Fear of supernatural punishment can harmonize human societies with nature: an evolutionary game-theoretic approach

Shota Shibasaki[1, †], Yo Nakawake[2,3], Wakaba Tateishi[4], Shuhei Fujii[5], and Ryosuke Nakadai[6,7]

[1]Center for Frontier Research, National Institute of Genetics, 1111 Yata, Mishima, Shizuoka 411-8540, Japan

[†]Current affiliation: Faculty of Culture and Information Science, Doshisha University, Tataramiyakodani 1-3, Kyotanabe, Kyoto, 610-0394,Japan

[2]Graduate School of Advanced Science and Technology, Japan Advanced Institute of Science and Technology, 1-1 Asahidai, Nomi, Ishikawa, 923-1211, Japan

[3]School of Anthropology and Museum Ethnography, University of Oxford, OX26PE, UK

[4]Hokkaido Musashi Women's University,Kita 23-Jo Nishi 14-Chome, Kita-ward, Sapporo, Hokkaido 001-0023, Japan

[5]The Institute for Japanese Culture and Classics, Kokugakuin University, 4-10-28 Higashi, Shibuya-Ward, Tokyo, 150-0011, Japan

[6]Faculty of Environment and Information Sciences, Yokohama National University, 79-7 Tokiwadai, Hodogaya-Ward, Yokohama, Kanagawa 240-8501, Japan

[7]Research Institute for Humanity and Nature, 457-4 Motoyama, Kamigamo, Kita-Ward, Kyoto, 603-8047, Japan

Keywords: environmental feedback game; evolutionary game theory; mathematical model; supernatural belief; tragedy of the commons.

Abstract

Human activities largely impact the natural environment negatively and radical changes in human societies would be required to achieve their sustainable relationship with nature. Although frequently overlooked, previous studies have suggested that supernatural beliefs can protect nature from human overexploitation via beliefs that supernatural entities punish people who harm nature. Studies of folklore and ethnology have shown that such supernatural beliefs are widely found. However, it remains unclear under which conditions such supernatural beliefs prevent people from harming nature, because overexploiting natural resources without supernatural beliefs produces the greatest benefits. The current study aimed to build a mathematical model based on the evolutionary game theory and derive the conditions under which supernatural beliefs can spread in society, thereby preserving natural resources. To maintain supernatural beliefs, the fear of supernatural punishment invoked by scarce natural environments would, on one hand, be strong enough to prevent overexploitation but, on the other, be weak enough for the supernatural belief to spread in society via missionary events. Our results supported that supernatural beliefs would facilitate sustainable relationships between human societies and nature. In particular, the study highlighted supernatural beliefs as an essential driver for achieving sustainability by altering peoples interaction with nature.

1 Introduction

Negative human impacts on natural environments have been widely recognised (Cardinale et al., 2012; Dirzo et al., 2014; Malhi et al., 2014), and fundamental changes in human societies are considered necessary to achieve a sustainable relationship with nature (McPhearson et al., 2021; Pascual et al., 2023). Beliefs in supernatural entities that punish people who harm nature may play an important role in harmonising human societies with nature (Purzycki et al., 2022). Previous folklore and ethnological studies have shown that such supernatural beliefs exist across human societies and may protect nature from human overexploitation (Hartberg et al., 2016). For example, Frazer (1890) recorded the taboos of plant abuse worldwide. Ethnographic data analysis revealed that Japanese folklore includes episodes where spirits of nature (e.g., mountains and trees) punish or avenge people who develop or overuse natural resources (Nakawake and Sato, 2022). Similarly, the Batak people of Palawan Island in the Philippines believe in the forest spirits that punish people who overexploit or waste forest resources (Eder, 1997). Itza' Maya, Guatemala, also views forest spirits as punitively protecting local forests against exploitation (Atran et al., 2002). It remains unclear, however, under which conditions human society can maintain the beliefs in supernatural punishment and when such beliefs can protect nature from human overexploitation.

The problem of overusing natural resources is referred to as the tragedy of the commons (Hardin, 1968) in the context of the evolution of cooperation. If a society is composed of cooperators who self-regulate the usage of nature, natural resources can remain abundant, and people can continue to earn great benefits therefrom. Such a society is, however, vulnerable to invasion by selfish individuals who overexploit natural resources since the selfish people gain more benefits than the cooperators. Previous studies have shown that cooperation can evolve if cooperative individuals interact more frequently with other cooperators than with selfish ones via kin selection, multi-level selection, direct reciprocity, and indirect reciprocity (Rand and Nowak, 2013; Apicella and Silk, 2019). For example, punishing selfish individuals is a form of direct or indirect reciprocity that facilitates the evolution of cooperation (Fowler, 2005; Brandt et al., 2006; Hauert et al., 2007). Although humans can spontaneously punish selfish individuals (Yamagishi, 1988; Fehr and Gächter, 2002; Fehr and Fischbacher, 2004; Henrich et al., 2006; Rand et al., 2009; Raihani and Bshary, 2019), such punishments are accompanied by the problem of costs. Punishers need to spend time or energy to monitor and punish others, and they may be retaliated upon by the punished individuals (Denant-Boemont et al., 2007; Janssen and Bushman, 2008). As a result, cooperation collapses due to the increase of individuals who do not contribute to the costly punishment (Sigmund, 2007). This remains a central problem in the evolution of cooperation and punishment.

In human societies, beliefs in supernatural punishment may solve the problem of costly punishments (Johnson and Krüger, 2004; Bourrat and Viciana, 2016; Lightner and Purzycki, 2021; Schloss and Murray, 2011; Fitouchi et al., 2023), although scholars have debated whether such beliefs drove the evolution of social complexity (Turchin et al., 2023a,b; Whitehouse et al., 2023). Supernatural punishment is advantageous over "real" punishment because people do not have to bear the costs of the punishments (Johnson and Bering, 2006). Thus, the fear of supernatural punishment can prevent believers from behaving selfishly; but see Lenfesty and Morgan

(2019) for an alternative hypothesis on how supernatural beliefs facilitate the evolution of human cooperation. The moralising gods hypothesis associates the cooperation in human societies with the belief in moralising gods, who monitor human activities and enforce moralistic behaviors (Johnson, 2005; Watts et al., 2015; Purzycki et al., 2016; Bayramoglu et al., 2018; Lang et al., 2019; Singh et al., 2021). Although the relationship between humans and nature was not considered in this hypothesis, some scholars argue that supernatural beliefs can also regulate human behaviors toward nature (Bendixen et al., 2023; Bendixen and Grant Purzycki, 2023). To clarify the conditions under which supernatural beliefs contribute to sustainability, we must examine the conditions that allow these beliefs to persist in human societies and regulate the human usage of natural resources.

Here, we built and investigated a mathematical model to reveal whether and how beliefs in supernatural punishment facilitate the sustainable relationship between human societies and nature. We used evolutionary game theory to analyze the co-evolutionary dynamics of three elements as follows: (i) the belief in supernatural punishment, (ii) the intensity of human exploitation of nature, and (iii) the amount of natural resources. Recent advances in the evolutionary game theory have introduced the environmental feedback games (Weitz et al., 2016; Tilman et al., 2020; Ito and Yamamichi, 2024), in which the payoffs depend on the individuals' strategies (for example, how many trees people cut) and current environmental states (for example, the abundance of trees in a forest). At the same time, the environment also changes depending on the strategies individuals apply. This is an ideal framework for investigating how the evolution of both human behaviors and beliefs affects natural resources as public goods. We mathematically derived two conditions under which the beliefs in supernatural punishment could spread in human society and protect nature from overuse. Intuitively, the first condition indicates that the fear of supernatural punishment should exceed the net benefits of overexploiting natural resources so that believers stop the overexploitation. The second condition represents that the fear of supernatural punishment should be small so that people can accept the supernatural beliefs through the missionary events. Our study could provide a theoretical foundation for how and when supernatural beliefs can facilitate the sustainable relationship between human societies and nature.

2 Model

In this study, we considered the public goods game, including the environmental feedback (Weitz et al., 2016; Tilman et al., 2020; Ito and Yamamichi, 2024) and the positive or negative missionary events (Fig. 1). We considered an infinite human population in which each individual was characterized by two binary independent aspects. The first aspect distinguishes the usage of natural resources (Fig. 1A). We call individuals cooperators if he/she exploits only a small amount of the natural resource so that the resource is conserved. In contrast, selfish people are those who exploit the natural resources more than the cooperators to earn larger benefits. The second aspect represents whether each individual believes in supernatural punishment when he/she overexploits natural resources (Fig. 1B). We assumed that selfish believers bore the cost of fearing supernatural punishment even when they were not really punished, because studies suggested that religious guilt can damage mental health. (see the meta-analysis by Aggarwal et al., 2023).

Combining the benefits of natural resources and the fear of supernatural punishment, the payoffs of the four strategies – cooperative believers (CS), selfish believers (SB), cooperative non-believers (CN), and selfish non-believers (SN) – were represented as follows:

$$f_{\mathrm{CB}}(R) = (aR)^w, \tag{1a}$$

$$f_{\mathrm{SB}}(R) = (bR)^w - P(R), \tag{1b}$$

$$f_{\mathrm{CN}}(R) = (aR)^w, \tag{1c}$$

$$f_{\mathrm{SN}}(R) = (bR)^w, \tag{1d}$$

where a and b were the rates of natural resource exploitation by cooperative and selfish individuals ($b > a > 0$), respectively, and $P(R)$ is the fear of supernatural punishment when the amount of natural resource is R. Here, $w > 0$ determines how the benefit increases over R; $w = 1$ corresponds to a linear function, $0 < w < 1$ corresponds to a concave function, and $w > 1$ corresponds to a convex function. Regardless of the value of w, the selfish strategy always had greater benefits than the cooperative strategy, which led to the tragedy of the commons (Hardin, 1968). The fear of supernatural punishment $P(R)$ differs in two aspects from the models of real punishments. First, no individual in our model pays costs for punishing others since no one punishes selfish individuals in our model. Second, selfish people bear the costs of the fear of supernatural punishment $P(R)$ only if he/she believes in supernatural punishment; selfish non-believers do not bear this cost. This highlights the difference from typical real punishment systems, in which all selfish individuals are punished.

The payoff functions (Eqs. 1a-1d) clarified that we did not assume that believing in supernatural punishment was adaptive, although supernatural beliefs could motivate individuals to cooperate. The payoffs of the cooperators did not change regardless of whether they believed in supernatural punishment. For selfish individuals, on the other hand, believing in supernatural punishment decreased their payoffs since the fear of supernatural punishment damaged their mental health.

The strength of supernatural beliefs was assumed to positively correlate with the extent of nature. In other words, we assumed that people may be more likely to perceive spiritual entities in richer natural elements, fostering religious beliefs grounded in awe and fear (Frazer, 1890). Based on this idea, we assumed that the amount of natural resources increased the perceived fear of the supernatural punishment (and the associated costs) P. Similar to the benefits obtained from natural resources, the fear of supernatural punishment was formulated using the following equations:

$$P(R) \equiv (pR)^u \tag{2}$$

where p^u is the fear of supernatural punishment when $R = 1$, and $u > 0$ determines the shape of the function $P(R)$ over a natural resource. In SI 6, we relaxed this assumption and analyzed the cases when the fear of supernatural punishment decreased over R, which did not qualitatively alter our findings.

Our model considered the public goods game with environmental feedback so that the dynamics of the natural resources were explicitly represented (Estrela et al., 2019). This allowed us to incorporate the difference in time scales between the evolution of human behavior and the recovery of natural resources. Here, we assumed that the natural resource was recovered following a logistic growth model, but was consumed by the local people whose exploitation rates were either a or b. The dynamics of the natural resource followed a classical consumer-resource model (MacArthur, 1970) with four parameters: the intrinsic growth rate (per-capita growth rate when the natural resource is scarce) μ, the carrying capacity (the maximum amount of natural resource) K, and the consumption rates by the cooperative or selfish individuals (a and b, respectively). Sethi and Somanathan (1996) analyzed a similar model that combined resource dynamics with the evolution of cooperation and real punishment. On the other hand, our model analyzed the role of belief in supernatural punishment and whether such beliefs could be maintained in a population.

Further, we assumed that whether an individual believed in the supernatural punishment changed due to the positive and negative missionary events (Figs. 1C and D); non-believers became believers when they frequently interacted with the latter at the rate v_+ (the positive missionary rate), and vice versa (the negative missionary rate v_-). This formulation follows a typical epidemiological analogy (Olsson and Galesic, 2024). Positive and negative missionary events can be justified by combining the positive frequency-dependent biases (i.e., mimicking the majority) and the content biases (i.e., difference in cognitive attractiveness) (Mesoudi, 2016). If most people believe in supernatural punishment and the beliefs are readily transmitted, for example, the non-believers can immediately become believers.

The governing dynamics of human behavior and natural resources are composed of the replicator dynamics (Nowak, 2006) with positive and negative missionary events and the consumer-resource model:

$$\epsilon\dot{x}_{\mathrm{CB}} = x_{\mathrm{CB}}\left(f_{\mathrm{CB}}(R) - \bar{f}(R)\right) + v_+\left(x_{\mathrm{CB}} + x_{\mathrm{SB}}\right)x_{\mathrm{CN}} - v_-\left(x_{\mathrm{CN}} + x_{\mathrm{SN}}\right)x_{\mathrm{CB}}, \tag{3a}$$

$$\epsilon\dot{x}_{\mathrm{SB}} = x_{\mathrm{SB}}\left(f_{\mathrm{SB}}(R) - \bar{f}(R)\right) + v_+\left(x_{\mathrm{CB}} + x_{\mathrm{SB}}\right)x_{\mathrm{SN}} - v_-\left(x_{\mathrm{CN}} + x_{\mathrm{SN}}\right)x_{\mathrm{SB}}, \tag{3b}$$

$$\epsilon\dot{x}_{\mathrm{CN}} = x_{\mathrm{CN}}\left(f_{\mathrm{CN}}(R) - \bar{f}(R)\right) + v_-\left(x_{\mathrm{CN}} + x_{\mathrm{SN}}\right)x_{\mathrm{CB}} - v_+\left(x_{\mathrm{CB}} + x_{\mathrm{SB}}\right)x_{\mathrm{CN}}, \tag{3c}$$

$$\epsilon\dot{x}_{\mathrm{SN}} = x_{\mathrm{SN}}\left(f_{\mathrm{SN}}(R) - \bar{f}(R)\right) + v_-\left(x_{\mathrm{CN}} + x_{\mathrm{SN}}\right)x_{\mathrm{SB}} - v_+\left(x_{\mathrm{CB}} + x_{\mathrm{SB}}\right)x_{\mathrm{SN}} \tag{3d}$$

$$\dot{R} = \mu R\left(1 - \frac{R}{K}\right) - R\left\{a\left(x_{\mathrm{CB}} + x_{\mathrm{CN}}\right) + b\left(x_{\mathrm{SB}} + x_{\mathrm{SN}}\right)\right\} \tag{3e}$$

where x_i is the fraction of strategy i (i = CB, SB, CN, SN), $\bar{f}(R) = \sum_i x_i f_i(R)$ is the average payoff in the population, ϵ changes the time scales of the dynamics of the human behavior and the natural resources: $1 > \epsilon > 0$ indicates that the evolutionary dynamics of the human behavior to be faster than that of the natural resources. In contrast, $\epsilon > 1$ represents that the evolution of human behavior to be slower than the dynamics of the natural resources. The dynamics of human behaviors and beliefs affect the dynamics of natural resources, while the amounts of natural resources affect the payoffs by changing the benefits from the natural resources and the fear of supernatural punishment; our model investigates the public goods game accompanying the environmental feedback and cultural evolution of supernatural beliefs (Fig 1E).

