

# NATURAL SELECTION AND THE PROBLEM OF EVIL: AN EVOLUTIONARY MODEL WITH APPLICATION TO AN ANCIENT DEBATE

by Robert K. Fleck

*Abstract.* Since Darwin, scholars have contemplated what our growing understanding of natural selection, combined with the fact that great suffering occurs, allows us to infer about the possibility that a benevolent God created the universe. Building on this long line of thought, I develop a model that illustrates how undesirable characteristics of the world (stylized “evils”) can influence long-run outcomes. More specifically, the model considers an evolutionary process in which each generation faces a risk from a “natural evil” (e.g., predation, disease, or a natural disaster) subsequent to a basic resource allocation game. This allows both resource allocation and the natural evil to influence the number of surviving offspring. As the model shows, when the risk from the natural evil can be mitigated through the benevolent behavior of neighbors, the population may have increasing benevolence as a result of (1) greater risk from the natural evil and (2) a greater degree to which selfish individuals transfer resources to themselves in the resource allocation game. The main implication is that a world with evolutionary processes (in contrast to a world of static design) can allow two factors that have traditionally been considered “evils”—namely, the indiscriminate cruelty of the natural world and the capacity for humans to harm each other—to promote desirable long-run outcomes.

*Keywords:* argument from evil; evolution; natural selection; problem of evil

---

Thus, from the war of nature, from famine and death, the most exalted object which we are capable of conceiving, namely, the production of the higher animals, directly follows. There is grandeur in this view of life, with its several powers, having been originally breathed by the Creator into a few forms or into one; and that, whilst this planet has gone cycling on according to the fixed law of gravity, from so simple a beginning endless forms most beautiful and most wonderful have been, and are being, evolved.

—Concluding sentences from *On the Origin of Species* (Darwin 1859)

Although the question of whether a benevolent purpose could plausibly explain the design of the universe is an ancient one, the development of

Robert K. Fleck is a professor of Economics in the Department of Agricultural Economics and Economics, Montana State University, Bozeman, MT 59717, USA; e-mail rfleck@montana.edu.

evolutionary theory in the nineteenth century cast the question in new light. Darwin himself doubted whether a “beneficent and omnipotent God” would have designed a world in which some species would evolve to prey on others in the particularly cruel ways we observe in nature (Darwin 1860), but he also suggested how events that cause suffering might bring about desirable things (Darwin 1859—see the quotation above). The points raised by Darwin still hold great interest among scientists, philosophers, theologians, and the general public. Dawkins (1995), for example, begins his discussion of “God’s utility function” (i.e., maximizing DNA survival, indifference to happiness) by quoting Darwin. In a recent example from philosophy, Draper leads off his “Evil and Evolution” debate with Plantinga by discussing Darwin’s views on what the scale of suffering in the world implies for theism (Draper 2007a, 2007b, 2007c; Plantinga 2007a, 2007b, 2007c). Along similar lines, recent theological perspectives that embrace evolution must reconcile the creator’s benevolence with the harsh realities of a world shaped by natural selection (e.g., Attfield 2006; Southgate 2008). Moreover, in the recent wave of publications by proponents and critics of the “new atheism,” a central question is what those harsh realities allow us to infer about the existence (or nonexistence) of a benevolent designer (e.g., Dawkins 2006; Dowd 2006).

This paper contributes to the literature on evil and natural selection by developing an evolutionary model that integrates complementary ideas from different disciplines. In the field of biology, the question of why altruism (as opposed to selfishness) can survive in a world of natural selection has been a major issue, and theoretical models that address that question can illuminate philosophical questions about evil.<sup>1</sup> Moreover, the distinction between natural evils and moral evils—a distinction that plays a central role in the philosophical literature on why evil would exist—has not previously been used to motivate the specification of an evolutionary game.<sup>2</sup> To bring these ideas together, I develop a model that examines how the evolutionary fitness of altruism can be affected by the *interaction* between a natural evil’s effects and a moral evil’s effects.

The model works as follows. As in Trivers (1971) and much subsequent work, there are two types of individuals: altruists, who have altruistic offspring, and selfish individuals, who have selfish offspring. With types defined this way, a simple zero sum game of resource allocation will yield higher payoffs for selfishness than for altruism. Thus, the starting point is obvious: If reproduction is positively related to payoffs in such a game (when played between unrelated individuals), selfishness will yield higher rates of reproduction. Hence, in the most basic theoretical settings, to the extent that selfishness and altruism are inherited or learned from parents, nature would tend to select selfishness over altruism.<sup>3</sup> Yet, that finding can be reversed when individuals play a resource allocation game and face a risk from a natural evil, where that natural evil could be a predator, disease,

or natural disaster.<sup>4</sup> More specifically, the model considers interactions between randomly paired neighbors (in every generation, each individual draws a neighbor randomly from the population) and focuses on how outcomes change when a threat from a stylized natural evil occurs in the period following the resource allocation game. When the likelihood of surviving an encounter with a natural evil can be improved by a neighbor's altruistic actions, the population may have increasing altruism as a result of changes in two exogenous parameters: (1) a greater frequency of encountering natural evils and (2) a greater degree to which selfish individuals transfer resources to themselves from altruists in the resource allocation game.<sup>5</sup>

These two parameters can have these effects—even when the resource allocation game is zero sum—because, by reallocating resources away from altruistic neighbors, selfish individuals reduce the resources their neighbors have available for altruistic behavior. That reduction in resources will decrease the neighbors' altruistic activities and, therefore, decrease the likelihood that the selfish individuals' offspring will survive if struck by a natural evil. Thus, if the natural evil strikes sufficiently frequently and if altruistic activities have a sufficiently large effect on the risk created by the natural evil, the reduction in altruistic activities by those who neighbor selfish individuals can lead to natural selection against selfishness.

It would be possible, of course, to make very simple assumptions that lead directly to the survival of altruism, and to the growth of altruism being enhanced by a natural evil. For example, the propensity of predators (or disease or natural disasters) to strike altruists could be less than their propensity to strike selfish individuals, or altruists could direct their efforts toward other altruists (e.g., through assortative mating or reciprocity). I assume away most such factors because modeling them would be uninformative—it is obvious, for example, that the evolutionary fitness of altruism would be enhanced by natural evils that target selfish individuals and/or by altruists who obtain mutual gains by intentionally pairing themselves with others of their own type. Thus, what the model shows is that altruism can be promoted by a greater frequency of natural evils and by a greater extent to which selfish individuals transfer resources to themselves from altruists—even when natural evils strike individuals randomly, when altruistic and selfish individuals are paired randomly to play a game of resource allocation and defense against natural evils, and when altruists treat both kinds of individuals altruistically.<sup>6</sup>

To understand this result, it is essential to recognize that the selection of altruism in the population relies on sufficiently great evolutionary success of altruists who have altruistic neighbors. (Note that altruist–selfish pairings of neighbors cannot drive the selection of altruism because when a selfish individual is paired with an altruist, the selfish player will have more

expected surviving offspring than will the altruist.) Naturally, when there are few altruists in the population, only a small fraction of those few altruists will—given the random pairing of neighbors—have altruistic neighbors. Therefore, as in the previous literature on frequency dependence (e.g., Boorman & Levitt 1973), whether nature selects altruism or selfishness depends on the proportions of selfish and altruistic individuals in the initial population.

The model's findings build on the large literature that considers how altruism can survive and, in particular, how a stylized "gene" for altruism (or some similar type of strategy or form of mimicking) can survive in a process of natural selection.<sup>7</sup> My paper makes new contributions by examining the interaction between the effects of resource allocation and the effects of a natural evil, and interpreting that interaction in a manner inspired by theology and philosophy. The closest antecedents to my model are those that emphasize the importance of altruists being grouped together in order for altruism to survive.<sup>8</sup> As will become clear in the next section, my model's random pairing of neighbors for the combination of a resource allocation game and the risk of a natural evil produces a situation in which the natural evil has an effect similar to pairing altruists disproportionately with other altruists.<sup>9</sup> Thus, the model is part of a more general class of models in which a variety of mechanisms (e.g., kin living together, assortative mating, repeat interaction with neighbors) can pair altruists with altruists and thereby allow altruism to survive.<sup>10</sup>

After presenting the model, I will explain how the paper's findings can be applied to topics from the literature on the "problem of evil." The debate centers around the "argument from evil," which can be summarized usefully (albeit loosely) as follows: An omnipotent, omniscient, morally perfect creator of the universe could create, and would want to create, a universe without great evil; therefore, the existence of great evil provides evidence against the existence of an omnipotent, omniscient, morally perfect creator.<sup>11</sup> To be clear here, I should state that my paper is not intended to address the issue of the universe's creation. Nor is my paper intended as a rebuttal to the argument from evil. Those issues are beyond the scope of my paper. Rather, my paper shows how two factors that have traditionally been considered "evils"—namely, the indiscriminate cruelty of the natural world and the degree to which human selfishness can harm others—may promote good character (in the form of an altruistic population) in the long run. And this has interesting implications for the ongoing debate over the problem of evil.

