.A., Aveyard, M., Brauer, M., Gritskov, V., Kiyonari, T., Kuhlman, D.M., Maitner, A.T., Manesi, Z., Molho, C., Peperkoorn, L.S., Rizwan, M., Stivers, A.W., Tian, Q., van Lange, P. A.M., Vartanova, I., Wu, J., Yamagishi, T., 2017. Cultural universals and cultural differences in meta-norms about peer punishment. Manage. Organiz. Rev. 13 (4), 851–870. https://doi.org/10.1017/mor.2017.42

Eriksson, K., Strimling, P., Gelfand, M., Wu, J., Abernathy, J., Akotia, C.S., Aldashev, A., Andersson, P.A., Andrighetto, G., Anum, A., Arikan, G., Aycan, Z., Bagherian, F., Barrera, D., Basnight-Brown, D., Batkeyev, B., Belaus, A., Berezina, E., Björnstjerna, M., Blumen, S., Boski, P., Bou Zeineddine, F., Bovina, I., Huyen, B. T.T., Cardenas, J.-C., Čekrlija, D., Choi, H.-S., Contreras-Ibáñez, C.C., Costa-Lopes, R., Barra, M., Zoysa, P., Dorrough, A., Dvoryanchikov, N., Eller, A., Engelmann, J.B., Euh, H., Fang, X., Fiedler, S., Foster-Gimbel, O.A., Fülöp, M., Gardarsdottir, R.B., Gill, C.M. H.D., Glöckner, A., Graf, S., Grigoryan, A., Gritskov, V., Growiec, K., Halama, P., Hartanto, A., Hopthrow, T., Hřebíčková, M., Iliško, D., Imada, H., Kapoor, H., Kawakami, K., Khachatryan, N., Kharchenko, N., Khoury, N., Kiyonari, T., Kohút, M., Linh, L.T., Leslie, L.M., Li, Y., Li, N.P., Li, Z., Liik, K., Maitner, A.T., Manhique, B., Manley, H., Medhioub, I., Mentser, S., Mohammed, L., Nejat, P., Nipassa, O., Nussinson, R., Onyedire, N.G., Onyishi, I.E., Özden, S., Panagiotopoulou, P., Perez-Floriano, L.R., Persson, M.S., Pheko, M., Pirttilä-Backman, A.-M., Pogosyan, M., Raver, J., Reyna, C., Rodrigues, R.B., Romanò, S., Romero, P.P., Sakki, I., San Martin, A., Sherbaji, S., Shimizu, H., Simpson, B., Szabo, E., Takemura, K., Tieffi, H., Mendes Teixeira, M.L., Thanomkul, N., Tiliouine, H., Travaglino, G.A., Tsirbas, Y., Wan, R., Widodo, S., Zein, R., Zhang, Q.-P., Zirganou-Kazolea, L., van Lange, P. A.M., 2021. Perceptions of the appropriate response to norm violation in 57 societies. Nat. Commun. 12 (1), 1481. https://doi.org/10.1038/s41467-021-21602-9

Espín, A.M., Brañas-Garza, P., Herrmann, B., Gamella, J.F., 2012. Patient and impatient punishers of free-riders. Proc. Royal Society B: Biol. Sci. 279 (1749), 4923–4928. https://doi.org/10.1098/rspb.2012.2043

ESS, 2024. European Social Survey Cumulative Data File, ESS 1–11. https://doi.org/10.21338/NSD-ESS-CUMULATIVE

EVS, 2022. EVS Trend File 1981-2017EVS Trend File 1981–2017. https://doi.org/10.4232/1.14021

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., Sunde, U., 2018. Global evidence on economic preferences. Q. J. Econ. 133 (4), 1645–1692. https://doi.org/10.1093/qje/qjy013

Falk, A., Becker, A., Dohmen, T., Huffman, D., Sunde, U., 2023. The preference survey module: a validated instrument for measuring risk, time, and social preferences. Manage. Sci. 69 (4), 1935–1950. https://doi.org/10.1287/mnsc.2022.4455

FAO, IIASA, 2025. Global Agro Ecological Zones version 4 (GAEZ v4). http://www.fao.org/gaez/.