For ease of analysis, we define the temptation to selfishness (i.e., the difference in the benefits between selfishness and cooperation) as follows:

$$\Delta(R) \equiv (bR)^w - (aR)^w \geq 0. \tag{4}$$

Table 1 lists the key variables and parameters.

A strategy is evolutionarily stable if it is not invaded by any other strategy (Maynard Smith and Price, 1973). The temptation to selfishness $\Delta(R)$, the fear of supernatural punishment $P(R)$, the positive missionary rate v_+, and the negative missionary rate v_- determined whether a strategy was evolutionarily stable in our model (see SI 1 for derivation).

Numerical simulations were performed by the solve_ivp function with the RK45 method in Scipy version 1.11.3 (Virtanen et al., 2020) in Python 3.11.5. To analyze how parameter values affected the dynamics, we fixed the step size as 0.01 so that the solve_ivp function would not change the step size depending on the parameter values. We evaluated the average density of each strategy and natural resource at time $T_f - 100 \leq t \leq T_f$, where the simulation finished at $t = T_f$. If the average density of a strategy was equal to or smaller than 10^{-4}, we regarded it as extinct; otherwise, it persisted. For a persistent strategy, we evaluated the coefficient of variation at time $T_f - 100 \leq t \leq T_f$. For strategies that went extinct, the coefficient of variation was set to 0. If the mean of the coefficient of variation across the four strategies exceeded 0.1, the dynamics were considered to be oscillating.

3 Results

3.1 Selfish non-believers are stable without positive and negative missionary events

We first began by analysing the simplest model without the positive and negative missionary events ($v_+ = v_- = 0$). SN was evolutionarily stable without the missionaries because the payoff of SN was the highest when $R > 0$. The amount of the natural resource, in this case, remained at its minimum value $R_b^* \equiv K(1 - b/\mu)$. By incorporating the positive and negative missionary events in the following subsections, we aimed to determine the conditions under which the cooperators evolved and the amount of the natural resources exceeded R_b^*.

3.2 Introduction of positive missionary events stabilises cooperative believers and conserves the natural resource

Next, the positive missionary events were introduced into the model ($v_+ > 0$) while the negative missionary events were not ($v_- = 0$). This led to the fixation of CB, and we investigated how the parameter values changed the evolutionary fate.

When the evolutionary dynamics of human behaviors were much faster than that of the natural resources $\epsilon \to 0$, we assumed that the amount of the natural resource R, the temptation to the selfishness Δ, and the fear of the supernatural punishment P were constant over time. Then, either the CB, SB, or SN was evolutionarily stable depending on the inequality among Δ, P, and v_+:

- $v_+ > P > \Delta$: CB was evolutionarily stable (Fig. 2A).
- $P > v_+ > \Delta$: Both CB and SN were evolutionarily stable (Figs. 2C and D).
- $v_+ > \Delta > P$ or $\Delta > v_+ > P$: SB was evolutionarily stable (Fig. 2E).
- $P > \Delta > v_+$ or $\Delta > P > v_+$: SN was evolutionarily stable (Figs. 2B and F).

In other words, the dynamics always converged to one of the three equilibria (fixation of CB, SB, or SN). The initial conditions and the parameter values determined the strategy that was ultimately fixed.

When the amount of natural resources changed over time, we derived similar conditions for the evolutionarily stable strategies by replacing Δ and P with $\Delta(R)$ and $P(R)$, respectively, at equilibrium (see SI 1 for details). In other words, the CB (Fig. 3A), SB (Fig. 3B), and SN (Fig. 3C) would be evolutionarily stable under evolving the amount of the natural resources. Fig. 3 shows that the amount of the natural resource at the equilibrium was the highest ($R^*_a \equiv K(1 - a/\mu)$) when the CB was evolutionarily stable.

Unlike the constant resource scenario (Fig. 2), none of these strategies could be evolutionarily stable when the amount of the natural resources changed over time. Furthermore, this scenario stabilised the coexistence of multiple strategies in two types of equilibria, one where CB coexisted with SB (Fig. 3D), and the other when CB coexisted with SB and SN (Figs. 4A and B). The local stability conditions of the equilibria were analytically derived assuming that a fraction of CN remained negligible ($x_{\mathrm{CN}} \approx 0$; see SI 2 for details). Remarkably, the time-scale parameter ϵ affected the stability of the equilibrium where CB, SB, and SN coexisted (Fig. 4C). We further observed this oscillatory dynamics when none of the equilibria was stable (Fig. 4D).

Conversely, CN could not coexist with any other strategy since the payoff of CB and CN were identical for any R and the negative missionary events were not allowed in the current setting (see SI 3 for mathematical details).

3.3 A small negative missionary rate allows the evolution of cooperation and the maintenance of the natural resources

The full model included the dynamics of the natural resources, positive missionary events, and negative missionary events ($v_+, v_- > 0$). Due to its high dimensionality and nonlinearity, it was challenging to derive complete analytical solutions for this model. However, we derived the conditions under which the selfish strategies cannot be fixed, resulting in the amount of natural resources exceeding its minimum ($R{*}_b$).

From the calculation in SI 1, neither SB nor SN is evolutionarily stable if and only if

$$v_+ - v_- > P(R^*_b) > \Delta(R^*_b). \tag{5}$$

Intuitively, this inequality means that the fear of the supernatural punishment needs to be stronger than the temptation to selfishness (i.e., allowing the CB to invade the SB) while it needs to be smaller than the positive missionary rate minus the negative missionary rate (so that the SN is not evolutionarily stable). It should be noted that the coexistence of SB and SN was unstable in the presence of the negative missionary events (see SI 4). When the inequalities (5) are satisfied, the cooperators can, therefore, evolve, and the amount of the natural resources can be higher than its minimum value R_b^*.

Fig. 5 shows how the negative missionary rate v_- and the exploitation rate of cooperators a affect the evolutionary dynamics. The horizontal and vertical dashed vertical lines in Fig. 5A represent the two thresholds $P(R_b^*) = \Delta(R_b^*)$ and $v_+ = v_- + P(R_b^*)$, respectively. When the negative missionary rate was sufficiently high, the dynamics converged to the fixation of SN (the black areas on the right in Fig. 5A) in most cases since it was evolutionarily stable. When the exploitation rate of the cooperators was large and close to that of the selfish strategies while the negative missionary rate remained high (the top-right skyblue area surrounded by the gray line in Fig. 5A), the CB could also be fixed; both CB and SN were evolutionarily stable in these parameter ranges and thus the initial conditions determined which strategy was fixed (Figs. 5C and D). When the negative missionary rate was low and the exploitation rate of the cooperators was below the threshold, the SB was fixed (the bottom-left orange areas in Fig. 5A) since the temptation to selfishness was so large that the fear of the supernatural punishment did not allow the invasion by CB. If the exploitation rate by the cooperators was sufficiently large and the negative missionary rate was low, CB could persist and potentially coexist with other strategies (the top left blue, green, or pink areas in Fig. 5A). Further, we also observed the oscillations when multiple strategies coexisted (the cross marks in Fig. 5A).

The average amounts of the natural resources at time $T_f - 100 \leq t \leq T_f$ are shown in Fig. 5B. The persistence of CB resulted in a higher amount of natural resources than its minimum R_b^*. In particular, the amount of natural resources reached its maximum R_a^* when CB was fixed. Although the parameter space where CB was fixed increased over the exploitation rate of the cooperators a, increasing a resulted in fewer resource availabilities since R_a^* decreased linearly over a. Overall, the highest natural resource was achieved when the negative missionary rate was lower than the threshold and when the exploitation rate of the cooperators was the lowest value that fixated CB.

Next, we also examined how the evolutionary fate changed when the temptation to selfishness or the fear of supernatural punishment was a nonlinear function of R. SI 5 shows that the nonlinear functions of the $P(R)$ decreased the area where the cooperative believers persisted. In contrast, the cooperative believers remained in a broader parameter space when the temptation $\Delta(R)$ was a convex function. We also examined the cases when the fear of supernatural punishment decreased over the amount of natural resources (SI 6). In all cases, inequality (5) provided the information on when the CB could persist and when the natural resource remained higher than the minimum.

4 Discussion

Previous studies have discussed how supernatural beliefs affect human activities, including achieving sustainability (Rolston, 2006; Rakodi, 2012). While the moralising gods hypothesis associates the norms in human relationships with complex human societies (Purzycki et al., 2016; Watts et al., 2015), supernatural beliefs may also impose norms on the relationship between humans and nature (Purzycki et al., 2022; Bendixen et al., 2023; Bendixen and Grant Purzycki, 2023). However, it remains unclear when beliefs in supernatural punishment can spread in human society and whether such a belief can harmonise human society with nature. We built a formal mathematical model to investigate the coevolutionary dynamics of human exploitation of natural resources, belief in supernatural punishment, and the amount of natural resources. The mathematical analysis revealed two conditions under which supernatural beliefs can be maintained in human society while sustaining abundant natural resources: the fear of supernatural punishment should be larger than the temptation to selfishness and be smaller than the positive missionary rate minus the negative one. While a previous study shows the natural impact on human beliefs (Nakadai, 2023), our results suggest how supernatural beliefs affect natural environments.

Inequality (5) show the two conditions under which beliefs in supernatural punishment could facilitate sustainability. These conditions clarified the similarity and difference between the systems with real punishment and beliefs in supernatural punishment. The first condition, $P(R_n^*) > \Delta(R_b^*)$, implies that the fear of supernatural punishment needs to be stronger than the temptation to selfishness so that cooperative believers are more adaptive than selfish believers. Studies on the evolution of cooperation under real punishment derived similar conditions; cooperation can evolve if the punishment is strong enough to make selfish behaviors maladaptive (p.283 Broom and Rychtar, 2013; Fowler, 2005; Nakamaru and Iwasa, 2006). Both real punishment and beliefs in supernatural punishment can facilitate the sustainable relationship between human society and nature if they invoke strong (fear of) punishment on those who harm nature. The second condition in inequality (5), $v_+ - v_- > P(R_b^*)$, argued that supernatural beliefs should spread more efficiently via positive missionary events than it is lost via negative missionary events. This condition highlights the difference between real punishment and beliefs in supernatural punishments. The systems of real punishment need to compensate the costs of punishing others (Boyd et al., 2003; Gardner and West, 2004; Dos Santos et al., 2011); otherwise, punishing systems collapse. In the current model, such conditions do not exist because no one pays the cost for punishing selfish individuals. Instead, supernatural beliefs should easily spread in human society because only believers would be afraid of supernatural punishment. If this condition is not satisfied, individuals selfishly behave and natural resources remain scarce (e.g., the bottom-right area in Figs. 5A and B). In short, real punishment and beliefs in supernatural punishment have different obstacles to facilitating sustainable relationships between human society and nature.

It is beyond the scope of this manuscript to formally test our theoretical prediction with empirical data. Nevertheless, ethnological and psychological studies suggest that inequality (5) seems reasonable. Consistent with our first condition, previous studies have shown that certain supernatural beliefs invoke a strong fear of supernatural punishment. Nakawake and Sato (2022) quantitatively showed that severe supernatural punishment (e.g., the death of members of a village, kinship, or family of individuals who harm nature) is typical in Japanese folklore. A global meta-analysis by Hartberg et al. (2016) show that one-third of supernatural punishment results in the death of those who harm nature.For example, cutting trees on a mountain was believed to cause a flood that could wash away all houses in a village (Sakurai, 1999). When and how could a belief in supernatural punishment spread efficiently in human society via missionaries and satisfy the second condition in inequality (5)? One possibility is that supernatural or religious beliefs are a by-product of cognitive adaptation and thus likely to be accepted (Boyer, 2003). For example, the minimally counterintuitive theory suggests that many religious concepts violate an optimal number of our expectations, which increases their memorability and helps them spread (Boyer, 2003; Barrett and Nyhof, 2001). Another possibility is the prestige bias; if prestigious people believe in supernatural punishment for any reason, other people would also start believing in the one-to-many transmission of supernatural beliefs. In fact, in many religious traditions, religious leaders tend to gain power in non-religious domains, such as political or juridical domains (Winkelman, 1990), which might strengthen their prestige as religious leaders. Further, costly religious rituals practised by religious leaders might also help spread religious beliefs (Sosis, 2003; Norenzayan et al., 2016). To test our theoretical prediction, future studies are encouraged to compare the degree of fear of supernatural punishments, transmission rates of such beliefs, and abundance of natural resources.

Our results may be generalized to non-supernatural belief systems that impose norms on human-nature relationships without punishment or sanction (e.g., environmental ethics). In such cases, $P(R)$ should be regarded as the strength of guilt when an individual feels in overexploiting nature, and v_+ and v_- represent the rates of acquisition and loss of such morality through interaction with others, respectively. In this framework, inequality (5) represents the conditions under which such moral systems – whether supernatural beliefs or ethics – can facilitate the sustainable relationship between human society and nature. This inequality suggests that moral systems invoking a too strong feeling of guilt fail in achieving a sustainable relationship because the second condition, $v_+ - v_- > P(R_b^*)$, cannot be met. Instead, increasing the transmission rate of the moral systems (v_+) would successfully lead to sustainable resource usage because $v_+ - v_-$ becomes larger. These conditions suggest how to design environmental moral systems to contribute to sustainability successfully. In a society with a belief in supernatural punishment, for example, incorporating environmental ethics in the supernatural belief (i.e., introducing $P(R)$) may be more effective than trying to spread the environmental ethics through secular approaches, because the supernatural belief would have a large v_+. Alternatively, using non-rational narratives that are easy to accept, including supernatural beliefs, may facilitate the transmission of environmental ethics among individuals because these narratives would increase v_+ of the environmental ethics.

Our model could also be extended to a quantitative model by formulating the evolution of the exploitation rate of natural resources and the strength of belief in supernatural punishment. However, we focused on the current qualitative strategies because, to the best of our knowledge, the current model is the first rigorous formulation of the coevolutionary dynamics of human behaviors, beliefs in supernatural punishment, and natural resources. Future studies should investigate whether the results in this manuscript are valid for quantitative models because such models would be easier to compare with empirical data than our current model.