#### THEORETICAL MODEL

*Assumptions.* The model describes generation-to-generation changes in a large population that has individuals of two types: altruistic and

selfish. In every generation, four events occur: (1) individuals are randomly paired so that each individual has one stylized neighbor, (2) each pair of neighbors plays a game of resource allocation, (3) each individual produces offspring of the individual's type and in proportion to the individual's resources, (4) some members of the population encounter a natural evil, with the probability of surviving the encounter higher in the presence of an altruistic neighbor (because an altruist is willing to incur costs in the effort to help a neighbor). The surviving offspring then repeat the four events.<sup>12</sup>

Formalizing these assumptions is straightforward. Let  $a_0$  represent the fraction of the initial generation that is altruistic, and  $s_0$  represent the fraction of the initial generation that is selfish, with  $0 < a_0 < 1$  and  $a_0 + s_0 = 1$ . When the resource allocation game has a homogeneous pair of players, each player receives half of the resources allocated. With a heterogeneous pair of players, the selfish player obtains a greater share than does the altruist. Let  $\tau$  represent the additional share going to the selfish player. (That is, let  $0.5 + \tau$  go to the selfish player and  $0.5 - \tau$  to the altruist, with  $0 < \tau \leq 0.5$ .) The joint endowment of resources to divide does not vary from one pair of players to another pair of players. Each individual faces the probability  $p$  that he or she (along with his or her offspring) will encounter a natural evil;  $0 < p < 1$ . Altruists and their offspring will attempt to help their neighbors if their neighbors encounter a natural evil. Attempting to help entails a risk of death: There is a probability of  $r$  that the altruist and his or her offspring will all die, and a probability of  $1 - r$  that the altruist and his or her offspring will all live ( $0 \leq r \leq 1$ ).<sup>13</sup> Selfish individuals will not attempt to help. If an individual paired with a selfish neighbor encounters a natural evil, he or she will have no surviving offspring. If paired with an altruistic neighbor, however, an individual's encounter with a natural evil will lead either to the offspring all dying or to the offspring all surviving, with a survival probability that increases with the altruistic neighbor's resources (equivalently, family size). To keep the assumption simple, let the survival probability be  $2m$  times the helpful individual's share of resources, with  $0 \leq m \leq 1$ .<sup>14</sup> If both individuals in a pair of neighbors encounter natural evils, I assume the encounters occur sequentially, and that the victim of the first encounter can, if and only if he or she survives, help his or her neighbor. For all the random factors (pairing of neighbors, encountering natural evils, and surviving natural evils), the random draws are based on statistically independent distributions, and population size is treated as a large sample.<sup>15</sup>

An additional consideration is that selfishness may shrink the pool of resources to allocate, perhaps through wasteful competition or a failure to obtain gains from cooperation. Including this type of "deadweight" loss in the model is not essential, but it is easy to do and highlights an important aspect of the model: The selection of altruism over selfishness does not

rest on selfishness reducing the total pool of resources at the resource allocation stage. To account for such deadweight losses, I will retain the assumption that all pairs of neighbors are homogeneous with respect to the endowment of resources they have to divide, but let  $d$  represent the proportional reduction in resources that occurs as a result of a player acting selfishly;  $0 \leq d \leq 1$ .<sup>16</sup> Setting  $d = 0$  thus characterizes a zero sum game at the resource allocation stage (the most interesting case).

*Implications.* The assumptions just described yield a simple characterization of generation-to-generation changes. For any pair of generations, the shares of the population that are altruistic and selfish can be stated, for generation  $t + 1$ , as functions of those shares in generation  $t$ :

$$a_{t+1} = [a_t a_t \pi_{aa} + a_t s_t \pi_{as}] [a_t a_t \pi_{aa} + a_t s_t \pi_{as} + s_t a_t \pi_{sa} + s_t s_t \pi_{ss}]^{-1}$$

$$s_{t+1} = [s_t a_t \pi_{sa} + s_t s_t \pi_{ss}] [a_t a_t \pi_{aa} + a_t s_t \pi_{as} + s_t a_t \pi_{sa} + s_t s_t \pi_{ss}]^{-1}$$

where:

$$\begin{aligned} \pi_{aa} = & [(1-p)^2 + (1-p)(p)(1-r) + (p)(1-p)(m) \\ & + (0.5p^2)(m)(1-r+r^2) + (0.5p^2)(m^2)(1-r)](0.5) \end{aligned}$$

$$\pi_{as} = [(1-p)^2 + (1-p)(p)(1-r)](0.5 - \tau)(1-d)$$

$$\begin{aligned} \pi_{sa} = & [(1-p)^2 + (1-p)(p) + (p)(1-p)(2m)(0.5 - \tau) \\ & + (0.5p^2)(2m)(0.5 - \tau)](0.5 + \tau)(1-d) \end{aligned}$$

$$\pi_{ss} = [(1-p)^2 + (1-p)(p)](0.5)(1-d)^2$$

The values of  $\pi$  indicate resource payoffs weighted by survival probabilities. Hence, they indicate relative rates of reproduction for an altruist encountering an altruistic neighbor ( $\pi_{aa}$ ) or selfish neighbor ( $\pi_{as}$ ) and for a selfish individual encountering an altruistic neighbor ( $\pi_{sa}$ ) or selfish neighbor ( $\pi_{ss}$ ). Given the initial population shares,  $a_0$  and  $s_0$ , one can use the equations for  $a_{t+1}$  and  $s_{t+1}$  to determine which type will grow or shrink (as a population share) over time.<sup>17</sup> See Appendix A for a more detailed explanation of how these equations are derived.

The next step is to consider seven propositions that illustrate the model's implications. As a starting point, I will discuss two propositions that describe outcomes when the frequency of the natural evil is low and when a neighbor's altruistic behavior has a small effect on the probability of surviving an encounter with a natural evil. This pair of propositions provides a useful baseline because the model's principal purpose is to

examine the opposite case (i.e., highly frequent natural evils when altruism has a large effect).

*Proposition 1.* In the case of the model when the resource allocation game is zero sum, a sufficiently low frequency of natural evils implies the following: Nature will select selfishness over altruism, and a greater effect of selfishness on resource allocation will cause more rapid growth of the selfish population. Stated more formally, when  $d = 0$ , sufficiently low  $p$  implies that (1)  $s_t < s_{t+1}$ , with  $s_t \rightarrow 1$  as  $t \rightarrow \infty$ , and (2) a higher  $\tau$  will cause a greater  $s_{t+1}$  for any given  $s_t$ . (Proof in Appendix B.)

*Proposition 2.* In the case of the model when the resource allocation game is zero sum, a sufficiently small effect of altruism on the risk from a natural evil implies the following: Nature will select selfishness over altruism, and a greater effect of selfishness on resource allocation will cause more rapid growth of the selfish population. Stated more formally, when  $d = 0$ , sufficiently low  $m$  implies that (1)  $s_t < s_{t+1}$ , with  $s_t \rightarrow 1$  as  $t \rightarrow \infty$ , and (2) a higher  $\tau$  will cause a greater  $s_{t+1}$  for any given  $s_t$ . (Proof in Appendix B.)

These propositions describe conditions under which nature selects selfishness and under which a greater effect of selfishness on resource allocation leads to more rapid growth of the selfish population. In view of these propositions, if the model shows nature selecting altruism, then it must be the result of one or both of the following two factors: (1) the resource allocation game is not zero sum, (2) the frequency of the natural evil and the effect of altruism on the risk from encountering a natural evil are sufficiently large. Similarly, if the model shows that a higher  $\tau$  causes more rapid growth in the altruistic population, the result must rest on one or both of those factors.

To examine the first of the two factors, consider the effects of dropping the condition that the resource allocation game is zero sum and, instead, assuming that selfishness decreases the sum of the payoffs.

*Proposition 3.* In the case of the model when selfishness causes a sufficiently large deadweight loss in resource allocation (i.e.,  $d$  is sufficiently high), nature will select altruism over selfishness ( $a_t < a_{t+1}$ , with  $a_t \rightarrow 1$  as  $t \rightarrow \infty$ ). (Proof in Appendix B.)

The logic underlying this result is very simple. When selfishness wastes a sufficiently large share of the resources that a pair of individuals has available to divide, selfishness will lead to low payoffs for selfish individuals as well as for altruists paired with those individuals. Hence, the rates of reproduction for pairs of altruists can be sufficiently great (relative to other types of pairs) that nature selects altruism.

Taking these first three propositions as a basis for comparison, we can now turn to the task of identifying the conditions under which nature will select altruism in the presence of a zero sum game of resource allocation. The next two propositions address this issue and, along with Propositions 6 and 7, present the paper's central results.

*Proposition 4.* In the case of the model when the resource allocation game is zero sum, the following holds true: A sufficiently high frequency of encountering a natural evil, combined with a sufficiently large effect of altruistic help, will cause nature to select altruism over selfishness. Stated more formally, when  $d = 0$ , the combination of sufficiently large  $p$  and sufficiently large  $m$  implies that (1)  $a_t < a_{t+1}$  and (2)  $a_t \rightarrow 1$  as  $t \rightarrow \infty$ . (Proof in Appendix B.)

*Proposition 5.* In the case of the model when the resource allocation game is zero sum, the following holds true: With a sufficiently high frequency of encountering a natural evil and a sufficiently large effect of altruistic help, a greater effect of selfishness on resource allocation will cause more rapid growth of the altruistic population. Stated more formally, when  $d = 0$ , the combination of sufficiently large  $p$  and sufficiently large  $m$  implies that, conditional on  $a_t$ , a higher  $\tau$  will lead to a greater  $a_{t+1}$ . (Proof in Appendix B.)

What is the mechanism driving these results? By reallocating resources away from altruists, selfish individuals reduce the resources used for altruistic behavior, and that can, in turn, reduce the selfish individuals' success in having surviving offspring. The key here is that, although altruists do not target their altruism toward other altruists, and while there is no reciprocity *per se*, the multiple interactions between neighbors mean that selfish individuals interact with poorer altruists (equivalently, fewer altruists) when they take more resources from altruists. Thus, when encounters with natural evils occur sufficiently frequently and altruism has a sufficiently great effect on mitigating the risk from a natural evil, the net effect of selfishness on evolutionary success is negative (Proposition 4). Furthermore, that negative effect is stronger when the effect of selfishness on resource allocation is greater (Proposition 5).

To understand more precisely what drives the model, consider which types of pairings of neighbors lead to greater reproductive success for selfish individuals and which lead to greater reproductive success for altruists. A selfish individual in a selfish–selfish pairing will have more expected surviving offspring than will an altruist in a selfish–altruist pairing. Furthermore, a selfish individual in a selfish–altruist pairing will have more expected offspring than will an altruist in a selfish–altruist pairing. Both of these factors will favor selfishness in terms of natural selection. However, an altruist in an altruist–altruist pairing will, for some sets of parameter values, have more expected surviving offspring than will a selfish individual in a selfish–altruist pairing. It is this last factor that drives the results stated in Propositions 4 and 5.