Food and Agriculture Organization of the United Nations (FAOSTAT), 2025. World Census of Agriculture / Structural data from agricultural censuses.

Fick, S.E., Hijmans, R.J., 2017. Worldclim 2: new 1-km spatial resolution climate surfaces for global land areas. Int. J. Climatol. 37 (12), 4302–4315. https://doi.org/10.1002/joc.5086

Fiske, A.P., 1992. The four elementary forms of sociality: framework for a unified theory of social relations. Psychol. Rev. 99 (4), 689–723. https://doi.org/10.1037//0033-295x.99.4.689

Garfield, Z.H., Ringen, E.J., Buckner, W., Medupe, D., Wrangham, R.W., Glowacki, L., 2023. Norm violations and punishments across human societies. Evol. Human Sci. 5, e11. https://doi.org/10.1017/ehs.2023.7

Gelfand, M.J., Severance, L., Lee, T., Bruss, C.B., Lun, J., Abdel-Latif, A.-H., Al-Moghazy, A.A., Moustafa Ahmed, S., 2015. Culture and getting to yes: the linguistic signature of creative agreements in the united states and egypt. J. Organ. Behav. 36 (7), 967–989. https://doi.org/10.1002/job.2026

Giuliano, P., Nunn, N., 2018. Ancestral characteristics of modern populations. Econ. History Devel. Regions 33 (1), 1–17. https://doi.org/10.1080/20780389.2018.1435267

Gorodnichenko, Y., Roland, G., 2017. Culture, institutions, and the wealth of nations. Rev. Econ. Stat. 99 (3), 402–416. https://doi.org/10.1162/REST_a_00599

Gouldner, A.W., 1960. The norm of reciprocity: a preliminary statement. Am. Sociol. Rev. 25 (2), 161. https://doi.org/10.2307/2092623

Gowdy, J., Krall, L., 2016. The economic origins of ultrasociality. Behav. Brain Sci. 39, e92. https://doi.org/10.1017/S0140525X1500059X

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S.P., Ditto, P.H., 2013. Chapter two - moral foundations theory: the pragmatic validity of moral pluralism. In: Devine, P., Plant, A. (Eds.), Advances in Experimental Social Psychology. Academic Press, pp. 55–130. https://doi.org/10.1016/B978-0-12-407236-7.00002-4

Greif, A., 1994. Cultural beliefs and the organization of society: a historical and theoretical reflection on collectivist and individualist societies. J. Political Econ. 102 (5), 912–950. https://doi.org/10.1086/261959

Greif, A., 2006. History lessons: the birth of impersonal exchange: the community responsibility system and impartial justice. J. Econ. Perspect. 20 (2), 221–236. https://doi.org/10.1257/jep.20.2.221

Greif, A., 2010. Institutions and the path to the modern economy: Lessons from medieval trade. Political Economy of Institutions and Decisions, Cambridge University Press, Cambridge. 1st ed. edition.

Greif, A., Tabellini, G., 2017. The clan and the corporation: sustaining cooperation in china and europe. J. Comp. Econ. 45 (1), 1–35. https://doi.org/10.1016/j.jce.2016.12.003

Gächter, S., Herrmann, B., 2009. Reciprocity, culture and human cooperation: previous insights and a new cross-cultural experiment. Philosoph. Transac. Royal Society B: Biol. Sci. 364 (1518), 791–806. https://doi.org/10.1098/rstb.2008.0275

Gächter, S., Schulz, J.F., 2016. Intrinsic honesty and the prevalence of rule violations across societies. Nature 531 (7595), 496–499. https://doi.org/10.1038/nature17160

Hackman, J.V., Kramer, K.L., 2021. Kin ties and market integration in a yucatec mayan village. Soc. Sci. 10 (6), 216. https://doi.org/10.3390/socsci10060216