In conclusion, our model provides a theoretical foundation for supernatural beliefs to facilitate sustainability. While the moralising gods hypothesis argues that some supernatural beliefs impose norms on human relationships, others regulate the relationship between humans and nature. Our mathematical model suggested two conditions under which such supernatural beliefs could prevent humans from overexploiting nature through the fear of supernatural punishment: the fear of supernatural punishment should be larger than the temptation to selfishness and be smaller than the positive missionary rate minus the negative one. Although believing in supernatural punishment is not adaptive, positive missionary events can stabilise cooperative individuals who believe in supernatural punishment and self-regulate the exploitation of nature. Even if they are not evolutionarily stable, cooperative believers can coexist with selfish believers and non-believers when the two conditions are met. Therefore, the current results supported the idea that supernatural beliefs harmonise human societies and nature, and that supernatural beliefs could play an important role in achieving sustainability. Future studies are encouraged to empirically test our theoretical prediction by examining the association among the degree of fear of supernatural punishments, transmission rates of such supernatural beliefs, and abundance of natural resources.

References

Aggarwal, S., Wright, J., Morgan, A., Patton, G., and Reavley, N. Religiosity and spirituality in the prevention and management of depression and anxiety in young people: a systematic review and meta-analysis. BMC Psychiatry, 23(1), 12 2023. ISSN 1471244X. doi: 10.1186/s12888-023-05091-2.

Apicella, C. L. and Silk, J. B. The evolution of human cooperation. Current Biology, 29(11):R447–R450, 6 2019. ISSN 09609822. doi: 10.1016/j.cub.2019.03.036. URL https://linkinghub.elsevier.com/retrieve/pii/S0960982219303343.

Atran, S., Medin, D., Ross, N., Lynch, E., Vapnarsky, V., Ek, E., Coley, J., Timura, C., and Baran, M. Folkecology, Cultural Epidemiology, and the Spirit of the Commons. Current Anthropology, 43(3):421–450, 6 2002. ISSN 0011-3204. doi: 10.1086/339528. URL http://www.journals.uchicago.edu/doi/10.1086/339528.

Barrett, J. L. and Nyhof, M. A. Spreading non-natural concepts: The role of intuitive conceptual structures in memory and transmission of cultural materials. Journal of Cognition and Culture, 1(1):69–100, 2001. ISSN 15685373. doi: 10.1163/156853701300063589.

Bayramoglu, Y., Harma, M., and Yilmaz, O. The Relationship between Attachment to God, Prosociality, and Image of God. Archive for the Psychology of Religion, 40(2-3):202–224, 12 2018. ISSN 0084-6724. doi: 10.1163/15736121-12341356. URL https://journals.sagepub.com/doi/10.1163/15736121-12341356.

Bendixen, T. and Grant Purzycki, B. Accounting for Cross-Cultural Variation in the Minds of Gods. In The Minds of Gods. Bloomsbury Academic, 2023. doi: 10.5040/9781350265738.ch-0013. URL https://www.bloomsburycollections.com/monograph-detail?docid=b-9781350265738&tocid=b-9781350265738-chapter13.

Bendixen, T., Lightner, A. D., and Purzycki, B. G. The Cultural Evolution of Religion and Cooperation. In The Oxford Handbook of Cultural Evolution. Oxford University Press, 4 2023. doi: 10.1093/oxfordhb/9780198869252.013.47. URL https://academic.oup.com/edited-volume/45648/chapter/402524801.

Bourrat, P. and Viciana, H. Supernatural Beliefs and the Evolution of Cooperation. In The Oxford Handbook of Evolutionary Psychology and Religion, pages 297–314. Oxford University Press, 6 2016. doi: 10.1093/oxfordhb/9780199397747.013.23. URL https://academic.oup.com/edited-volume/34242/chapter/290337927.

Boyd, R., Gintis, H., Bowles, S., and Richerson, P. J. The evolution of altruistic punishment. Proceedings of the National Academy of Sciences, 100(6):3531–3535, 3 2003. ISSN 0027-8424. doi: 10.1073/pnas.0630443100. URL https://pnas.org/doi/full/10.1073/pnas.0630443100.

Boyer, P. Religious thought and behaviour as by-products of brain function. Trends in Cognitive Sciences, 7(3):119–124, 2003. ISSN 13646613. doi: 10.1016/S1364-6613(03)00031-7.

Brandt, H., Hauert, C., and Sigmund, K. Punishing and abstaining for public goods. Proceedings of the National Academy of Sciences of the United States of America, 103(2):495–497, 1 2006. ISSN 00278424. doi: 10.1073/pnas.0507229103.

Broom, M. and Rychtar, J. Game-Theoretical Models in Biology. Chapman and Hall/CRC, 1 edition, 2013. ISBN 9781439853221. URL https://www.crcpress.com/Game-Theoretical-Models-in-Biology/Broom-Rychtar/9781439853214.

Cardinale, B. J., Duffy, J. E., Gonzalez, A., Hooper, D. U., Perrings, C., Venail, P., Narwani, A., Mace, G. M., Tilman, D., A.Wardle, D., Kinzig, A. P., Daily, G. C., Loreau, M., Grace, J. B., Larigauderie, A., Srivastava, D. S., and Naeem, S. Biodiversity loss and its impact on humanity. Nature, 489(7415):326–326, 2012. ISSN 0028-0836. doi: 10.1038/nature11373. URL http://www.nature.com/doifinder/10.1038/nature11373.

Denant-Boemont, L., Masclet, D., and Noussair, C. N. Punishment, counterpunishment and sanction enforcement in a social dilemma experiment. Economic Theory, 33(1):145–167, 10 2007. ISSN 0938-2259. doi: 10.1007/s00199-007-0212-0. URL https://link.springer.com/10.1007/s00199-007-0212-0.

Dirzo, R., Young, H. S., Galetti, M., Ceballos, G., Isaac, N. J. B., and Collen, B. Defaunation in the Anthropocene. Science, 345(6195):401–406, 7 2014. ISSN 0036-8075. doi: 10.1126/science.1251817. URL https://www.science.org/doi/10.1126/science.1251817.

Dos Santos, M., Rankin, D. J., and Wedekind, C. The evolution of punishment through reputation. Proceedings of the Royal Society B: Biological Sciences, 278(1704):371–377, 2 2011. ISSN 0962-8452. doi: 10.1098/rspb.2010.1275. URL https://royalsocietypublishing.org/doi/10.1098/rspb.2010.1275.

Eder, J. F. Batak Resource Management. IUCN, 1 1997. ISBN 978-2-8317-0366-4.

Estrela, S., Libby, E., Van Cleve, J., Débarre, F., Deforet, M., Harcombe, W. R., Peña, J., Brown, S. P., and Hochberg, M. E. Environmentally Mediated Social Dilemmas. Trends in Ecology & Evolution, 34(1):6–18, 1 2019. ISSN 01695347. doi: 10.1016/j.tree.2018.10.004. URL https://linkinghub.elsevier.com/retrieve/pii/S0169534718302490.

Fehr, E. and Fischbacher, U. Third-party punishment and social norms. Evolution and Human Behavior, 25(2):63–87, 3 2004. ISSN 10905138. doi: 10.1016/S1090-5138(04)00005-4.

Fehr, E. and Gächter, S. Altruistic punishment in humans. Nature, 415(6868):137–140, 1 2002. ISSN 0028-0836. doi: 10.1038/415137a. URL https://www.nature.com/articles/415137a.

Fitouchi, L., Singh, M., André, J.-B., and Baumard, N. Prosocial religions as folk-technologies of mutual policing. PsyArXiv, 2023. doi: https://doi.org/10.31234/osf.io/qdhka.

Fowler, J. H. Altruistic punishment and the origin of cooperation. Proceedings of the National Academy of Sciences, 102(19):7047–7049, 5 2005. ISSN 0027-8424. doi: 10.1073/pnas.0500938102. URL https://pnas.org/doi/full/10.1073/pnas.0500938102.

Frazer, J. G. The golden bough: a study in comparative religion. Macmillan, 1890.

Gardner, A. and West, S. A. Cooperation and punishment, especially in humans. American Naturalist, 164(6): 753–764, 12 2004. ISSN 00030147. doi: 10.1086/425623.

Hardin, G. The Tradegy of the Commons. Science, 162(June):1243–1248, 1968. ISSN 0036-8075. doi: 10.1126/science.162.3859.1243.

Hartberg, Y., Cox, M., and Villamayor-Tomas, S. Supernatural monitoring and sanctioning in community-based resource management. Religion, Brain & Behavior, 6(2):95–111, 4 2016. ISSN 2153-599X. doi: 10.1080/2153599X.2014.959547. URL http://www.tandfonline.com/doi/full/10.1080/2153599X.2014.959547.

Hauert, C., Traulsen, A., Brandt, H., Nowak, M. A., and Sigmund, K. Via freedom to coercion: The emergence of costly punishment. Science, 316(5833):1905–1907, 6 2007. ISSN 00368075. doi: 10.1126/science.1141588.

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., and Ziker, J. Costly Punishment Across Human Societies. Science, 312(5781):1767–1770, 6 2006. ISSN 0036-8075. doi: 10.1126/science.1127333. URL https://www.science.org/doi/10.1126/science.1127333.

Ito, H. and Yamamichi, M. A complete classification of evolutionary games with environmental feedback. PNAS Nexus, 10 2024. ISSN 2752-6542. doi: 10.1093/pnasnexus/pgae455. URL https://academic.oup.com/pnasnexus/advance-article/doi/10.1093/pnasnexus/pgae455/7817389.

Janssen, M. A. and Bushman, C. Evolution of cooperation and altruistic punishment when retaliation is possible. Journal of Theoretical Biology, 254(3):541–545, 10 2008. ISSN 00225193. doi: 10.1016/j.jtbi.2008.06.017.

Johnson, D. and Bering, J. Hand of God, Mind of Man: Punishment and Cognition in the Evolution of Cooperation. Evolutionary Psychology, 4(1):147470490600400, 1 2006. ISSN 1474-7049. doi: 10.1177/147470490600400119. URL http://journals.sagepub.com/doi/10.1177/147470490600400119.

Johnson, D. and Krüger, O. The Good of Wrath: Supernatural Punishment and the Evolution of Cooperation. Political Theology, 5(2):159–176, 2 2004. ISSN 1462-317X. doi: 10.1558/poth.2004.5.2.159. URL https://www.tandfonline.com/doi/full/10.1558/poth.2004.5.2.159.

Johnson, D. D. P. Gods punishment and public goods. Human Nature, 16(4):410–446, 12 2005. ISSN 1045-6767. doi: 10.1007/s12110-005-1017-0. URL http://link.springer.com/10.1007/s12110-005-1017-0.

Lang, M., Purzycki, B. G., Apicella, C. L., Atkinson, Q. D., Bolyanatz, A., Cohen, E., Handley, C., Kundtová Klocová, E., Lesorogol, C., Mathew, S., McNamara, R. A., Moya, C., Placek, C. D., Soler, M., Vardy, T., Weigel, J. L., Willard, A. K., Xygalatas, D., Norenzayan, A., and Henrich, J. Moralizing gods, impartiality and religious parochialism across 15 societies. Proceedings of the Royal Society B: Biological Sciences, 286(1898):20190202, 3 2019. ISSN 0962-8452. doi: 10.1098/rspb.2019.0202. URL https://royalsocietypublishing.org/doi/10.1098/rspb.2019.0202.

Lenfesty, H. L. and Morgan, T. J. By Reverence, Not Fear: Prestige, Religion, and Autonomic Regulation in the Evolution of Cooperation. Frontiers in Psychology, 10, 12 2019. ISSN 16641078. doi: 10.3389/fpsyg.2019.02750.

Lightner, A. D. and Purzycki, B. G. Game Theoretical Aspects of the Minds of Gods. PsyArxiv, 2021. doi: https://doi.org/10.31234/osf.io/ybwcd.

MacArthur, R. Species packing and competitive equilibrium for many species. Theoretical Population Biology, 1 (1):1–11, 5 1970. ISSN 00405809. doi: 10.1016/0040-5809(70)90039-0. URL https://linkinghub.elsevier.com/retrieve/pii/0040580970900390.

Malhi, Y., Gardner, T. A., Goldsmith, G. R., Silman, M. R., and Zelazowski, P. Tropical Forests in the Anthropocene. Annual Review of Environment and Resources, 39(1):125–159, 10 2014. ISSN 1543-5938. doi: 10.1146/annurev-environ-030713-155141. URL https://www.annualreviews.org/doi/10.1146/annurev-environ-030713-155141.

Maynard Smith, J. and Price, G. R. The Logic of Animal Conflict. Nature, 246(5427):15–18, 11 1973. ISSN 0028-0836. doi: 10.1038/246015a0. URL https://www.nature.com/articles/246015a0.

McPhearson, T., M. Raymond, C., Gulsrud, N., Albert, C., Coles, N., Fagerholm, N., Nagatsu, M., Olafsson, A. S., Soininen, N., and Vierikko, K. Radical changes are needed for transformations to a good Anthropocene. npj Urban Sustainability, 1(1):5, 2 2021. ISSN 2661-8001. doi: 10.1038/s42949-021-00017-x. URL https://www.nature.com/articles/s42949-021-00017-x.

Mesoudi, A. Cultural Evolution: A Review of Theory, Findings and Controversies. Evolutionary Biology, 43 (4):481–497, 12 2016. ISSN 00713260. doi: 10.1007/s11692-015-9320-0.

Nakadai, R. Macroecological processes drive spiritual ecosystem services obtained from giant trees. Nature Plants, 9(2):209–213, 2 2023. ISSN 2055-0278. doi: 10.1038/s41477-022-01337-1. URL https://www.nature.com/articles/s41477-022-01337-1.

Nakamaru, M. and Iwasa, Y. The coevolution of altruism and punishment: Role of the selfish punisher. Journal of Theoretical Biology, 240(3):475–488, 6 2006. ISSN 10958541. doi: 10.1016/j.jtbi.2005.10.011.

Nakawake, Y. and Sato, K. Does nature take revenge? a quantitative analysis of japanese folklore on supernatural revenges. Proceedings of Jinmoncom 2022, 2022:119–124, 12 2022. URL https://cir.nii.ac.jp/crid/1050294643541504896. in Japanese.

Norenzayan, A., Shariff, A. F., Gervais, W. M., Willard, A. K., McNamara, R. A., Slingerland, E., and Henrich, J. The cultural evolution of prosocial religions. Behavioral and Brain Sciences, 39:e1, 12 2016. ISSN 0140-525X. doi: 10.1017/S0140525X14001356. URL https://www.cambridge.org/core/product/identifier/S0140525X14001356/type/journal_article.

Nowak, M. A. Evolutionary dynamics exploring the equations of life. Belknap Press of Harvard University Press, Cambridge, 2006. ISBN 9780674023383.

Olsson, H. and Galesic, M. Analogies for modeling belief dynamics. Trends in Cognitive Sciences, 7 2024. ISSN 13646613. doi: 10.1016/j.tics.2024.07.001. URL https://linkinghub.elsevier.com/retrieve/pii/S1364661324001724.