Given the importance of altruist–altruist pairings to altruism's potential evolutionary advantage over selfishness, an obvious point to consider is the relative frequency of those pairings. The relative frequency depends, of course, on the proportions of altruists and selfish individuals in the population. This matters for interpreting Propositions 4 and 5, because both propositions describe the implications of *sufficiently large* parameter



values reflecting the frequency of a natural evil and the effects of altruistic help: What constitutes sufficiently large here will depend on the composition of the population. Thus, the composition of the initial population is critical, as is often the case in evolutionary models (e.g., Axelrod 1984; Boorman & Levitt 1973; Dawkins 1989; Fudenberg & Levine 1998). More formally:

*Proposition 6. Which type of individual will have the greater expected number of surviving offspring can depend on the relative numbers of altruistic and selfish individuals in the initial period. In particular, even for a given set of parameter values, nature may select altruism if there are sufficiently many altruists initially, yet select selfishness if there are sufficiently few altruists initially. (Proof in Appendix B.)*

The logic of this proposition is straightforward. Because the selection of altruism depends on altruists being paired with altruists, when too few altruists exist, nature will select selfishness. Thus, if one views the two types of individuals as the two possible strategies, then a single set of parameter values may allow each type to be an evolutionarily stable strategy.<sup>18</sup>

Again consider the composition of the initial population, but this time consider its interaction with the degree to which selfishness affects resource allocation.

*Proposition 7. The relative numbers of altruistic and selfish individuals in the initial period can determine whether a greater effect of selfishness on resource allocation (i.e., a higher  $\tau$ ) will cause more rapid growth of the altruistic population (i.e., a greater  $a_{t+1}$  for any given  $a_t$ ). In particular, even for a given set of parameter values, it can be that a greater effect of selfishness on resource allocation accelerates the growth of the altruistic population if there are sufficiently many altruists initially, yet accelerates the growth of the selfish population if there are sufficiently few altruists initially. (Proof in Appendix B.)*

Thus, whether a greater effect of selfishness on resource allocation works for or against the selection of altruism depends on the composition of the population, as well as on the parameter values.

#### INTERPRETING THE EVOLUTIONARY GAME

This section discusses how the model's structure and parameters relate to real world factors. The main purpose here is to illustrate the intuition behind the model and, by doing so, set up the explanation (in the "Implications for the Argument from Evil and the Free Will Defense" section) of how the model illuminates the problem of evil. As a preliminary point, it is essential to recognize that the propositions derived in the previous section come from a highly stylized characterization of behavior—as do, of course, the implications of other evolutionary models examining altruism. Notably, the model allows only two possible types of individuals, and the particular form of altruism considered cannot be targeted toward

kin or those who will reciprocate. Although the real world is obviously more complex, these abstractions are valuable because they allow a tractable model in which a natural evil can influence the evolutionary fitness of a form of altruism that corresponds quite reasonably to benevolence on the part of an individual.<sup>19</sup> That said, the reader must keep in mind that a model incorporating such abstractions cannot logically be interpreted as an attempt to provide a *general* explanation of why altruism persists in a population. Rather, the model identifies factors that may *contribute* to the selection and persistence of altruistic traits.

*Real-world analogues to  $p$  and  $\tau$ .* The model illustrates how highly frequent encounters with natural evils ( $p$ ) and highly effective transfers of resources to selfish individuals from altruists ( $\tau$ ) may lead to a more altruistic population. What real world phenomena might resemble the theoretical parameters? Consider, for example, the scarcity of hunting grounds and other food supplies—these have long created a resource allocation game that determines the number of offspring an individual (or couple or family) can support. And attacks by large predators, as well as disease, natural disasters, and accidents, have long created circumstances under which altruistic help can be costly to the helper and potentially life saving for the recipient of that help. When such circumstances arise and individuals (and/or their offspring) remain in the same geographical location long enough to experience resource allocation and some threat analogous to encountering a natural evil, factors analogous to high  $p$  and high  $\tau$  may, through the logic of Propositions 4 and 5, promote an altruistic population in a manner that complements the effects of other phenomena (e.g., kin selection, reciprocal altruism, group selection, and mimicry of neighbors).

Interestingly, the scientific literature suggests that natural disasters may have—essentially by increasing  $p$  to high values—played a major role in the evolution of human cooperation. For example, the Toba volcanic eruption (approximately 73,500 years ago) may have reduced the human population dramatically, perhaps to fewer than 10,000 individuals (e.g., Ambrose 1998; Rampino & Ambrose 2000). As Ambrose (1998, 2003) hypothesizes, the huge eruption created conditions (including an ice age) that affected human behavior by causing strong selection for social cooperation. If so, the effects of Toba could have been similar to what my model illustrates for the case of  $p$  near one and  $m$  very high.

Several factors may contribute to circumstances analogous to a large value of  $\tau$ . If individuals of the selfish type are particularly aggressive in their efforts to obtain resources at the expense of altruists and/or if individuals of the altruistic type are particularly meek when competing against others for resources, that will cause greater inequality in the allocation of resources in altruist–selfish pairings. That is,  $\tau$  will be large. Conditional on the level of

aggressiveness and meekness, the size of  $\tau$  would also depend on the nature of the resources to be allocated. If the principal resource at stake is, say, wild nuts and berries that individuals consume hand-to-mouth by foraging in a large area that has only slight heterogeneity in productivity, the resource may be allocated nearly equally in altruist–selfish pairings. Thus,  $\tau$  will be small.<sup>20</sup>

What do these influences on  $\tau$  imply for interpreting the model? As Proposition 5 shows, with sufficiently frequent encounters with natural evils ( $p$ ) and a sufficiently large effect of altruistic help ( $m$ ), a greater value of  $\tau$  will accelerate the growth of the altruistic population. Thus, when competing for resources, greater aggressiveness on the part of the selfish type and/or greater meekness on the part of the altruistic type can lead to more rapid selection of altruists. In essence, if the altruistic meek are to inherit the earth, they may inherit it faster if they get very few resources when paired with selfish individuals.<sup>21</sup>

#### IMPLICATIONS FOR THE ARGUMENT FROM EVIL AND THE FREE WILL DEFENSE

Although there have been many efforts to explain why a benevolent God would create a world that allows great evil, the most influential arguments in the scholarly literature appeal to the value of free will.<sup>22</sup> The main idea of the “free will defense” against the argument from evil is that free will—if it is to be meaningful—must allow the existence of evil. My model does not incorporate a concept of free will (though, as noted later, the model could be modified to do so), but it does address a key question raised by critics of the free will defense: Even if the existence of free will requires that humans have the ability to engage in moral evils, what would be the purpose of natural evils? To show the relevance of my model, I will discuss how moving from a static framework to a dynamic framework (in which frequency dependence plays a key role) can fundamentally change the effects of evil.

*Good and evil in a dynamic, evolutionary world.* Recall that my main theoretical results arise from the dynamic features of the model: It is the generation-to-generation changes that allow the conclusion that a combination of evils can lead to the survival and faster growth of altruism.

Moreover, for a given set of parameter values, the characteristics of the initial population may determine both (1) whether nature selects altruism or selfishness and (2) whether selfish individuals transferring resources to themselves from altruists works for or against the selection of altruism. In other words, for a given set of rules for the world (i.e., given parameter values), there can be two populations (e.g., isolated islands on a planet) with different trajectories: one population on the path to greater selfishness with

growth in selfishness accelerated by a large effect of selfishness on resource allocation, the other population on the path to greater altruism with growth in altruism accelerated by a large effect of selfishness on resource allocation. In a static world, such effects (and indeed all of the interesting effects illustrated in my model) would not exist.<sup>23</sup> Thus, this paper demonstrates how moving from a static view of the world to a dynamic framework with natural selection can provide new insight into the potential for natural and moral evils to promote desirable long-run outcomes.

For understanding my contribution to the literature, it is important to recognize that rebuttals to the argument from evil sometimes have a dynamic nature. An obvious point here is that some religious traditions emphasize a long-run approach rather than a static approach: Good eventually will prevail over evil, but only after a long struggle. Building on that traditional view, some recent theological perspectives on evolution and evil call upon humans to reduce the suffering that necessarily arises in an evolutionary world (e.g., Attfield 2006; Southgate 2008).

It is also important to consider the views of the scholars who set out the ideas on which my model, like other dynamic models in biology and economics, builds. Darwin provides perhaps the most obvious case in point (see the quotation used to lead off this paper). Other examples come from Malthus and Mill. After explaining his famous “principle of population,” Malthus (1798, 114) writes, “This general law [that population will increase faster than food production] undoubtedly produces much partial evil, but a little reflection may, perhaps, satisfy us, that it produces a great overbalance of good.” In essence, Malthus argues that although humankind’s potential for rapid population growth (relative to food supply) causes suffering, the scarcity of food provides strong incentives for humans to work hard, both physically and mentally, to reach desirable outcomes. Mill’s (1879) approach to the issue is similar, but more focused on optimistic expectations of economic growth.<sup>24</sup> Arguments along the lines of Darwin, Malthus, and Mill are important intellectual predecessors to what I argue here.<sup>25</sup>

That said, my model provides a much more precise framework to see how the existence of evil may shape the process through which the prevalence of good increases while the prevalence of evil declines. Consider, for example, the following scenario: a one-time (i.e., single generation) plague that strikes the population indiscriminately, so that “good” people—analogueous to altruists in my model—are not immune. (As an alternative to a plague, one can substitute a giant flood, a Toba-scale volcanic eruption, or some other cataclysmic event of one’s choice.) My model demonstrates a mechanism through which such an event could generate a *long-run* result that arguably might be valued by a benevolent creator.<sup>26</sup> More specifically, if goodness of character helps others survive a plague, then a particularly virulent plague may increase the share of good people in the population.

This is true (as Proposition 4 shows) even if being a good person means that one would help any neighbor (good or bad), even if good people risk death when helping bad people (as well as when helping good people), even if bad people never help good people, and even if good and bad people encounter each other randomly, as opposed to a situation where good people are separated from bad people. And (as Proposition 5 shows) the worse the bad people have been in the sense of how much they have harmed their good neighbors, the more strongly a plague may select good people over bad people for survival.