Haerpfer, C., Inglehart, R., Moreno, A., Welzel, C., Kizilova, K., Diez-Medrano, J., Lagos, M., Norris, P., Ponarin, E., Puranen, B., 2021. World Values Survey Time-Series (1981–2020) Cross-National Data-Set. Technical Report. World Values Survey Association. https://doi.org/10.14281/18241.15

Hajikhameneh, A., 2024. Reputation or court: individualism, collectivism, and the choice of enforcement mechanism in exchange. J. Econ. Behav. Organiz. 217, 184–206. https://doi.org/10.1016/j.jebo.2023.10.025

Haley, K.J., Fessler, D. M.T., 2005. Nobody'S watching?: subtle cues affect generosity in an anonymous economic game. Evol. Human Behav. 26 (3), 245–256. https://doi.org/10.1016/j.evolhumbehav.2005.01.002

Henrich, J., 2015. Culture and social behavior. Curr Opin Behav Sci 3, 84–89. https://doi.org/10.1016/j.cobeha.2015.02.001

Henrich, J., 2018. The secret of our success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter. Princeton University Press, Princeton.

Henrich, J., 2020. The weirdest people in the world: How the West became psychologically peculiar and particularly prosperous. Allen Lane, London.

Henrich, J., Boyd, R., 2001. Why people punish defectors. weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas. J. Theor. Biol. 208 (1), 79–89. https://doi.org/10.1006/jtbi.2000.2202

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Cardenas, J.C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D., Ziker, J., 2006. Costly punishment across human societies. Science 312 (5781), 1767–70. https://doi.org/10.1126/science.1127333

Henrich, N., Henrich, J., 2007. Why Humans Cooperate: A Cultural and Evolutionary Explanation. Oxford University Press, New York.

Herkenrath, M., König, C., Scholtz, H., Volken, T., 2005. Convergence and divergence in the contemporary world system: an introduction. Int J Comp Sociol 46 (5–6), 363–382. https://doi.org/10.1177/0020715205060227

Herrmann, B., Thöni, C., Gächter, S., 2008. Antisocial punishment across societies. Science 319 (5868), 1362–1367. https://doi.org/10.1126/science.1153808

Hill, K.R., Walker, R.S., Božičević, M., Eder, J., Headland, T., Hewlett, B., Hurtado, A.M., Marlowe, F., Wiessner, P., Wood, B., 2011. Co-residence patterns in hunter-gatherer societies show unique human social structure. Science 331 (6022), 1286–1289. https://doi.org/10.1126/science.1199071

Homans, G.C., 1974. Social behavior: Its elementary forms. Harcourt Brace, New York.

Horak, S., Klein, A., Svirina, A., 2021. Trust, reciprocity and reputation in informal networks in post-soviet russia. European Journal of International Management 15 (4), 657–685. https://doi.org/10.1504/EJIM.2021.114623

International Labour Organization, I., 2025. ILO modelled estimates database. https://ilostat.ilo.org/data/.

Kashima, Y., Shafa, S., Blake, K., 2025. Cultural logics and individualism-collectivism: a conceptualization of the two frameworks from a cultural dynamical perspective. Current Research in Ecological and Social Psychology 9, 100232. https://doi.org/10.1016/j.cresp.2025.100232

Keuschnigg, M., Wolbring, T., 2015. Disorder, social capital, and norm violation: three field experiments on the broken windows thesis. Rationality and Society 27 (1), 96–126. https://doi.org/10.1177/1043463114561749

Kiley, K., Vaisey, S., 2020. Measuring stability and change in personal culture using panel data. Am Sociol Rev 85 (3), 477–506. https://doi.org/10.1177/0003122420921538