Pascual, U., Balvanera, P., Anderson, C. B., Chaplin-Kramer, R., Christie, M., González-Jiménez, D., Martin, A., Raymond, C. M., Termansen, M., Vatn, A., Athayde, S., Baptiste, B., Barton, D. N., Jacobs, S., Kelemen, E., Kumar, R., Lazos, E., Mwampamba, T. H., Nakangu, B., OFarrell, P., Subramanian, S. M., van Noordwijk, M., Ahn, S., Amaruzaman, S., Amin, A. M., Arias-Arévalo, P., Arroyo-Robles, G., Cantú-Fernández, M., Castro, A. J., Contreras, V., De Vos, A., Dendoncker, N., Engel, S., Eser, U., Faith, D. P., Filyushkina, A., Ghazi, H., Gómez-Baggethun, E., Gould, R. K., Guibrunet, L., Gundimeda, H., Hahn, T., Harmáčková, Z. V., Hernández-Blanco, M., Horcea-Milcu, A.-I., Huambachano, M., Wicher, N. L. H., Aydn, C. ., Islar, M., Koessler, A.-K., Kenter, J. O., Kosmus, M., Lee, H., Leimona, B., Lele, S., Lenzi, D., Lliso, B., Mannetti, L. M., Merçon, J., Monroy-Sais, A. S., Mukherjee, N., Muraca, B., Muradian, R., Murali, R., Nelson, S. H., Nemogá-Soto, G. R., Ngouhouo-Poufoun, J., Niamir, A., Nuesiri, E., Nyumba, T. O., Özkaynak, B., Palomo, I., Pandit, R., Pawłowska-Mainville, A., Porter-Bolland, L., Quaas, M., Rode, J., Rozzi, R., Sachdeva, S., Samakov, A., Schaafsma, M., Sitas, N., Ungar, P., Yiu, E., Yoshida, Y., and Zent, E. Diverse values of nature for sustainability. Nature, 620(7975):813–823, 8 2023. ISSN 0028-0836. doi: 10.1038/s41586-023-06406-9. URL https://www.nature.com/articles/s41586-023-06406-9.

Purzycki, B. G., Apicella, C., Atkinson, Q. D., Cohen, E., McNamara, R. A., Willard, A. K., Xygalatas, D., Norenzayan, A., and Henrich, J. Moralistic gods, supernatural punishment and the expansion of human sociality. Nature, 530(7590):327–330, 2 2016. ISSN 14764687. doi: 10.1038/nature16980.

Purzycki, B. G., Bendixen, T., Lightner, A. D., and Sosis, R. Gods, games, and the socioecological landscape. Current Research in Ecological and Social Psychology, 3, 1 2022. ISSN 26666227. doi: 10.1016/j.cresp.2022.100057.

Raihani, N. J. and Bshary, R. Punishment: one tool, many uses. Evolutionary Human Sciences, 1:e12, 11 2019. ISSN 2513-843X. doi: 10.1017/ehs.2019.12. URL https://www.cambridge.org/core/product/identifier/S2513843X19000124/type/journal_article.

Rakodi, C. A framework for analysing the links between religion and development. Development in Practice, 22(5-6):634–650, 8 2012. ISSN 0961-4524. doi: 10.1080/09614524.2012.685873. URL https://www.tandfonline.com/doi/full/10.1080/09614524.2012.685873.

Rand, D. G. and Nowak, M. A. Human cooperation. Trends in cognitive sciences, 17(8):413–25, 8 2013. ISSN 1879-307X. doi: 10.1016/j.tics.2013.06.003. URL http://www.ncbi.nlm.nih.gov/pubmed/23856025.

Rand, D. G., Dreber, A., Ellingsen, T., Fudenberg, D., and Nowak, M. A. Positive Interactions Promote Public Cooperation. Science, 325(5945):1272–1275, 9 2009. ISSN 0036-8075. doi: 10.1126/science.1177418. URL https://www.science.org/doi/10.1126/science.1177418.

Rolston, H. Science and Religion in the Face of the Environmental Crisis. In The Oxford Handbook of Religion and Ecology, pages 376–397. Oxford University Press, 11 2006. ISBN 9780199892136. doi: 10.1093/oxfordhb/9780195178722.003.0018. URL https://academic.oup.com/edited-volume/34392/chapter/291660755.

Sakurai, T. The potential of folktales for the protection of the natural environment and development. Wildlife Conservation Japan, 4(2):63–92, 1999. doi: https://doi.org/10.4200/jjhg1948.36.473. in Japanese.

Schloss, J. P. and Murray, M. J. Evolutionary accounts of belief in supernatural punishment: a critical review. Religion, Brain & Behavior, 1(1):46–99, 2 2011. ISSN 2153-599X. doi: 10.1080/2153599X.2011.558707. URL http://www.tandfonline.com/doi/abs/10.1080/2153599X.2011.558707.

Sethi, R. and Somanathan, E. The Evolution of Social Norms in Common Property Resource Use. The American Economic Review, 86(4):766–788, 1996.

Sigmund, K. Punish or perish? Retaliation and collaboration among humans. Trends in Ecology and Evolution, 22(11):593–600, 11 2007. ISSN 01695347. doi: 10.1016/j.tree.2007.06.012.

Singh, M., Kaptchuk, T. J., and Henrich, J. Small gods, rituals, and cooperation: The Mentawai water spirit Sikameinan. Evolution and Human Behavior, 42(1):61–72, 1 2021. ISSN 10905138. doi: 10.1016/j.evolhumbehav.2020.07.008.

Sosis, R. Why aren't we all hutterites? Human Nature, 14(2):91–127, 6 2003. ISSN 1045-6767. doi: 10.1007/s12110-003-1000-6.

Tilman, A. R., Plotkin, J. B., and Akçay, E. Evolutionary games with environmental feedbacks. Nature Communications, 11(1):915, 12 2020. ISSN 2041-1723. doi: 10.1038/s41467-020-14531-6. URL http://www.nature.com/articles/s41467-020-14531-6.

Turchin, P., Whitehouse, H., Larson, J., Cioni, E., Reddish, J., Hoyer, D., Savage, P. E., Covey, R. A., Baines, J., Altaweel, M., Anderson, E., Bol, P., Brandl, E., Carballo, D. M., Feinman, G., Korotayev, A., Kradin, N., Levine, J. D., Nugent, S. E., Squitieri, A., Wallace, V., and François, P. Big Gods and big science: further reflections on theory, data, and analysis. Religion, Brain & Behavior, 13(2):218–231, 4 2023a. ISSN 2153-599X. doi: 10.1080/2153599X.2022.2065354. URL https://www.tandfonline.com/doi/full/10.1080/2153599X.2022.2065354.

Turchin, P., Whitehouse, H., Larson, J., Cioni, E., Reddish, J., Hoyer, D., Savage, P. E., Covey, R. A., Baines, J., Altaweel, M., Anderson, E., Bol, P., Brandl, E., Carballo, D. M., Feinman, G., Korotayev, A., Kradin, N., Levine, J. D., Nugent, S. E., Squitieri, A., Wallace, V., and François, P. Explaining the rise of moralizing religions: a test of competing hypotheses using the Seshat Databank. Religion, Brain &

Behavior, 13(2):167–194, 4 2023b. ISSN 2153-599X. doi: 10.1080/2153599X.2022.2065345. URL https://www.tandfonline.com/doi/full/10.1080/2153599X.2022.2065345.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods, 17:261–272, 2020. doi: 10.1038/s41592-019-0686-2.

Watts, J., Greenhill, S. J., Atkinson, Q. D., Currie, T. E., Bulbulia, J., and Gray, R. D. Broad supernatural punishment but not moralizing high gods precede the evolution of political complexity in Austronesia. Proceedings of the Royal Society B: Biological Sciences, 282(1804), 3 2015. ISSN 14712954. doi: 10.1098/rspb.2014.2556.

Weitz, J. S., Eksin, C., Paarporn, K., Brown, S. P., and Ratcliff, W. C. An oscillating tragedy of the commons in replicator dynamics with game-environment feedback. Proceedings of the National Academy of Sciences, 113(47):E7518–E7525, 2016. ISSN 0027-8424. doi: 10.1073/pnas.1604096113. URL http://www.pnas.org/lookup/doi/10.1073/pnas.1604096113.

Whitehouse, H., François, P., Savage, P. E., Hoyer, D., Feeney, K. C., Cioni, E., Purcell, R., Larson, J., Baines, J., Haar, B. t., Covey, A., and Turchin, P. Testing the Big Gods hypothesis with global historical data: a review and ”retake”. Religion, Brain and Behavior, 13(2):124–166, 2023. ISSN 21535981. doi: 10.1080/2153599X.2022.2074085.

Winkelman, M. J. Shamans and Other ”Magico-Religious” Healers: A Cross-Cultural Study of Their Origins, Nature, and Social Transformations. Ethos, 18(3):308–352, 9 1990. ISSN 0091-2131. doi: 10.1525/eth.1990.18.3.02a00040. URL http://doi.wiley.com/10.1525/eth.1990.18.3.02a00040.

Yamagishi, T. Seriousness of Social Dilemmas and the Provision of a Sanctioning System. Social Psychology Quarterly, 51(1):32–42, 1988.

Figure legend

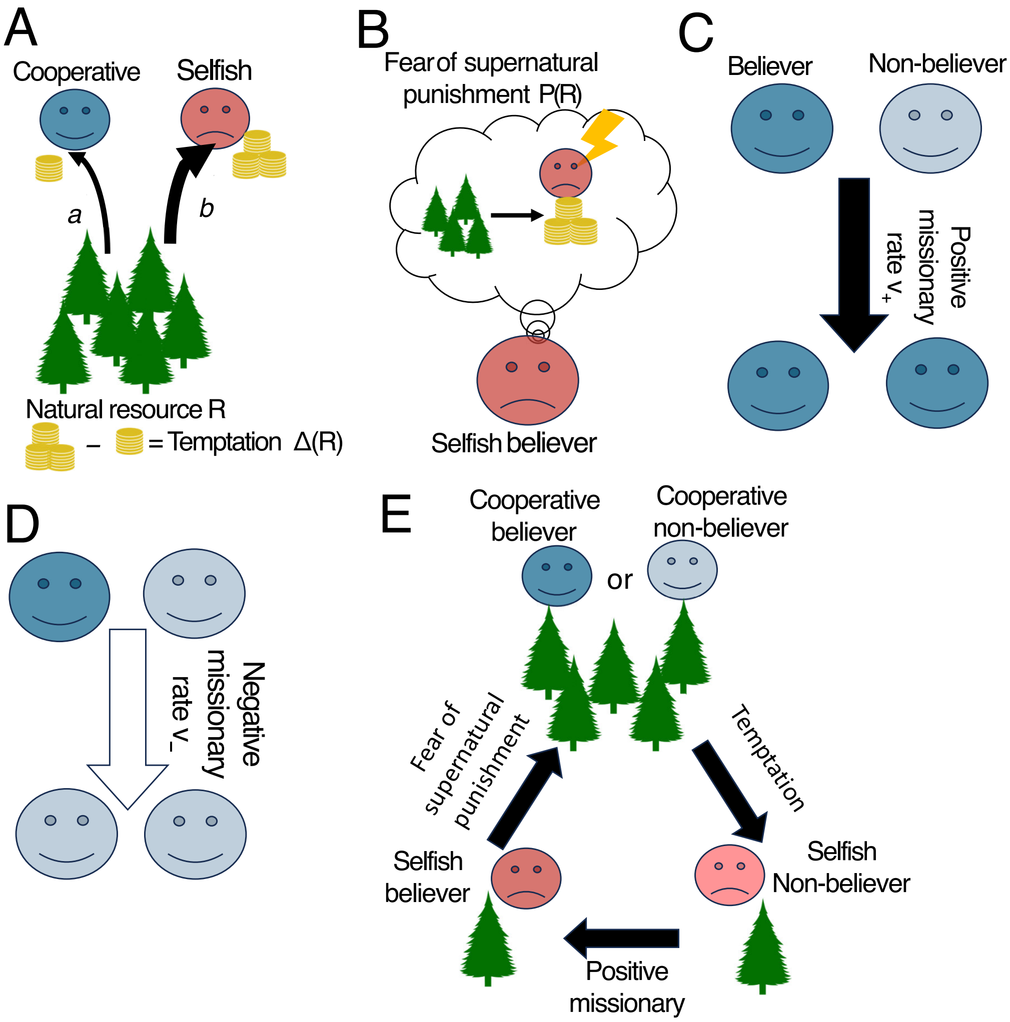

Figure 1: Schematic representation of the model

A: Local people play the public goods game by exploiting natural resources (e.g., woods). Cooperative strategies regulate the exploitation of natural resources (cooperative exploitation rate a), wheares selfish strategies do not (selfish exploitation rate $b > a$). As a result, the selfish strategies yield more benefits than the cooperative ones; the difference in the benefits represents the temptation to selfishness $\Delta(R)$. B: The selfish believers (SB), however, are afraid of supernatural punishment, which damages their health and decreases their payoffs by $P(R)$. C and D: The individuals change whether they believe in supernatural punishment or not, following the positive and negative missionary rates, v_+ and v_-, respectively. The events occurred in accordance with the proportions of believers and non-believers. Due to the environmental feedback, the amount of natural resources depends on the fractions of the four strategies. If either cooperative believers (CB) or non-believers (CN) dominate, the number of natural resources remains high. This leads to a strong temptation to selfishness, and selfish non-believers (SN) can become dominant. Once this occurs, the amount of natural resources declines due to overexploitation. However, the SN may be replaced by SB via positive missionary events. Although SB has an identical exploitation rate to SN, the fear of supernatural punishment can turn SB into CB (or CN), which can then recover the amount of natural resources.