Also note that, because natural selection is a dynamic process, even a one-time plague may permanently reverse the course along which a population is heading. To see why, consider a population on a path of declining goodness—that is, the altruistic share of the population is declining. For the reasons discussed in the previous paragraph (i.e., the logic of Propositions 4 and 5), a plague striking this population may, over the course of the generation in which it strikes, increase the share of altruistic people in the population. But the key point is this: The population may not return to the path of declining goodness, even if the plague leaves other things unchanged (i.e., the post-plague world and the pre-plague world have the same parameter values). The mechanism allowing a transitory plague to have permanent effects is frequency dependence (Proposition 6): A one-time event may increase the altruistic share of the population from below to above the critical number necessary for the evolutionary survival of altruism. In sum, a one-time plague that kills many people may change the long-run outcome dramatically: from one in which good character is heading toward extinction, to one in which bad character is heading toward extinction.

*How does the model relate to criticisms of the free will defense?* To understand my point here, it is essential to recognize from the start that the arguments in this paper cannot logically be interpreted as a defense against the argument from evil. More specifically, showing that stylized evils may increase the frequency of altruism in a population is, at most, an argument that evils can have some desirable effects, not that evils are on the whole desirable. And evils must, of course, have some undesirable effects—otherwise, they would not really be evil. Even if one accepts that these desirable effects are extremely valuable, there remains the question of whether the specific types of evil we see in the world are the best of all possible ways to obtain those benefits (e.g., Attfield 2006; Southgate 2008). Moreover, even if one accepts that a benevolently designed world may have moral and natural evils, there remains the question of whether a benevolently designed world would allow such a disproportionately large share of the suffering caused by those evils to fall on some particular individuals. In this light, the reader should keep in mind that my model is not designed to consider, and thus cannot assess, the *net* desirability of evil.

Nevertheless, the model's findings complement the free will defense. As noted earlier, even if we operate on the assumption that free will requires humans to have the ability to engage in moral evils, that leaves open the question of why we would see natural evils, especially those of the magnitude we see (e.g., epidemics of deadly diseases, massive natural disasters).<sup>27</sup> A variety of responses have been offered, including the idea that natural evils have origins in humankind's (biblical) fall from grace, or that they are essential for humankind to have sufficient knowledge to exercise free will.<sup>28</sup> My model suggests a different response: In the presence of a moral evil (selfishness in my model), a greater prevalence of natural evils can reduce the evolutionary success of that moral evil. Indeed, a natural evil of sufficient magnitude may even reverse the evolutionary path of the population from one with growing evil to one with declining evil.<sup>29</sup>

*How the model differs from "virtuous response to suffering" arguments.* My model works in a manner that differs starkly from the concept of a "virtuous response to suffering"—a major line of thought on why evil may enable greater good (e.g., Schlesinger 1983). More specifically, the survival and growth of altruism in the model do not rest on moral evil delivering any benefits to those who encounter it. Rather, the model's altruists who encounter selfishness *always* suffer harm in terms of reduced resource allocations and reduced survival probabilities. Furthermore, in no case of the model would being selfish (as opposed to being altruistic) lead to more altruism in the next generation. Indeed, the greater the size of the selfish population today, the greater the size of the selfish population for all future periods. To put this in plain language and relate it to the context of Nazi Germany (a ubiquitous example in objections to arguments based on a virtuous response to suffering): My model's results neither depend on, nor suggest a reason for, anything good coming from the existence of Nazis.<sup>30</sup>

*Adding "free will" to the model.* Given its set-up as a standard evolutionary game, the model does not incorporate anything along the lines of free will, but a stylized choice between good and evil could be added to the model. The simplest approach would be to modify the replicator dynamic so that each individual chooses between altruism and selfishness. In doing this, the basic logic of the evolutionary process could be retained by setting a higher probability of choosing altruism among the offspring of altruists, and a lower probability among the offspring of the selfish. With such an extension, the logic of the model presented in this paper would be preserved.<sup>31</sup> Another possible extension would be to assume that individuals have objective (or utility) functions of some form, then allow individuals to choose between altruism and selfishness based on some

signal conveying information about the characteristics of the population.<sup>32</sup> This would allow choices between altruism and selfishness to be frequency dependent.<sup>33</sup> Such extensions of the model are potential directions for future research.

## CONCLUSION

By integrating ideas from evolutionary biology with ideas from religion and philosophy, this paper develops a model that combines a simple game of resource allocation with the risk from a natural evil (predation, disease, or natural disaster). The model assumes that, in every generation, each individual draws a neighbor randomly from the population. It also assumes that, by drawing an altruistic neighbor rather than a selfish neighbor, an individual will have more resources and a greater likelihood of surviving the natural evil. The expected number of surviving offspring an individual will have depends on the allocation of resources between the individual and a neighbor, the likelihood of encountering (and being killed by) a natural evil, and the likelihood of being killed while providing altruistic help during a neighbor's encounter with a natural evil.

The most interesting finding is that, for some parameter values and initial conditions, the population will have greater altruism as a result of (1) a greater risk from a natural evil and (2) a greater degree to which selfish individuals transfer resources to themselves from altruists. In the context of the ancient debate over the problem of evil (and in particular the recent debate over the free will defense), this finding makes a new contribution by demonstrating how a population can, in the long run, develop greater altruism as a result of two factors that have traditionally been considered evils. Furthermore, because it is the dynamic nature of the model that generates the interesting findings, the model shows how the effect of the two evils may vary between populations, even when the same rules and parameter values apply to both populations. In essence, the model explains why natural selection might lead to meek altruists inheriting the earth, how a single cataclysmic event can prevent the extinction of meek altruists (even if they help selfish individuals survive the event, yet receive no help from the selfish), and why meek altruists might inherit the earth even faster because of their meekness and because of particularly aggressive selfishness in other individuals. In this light, the idea of a benevolent God creating a world that allows great evil is made more plausible—not less plausible—by the fact that we live in a world of evolutionary change.

## ACKNOWLEDGMENTS

For helpful comments, I thank an anonymous referee, Beth Davenport, Ross Emmett, Dino Falaschetti, Andy Hanssen, and Chris Stoddard.

## APPENDIX A: CALCULATING CHANGES IN POPULATION SHARES

As explained in the “Theoretical Model” section, the values of  $\pi$  indicate resource payoffs weighted by survival probabilities, where

$$\begin{aligned}\pi_{aa} &= [(1-p)^2 + (1-p)(p)(1-r) + (p)(1-p)(m) \\ &\quad + (0.5p^2)(m)(1-r+r^2) + (0.5p^2)(m^2)(1-r)](0.5) \\ \pi_{as} &= [(1-p)^2 + (1-p)(p)(1-r)](0.5-\tau)(1-d) \\ \pi_{sa} &= [(1-p)^2 + (1-p)(p) + (p)(1-p)(2m)(0.5-\tau) \\ &\quad + (0.5p^2)(2m)(0.5-\tau)](0.5+\tau)(1-d) \\ \pi_{ss} &= [(1-p)^2 + (1-p)(p)](0.5)(1-d)^2\end{aligned}$$

To explain these formulas, I will discuss how to calculate the survival probabilities and resource payoffs for each possible combination of encountering or not encountering a natural evil. For each individual, the probability of an encounter is  $p$ . Hence, the formulas for the various  $\pi$  values reflect, for an individual and his or her neighbor, the probabilities of five different outcomes:  $(1-p)^2$  for neither encountering a natural evil;  $(1-p)p$  for only the neighbor having an encounter;  $p(1-p)$  for only the individual having an encounter;  $0.5p^2$  for both having encounters, with the individual having the first encounter;  $0.5p^2$  for both having encounters, with the neighbor having the first encounter. I will now consider the survival probabilities, conditional on each of those encounters-with-natural-evil scenarios, for each type of pairing of individuals.

For an altruistic individual paired with an altruistic neighbor, the survival probabilities are as follows:  $(1-r)$  if only the neighbor has an encounter,  $m$  (which equals  $2m$  times the neighbor’s resources) if only the individual has an encounter,  $m(1-r+r^2)$  if the individual has the first encounter and the neighbor has a second, and  $m^2(1-r)$  if the neighbor has the first encounter and the individual has the second.<sup>34</sup> In each case, if the individual survives, the individual’s share of resources is 0.5. Weighting the probabilities of the scenarios by the survival probabilities, then multiplying by the share of resources, yields the formula for  $\pi_{aa}$ .

To obtain the formula for  $\pi_{as}$ , there are two key changes from  $\pi_{aa}$ . First, because the neighbor is selfish, the individual’s share of resources, conditional on survival, is reduced to  $(0.5-\tau)(1-d)$ . Second, the individual will survive if and only if the individual does not encounter a natural evil.

To obtain the formula for  $\pi_{sa}$ , there are five key changes from  $\pi_{aa}$ . First, because the individual is selfish, the individual’s share of resources, conditional on survival, is changed to  $(0.5+\tau)(1-d)$ . Second, if only the neighbor encounters a natural evil, the individual’s survival probability is higher, at 1 instead of  $(1-r)$ , because a selfish individual will not incur



the risk involved in attempting to help the neighbor. Third, if only the individual encounters the natural evil, the individual's survival probability is lower, at  $(2m)(0.5 - \tau)$  instead of  $m$ , because a selfish individual leaves the neighbor with fewer resources. Fourth, if the individual has the first encounter with a natural evil and the neighbor has a second encounter, the survival probability changes to  $(2m)(0.5 - \tau)$  from  $m(1 - r + r^2)$  because (1) the selfish individual leaves the neighbor with fewer resources and (2) the selfish individual will not incur the risk involved in attempting to help the neighbor. Fifth, if the neighbor has the first encounter and the individual has the second encounter, the individual will not survive, because the neighbor will not be alive to provide help.

To obtain the formula for  $\pi_{ss}$ , there are two key changes from  $\pi_{as}$ . First, because both individuals, rather than just the neighbor, are selfish, the individual's share of resources changes from  $(0.5 - \tau)(1 - d)$  to  $(0.5)(1 - d)^2$ . Second, if only the neighbor has an encounter, the individual's survival probability is higher, at 1 instead of  $(1 - r)$ , because a selfish individual will not incur the risk involved in attempting to help the neighbor.