Klein Goldewijk, K., Beusen, A., van Drecht, G., de Vos, M., 2011. The HYDE 3.1 spatially explicit database of human-induced global land-use change over the past 12,000 years. Global Ecol. Biogeogr. 20 (1), 73–86. https://doi.org/10.1111/j.1466-8238.2010.00587.x
Kostov, P., Lingard, J., 2004. Subsistence agriculture in transition economies: its roles and determinants. Journal of Agricultural Economics 55 (3), 565–579. https://doi.org/10.1111/j.1477-9552.2004.tb00115.x
Leimar, O., Hammerstein, P., 2001. Evolution of cooperation through indirect reciprocity. Proceedings of the Royal Society B: Biological Sciences 268 (1468), 745–53. https://doi.org/10.1098/rspb.2000.1573
Leung, A. K.-Y., Cohen, D., 2011. Within- and between-culture variation: individual differences and the cultural logics of honor, face, and dignity cultures. J Pers Soc Psychol 100 (3), 507–526. https://doi.org/10.1037/a0022151
Licht, A.N., Goldschmidt, C., Schwartz, S.H., 2007. Culture rules: the foundations of the rule of law and other norms of governance. J Comp Econ 35 (4), 659–688. https://doi.org/10.1016/j.jce.2007.09.001
Lightner, A.D., Hagen, E.H., 2021. Acculturation and market integration are associated with greater trust among tanzanian maasai pastoralists. Evolutionary Human Sciences 3, e15. https://doi.org/10.1017/ehs.2021.10
Lin, Y., Caluori, N., Öztürk, E.B., Gelfand, M.J., 2022. From virility to virtue: the psychology of apology in honor cultures. Proceedings of the National Academy of Sciences 119 (41), e2210324119. https://doi.org/10.1073/pnas.2210324119
Liu, H.-C.W., 1959. The traditional Chinese clan rules. J.J. Augustin Incorporated Publishers, New York.
Luttrell, A., Togans, L.J., 2021. The stability of moralized attitudes over time. Personality and Social Psychology Bulletin 47 (4), 551–564. https://doi.org/10.1177/0146167220935737
Mahdi, N.Q., 1986. Pukhtunwali: ostracism and honor among the pathan hill tribes. Ethol Sociobiol 7 (3–4), 295–304. https://doi.org/10.1016/0162-3095(86)90055-5
Marlowe, F.W., Berbesque, J.C., Barr, A., Barrett, C., Bolyanatz, A., Cardenas, J.C., Ensminger, J., Gurven, M., Gwako, E., Henrich, J., Henrich, N., Lesorogol, C., McElreath, R., Tracer, D., 2008. More 'altruistic' punishment in larger societies. Proceedings. Biological sciences 275 (1634), 587–90. https://doi.org/10.1098/rspb.2007.1517
Michalopoulos, S., Xue, M.M., 2021. Folklore. Q J Econ 136 (4), 1993–2046. https://doi.org/10.1093/qje/qjab003
Molho, C., De Petrillo, F., Garfield, Z.H., Slewe, S., 2024. Cross-societal variation in norm enforcement systems. Philosophical Transactions of the Royal Society B: Biological Sciences 379 (1897), 20230034. https://doi.org/10.1098/rstb.2023.0034
Murdock, G.P., Ford, C.S., Hudson, A.E., Kennedy, R., Simmons, L.W., Whiting, J. W.M., 2008. Outline of Cultural Materials. 6th Revise. Human Relations Area Files, New Haven. ERIC Number: ED044998.
Murdock, G.P., Provost, C., 1973. Measurement of cultural complexity. Ethnology 12 (4), 379–392. Publisher: University of Pittsburgh- Of the Commonwealth System of Higher Education. https://doi.org/10.2307/3773367
Murdock, G.P., White, D.R., 2024. D-PLACE dataset derived from Murdock and White 1969 'Standard Cross-Cultural Sample'. https://doi.org/10.5281/zenodo.13318864
Nowak, A., Gelfand, M.J., Borkowski, W., Cohen, D., Hernandez, I., 2016. The evolutionary basis of honor cultures. Psychol Sci 27 (1), 12–24. https://doi.org/10.1177/0956797615602860
Nowak, M.A., Sigmund, K., 1992. Tit for tat in heterogeneous populations. Nature 355 (6357), 250–253. https://doi.org/10.1038/355250a0
Nowak, M.A., Sigmund, K., 2005. Evolution of indirect reciprocity. Nature 437 (7063), 1291–1298. https://doi.org/10.1038/nature04131
Nunn, N., 2021. Chapter 3 - history as evolution. In: Bisin, A. (Ed.), The Handbook of Historical Economics(ebook). Academic Press, -, pp. 41–91. https://doi.org/10.1016/B978-0-12-815874-6.00010-1
Nunn, N., Puga, D., 2012. Ruggedness: the blessing of bad geography in Africa. Rev Econ Stat 94 (1), 20–36. https://doi.org/10.1162/REST_a_00161
Ohtsuki, H., Iwasa, Y., 2004. How should we define goodness?-reputation dynamics in indirect reciprocity. J. Theor. Biol. 231 (1), 107–120. https://doi.org/10.1016/j.jtbi.2004.06.005