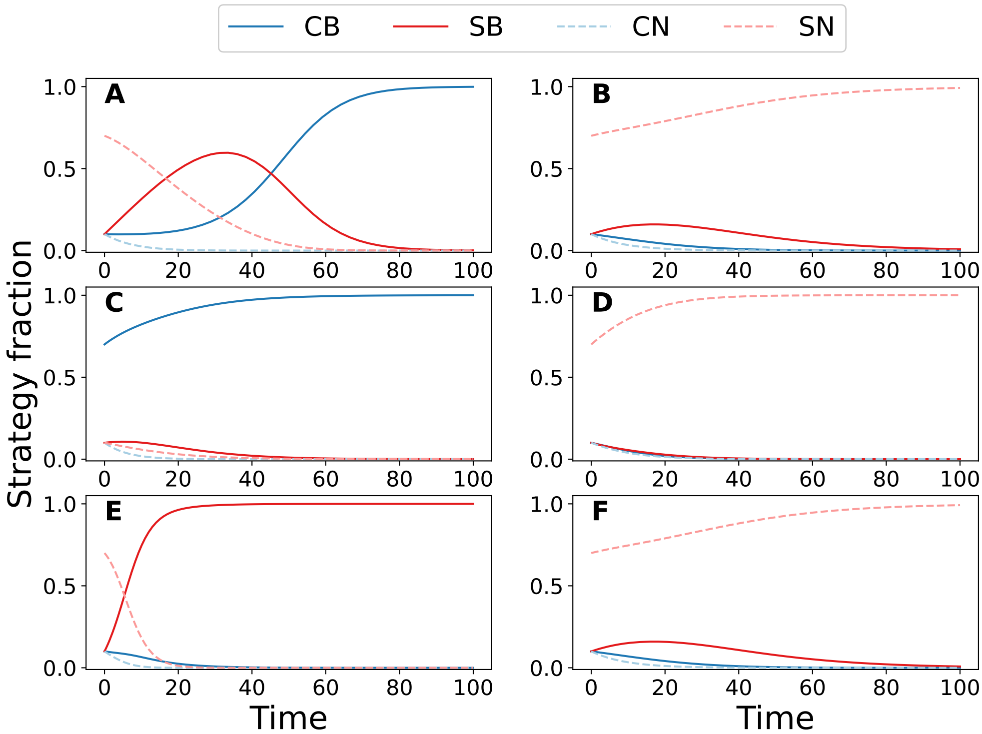

Figure 2: Each of the three strategies was fixed under the constant natural resources and positive missionary events

Examples of human behavior dynamics under constant natural resources and positive missionary events. Negative missionary events were not allowed in these examples ($v_- = 0$). The dynamics depended on the inequality across the temptation to the selfishness $\Delta = (bR)^w - (aR)^w$, of the supernatural punishment $P = (pR)^u$, and the positive missionary rate v_+. Each panel differed in the values of p and v_+, resulting in changes in the relationship among the three parameters. The remaining parameter values were fixed as follows: $a = 0.5$, $b = 0.8$, $w = 1.7$ (thus $\Delta \approx 0.116$), $u = 2$, and $R = 0.5$. A: $p = 1$ and $v_+ = 0.3$ result in $v_+ > P > \Delta$. The CB was fixed and evolutionarily stable in this condition. B: $p = 1$ and $v_+ = 0.1$ resulted in $P > \Delta > v$. SN was then fixed and evolutionarily stable. C and D: $p = 1$ and $v_+ = 0.2$ resulted in $P > v_+ > \Delta$. This condition stabilised both the CB (C) and SN (D). These two panels differed in the initial fractions of the four strategies. E: $p = 0.1$ and $v_+ = 0.3$ resulted in $v_+ > \Delta > P$. In this case, the SB was evolutionarily stable. $\Delta > P > v_+$ also stabilised the SB. F: $p = 0.1$ and $v_+ = 0.1$ resulted in $\Delta > v_+ > P$. In this case, SN was evolutionarily stable.

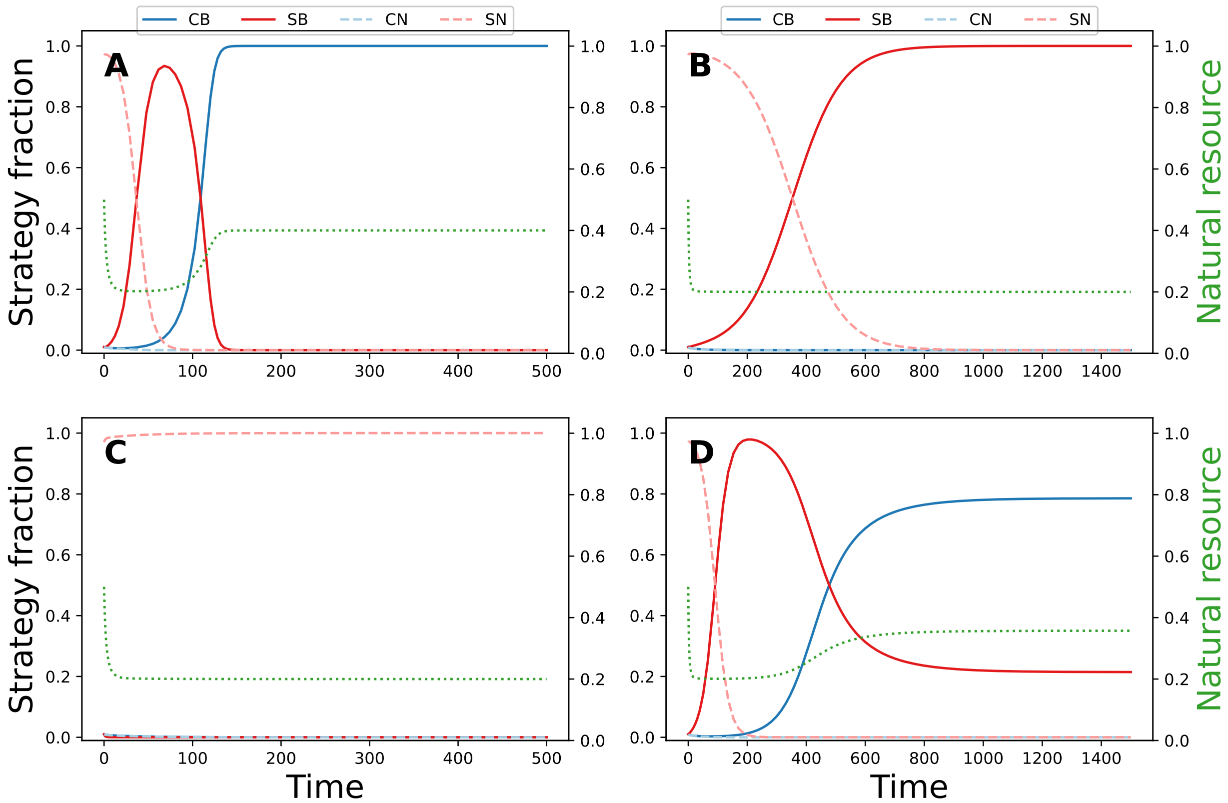

Figure 3: The positive missionary events increased cooperative believers and selfish believers

Four examples of the dynamics of human behaviors and beliefs co-evolving with the amount of natural resources are shown (the dotted green lines). Here, only positive missionary events occur ($v_+ > 0$); no negative missionary events ($v_- = 0$). The parameter values changed the evolutionary fate, although all four dynamics started from the identical initial condition $(R, x_{\text{CB}}, x_{\text{SB}}, x_{\text{CN}}, x_{\text{SN}}) = (0.01, 0.01, 0.01, 0.97, 0.5)$. (A): CB was fixed. (B): SB was fixed. (C): SN was fixed. (D): CB stably coexisted with SB. The four panels differed in the values of (v_+, p, u). (A): $(v_+, p, u) = (0.1, 1, 2)$. (B): $(v_+, p, u) = (0.01, 0.02, 1)$. (C) $(v_+, p, u) = (0.01, 1, 1)$. (D): $(v_+, p, u) = (0.042, 0.1, 1)$. The remaining parameter values were fixed at: $a = 0.6$, $b = 0.8$, $\mu = 1$, $K = 1$, $v_- = 0$, $w = 2$, $u = 2$, and $\epsilon = 0.5$.

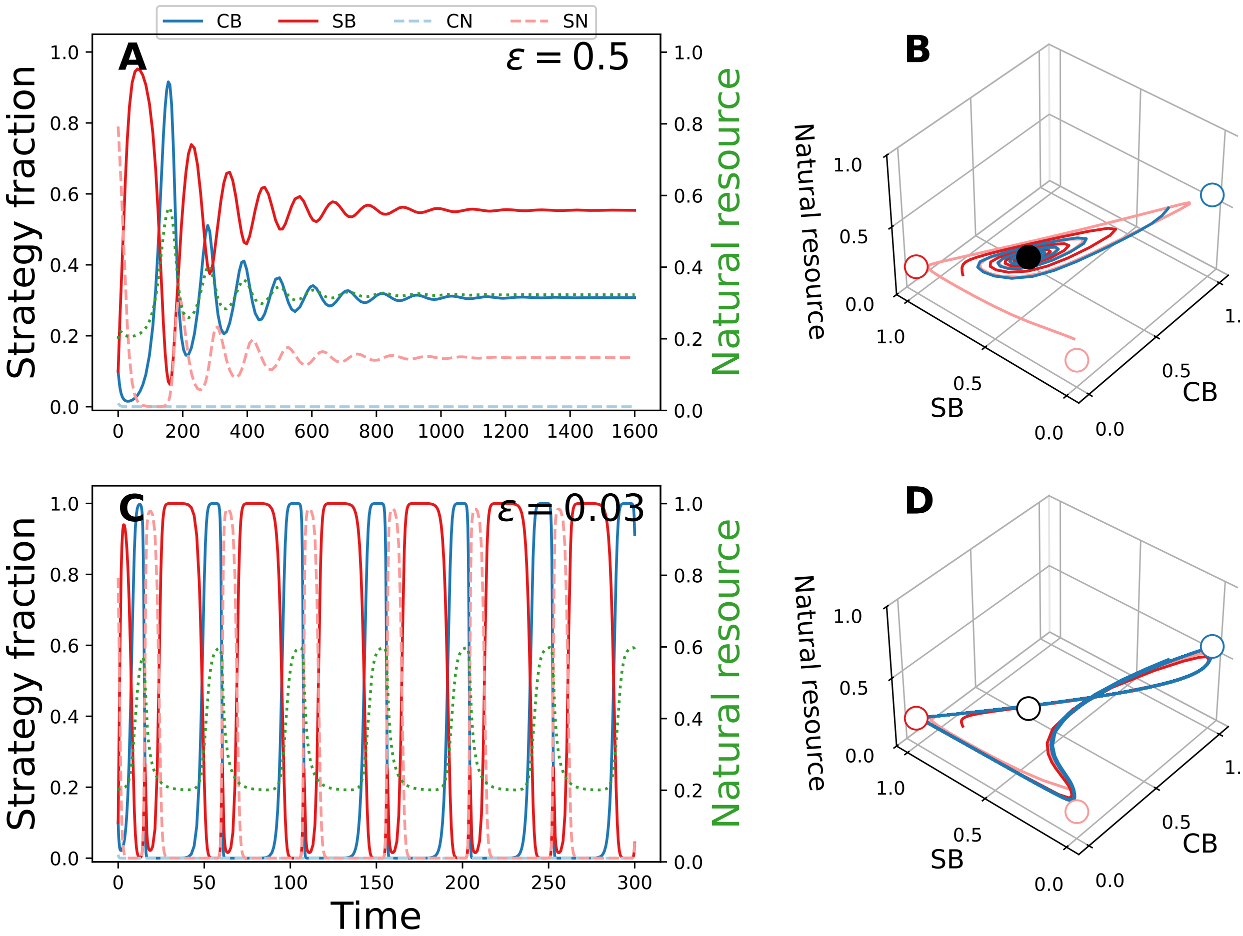

Figure 4: Faster evolution of the human behavior destabilised the coexistence of the three strategies

The CB, SB, and SN can coexist when human behavior evolves slowly; however, their coexistence is unstable under rapid human evolution. (A): When the evolution of human behavior was slow ($\epsilon = 0.5$), the dynamics converged to the equilibrium where CB (the solid blue line), SB (the solid red line), and SN (the dashed pink line) coexisted. The CN (the dashed sky-blue line) remained small, wheares the dynamics of the natural resource (the dotted green line) converged to a moderate value. (B) A phase-space diagram of the system is shown. Since the fraction of CN remained small, we omitted its dynamics and simplified the phase-space diagram into three dimensions. In the current parameter values, either CB (the open blue dot), SB (the open red dot), or SN (the open pink dot) was not evolutionarily stable. The three dynamics, starting from different initial conditions (shown in different colours), converged to the coexistence of the three strategies (the black dot). (C) The evolution of human behavior and beliefs was faster ($\epsilon = 0.03$) in this panel than in panel A, while maintaining the rest of the parameter values. The dynamics exhibited the oscillations. (D) The phase-space diagram and the three examples of the dynamics started from different initial conditions (shown by different colours) under the fast human evolution. Since all four equilibria were unstable, the dynamics oscillated regardless of the initial conditions. Parameter values were as follows: $a = 0.4$, $b = 0.8$, $p = 0.5$, $\mu = 1$, $K = 1$, $v_+ = 0.15$, $v_- = 0$, $w = 1$, $u = 1$, and $\epsilon = 0.5$ (panels A and B) or $\epsilon = 0.03$ (panels C and D). In panels A and C, the initial condition is $(R, x_{\text{CB}}, x_{\text{SB}}, x_{\text{CN}}, x_{\text{SN}}) = (0.2, 0.1, 0., 10.01, 0.79)$. In panels B and D, the initial conditions were as follows: $(R, x_{\text{CB}}, x_{\text{SB}}, x_{\text{CN}}, x_{\text{SN}}) = (0.2, 0.1, 0.1, 0.01, 0.79,)$ for the pink lines, $(0.1, 0.79, 0.01, 0.1, 0.2)$ for the red lines, and $(0.2, 0.79, 0.1, 0.01, 0.1)$ for the blue lines.

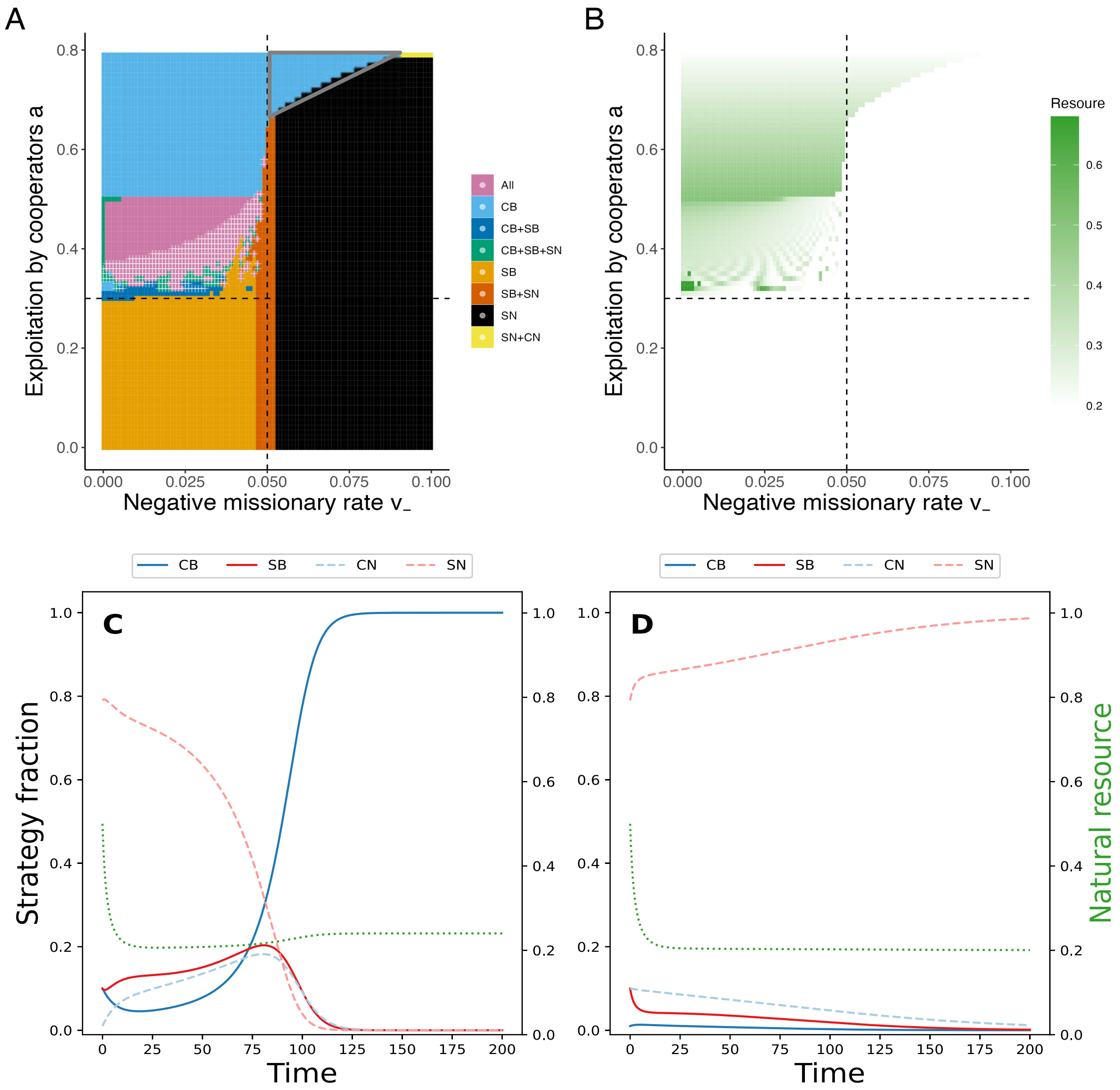

Figure 5: Negative missionary rate and the exploitation by cooperators affected the evolutionary fates and natural resources