Now, with  $\pi_{aa}$ ,  $\pi_{as}$ ,  $\pi_{sa}$ , and  $\pi_{ss}$  as defined above, one can write concise expressions for the period-to-period changes in the population shares for each type. Because  $a_t$  and  $s_t$  measure population shares in a large population,  $a_t$  indicates the probability that any given individual will draw an altruist for a neighbor, while  $s_t$  indicates the probability that any individual will draw a selfish neighbor. Thus,  $a_t a_t$  indicates the share of the population made up of altruists with altruistic neighbors,  $a_t s_t$  indicates the share of the population made up of altruists with selfish neighbors,  $s_t a_t$  indicates the share of the population made up of selfish individuals with altruistic neighbors, and  $s_t s_t$  indicates the share of the population made up of selfish individuals with selfish neighbors. Multiplying each of those shares by its corresponding value of  $\pi$  thus indicates the resources going to altruists ( $a_t a_t \pi_{aa} + a_t s_t \pi_{as}$ ) and to selfish individuals ( $s_t a_t \pi_{sa} + s_t s_t \pi_{ss}$ ). Finally, given the assumption that reproduction occurs in proportion to resources, one can obtain population shares for each type by dividing the survival-weighted resources going to each type by the total survival-weighted resources allocated:

$$a_{t+1} = [a_t a_t \pi_{aa} + a_t s_t \pi_{as}] [a_t a_t \pi_{aa} + a_t s_t \pi_{as} + s_t a_t \pi_{sa} + s_t s_t \pi_{ss}]^{-1}$$

$$s_{t+1} = [s_t a_t \pi_{sa} + s_t s_t \pi_{ss}] [a_t a_t \pi_{aa} + a_t s_t \pi_{as} + s_t a_t \pi_{sa} + s_t s_t \pi_{ss}]^{-1}$$

Equivalently, for  $a_t > 0$  and  $s_t > 0$ :

$$a_{t+1}/a_t = [a_t \pi_{aa} + s_t \pi_{as}] [a_t a_t \pi_{aa} + a_t s_t \pi_{as} + s_t a_t \pi_{sa} + s_t s_t \pi_{ss}]^{-1}$$

$$s_{t+1}/s_t = [a_t \pi_{sa} + s_t \pi_{ss}] [a_t a_t \pi_{aa} + a_t s_t \pi_{as} + s_t a_t \pi_{sa} + s_t s_t \pi_{ss}]^{-1}$$

## APPENDIX B: PROOFS OF PROPOSITIONS 1–7

*Proof of Proposition 1.* With  $d = 0$ , as  $p \rightarrow 0$ , the limits of  $\pi$  values are as follows:

$$\pi_{aa} \rightarrow (0.5)$$

$$\pi_{as} \rightarrow (0.5 - \tau)$$

$$\pi_{sa} \rightarrow (0.5 + \tau)$$

$$\pi_{ss} \rightarrow (0.5)$$

Now consider what these limits imply about the selection of selfishness. Because the model assumes  $\tau > 0$ , it is obvious that  $\pi_{aa} < \pi_{sa}$  and  $\pi_{as} < \pi_{ss}$ . Given that  $\pi_{aa} < \pi_{sa}$  and  $\pi_{as} < \pi_{ss}$ , along with the model's general assumption that  $a_0 > 0$  and  $s_0 > 0$ , it follows immediately from the formulas for  $a_{t+1}/a_t$  and  $s_{t+1}/s_t$  (see Appendix A) that  $a_{t+1}/a_t < 1$  and  $s_{t+1}/s_t > 1$ . Thus, when  $p$  is sufficiently small, nature selects selfishness. That, combined with the fact that  $(a_{t+1}/a_t)/(s_{t+1}/s_t)$  does not approach 1 (even as  $t \rightarrow \infty$ ), implies that  $s_t \rightarrow 1$  as  $t \rightarrow \infty$ . Furthermore, because of the negative effect of  $\tau$  on  $\pi_{as}$  and the positive effect of  $\tau$  on  $\pi_{sa}$ , a greater value of  $\tau$  leads to a greater  $s_{t+1}$  for any given  $s_t$  (i.e., it leads to more rapid growth of the selfish population).

*Proof of Proposition 2.* With  $d = 0$ , as  $m \rightarrow 0$ , the limits of  $\pi_{aa}$  and  $\pi_{sa}$ , and the values of  $\pi_{as}$  and  $\pi_{ss}$ , are as follows:

$$\pi_{aa} \rightarrow [(1 - p)^2 + (1 - p)(p)(1 - r)](0.5)$$

$$\pi_{as} = [(1 - p)^2 + (1 - p)(p)(1 - r)](0.5 - \tau)$$

$$\pi_{sa} \rightarrow [(1 - p)^2 + (1 - p)(p)](0.5 + \tau)$$

$$\pi_{ss} = [(1 - p)^2 + (1 - p)(p)](0.5)$$

Now consider what these limits and values imply about the selection of selfishness. Because the model assumes  $\tau > 0$ , it is obvious that  $\pi_{aa} < \pi_{sa}$  and  $\pi_{as} < \pi_{ss}$ . Given that  $\pi_{aa} < \pi_{sa}$  and  $\pi_{as} < \pi_{ss}$ , along with the model's general assumption that  $a_0 > 0$  and  $s_0 > 0$ , it follows immediately from the formulas for  $a_{t+1}/a_t$  and  $s_{t+1}/s_t$  (see Appendix A) that  $a_{t+1}/a_t < 1$  and  $s_{t+1}/s_t > 1$ . Thus, when  $m$  is sufficiently small, nature selects selfishness. That, combined with the fact that  $(a_{t+1}/a_t)/(s_{t+1}/s_t)$  does not approach 1 (even as  $t \rightarrow \infty$ ), implies that  $s_t \rightarrow 1$  as  $t \rightarrow \infty$ . Furthermore, because of the negative effect of  $\tau$  on  $\pi_{as}$  and the positive effect of  $\tau$  on  $\pi_{sa}$ , a greater value of  $\tau$  leads to a greater  $s_{t+1}$  for any given  $s_t$  (i.e., it leads to more rapid growth of the selfish population).

*Proof of Proposition 3.* As  $d \rightarrow 1$ , the value of  $\pi_{aa}$  and the limits of  $\pi_{as}$ ,  $\pi_{sa}$ , and  $\pi_{ss}$  are as follows:

$$\begin{aligned} \pi_{aa} &= [(1 - p)^2 + (1 - p)(p)(1 - r) + (p)(1 - p)(m) \\ &\quad + (0.5p^2)(m)(1 - r + r^2) + (0.5p^2)(m^2)(1 - r)](0.5) \\ \pi_{as} &\rightarrow 0 \\ \pi_{sa} &\rightarrow 0 \\ \pi_{ss} &\rightarrow 0 \end{aligned}$$

Now consider what these limits and values imply about the selection of selfishness. Given that  $\pi_{aa} > \pi_{sa}$  and  $\pi_{as} = \pi_{ss}$ , from the formulas for  $a_{t+1}/a_t$  and  $s_{t+1}/s_t$  (see Appendix A) it follows that, for any  $a_t > 0$  and  $s_t > 0$ , it must be that  $a_{t+1}/a_t > 1$  and  $s_{t+1}/s_t < 1$ . This implies (given the model's general assumption that  $a_0 > 0$  and  $s_0 > 0$ ) that nature will select altruism when  $d$  is sufficiently large (i.e., sufficiently near 1). The  $\pi$  values also imply that  $a_t \rightarrow 1$  as  $t \rightarrow \infty$ .

*Proof of Proposition 4.* With  $d = 0$ , as  $p \rightarrow 1$  and  $m \rightarrow 1$ , the limits of the  $\pi$  values are as follows:

$$\begin{aligned} \pi_{aa} &\rightarrow 0.5 - 0.5r + 25r^2 \\ \pi_{as} &\rightarrow 0 \\ \pi_{sa} &\rightarrow 0.25 - \tau^2 \\ \pi_{ss} &\rightarrow 0 \end{aligned}$$

Now consider what these limits imply about the selection of altruism. The model assumes  $\tau > 0$  and  $0 \leq r \leq 1$ . Therefore, the formulas above show that  $\pi_{aa} > \pi_{sa}$  and  $\pi_{as} = \pi_{ss}$ .<sup>35</sup> And, hence, for any  $a_t > 0$  and  $s_t > 0$ , it follows immediately from the formulas for  $a_{t+1}/a_t$  and  $s_{t+1}/s_t$  (see Appendix A) that  $a_{t+1}/a_t > 1$  and  $s_{t+1}/s_t < 1$ . This implies (given the model's general assumption that  $a_0 > 0$  and  $s_0 > 0$ ) that nature will select altruism when  $p$  is sufficiently large (i.e., sufficiently near 1) and  $m$  is sufficiently large (i.e., sufficiently near 1). The  $\pi$  values also imply that  $a_t \rightarrow 1$  as  $t \rightarrow \infty$ .

*Proof of Proposition 5.* The assumed conditions are the same as in Proposition 4. Again, consider what the limits imply about the selection of altruism. As is easily seen from the proof of Proposition 4, an increase in  $\tau$  will decrease  $\pi_{sa}$ , while leaving  $\pi_{aa}$ ,  $\pi_{as}$ , and  $\pi_{ss}$  unchanged. This means

that, for any given  $a_t > 0$  and  $s_t > 0$ , an increase  $\tau$  will cause an increase in  $a_{t+1}$ . This implies a greater  $a_{t+1}$  for any given  $a_t$ . In other words, it implies more rapid growth of the altruistic population.