Ohtsuki, H., Iwasa, Y., 2006. The leading eight: social norms that can maintain cooperation by indirect reciprocity. J. Theor. Biol. 239 (4), 435–444. https://doi.org/10.1016/j.jtbi.2005.08.008
Ostrom, E., 2005. Policies that crowd out reciprocity and collective action. In: Gintis, H. (Ed.), Moral Sentiments and Material Interests: the Foundations of Cooperation in Economic Life. MIT Press, Cambridge, Mass. Vol. 6 of *Economic learning and social evolution*, pp. 253–275.
Oyserman, D., 2017. Culture three ways: culture and subcultures within countries. Annu Rev Psychol 68, 435–463. https://doi.org/10.1146/annurev-psych-122414-033617
Panchanathan, K., Boyd, R., 2004. Indirect reciprocity can stabilize cooperation without the second-order free rider problem. Nature 432 (7016), 499–502. https://doi.org/10.1038/nature02978
Pearl, J., et al., 2011. Invited commentary: understanding bias amplification. Am. J. Epidemiol. 174 (11), 1223–1227. https://doi.org/10.1093/aje/kwr352
Pryor, F.L., 1985. The invention of the plow. Comp Stud Soc Hist 27 (4), 727–743. https://doi.org/10.1017/S0010417500011749
Rand, D.G., Nowak, M.A., 2013. Human cooperation. Trends Cogn. Sci. (Regul. Ed.) 17 (8), 413–425. https://doi.org/10.1016/j.tics.2013.06.003
Rand, D.G., Ohtsuki, H., Nowak, M.A., 2009. Direct reciprocity with costly punishment: generous tit-for-tat prevails. J. Theor. Biol. 256 (1), 45–57. https://doi.org/10.1016/j.jtbi.2008.09.015
Reeskens, T., Hooghe, M., 2008. Cross-cultural measurement equivalence of generalized trust. evidence from the european social survey (2002 and 2004). Soc Indic Res 85 (3), 515–532. https://doi.org/10.1007/s11205-007-9100-z
Richerson, P.J., Boyd, R., 2008. Not By Genes Alone: How Culture Transformed Human Evolution. University of Chicago Press.
Rothstein, B., Stolle, D., 2008. The state and social capital: an institutional theory of generalized trust. Comp Polit 40 (4), 441–459. https://doi.org/10.5129/001041508X12911362383354
Salali, G.D., Migliano, A.B., 2015. Future discounting in congo basin hunter-gatherers declines with socio-economic transitions. PLoS ONE 10 (9), e0137806. Publisher: Public Library of Science. https://doi.org/10.1371/journal.pone.0137806
dos Santos, M., Rankin, D.J., Wedekind, C., 2011. The evolution of punishment through reputation. Proceedings of the National Academy of Sciences, Biological Sciences 278 (1704), 371–7. https://doi.org/10.1098/rspb.2010.1275
Schiavina, M., Melchiorri, M., Pesaresi, M., 2023. GHS-SMOD R2023A - GHS settlement layers, application of the Degree of Urbanisation methodology (stage I) to GHS-POP R2023A and GHS-BUILT-S R2023A, multitemporal (1975–2030). https://doi.org/10.2905/A0DF7A6F-49DE-46EA-9BDE-563437A6E2BA
Schläpfer, A., 2024. Cross-cultural differences in retaliation: evidence from the soccer field. J. Econ. Behav. Organiz. 223, 216–233. https://doi.org/10.1016/j.jebo.2024.05.010
Schmid, L., Chatterjee, K., Hilbe, C., Nowak, M.A., 2021. A unified framework of direct and indirect reciprocity. Nat. Hum. Behav. 5 (10), 1292–1302. https://doi.org/10.1038/s41562-021-01114-8
Schneider, J., 1971. Of vigilance and virgins: honor, shame and access to resources in mediterranean societies. Ethnology 10 (1), 1–24. https://doi.org/10.2307/3772796
Schulz, J.F., 2022. Kin networks and institutional development. Econ. J. 132 (647), 2578–2613. https://doi.org/10.1093/ej/ueac027
Schulz, J.F., Bahrami-Rad, D., Beauchamp, J.P., Henrich, J., 2019. The church, intensive kinship, and global psychological variation. Science 366 (6466). https://doi.org/10.1126/science.aau5141
Sibhatu, K.T., Qaim, M., 2017. Rural food security, subsistence agriculture, and seasonality. PLoS ONE 12 (10), e0186406. https://doi.org/10.1371/journal.pone.0186406
Sommers, T., 2009. The two faces of revenge: moral responsibility and the culture of honor. Biol. Phil. 24 (1), 35–50. https://doi.org/10.1007/s10539-008-9112-3
Steiner, P.M., Kim, Y., 2016. The mechanics of omitted variable bias: bias amplification and cancellation of offsetting biases. J. Causal Infer. 4 (2), 20160009. https://doi.org/10.1515/jci-2016-0009
Sunde, U., Dohmen, T., Enke, B., Falk, A., Huffman, D., Meyerheim, G., 2022. Patience and comparative development. Rev. Econ. Stud. 89 (5), 2806–2840. https://doi.org/10.1093/restud/rdab084