A: Increasing the negative missionary rate v_- and the exploitation rate by the cooperators a (i.e., decreasing the temptation to selfishness Δ(R)) affect the evolutionary fates (represented by different colours). When the negative missionary rate is high (the right area of the vertical dashed line $v_- = v_+ - P(R_b^*)$), the SN is fixed unless the exploitation by the cooperators is close to that of the selfish strategies. When the negative missionary rate is low (the left area of the vertical dashed line), the cooperators' exploitation rate a determines the evolutionary fate. The SB is fixed below the horizontal dashed line $P(R_b^*) = \Delta(R_b^*)$. Above the horizontal dashed line, the CB is maintained alone or with other strategies. Cross symbols in the panel indicate the oscillations. B: The average natural resource availability at time $T_f - 100 \leq t \leq T_f$ is shown over the negative missionary rate v_- and exploitation rate by the cooperators a. The extinction of cooperative strategies resulted in minimum natural resource availability $R_b^* = 0.2$ (i.e., the white areas). However, the persistence of the cooperators increased the natural resources. Greener areas retain more natural resources. C and D: When $v_- > v_+ - P(R_b^*)$ but a is close to b (the skyblue area surrounded by the gray line in panel A), both CB and SN are evolutionarily stable (panels C and D, respectively). In such cases, the initial condition determines which strategy is fixed. The values of fixed parameters are as follows: $b = 0.8$, $p = 0.5$, $\mu = 1$, $K = 1$, $v_+ = 0.15$, $w = 1$, $u = 1$, $\epsilon = 0.5$, and $T_f = 1600$. All simulations started from $(R, x_{\text{CB}}, x_{\text{SB}}, x_{\text{CN}}, x_{\text{SN}}) = (0.5, 0.1, 0.1, 0.01, 0.79)$ in panels A, B, and C, while panel D started from $(R, x_{\text{CB}}, x_{\text{SB}}, x_{\text{CN}}, x_{\text{SN}}) = (0.5, 0.01, 0.1, 0.1, 0.79)$. See also Figs S1 – S4 for the cases where either the temptation to selfishness or the fear of the supernatural punishment is a nonlinear function of R.

Tables

Table 1: List of variables and parameters

Symbol	Description
$x_i(t)$	Fraction of strategy i in a local population at time t
$R(t)$	Amount of natural resources at time t
a	Exploitation rates of cooperative strategies
b	Exploitation rates of selfish strategies
$\Delta(R)$	Temptation to selfishness, see Eq (4)
μ	Intrinsic growth rate of the natural resource
K	Carrying capacity of the natural resource
$P(R)$	Fear of supernatural punishment, see Eq (2)
v_+	Positive missionary rate
v_-	Negative missionary rate
ϵ	Time-scale parameter

Data availability

The codes and simulation data used in this study will be publicly available at the repository server upon publication.

Competing interests

The authors declare no competing interests.

Ethical statements

This article does not contain any studies with human participants performed by any of the authors.

Supporting Information

SI 1 Derivation of the evolutionarily stable strategy

An evolutionarily stable strategy (ESS) is a strategy that is not invaded by any other strategies when it is dominant in the population. This section derives the conditions under which each of the four strategies, the cooperative believers (CB), selfish believers (SB), cooperative non-believers (CN), and selfish non-believers (SN), is evolutionarily stable in our general model in the main text. When cooperative or selfish strategies are fixed, the amount of the natural resources is $R_a^* = K(1 - a/\mu)$ or $R_b^* = K(1 - b/\mu)$, respectively. Note that $R_a^* > R_b^*$ because $b > a > 0$.

SI 1.1 CB can be an ESS

SB cannot invade the population of CB if and only if the fear of the supernatural punishment is larger than the temptation to selfishness:

$$P(R_a^*) > \Delta(R_a^*). \tag{S1}$$

CN cannot invade the population of the CB if the positive missionary rate is larger than the negative missionary rate because their payoffs are identical for any R:

$$v_+ > v_-. \tag{S2}$$

SN cannot invade the population of CB if the positive missionary rate is larger than the temptation:

$$v_+ > \Delta(R_a^*). \tag{S3}$$

Combining the above three inequalities results in the necessary and sufficient conditions for CB to be an ESS.

When the amount of natural resources is constant over time, the temptation to selfishness and the fear of supernatural punishment become constant Δ and P, respectively. If the negative missionary events do not occur ($v_- = 0$), CB is evolutionarily stable if and only if

$$\begin{cases} P > \Delta \\ v_+ > \Delta. \end{cases} \tag{S4}$$

Figs. 2A and C are examples that satisfy the above conditions.

SI 1.2 SB can be an ESS

CB cannot invade the population of SB if the temptation to selfishness is larger than the fearness of the supernatural punishment:

$$\Delta(R_b^*) > P(R_b^*). \tag{S5}$$

CN cannot invade the population of SB if the temptation to selfishness plus the positive missionary rate is larger than the fearness of the supernatural punishment:

$$\Delta(R_b^*) + v_+ > P(R_b^*). \tag{S6}$$

SN cannot invade the population of SB if the difference between the positive missionary and negative missionary rates is larger than the fear of supernatural punishment:

$$v_+ - v_- > P(R_b^*). \tag{S7}$$

Because the CN cannot invade SB when CB cannot invade SB, SB is evolutionarily stable if and only if

$$\begin{cases} \Delta(R_b^*) > P(R_b^*) \\ v_+ - v_- > P(R_b^*). \end{cases} \tag{S8}$$

Assuming a constant amount of natural resources and no negative missionary events simplifies the conditions as follows:

$$\begin{cases} \Delta > P \\ v_+ > P. \end{cases} \tag{S9}$$

See Fig. 2E as an example.

SI 1.3 CN cannot be an ESS

CN cannot be evolutionarily stable because SN always invades the population of CN. However, CB cannot invade the population of CN if and only if

$$v_+ < v_-. \tag{S10}$$

In addition, SB cannot invade the population of CN if and only if

$$\Delta(R_a^*) < P(R_a^*) + v_+ \tag{S11}$$

SI 1.4 SN can be an ESS

CB and CN cannot invade the SN population because SN always has a larger payoff than them. The SN is evolutionarily stable if and only if SB cannot invade the population of SN:

$$v_+ - v_- < P(R_b^*) \tag{S12}$$

In other words, SN is evolutionarily stable if the fear of supernatural punishment is greater than the difference between positive and negative missionary rates.

These analyses also clarify that $v_+ - v_- > 0$ is a necessary condition for CB to be evolutionarily stable. If $v_+ - v_- < 0$ SN is a unique ESS (e.g., Section 3.1 in the main text).

SI 2 Local stability analysis of the coexistence of multiple strategies without negative missionary

This section shows the local stability analysis when two or all of CB, SB, and SN coexist without negative missionary events. Here, we assume $x_{\mathrm{CN}} = 0$ because CN obtains a lower payoff than SN, and CN changes into CB due to the positive missionary events. Then, because $x_{\mathrm{SN}} = 1 - x_{\mathrm{CB}} - x_{\mathrm{SB}}$, the system becomes simplified as follows.

$$\dot{R} = \mu R \left(1 - \frac{R}{K}\right) - R\left\{a x_{\mathrm{CB}} + b(1 - x_{\mathrm{CB}})\right\} \tag{S13a}$$

$$\epsilon \dot{x}_{\mathrm{CB}} = x_{\mathrm{CB}}\left\{f_1(R) - \bar{f}(R)\right\} \tag{S13b}$$

$$\epsilon \dot{x}_{\mathrm{SB}} = x_{\mathrm{SB}}\left\{f_{\mathrm{SB}}(R) - \bar{f}(R)\right\} + v_+(x_{\mathrm{CB}} + x_{\mathrm{SB}})(1 - x_{\mathrm{CB}} - x_{\mathrm{SB}}) \tag{S13c}$$

The Jacobian matrix J is then written as follows:

$$J = (J_1, J_2, J_3) \tag{S14}$$

where

$$J_1 = \begin{pmatrix} \mu(1-\frac{2R}{K}) - ax_{\mathrm{CB}} - b(1-x_{\mathrm{CB}}) \\ \frac{x_{\mathrm{CB}}}{\epsilon}\left(\frac{df_{\mathrm{CB}}}{dR} - \frac{\bar{f}}{dR}\right) \\ \frac{x_{\mathrm{SB}}}{\epsilon}\left(\frac{df_{\mathrm{SB}}}{dR} - \frac{d\bar{f}}{dR}\right) \end{pmatrix}, \tag{S15}$$

$$J_2 = \begin{pmatrix} R(b-a) \\ \left\{(1-2x_{\mathrm{CB}})f_{\mathrm{CB}} - x_{\mathrm{SB}}f_{\mathrm{SB}} - (1-2x_{\mathrm{CB}}-x_{\mathrm{SB}})f_{\mathrm{SN}}\right\}/\epsilon \\ \left[-x_{\mathrm{SB}}(f_{\mathrm{CB}}-f_{\mathrm{SN}}) + v_+\{1-2(x_{\mathrm{CB}}+x_{\mathrm{SB}})\}\right]/\epsilon \end{pmatrix}, \tag{S16}$$

$$J_3 = \begin{pmatrix} 0 \\ -x_{\mathrm{CB}}(f_{\mathrm{SB}}-f_{\mathrm{SN}})/\epsilon \\ \left[-x_{\mathrm{CB}}f_{\mathrm{CB}} + (1-2x_{\mathrm{SB}})f_{\mathrm{SB}} - (1-x_{\mathrm{CB}}-2x_{\mathrm{SB}})f_{\mathrm{SN}} + v_+\{1-2(x_{\mathrm{CB}}+x_{\mathrm{SB}})\}\right]/\epsilon \end{pmatrix}. \tag{S17}$$

According to the Routh-Hurwitz criteria, the coexistence of multiple strategies is locally stable if and only if

$$\begin{cases} \mathrm{trJ} < 0 \\ \mathrm{detJ} > 0 \\ \sum_{i=1}^{3} M_{ii} > 0. \end{cases} \tag{S18}$$

where M_{ii} represents the (i,i) minor of the Jacobian matrix.

SI 2.1 Coexistence of CB with SB

When CB coexists with SB, no positive missionary events occur. The amount of the natural resources at the equilibrium is, therefore, given by a root of

$$\Delta(R^*) = P(R^*). \tag{S19}$$

In other words, the temptation to selfishness and the fear of supernatural punishment are balanced at equilibrium. Once the root R^* is obtained, the fractions of CB and SB are written as follows, respectively:

$$x^*_{\mathrm{CB}} = \frac{b - \mu(1-R^*/K)}{b-a} \tag{S20}$$

$$x^*_{\mathrm{SB}} = 1 - x^*_{\mathrm{CB}} \tag{S21}$$

Before analysing the Routh-Hurwitz criteria, it should be noted that

$$f_1(R^*) = f_2(R^*) = \bar{f}(R^*) \equiv f^* \tag{S22}$$

for the convenience of further calculation. The Jacobian matrix at this equilibrium is written as follows:

$$J = \begin{pmatrix} -\mu R^* & R^(b-a) & 0 \\ J_{21} & J_{22} & J_{22} \\ -J_{21} & J_{32} & J_{32} \end{pmatrix} \tag{S23}$$

where

$$J_{21} = \frac{x^*_{\text{CB}}}{\epsilon}(1 - x^*_{\text{CB}})\left(\frac{df_{\text{CB}}}{dR} - \frac{df_{\text{SB}}}{dR}\right) \tag{S24a}$$

$$J_{22} = \frac{x^*_{\text{CB}}(f_{\text{SN}} - f^*)}{\epsilon} \tag{S24b}$$

$$J_{32} = \frac{(1 - x_{\text{CB}})(f_{\text{SN}} - f^*) - v_+}{\epsilon} \tag{S24c}$$

The Routh-Hurwitz criteria (S18) reduce to

$$\begin{cases} \text{trJ} < 0 \\ \text{detJ} < 0 \\ \sum_i^3 M_{ii} > 0 \end{cases} \Leftrightarrow \begin{cases} -\mu R^* + J_{22} + J_{32} < 0 \\ J_{21}(J_{22} + J_{32}) > 0 \\ -\mu R(J_{22} + J_{32}) - R(b-a)J_{21} > 0 \end{cases} \tag{S25}$$

$$\Leftrightarrow \begin{cases} J_{21} < 0 \\ J_{22} + J_{33} < 0 \end{cases} \tag{S26}$$

$$\Leftrightarrow \begin{cases} \frac{df_{\text{CB}}}{dR}|_{R=R^*} < \frac{df_{\text{SB}}}{dR}|_{R=R^*} \\ P(R^*) < v_+ \end{cases} \tag{S27}$$

The first inequality argues that the fitness gradient of SB is larger than that of CB, and the second one argues that SN cannot take the place of SB.

SI 2.2 CB cannot coexist with SN

Next, we consider the coexistence of CB with SN $(R, x1_, x_2) = (R^*, x_1^*, 0)$ where

$$0 < x^*_{\text{CB}} < 1 \tag{S28a}$$

$$0 < R^* \tag{S28b}$$

At this equilibrium, the following equation should be satisfied:

$$\epsilon \dot{x}_{\text{CB}} = 0$$

$$\Leftrightarrow (1 - x^*_{\text{CB}})\{f_{\text{CB}}(R) - f_{\text{SN}}(R^*)\} = 0$$

$$\Leftrightarrow (1 - x^*_{\text{CB}})\,\Delta(R^*) = 0 \tag{S29}$$

$$\Leftrightarrow x^*_{\text{CB}} = 1 \quad \text{or} \quad \text{R}^* = 0. \tag{S30}$$

This contradict with inequalitiws (S28a) and (S28b). Therefore, CB cannot coexist with SN.