*Proof of Proposition 6.* One can prove the proposition with a suitable set of parameter values. Consider the following:  $p = 0.9$ ,  $\tau = 0.4$ ,  $m = 1$ ,  $r = 0$ ,  $d = 0$ .<sup>36</sup> These parameters generate  $\pi_{aa} = 0.5$ ,  $\pi_{as} = 0.01$ ,  $\pi_{sa} = 0.1791$ ,  $\pi_{ss} = 0.05$ . Basic algebra (details available) shows that  $a_0 \geq 0.1109$  yields  $a_1/a_0 > 1$ , which implies that  $a_{t+1}/a_t > 1$  for all periods, while  $a_0 \leq 0.1108$  yields  $a_1/a_0 < 1$ , which implies that  $a_{t+1}/a_t < 1$  for all periods.<sup>37</sup> Thus, with the chosen set of parameters, nature will select altruism if there are sufficiently many altruists initially ( $a_0 \geq 0.1109$ ), yet select selfishness if there are sufficiently few altruists initially ( $a_0 \leq 0.1108$ ). That, combined with the fact that  $(a_{t+1}/a_t)/(s_{t+1}/s_t)$  decreases as  $a_t$  decreases and the fact that  $(s_{t+1}/s_t)/(a_{t+1}/a_t)$  decreases as  $s_t$  decreases, demonstrates convergence over time:  $a_t \rightarrow 1$  as  $t \rightarrow \infty$  when there are sufficiently many altruists initially, while  $s_t \rightarrow 1$  as  $t \rightarrow \infty$  when there are sufficiently few altruists initially.

*Proof of Proposition 7.* One can prove the proposition with an example that shows how a change in the value of  $a_t$  can reverse the effect that a greater  $\tau$  has on  $a_{t+1}/a_t$ . Consider again the example used to prove Proposition 6 (i.e.,  $p = 0.9$ ,  $\tau = 0.4$ ,  $m = 1$ ,  $r = 0$ ,  $d = 0$ ), but now consider increasing  $\tau$  from  $\tau = 0.4$  to any point in the range  $0.4 < \tau \leq 0.5$  (recall that  $\tau \leq 0.5$  in the model because  $\tau = 0.5$  corresponds to transferring all resources from the altruist to the selfish individual).<sup>38</sup> Basic calculus (details available) shows that, conditional on  $a_t$ , it must be that  $\partial a_{t+1}/\partial \tau > 0$  for  $0.1206 \leq a_t < 1$  when evaluated at  $0.4 < \tau \leq 0.5$ . This, combined with the fact that  $a_{t+1}/a_t > 1$  for  $a_t \geq 0.1109$  (see Proof of Proposition 6), implies that when  $0.1206 \leq a_t < 1$ , it must be that  $\partial a_{t+1}/\partial \tau > 0$  (for a given  $a_t$ ) for all periods. Hence, as the altruistic share of the population expands past 0.1206, a higher value of  $\tau$  will have a positive effect on the growth of the altruistic population. Similarly, basic algebra and calculus (details available) show that, conditional on  $a_t$ , it must be that  $\partial a_{t+1}/\partial \tau < 0$  for  $0 \leq a_t \leq 0.1054$  when evaluated at  $0.4 < \tau \leq 0.5$ . This, combined with the fact that  $a_{t+1}/a_t < 1$  for  $a_t \leq 0.1108$  (see Proof of Proposition 6), implies that when  $a_0 \leq 0.1054$ , increasing  $\tau$  to any value in the range  $0.4 < \tau \leq 0.5$  will decrease  $a_{t+1}/a_t$ . Hence, as the altruistic share of the population shrinks below 0.1054, a higher value of  $\tau$  will exacerbate the decline of the altruistic population. In sum, with the chosen set of parameters, a greater  $\tau$  accelerates the growth of the altruistic population if there are sufficiently many altruists initially ( $a_0 \geq 0.1206$ ), yet accelerates the growth of the selfish population if there are sufficiently few altruists initially ( $a_0 \leq 0.1054$ ).

## NOTES

1. Seminal articles in the biological literature include Hamilton (1964) on kin selection and Trivers (1971) on reciprocal altruism. Maynard Smith (1982) and Dawkins (1989) review the biological literature. Recent work on the evolutionary dynamics of cooperation (e.g., Nowak 2006) is especially relevant to questions of theology: As Coakley (2009) argues, an evolutionary process that supports widespread cooperation, which is what we observe, appears to be consistent with what a benevolent creator would set in motion.
2. Much of the relevant philosophical literature is part of the centuries of debate over the “problem of evil” (e.g., Mackie 1955; Plantinga 1974; Rowe 1979). I will discuss this in more detail later.
3. This point is widely recognized and often used as a starting point for modeling the evolution of altruism. See, for example, Dawkins (1989), Bergstrom (2002), and Nowak (2006).
4. Predation, disease, and natural disasters are all major factors analyzed in the biological literature on the evolution of altruism. Predation is widely recognized as one of the principal factors that creates an evolutionary advantage for cooperative and altruistic behavior among predators and among their prey; see, for example, Trivers (1971), Dawkins (1989), and Wilson (2000). The literature also considers how the evolutionary process is shaped by diseases, natural disasters, and population bottlenecks (i.e., periods with small populations) that result from catastrophic events.
5. The process of natural selection in the model could be genetic, but it may also reflect learning from, or conditioning by, parents. Of course, genetics and conditioning do not have completely independent effects, because genes influence the scope of the effects of conditioning (e.g., Dawkins 1989; Niehoff 1999; Pinker 2002) and the organization of society (e.g., Wilson 2000). See, for example, Dawkins (1989) on how the principles that operate with respect to genes can apply to cultural phenomena, Fudenberg and Levine (1998) on the relationship between evolutionary models and learning in games, Bisin and Verdier (2001) on the differences between cultural transmission mechanisms and evolutionary selection mechanisms, and Nowak and Sigmund (2005) on indirect reciprocity. Sober and Wilson (1998) analyze both evolutionary altruism and psychological altruism.
6. Trivers’s (1971) analysis of reciprocal altruism considers, but does not focus on, the case in which altruists act altruistically toward others without regard to the other individual’s type. In that case of Trivers’ model, altruism does not survive. The altruists in my model have a general, rather than conditional, disposition to benefit others—an altruist in my model will, even in the last period, act altruistically toward a selfish individual whose selfish behavior reduced the altruist’s share of resources in the first period. With either model, altruism survives only if the benefits of altruism go disproportionately to altruists. In Trivers’s model, reciprocity is the mechanism that causes the disproportionate distribution of the benefits of altruism. In my model, the mechanism (explained later) depends on the resource allocation game.
7. On the biological and game-theoretic reasons why altruism and cooperation can persist, see, for example, Hamilton (1964); Trivers (1971); Becker (1976); Maynard Smith (1982); Axelrod (1984); Frank (1987); Dawkins (1989); Bergstrom and Stark (1993); Bergstrom (1995, 2002); Bester and Güth (1998); Eshel, Samuelson, and Shaked (1998); Bisin and Verdier (2001); Nowak (2006). The literature on reciprocity is also extensive; see, for example, Hoffman, McCabe, and Smith (1998); Gintis (2000); Sethi and Somanathan (2001); Gintis, Bowles, Boyd, and Fehr (2003); Bowles and Gintis (2004); Nowak and Sigmund (2005); Friedman and Singh (2009). On the related topic of altruistic punishment, see Fehr and Gächter (2002), and de Quervain et al. (2004). Recently, the scientific literature has identified neurochemical mechanisms underlying the psychological basis of love, violence, and other types of behavior. For reviews, see, for example, Niehoff (1999) on violence, Pedersen (2004) on love, and Pinker (2002) for a broad discussion of human nature. Also see de Waal (2004) and Sapolsky and Share (2004) for related research on primates.
8. See, for example, Eshel (1972), Becker (1976), Eshel et al. (1998). Also see Darwin (1871), especially chapter 5.
9. In short, the resource allocation game leaves altruists wealthier (and, consequently, better able to provide altruistic help) when paired with other altruists than when paired with selfish individuals.
10. See, for example, Trivers (1971), Eshel (1972), Becker (1976), Bergstrom (1995, 2002), Eshel et al. (1998), and Maynard Smith and Harper (2003). Nowak (2006) reviews recent

findings and develops a general framework to summarize the conditions under which various mechanisms will allow the evolution of cooperation.

11. The literature on the problem of evil dates back at least to the ancient Greek philosopher Epicurus. Plantinga (1974) and Tooley (2010) review various forms of the argument from evil, as well as the main types of rebuttals.

12. Assuming asexual reproduction with the number of offspring in proportion to the payoffs from a game follows the replicator dynamic commonly used in the previous literature. The model's focus on stylized neighbors and random pairing also follows earlier work. See, for example, Eshel (1972) on the "neighbor effect" and why low demographic mobility can be critical to the evolution of altruistic behavior.

13. The substantive issue here is that, by providing altruistic help, an individual takes a risk of killing at least some of his or her offspring. The assumption that all or none survive is merely for simplicity.

14. The parameter ranges are set to generate an appropriate range of survival probabilities. Note that with the maximum value of  $m$  (i.e.,  $m = 1$ ) and the altruist's resources at the maximum they would ever be (i.e., equal to 0.5), the probability of survival is 1. When  $m = 0$  and/or the altruistic helper's resources are zero, the probability of survival is zero.

15. Notably, (1) an altruist's probability of being paired with an altruist equals a selfish individual's probability of being paired with an altruist, (2) whether an individual encounters a natural evil is independent of whether his or her neighbor encounters a natural evil and independent of whether his or her neighbor survives such an encounter, (3) when one neighbor assists another during an encounter with a natural evil, each individual's survival is independent of the other's survival, (4) when both individuals encounter a natural evil, the order of the encounters is random and independent of the individuals' types. Note that when both individuals in a pair of neighboring altruists encounter natural evils sequentially, the one who first encounters the natural evil will not incur costs defending his or her neighbor if that neighbor is killed while helping during the first encounter.

16. This implies that (1) with one player in a pair acting selfishly, resources are  $(1 - d)$  times what they would be with no selfish players and (2) with both players in a pair acting selfishly, resources are  $(1 - d)^2$  times what they would be with no selfish players.