Tabellini, G., 2008. The scope of cooperation: values and incentives. Q. J. Econ. 123 (3), 905–950. Publisher: Oxford Academic. https://doi.org/10.1162/qjec.2008.123.3.905
Tuden, A., Marshall, C., 1972. Political organization: cross-cultural codes 4. Ethnology 11 (4), 436–464. https://doi.org/10.2307/3773074
Uskul, A.K., Cross, S.E., 2020. Socio-ecological roots of cultures of honor. Curr. Opin. Psychol. 32, 177–180. https://doi.org/10.1016/j.copsyc.2019.11.001
Uskul, A.K., Cross, S.E., Günsoy, C., 2023. The role of honour in interpersonal, intrapersonal and intergroup processes. Soc. Personal Psychol. Compass 17 (1), e12719. https://doi.org/10.1111/spc3.12719
Vandello, J.A., Cohen, D., 2003. Male honor and female fidelity: implicit cultural scripts that perpetuate domestic violence. J. Pers. Soc. Psychol. 84 (5), 997. https://doi.org/10.1037/0022-3514.84.5.99
Vollan, B., 2012. Weird reciprocity? a 'within-culture across-country' trust experiment and methodological implications. J. Inst. Econ. 8 (3), 371–397. https://doi.org/10.1017/S1744137412000033
Vonasch, A.J., Reynolds, T., Winegard, B.M., Baumeister, R.F., 2018. Death before dishonor. Soc. Psychol. Personal. Sci. 9 (5), 604–613. https://doi.org/10.1177/1948550617720271
Wang, M., Rieger, M.O., Hens, T., 2016. How time preferences differ: evidence from 53 countries. J. Econ. Psychol. 52, 115–135. https://doi.org/10.1016/j.joep.2015.12.001
Wheeler, M.A., McGrath, M.J., Haslam, N., 2019. Twentieth century morality: the rise and fall of moral concepts from 1900 to 2007. PLoS ONE 14 (2), e0212267. https://doi.org/10.1371/journal.pone.0212267
World Bank, W., 2024. Urban population (% of total population). https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS.
World Bank, W., 2025. Poverty and Inequality Platform (version 20250401_2021_01_02_PROD). https://pip.worldbank.org.
Özak, O., 2015. Caloric Suitability Index - Data. https://doi.org/10.5281/zenodo.14714917