SI 2.3 SB cannot coexist with SN

This subsection shows that SB and SN cannot coexist ($R_2, 0, x_2^*$ where $0 < x_2^* < 1$). In this case,

$$\dot{x}_{\mathrm{SB}} = 0 \Leftrightarrow \{f_{\mathrm{SB}}(R^*) - f_{\mathrm{SN}}(R^*) + v_+\} = 0 \tag{S31}$$

Then, the Routh-Hurwitz criteria cannot be satisfied because

$$\mathrm{detJ} = 0. \tag{S32}$$

This is because $J_3 = \vec{0}$. SB cannot, therefore, stably coexist with SN.

SI 2.4 Coexistence of CB, SB, and SN

The three strategies, CB, SB, and SN, can stably coexist. At a such equilibrium, $(R, x_1, x_2) = (R^*, x_1^*, x_2^*)$ satisfied the following inequalities.

$$0 < R^* < K \tag{S33}$$

$$0 < x_{\mathrm{CB}}^* < 1 \tag{S34}$$

$$0 < x_{\mathrm{SB}}^* < 1 \tag{S35}$$

$$0 < x_{\mathrm{CB}}^* + x_{\mathrm{SB}}^* < 1 \tag{S36}$$

The equilibrium should satisfy the following equations:

$$x_1^* = \frac{\Delta^*(1 - P^*/v_+)}{\Delta^* - P^*} \tag{S37a}$$

$$x_2^* = \frac{\Delta^*(\Delta^*/v_* - 1)}{\Delta^* - P^*} \tag{S37b}$$

$$x_1^* + x_2^* = \frac{\Delta^*}{v_*} \tag{S37c}$$

$$R^* = K\left(1 - \frac{ax_1^* + b(1 - x_1^*)}{\mu}\right) \Leftrightarrow x_1^* = \frac{b - \mu(1 - R^*/K)}{b - a} \tag{S37d}$$

where $\Delta^* = \Delta(R^*)$ and $P^* = P(R^*)$, respectively. Eq (S37a) = Eq (S37d) derives the equilibrium, but it is challenging to solve this equation due to the nonlinearity of $\Delta(R)$ and $P(R)$.

Below, we continue the local stability analysis. Here we aim to show that the time scale parameter ϵ affects the stability without changing the equilibrium. For the rest of the types of equilibria, we have already shown that ϵ does not affect the stability. Notice that

$$\begin{cases} 1 > x^*_{\rm CB} > 0 \\ 1 > x^*_{\rm SB} > 0 \\ 1 > x^*_{\rm CB} + x^*_{\rm SB} > 0 \\ K > R^* > 0 \end{cases} \Leftrightarrow P^* > v_+ > \Delta^* \tag{S38a}$$

because $\Delta^* x_{\rm CB} + P^* x_{\rm SB} = \Delta^*$. The Jacobian matrix at this equilibrium is written as follows:

$$J = \begin{pmatrix} -\frac{\mu R^*}{K} & R^*(b-a) & 0 \\ J^*_{21} & \Delta^* x_{\rm CB}/\epsilon & P^* x_{\rm CB}/\epsilon \\ J^*_{31} & (\Delta^* x_{\rm SB} + v_+ - 2\Delta^*)/\epsilon & (P^* x_{\rm SB} - (P^* + \Delta^* - v_+))/\epsilon \end{pmatrix}. \tag{S39}$$

where

$$J^*_{21} = \frac{x^*_1}{\epsilon}\left\{-(1-x^*_1)\left.\frac{d\Delta}{dR}\right|_{R=R^*} + x^*_2 \left.\frac{dP}{dR}\right|_{R=R^*}\right\} \tag{S40a}$$

$$J^*_{31} = \frac{x^*_2}{\epsilon}\left\{-(1-x^*_2)\left.\frac{dP}{dR}\right|_{R-R^*} + x^*_1 \left.\frac{d\Delta}{dR}\right|_{R=R^*}\right\} \tag{S40b}$$

Note that $J_{22}, J_{23}, J_{33} > 0$ because $x^*_1, x^*_2 > 0$. Now, we consider the Routh-Hurwitz criteria. The trace of the Jacobian matrix at the equilibrium is always negative because $v_+ < P^*$:

$$\mathrm{trJ} < 0 \Leftrightarrow \frac{\mu R^*}{K} > \underbrace{\frac{v_+ - P^*}{\epsilon}}_{<0}. \tag{S41}$$

To investigate whether the equilibrium is stable or not, we need to evaluate the other two Routh-Hurwitz criteria:

$$\mathrm{detJ} < 0 \Leftrightarrow J_{11}M_{11} + J_{12}(J_{31}J_{23} - J_{21}J_{33}) < 0 \tag{S42}$$

$$\sum_i M_{ii} > 0 \Leftrightarrow \underbrace{M_{11}}_{\mathcal{O}(\epsilon^{-2})} + \underbrace{M_{22}}_{\mathcal{O}(\epsilon^{-1})} + \underbrace{M_{33}}_{\mathcal{O}(\epsilon^{-1})} > 0$$

$$\Leftrightarrow M_{11} + \frac{\mu R^*(P^* - v_+)}{K\epsilon} - J_{12}J_{21} > 0 \tag{S43}$$

Although it is difficult to continue the further analysis, the above equations clarify that the time scale parameter ϵ affects the stability of the equilibrium. Below, we illustrate an example when $\Delta(R)$ and $P(R)$ are linear functions of R.

SI 2.4.1 Simple example: linear temptation and fearness

For illustration, we consider the case when both the temptation to the selfishness $\Delta(R)$ and the fearness of the supernatural punishment $P(R)$ are linear functions (i.e., $w = u = 1$). In this case, the equilibrium where

CB, SB, and SN coexist is unique because Eq (S37a) = Eq (S37d) results in a linear equation of R^*. Once the equilibrium is derived, its stability is analyzed as follows: Because $\Delta x_1^* + P x_2^* = \Delta$ at this equilibrium,

$$(b-a)x_{\mathrm{CB}}^* + p x_{\mathrm{SB}}^* = b - a \Leftrightarrow x_{\mathrm{Cb}}^* \frac{d\Delta}{dR} + x_{\mathrm{SB}}^* \frac{dP}{dR} = \frac{d\Delta}{dR} \tag{S44}$$

$$\Leftrightarrow J_{21} = 0 \tag{S45}$$

Similarly,

$$J_{31} = x_{\mathrm{SB}}(b-a-p)/\epsilon < 0 \tag{S46}$$

Then,

$$M_{11} = -\frac{x_{\mathrm{CB}}^*(v_+ - \Delta^*)(P^* - \Delta^*)}{\epsilon^2} < 0. \tag{S47}$$

Now, the Routh-Hurwitz conditions are written as follows:

$$\mathrm{trJ} < 0 \Leftrightarrow \frac{\mu R^*}{K} > \frac{v_+ - P^*}{\epsilon} \tag{S48}$$

$$\mathrm{detJ} < 0 \Leftrightarrow \frac{(b-a)x_{\mathrm{SB}}^*}{R^*} > \frac{\mu}{K}(v_+ - \Delta^*) \tag{S49}$$

$$\sum_{i=1}^{3} M_{ii} > 0 \Leftrightarrow \underbrace{-\frac{x_{\mathrm{CB}}(v_+ - \Delta^*)(P^* - \Delta^*)}{\epsilon^2}}_{<0} + \underbrace{\frac{\mu R^*(P^* - v_*)}{K\epsilon}}_{>0} > 0 \tag{S50}$$

The first inequality always holds because $P^* > v_+ > \Delta^*$ should be satisfied if this equilibrium exists. One can easily evaluate the second inequality once the equilibrium is obtained. The third inequality argues that the stability changes over ϵ even when the other parameter values are fixed. When $\epsilon \gg 1$ (i.e., the evolution of human behavior is small compared to the dynamics of natural resources), $\sum_{i=1}^{3} M_{ii} > 0$ because the first term in the equality can be omitted. In this case, the coexistence of the three strategies is stable if detJ is positive. When $\epsilon \ll 1$ (i.e., the rapid evolution of human behavior), the equilibrium is unstable because $\sum_{i=1}^{3} M_{ii} \approx M_{11} < 0$.

Note that $w = u = 1$ indicates that CB cannot stably coexist with SB. When the two strategies coexist, the following equation should be satisfied.

$$\Delta(R) = P(R) \Leftrightarrow b - a = p. \tag{S51}$$

Recall that this equilibrium is stable if and only if

$$\frac{df_{\mathrm{CB}}}{dR}|_{R=R^*} < \frac{df_{\mathrm{SB}}}{dR}|_{R=R^*} \Leftrightarrow p < b - a. \tag{S52}$$

Therefore, the case of $w = u = 1$ has at most four equilibria: fixation of CB, fixation of SB, fixation of SN, or coexistence of the three strategies.

SI 3 CN cannot stably coexist with the other strategies

This section proves that CN cannot stably coexist with the other strategies without the negative missionary events ($v_- = 0$). We begin the analysis by examining whether CN coexists with one of the three states. First, CN cannot stably coexist with CB because their payoffs are identical for any R (i.e., $f_{\mathrm{CB}}(R) = (f_{\mathrm{CN}}(R))$ but the positive missionary events alter CN to CB. Second, CN cannot coexist with SB stably. Suppose SB and CN coexist. CB can, however, invade this coexistence due to the positive missionary events. CN cannot coexist with SN because the temptation to the selfishness $\Delta(R)$ alters CN to SN.

Next, we consider the coexistence of three strategies. Suppose CN coexist with CB and SB. At an equilibrium point, the fractions of CB and CN should satisfy the following equations:

$$\begin{cases} x^*_{\mathrm{CB}}\left(f_{\mathrm{CB}}(R^*) - \bar{f}(R^*)\right) + v_+\left(x^*_{\mathrm{CB}} + x^*_{\mathrm{SB}}\right)x^*_{\mathrm{CN}} = 0 \\ x^*_{\mathrm{CN}}\underbrace{\left(f_{\mathrm{CN}}(R^*) - \bar{f}(R^*)\right)}_{=f_{\mathrm{CB}}(R^*)-\bar{f}(R^*)} - v_+\left(x^*_{\mathrm{CB}} + x^*_{\mathrm{SB}}\right)x^*_{\mathrm{CN}} = 0 \end{cases} \tag{S53}$$

$$\Leftrightarrow x^*_{\mathrm{CB}} = -\frac{v_+\left(x^*_{\mathrm{CB}} + x^*_{\mathrm{SB}}\right)x^*_{\mathrm{CN}}}{\left(f_{\mathrm{CB}}(R^*) - \bar{f}(R^*)\right)} = -x_{\mathrm{CN}} \tag{S54}$$

where the asterisks represent the values at the equilibrium point. Because x^*_{CB} and x^*_{CN} should be positive, the coexistence of CB, CN, and SB is not feasible. The coexistence of all four strategies is not feasible for the same reason.

When CN coexists with SB and SN, this coexistence is not stable because CB can invade:

$$\dot{x}_{\mathrm{CN}} = x^*_{\mathrm{CN}}\underbrace{\left(f_{\mathrm{CN}}(R^*) - \bar{f}(R^*)\right)}_{=f_{\mathrm{CB}}(R^*)-\bar{f}(R^*)} - v_+ x^*_{\mathrm{SB}} x^*_{\mathrm{CN}} = 0 \tag{S55}$$

$$\Rightarrow \dot{x}_{\mathrm{CB}} = v_+ x^*_{\mathrm{SB}} x^*_{\mathrm{CN}} > 0. \tag{S56}$$

Therefore, CN cannot coexist with any of the other three strategies without the negative missionary.

SI 4 SB cannot stably coexist with SN in the full model

SB can coexist with the SN in the presence of the negative missionary (c.f., SI 2.3). At the equilibrium,

$$\begin{aligned} &\dot{x}_{\mathrm{SB}} = 0 \\ \Leftrightarrow &P(R^*_b) = v_- - v_+. \end{aligned} \tag{S57}$$

This equilibrium is feasible only if the negative missionary rate v_- is equal to or higher than the positive missionary rate v_+ because the fearness of the supernatural punishment is non-negative. However, this equilibrium

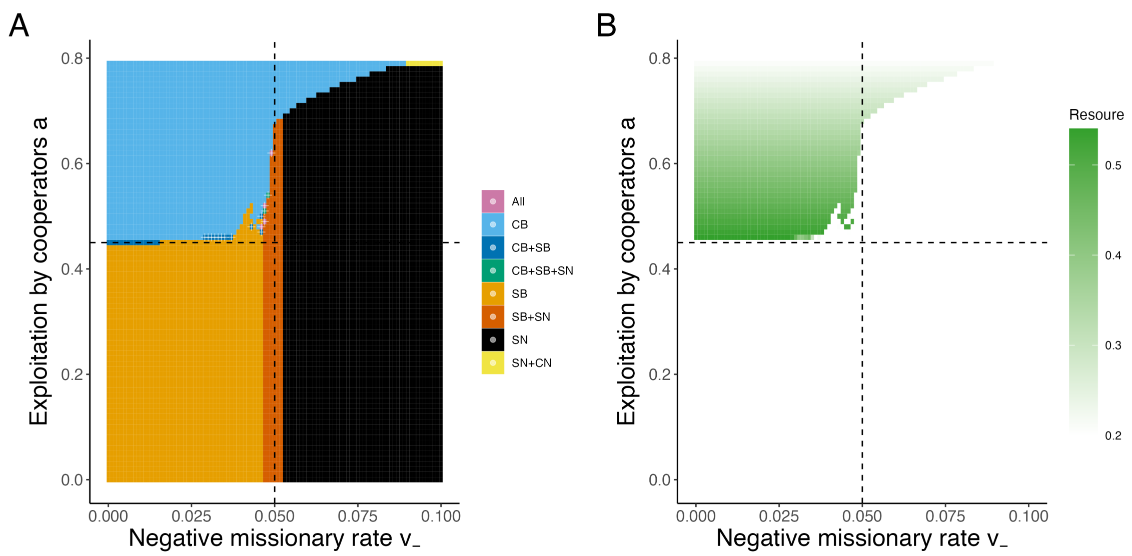

Figure S1: Parameter space when the temptation is a concave function

Similar to Fig. 5 in the main text, but the temptation to selfishness was a concave function in this figure ($w = 0.5$). The remaining parameter values were identical to Fig. 5.

is not stable because

$$\frac{\partial}{\partial x_{\text{SB}}} \dot{x}_{\text{SB}} = \{P(R_b^*) - v_- + v_+\} (1 - 2x_{\text{SB}}) = 0. \tag{S58}$$

In other words, when the fraction of SB changes from the equilibrium due to a small perturbation, the fraction cannot return to its original. The coexistence of SB with SN is, therefore, unstable.