17. Note that, because the shares sum to one (i.e.,  $a_{t+1} + s_{t+1} = 1$  and  $a_t + s_t = 1$ ), either of the two equations for population shares is sufficient to describe the changes in the frequency of altruism and selfishness. If the quantity of resources divided among the entire population remains constant over generations, the equations above can be scaled to measure the numbers (not just shares) of different types of individuals in the population.

18. This finding adds to previous work emphasizing the importance of altruists being grouped together in order for altruism to survive (e.g., Becker 1976; Eshel 1972; Eshel et al. 1998).

19. Although Trivers's (1971) model considers only two types of individuals, his discussion of the model's applications emphasizes the fact that individual humans have a mix of altruistic and cheating tendencies, and that the mix depends, in a highly complex way, on interactions with others. Recent work on the nature of indirect reciprocity (e.g., Nowak & Sigmund 2005) and "strong reciprocity" (e.g., Bowles & Gintis 2004; Gintis 2000; Gintis et al. 2003) considers more complex relationships and, specifically, reciprocity of a form that applies to nonkin and is more clearly "benevolent" than that considered by Trivers (1971).

20. In other words, even if selfishness leads to control of the best land, that land may be only marginally better than the land left to the altruists. By contrast, if the principal resource at stake is a particularly productive hunting or fishing location (e.g., an ideal place for catching salmon swimming upstream), an aggressive selfish individual may, when paired with a meek altruist for a neighbor, maintain sole control of the resource. Thus,  $\tau$  will be large.

21. Although the focus of this paper departs from that of the biological literature, it is easy to see how some biological discussions of altruism fit into the framework of the model. Consider the famous example of the symbiotic relationship between cleaner shrimp and the fish they clean (e.g., Trivers 1971, 39–43). To apply the model to the case of these two species, simply assume that all cleaner shrimp are altruistic (this avoids having to consider altruistic and selfish shares of two distinct populations). Placing Trivers's discussion of cleaner shrimp in the context of my model: (1) the assumption that a resource allocation game occurs before predation/disease is reasonable because a fish who has the opportunity to eat a cleaner shrimp may later require that particular shrimp's assistance in defense against parasites, (2)  $\tau$  is high (essentially,

$\tau = 0.5$ ) because eating the cleaner shrimp would reduce to zero the shrimp's resources available for future consumption and altruism, (3)  $p$  is high because parasites attack fish frequently, (4)  $m$  is high because the cleaners effectively protect the fish from parasites, and (5)  $r$  is low (reasonably viewed as zero or negative) because the cleaners eat the parasites. When these conditions hold, nature selects fish that refrain from eating cleaner shrimp. Of course, if  $p$  fell sufficiently (e.g., parasites became very rare) or  $m$  fell sufficiently (e.g., parasites became quite resistant to the work of cleaner shrimp), then there would be less reason for nature to select the combination of altruistic shrimp with fish who refrain from eating them.

22. See, for example, Plantinga (1974) and the literature review by Tooley (2010).

23. Consider the roles of  $p$  and  $\tau$  in a static framework. If one looks merely at a single round of the stage in which natural evils may be encountered, a higher value of  $p$  will cause harm to altruists and selfish individuals alike, with no apparent role in promoting desirable outcomes. Similarly, if one looks merely at one round of the resource allocation stage, a higher value of  $\tau$  will effectively punish altruism and reward selfishness, with no apparent role in promoting desirable outcomes.

24. As Mill (1879, 17) writes: "Yet no one whose opinion deserves a moment's consideration can doubt that most of the great positive evils of the world are in themselves removable, and will, if human affairs continue to improve, be in the end reduced within narrow limits. Poverty, in any sense implying suffering, may be completely extinguished by the wisdom of society, combined with the good sense and providence of individuals. Even that most intractable of enemies, disease, may be indefinitely reduced in dimensions by good physical and moral education, and proper control of noxious influences; while the progress of science holds out a promise for the future of still more direct conquests over this detestable foe."

25. From Darwin to the present, the idea that competition between self-interested individuals can have desirable effects—along the lines famously explained by Smith (1776)—has been highly influential in evolutionary biology. It is important to recognize, however, that the relevant concept of self-interest does not assume that individuals lack concern for others. Indeed, in Smith's view, concern for others played such a critical role in human behavior and welfare that, "The character of every individual, so far as it can affect the happiness of other people, must do so by its disposition either to hurt or to benefit them" (Smith 1790, VI.ii.intro.1).

26. An argument about whether plagues (or other deadly phenomena) could be benevolently motivated is beyond the scope of my model. Note, for example, that even if one were willing to operate on the premise that mass killing could serve a benevolent goal, there would remain the question of why a benevolent creator would rely on a method of killing that caused much suffering. For theological perspectives on how this issue relates to evolution, see Atfield (2006) and Southgate (2008).

27. In his review article, Beebe (2005) summarizes this widely recognized critique of the free will defense: "It seems that, although Plantinga's Free Will Defense may be able to explain why God allows moral evil to occur, it cannot explain why he allows natural evil. If God is going to allow people to be free, it seems plausible to claim that they need to have the capacity to commit crimes and to be immoral. However, it is not clear that human freedom requires the existence of natural evils like deadly viruses and natural disasters. How would my free will be compromised if tomorrow God completely eliminated cancer from the face of the Earth? Do people really need to die from heart disease and flash floods in order for us to have morally significant free will? It is difficult to see that they do. So, the objection goes, even if Plantinga's Free Will Defense explains why God allows moral evil, it does not explain why he allows natural evil." See, for example, Tooley (1980) and Draper (1989) for related critiques, Tooley (2010) for a review of the recent literature, Adams and Sutherland (1989) on horrendous evil, Harrison (1989) on animal suffering, and Southgate (2002, 2008) on evolution-caused suffering (evolutionary evil). Also see Rowe (1979) and the subsequent debate in Wykstra (1984), Alston (1991), Plantinga (1998), and Rowe (1984, 1991, 1998).

28. See, for example, Swinburne (1978), Schlesinger (1983), Moser (1984), and van Inwagen (1991). Beebe (2005) and Tooley (2010) provide relevant reviews.

29. Related to the question of why the existence of free will would require natural evils that cause extreme suffering is the question of why free will would require that humans have the ability to engage in moral evils that cause extreme suffering; see, for example, Plantinga (1974), Tooley (1980), Draper (1989), and the exchange between Draper (2007a, 2007b, 2007c) and Plantinga (2007a, 2007b, 2007c). For the most part, this issue is beyond the scope of my paper—because

my paper has nothing to say about why extreme free will (or indeed any level of free will) would be desirable for its own sake. That said, Proposition 5 does demonstrate how the short-run harm done by moral evil may influence long-run outcomes: When individuals who exhibit moral evil do great harm (as opposed to mild harm) to their altruistic neighbors, that can lead to the more rapid evolutionary success of altruism. Let me repeat, however that this is not to say that moral evil (extreme or otherwise) will ever be desirable.

30. Efforts to defend Nazi-scale evil as an opportunity for virtuous behavior are strained and, of course, contentious. See, for example, Dawkins' (2006, 88–89) account of the remarks made by Swinburne at a public debate that included both men.

31. The model would thus describe a simple Markov process, with the transition probabilities determining the mix of types to which the population would converge over time. One special case of these transition probabilities would be the replicator dynamic employed in the "Theoretical Model" section of this paper.

32. The application of this stylized choice between good and evil need not involve an individual literally choosing a type. Instead, it could pertain to the evolutionary fitness of individuals being conditioned (in the psychological sense) to become more altruistic or more selfish depending on the frequency with which they encounter individuals of those types. Indeed, humans and animals can be conditioned in this manner (e.g., Niehoff 1999; Sapolsky & Share 2004). Also note that developing an extended model would require an appropriate objective function for individuals. More specifically, to assume that individuals make rational choices (in any standard interpretation of rationality) would call for something different from assuming that individuals make decisions that maximize their success in passing along their DNA. On this style of objective function, see, for example, Galor and Moav (2002).

33. This approach would build on the models developed by Fudenberg and Levine (1998), and Friedman and Singh (2009).

34. If the individual has the first encounter and the neighbor the second, the factor is  $m(1 - r + r^2)$  because  $m$  reflects the probability that the individual survives the first encounter and  $(1 - r + r^2)$  reflects the risk of the individual not surviving an effort to help the neighbor. To understand the  $(1 - r + r^2)$  term, note that there is a  $(1 - r)$  probability that the neighbor will survive his or her efforts to help during the first encounter. Thus, there is a probability of  $m(1 - r)$  that the first individual will survive the first encounter and have a live neighbor to help during the second encounter, and there is a probability of  $mr$  that the first individual will survive the first encounter and not have a live neighbor to help during the second encounter. In the former case, the probability of the first individual surviving his or her effort to help the neighbor in the second encounter is  $(1 - r)$ . Thus, the survival probability is  $[m(1 - r)(1 - r) + mr]$ , which simplifies to  $m(1 - r + r^2)$ . If the neighbor has the first encounter and the individual has the second, the factor is  $m^2(1 - r)$  because: (1) the probability of the individual surviving the effort to help the neighbor in the first encounter is  $(1 - r)$ , (2) the probability that the neighbor survives the first encounter is  $m$ , (3) the probability (conditional on help by the neighbor) that the first individual survives the second encounter is  $m$ . Given the assumed statistical independence of the random factors, multiplying these three probabilities together yields  $m^2(1 - r)$ .

35. To see why, note that  $\pi_{aa}$  reaches a minimum when  $r = 1$ , and that minimum value of  $\pi_{aa}$  is 0.25, which always exceeds  $\pi_{sa}$  when  $\tau > 0$ .

36. Note that I set  $m = 1$ ,  $r = 0$ , and  $d = 0$  merely to simplify the arithmetic. These parameter values are not necessary to prove the proposition.

37. The exceptions here are, of course, the uninteresting cases of  $a_t = 0$  (which leads to  $a_{t+1} = 0$ ) and  $a_t = 1$  (which leads to  $a_{t+1} = 1$ ). Thus, the issue to address is, with  $0 < a_t < 1$ , what values of  $a_t$  cause  $a_{t+1}/a_t > 1$  and what values cause  $a_{t+1}/a_t < 1$ .