SI 5 Parameter space when the temptation and the fearness are nonlinear functions of the natural resource

In the main text, we investigated how the negative missionary rate v_- and the exploitation rate by the cooperators a affect the evolutionary fate of the human behaviors and the average natural resource availability at the end when the temptation to selfishness $\Delta(R)$ and the fearness of the supernatural punishment $P(R)$ are linear functions of R (i.e., $w = u - 1$). In this section, we analyzed the cases when either of the two functions was nonlinear. Remarkably, we investigated the instances where the temptation was a concave ($w = 0.5$) or convex ($w = 2$) function while the fearness remained the linear function ($u = 1$). We also investigated cases where the fearness was a concave ($u = 0.5$) or convex ($u = 2$) function while the temptation was linear ($w = 1$).

The nonlinearity of the temptation w changed thetreshold of $P(R_b^*) = \Delta(R_b^*)$. When the temptation was a concave function of R (Fig. S1), the areas where the cooperative believers coexisted with other strategies disappeared, and SB was fixed instead. The concave function shrunk the parameter spaces in which the R was larger than R_b^* at the end of the simulations. When the temptation was a convex function of R (Fig. S2),

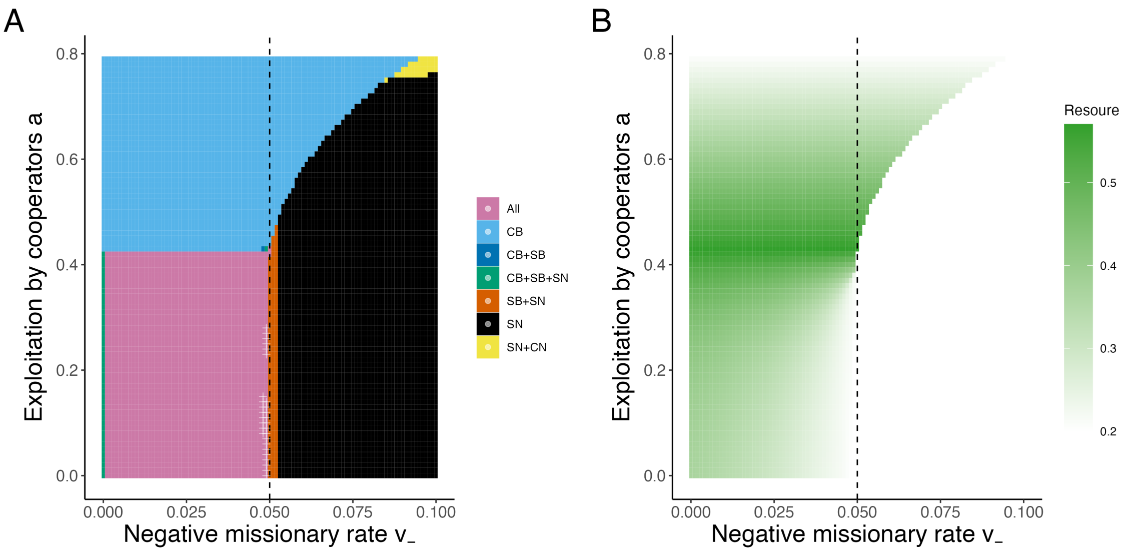

Figure S2: Parameter space when the temptation is a convex function

Similar to Fig. 5 in the main text, but the temptation to selfishness was a convex function in this figure ($w = 2$). The remaining values were identical to Fig. 5. The selfish believer cannot be evolutionarily stable in this case because no real a satisfies $P(R_b^*) = \Delta(R_b^*)$. The horizontal dashed line vanished for this reason.

SB was fixed because no real a satisfied $P(R_b^*) = \Delta(R_b^*)$. CB persisted in broader parameter ranges, typically coexisting with the three other strategies.

The nonlinerity of the fearness u, on the other hand, changed the two threshodls $P(R_b^*) = \Delta(R_b^*)$ and $v_- = v_+ - P(R_b^*)$. The nonlinearity resulted in a decrease in the parameter space where CB persisted. When the fear of supernatural punishment was a concave function (Fig. S3), the two thresholds became negative, leading to the fixation of SN in most cases. CB fixated only when the exploitation rates by the cooperatives and the selfish strategies were close. When the fearness of the supernatural punishment was a convex function (Fig. S4), the two thresholds became larger than in the linear case (Fig. 5). As a result, SB fixated in the population unless the exploitation rate by the cooperators was close to that of the selfish strategies. These two cases show limited areas where the resource availability remained higher than the minimum value R_b^*.

In short, while the nonlinearity in the fear of the supernatural punishment decreased the parameter space where CB can persist, a convex function of the temptation increased such parameter space.

SI 6 Cases of negative correlation between the fear of supernatural punishment and the amount of natural resources

In the main text, we assumed that the fear of supernatural punishment increased with the amount of natural resources. In this section, we analyzed our model under the opposite assumption: the fear of supernatural punishment decreased with the amount of natural resources because individuals may regard depleting them as

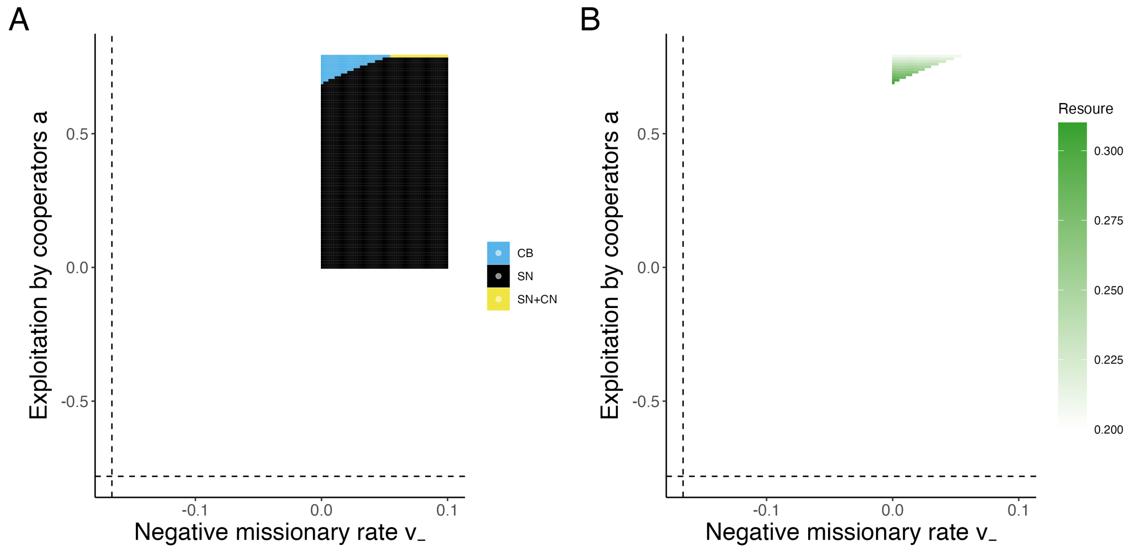

Figure S3: Parameter space when the fearness is a concave function

Similar to Fig. 5 in the main text, but the fearness of the supernatural punishment was a concave function in this figure ($u = 0.5$). The remaining parameter values were identical to Fig. 5. As in the main text, we analyzed the parameter ranges $0 \leq v_- \leq 0.1$ and $a \leq a \leq 0.79$.

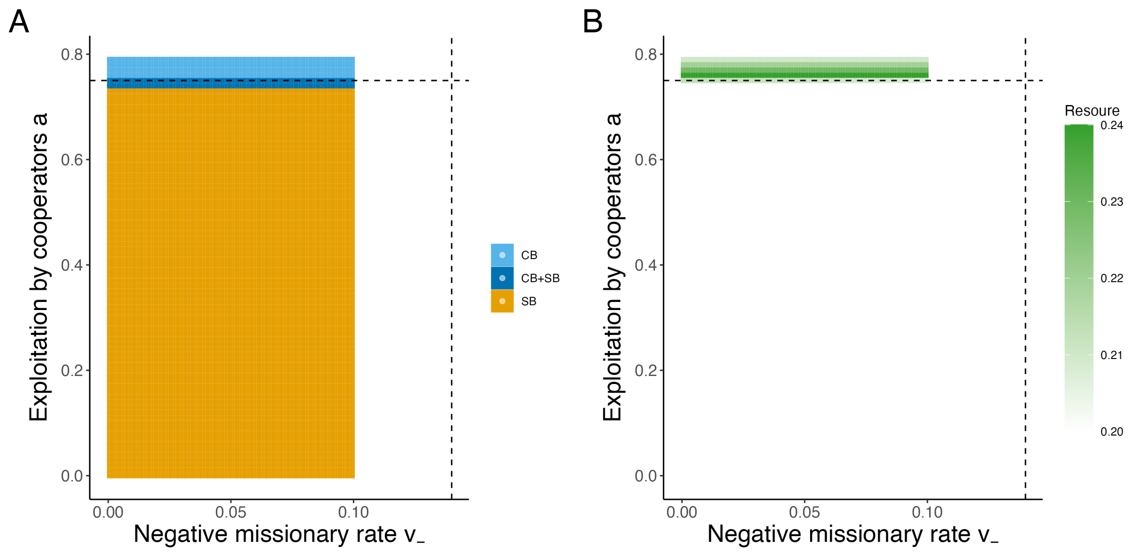

Figure S4: Parameter space when the fearness was a convex function

Similar to Fig. 5 in the main text, but the fearness of the supernatural punishment was a convex function in this figure ($u = 2$). The rest of the parameter values were identical to Fig. 5. As in the main text, we analyzed the parameter ranges $0 \leq v_- \leq 0.1$ and $a \leq a \leq 0.79$.

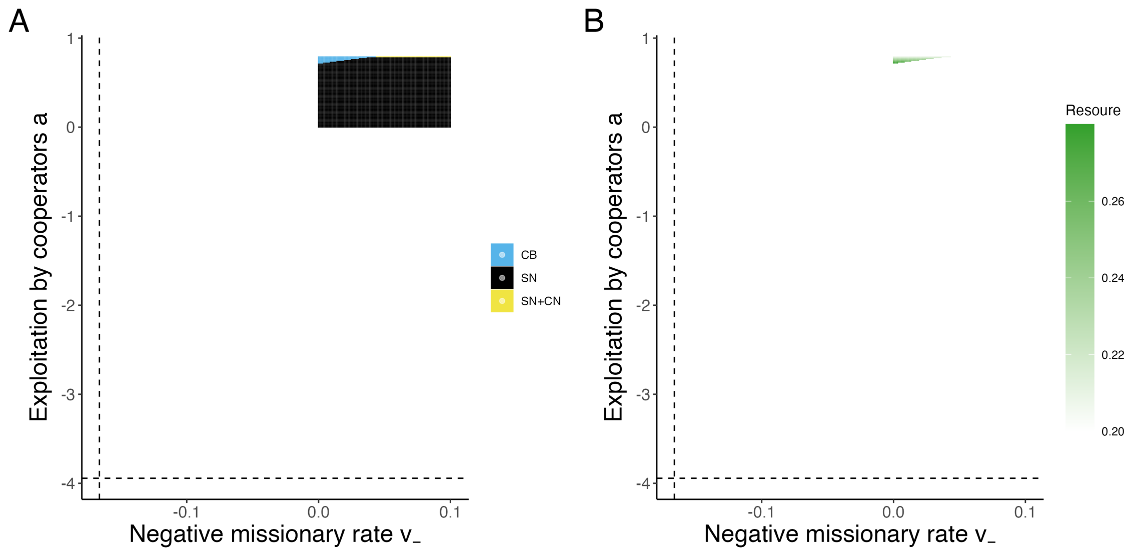

Figure S5: Parameter space when the fearness decreased over resource

Similar to Fig. 5 in the main text, but the fearness of the supernatural punishment was $P(R) = 1 - (pR)$ in this figure. The rest of the parameter values were identical to Fig. 5. As in the main text, we analyzed the parameter ranges $0 \leq v_{-} \leq 0.1$ and $a \leq a \leq 0.79$.

more "sinful." Mathematically, we assume the following function of $P(R)$ in this section:

$$P(R) = 1 - (pR)^u. \tag{S59}$$

In this case, dP/dR becomes negative. Nevertheless, our analyses in the main text and Supporting Information hold (except for the example in SI 2.4.1) because our conclusion does not depend on the sign of dP/dR.

Fig. S5 shows the evolutionary fate when $P(R) = 1 - (pR)$ (i.e., $u = 1$) over two parameter values a and v_{-} while the rest of the parameter values were identical to Fig. 5. Because $P(R_b^*)$ in this case was larger than in the main text, inequality (5) cannot hold in our parameter ranges. As a result, SN were evolutionarily stable in most cases. Although we found small areas where CB was fixed, both CB and SN were evolutionarily stable there. Fig. S6 shows that the initial conditions altered the evolutionary fate at such parameter values.

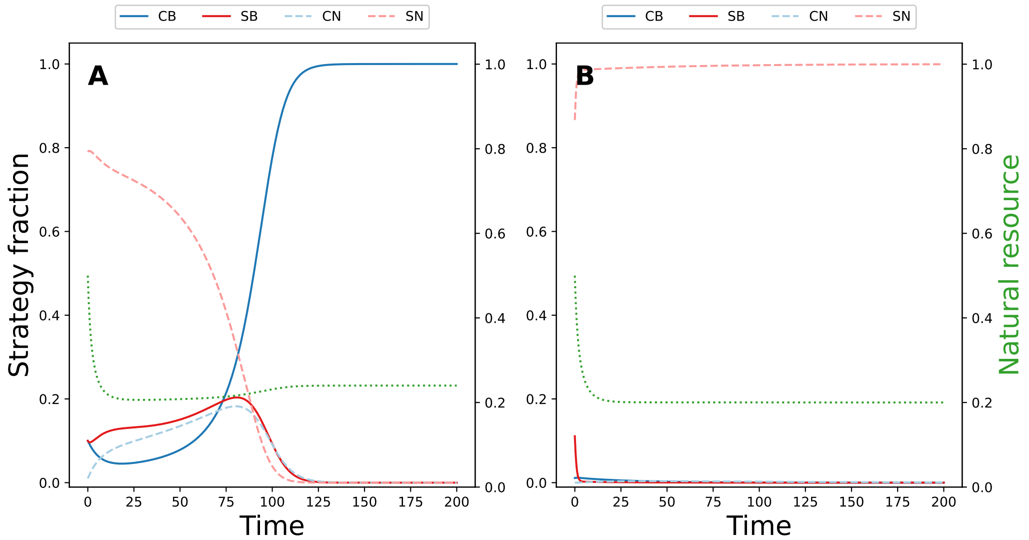

Figure S6: Example of dynamics when the fearness decreased over resource

Examples of the dynamics in Fig. S5. Panel A starte from $(R, x_{\mathrm{CB}}, x_{\mathrm{SB}}, x_{\mathrm{CN}}, x_{\mathrm{SN}}) = (0.5, 0.1, 0.1, 0.01, 0.79)$ while B strted from $(R, x_{\mathrm{CB}}, x_{\mathrm{SB}}, x_{\mathrm{CN}}, x_{\mathrm{SN}}) = (0.5, 0.01, 0.1, 0.1, 0.79)$