38. Again, I set  $m = 1$ ,  $r = 0$ , and  $d = 0$  to simplify the arithmetic. These parameter values are not necessary to prove the proposition.

## REFERENCES

- Adams, Marilyn McCord, and Stewart Sutherland. 1989. "Horrendous Evils and the Goodness of God." *Proceedings of the Aristotelian Society* 63:297–323.
- Alston, William P. 1991. "The Inductive Argument from Evil and the Human Cognitive Condition." *Philosophical Perspectives* 5:29–67.



- Ambrose, Stanley H. 1998. "Late Pleistocene Human Population Bottlenecks, Volcanic Winter, and Differentiation of Modern Humans." *Journal of Human Evolution* 34: 623–51.
- . 2003. "Did the Super-Eruption of Toba Cause a Human Population Bottleneck? Reply to Gathorne-Hardy and Harcourt-Smith." *Journal of Human Evolution* 45:231–37.
- Attfield, Robin. 2006. *Creation, Evolution, and Meaning*. Aldershot, UK: Ashgate Publishing Limited.
- Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York: Basic Books.
- Becker, Gary S. 1976. "Altruism, Egoism, and Genetic Fitness: Economics and Sociobiology." *Journal of Economic Literature* 14:817–26.
- Beebe, James R. 2005. "Logical Problem of Evil." *Internet Encyclopedia of Philosophy*. <http://www.iep.utm.edu/evil-log/>.
- Bergstrom, Theodore C. 1995. "On the Evolution of Altruistic Ethical Rules for Siblings." *American Economic Review* 85:58–81.
- . 2002. "Evolution of Social Behavior: Individual and Group Selection." *Journal of Economic Perspectives* 16:67–88.
- Bergstrom, Theodore C., and Oded Stark. 1993. "How Altruism Can Prevail in an Evolutionary Environment." *American Economic Review* (Papers and Proceedings) 83:149–55.
- Bester, Helmut, and Werner Güth. 1998. "Is Altruism Evolutionarily Stable?" *Journal of Economic Behavior & Organization* 34:193–209.
- Bisin, Alberto, and Thierry Verdier. 2001. "The Economics of Cultural Transmission and the Dynamics of Preferences." *Journal of Economic Theory* 97:298–319.
- Boorman, Scott A., and Paul R. Levitt. 1973. "A Frequency-Dependent Natural Selection Model for the Evolution of Social Cooperation Networks." *Proceedings of the National Academy of Sciences of the United States of America* 70:187–89.
- Bowles, Samuel, and Herbert Gintis. 2004. "The Evolution of Strong Reciprocity: Cooperation in Heterogeneous Populations." *Theoretical Population Biology* 65:17–28.
- Coakley, Sarah. 2009. "Evolution and Sacrifice: Cooperation as a Scientific Principle." *Christian Century* 126. <http://www.christiancentury.org/archives/vol126-issue21>.
- Darwin, Charles. 1859. *On the Origin of Species*. Hazelton, PA: Pennsylvania State University (*Electronic Classics Series*, 2001).
- . 1860. Letter to Asa Gray (May 22, 1860). Darwin Correspondence Project Database. Letter no. 2814. <http://www.darwinproject.ac.uk/entry-2814>.
- . 1871. *The Descent of Man, and Selection in Relation to Sex*. London: John Murray.
- Dawkins, Richard. 1989. *The Selfish Gene, New Edition*. Oxford: Oxford University Press.
- . 1995. *River Out of Eden*. New York: Basic Books.
- . 2006. *The God Delusion*. New York: Houghton Mifflin Company.
- de Quervain, Dominique J.-F., Urs Fischbacher, Valerie Treyer, Melanie Schellhammer, Ulrich Schnyder, Alfred Buck, and Ernst Fehr. 2004. "The Neural Basis of Altruistic Punishment." *Science* 305:1254–58.
- de Waal, Frans B. M. 2004. "Peace Lessons from an Unlikely Source." *Public Library of Science Biology* 2:434–36.
- Dowd, Michael. 2006. *Thank God for Evolution*. New York: Penguin Group.
- Draper, Paul. 1989. "Pain and Pleasure: An Evidential Problem for Theists." *Noûs* 23: 331–50.
- . 2007a. "In Defense of Sensible Naturalism." In *God or Blind Nature? Philosophers Debate the Evidence*, ed. Paul Draper. Internet Infidels. <http://www.infidels.org/library/modern/debates/great-debate.html>.
- . 2007b. "Natural Selection and the Problem of Evil." In *God or Blind Nature? Philosophers Debate the Evidence*, ed. Paul Draper. Internet Infidels. <http://www.infidels.org/library/modern/debates/great-debate.html>.
- . 2007c. "On the Plausibility of Naturalism and the Seriousness of the Argument From Evil." In *God or Blind Nature? Philosophers Debate the Evidence*, ed. Paul Draper. Internet Infidels. <http://www.infidels.org/library/modern/debates/great-debate.html>.
- Eshel, Ilan. 1972. "On the Neighbor Effect and the Evolution of Altruistic Traits." *Theoretical Population Biology* 3:258–77.
- Eshel, Ilan, Larry Samuelson, and Avner Shaked. 1998. "Altruists, Egoists, and Hooligans in a Local Interaction Model." *American Economic Review* 88:157–79.

- Fehr, Ernst, and Simon Gächter. 2002. "Altruistic Punishment in Humans." *Nature* 415:137–40.
- Frank, Robert H. 1987. "If *Homo Economicus* Could Choose His Own Utility Function, Would He Want One with a Conscience?" *American Economic Review* 77:593–604.
- Friedman, Daniel, and Nirvikar Singh. 2009. "Equilibrium Vengeance." *Games and Economic Behavior* 66:813–29.
- Fudenberg, Drew, and David K. Levine. 1998. *The Theory of Learning in Games*. Cambridge, MA: The MIT Press.
- Galor, Oded, and Omer Moav. 2002. "Natural Selection and the Origin of Economic Growth." *Quarterly Journal of Economics* 117:1133–91.
- Gintis, Herbert. 2000. "Strong Reciprocity and Human Sociality." *Journal of Theoretical Biology* 206:169–79.
- Gintis, Herbert, Samuel Bowles, Robert Boyd, and Ernst Fehr. 2003. "Explaining Altruistic Behavior in Humans." *Evolution and Human Behavior* 24:153–72.
- Hamilton, William D. 1964. "The Genetical Evolution of Social Behavior." *Journal of Theoretical Biology* 7:1–52.
- Harrison, Peter. 1989. "Theodicy and Animal Pain." *Philosophy* 64:79–92.
- Hoffman, Elizabeth, Kevin A. McCabe, and Vernon L. Smith. 1998. "Behavioral Foundations of Reciprocity: Experimental Economics and Evolutionary Psychology." *Economic Inquiry* 36:335–52.
- Mackie, J. L. 1955. "Evil and Omnipotence." *Mind* 64:200–12.
- Malthus, Thomas. 1798. *An Essay on the Principle of Population*. Electronic Scholarly Publishing Project (1998). <http://www.esp.org/books/malthus/population/malthus.pdf>.
- Maynard Smith, John. 1982. *Evolution and the Theory of Games*. Cambridge, UK: Cambridge University Press.
- Maynard Smith, John, and David Harper. 2003. *Animal Signals*. Oxford: Oxford University Press.
- Mill, John Stuart. 1879. *Utilitarianism*. Hazelton, PA: Pennsylvania State University (*Electronics Classics Series*, 2004).
- Moser, Paul K. 1984. "Natural Evil and the Free Will Defense." *International Journal for Philosophy of Religion* 15:49–56.
- Niehoff, Debra. 1999. *The Biology of Violence*. New York: The Free Press.
- Nowak, Martin A. 2006. "Five Rules for the Evolution of Cooperation." *Science* 314:1560–63.
- Nowak, Martin A., and Karl Sigmund. 2005. "Evolution of Indirect Reciprocity." *Nature* 437:1291–98.
- Pedersen, Cort A. 2004. "How Love Evolved from Sex and Gave Birth to Intelligence and Human Nature." *Journal of Bioeconomics* 6:39–63.
- Pinker, Steven. 2002. *The Blank Slate: The Modern Denial of Human Nature*. New York: Viking Penguin.
- Plantinga, Alvin. 1974. *God, Freedom, and Evil*. New York: Harper and Row.
- . 1998. "Degenerate Evidence and Rowe's New Evidential Argument from Evil." *Noûs* 32:531–44.
- . 2007a. "Against 'Sensible' Naturalism." In *God or Blind Nature? Philosophers Debate the Evidence*, ed. Paul Draper. Internet Infidels. <http://www.infidels.org/library/modern/debates/great-debate.html>.
- . 2007b. "Naturalism vs. Evolution: A Religion/Science Conflict?" In *God or Blind Nature? Philosophers Debate the Evidence*, ed. Paul Draper. Internet Infidels. <http://www.infidels.org/library/modern/debates/great-debate.html>.
- . 2007c. "Objections to Draper's Argument from Evil." In *God or Blind Nature? Philosophers Debate the Evidence*, ed. Paul Draper. Internet Infidels. <http://www.infidels.org/library/modern/debates/great-debate.html>.
- Rampino, Michael R., and Stanley H. Ambrose. 2000. "Volcanic Winter in the Garden of Eden: The Toba Supereruption and the Late Pleistocene Human Population Crash." In *Volcanic Hazards and Disasters in Human Antiquity*, eds. F. W. McCoy and G. Heiken. Boulder, CO: Geological Society of America (Special Paper 345).
- Rowe, William L. 1979. "The Problem of Evil and Some Varieties of Atheism." *American Philosophical Quarterly* 16:335–41.
- . 1984. "Evil and the Theistic Hypothesis: A Response to Wykstra." *International Journal for Philosophy of Religion* 16:95–100.

- . 1991. "Ruminations About Evil." *Philosophical Perspectives* 5:69–88.
- . 1998. "Reply to Plantinga." *Notus* 32:545–52.
- Sapolsky, Robert M., and Lisa J. Share. 2004. "A Pacific Culture among Wild Baboons: Its Emergence and Transmission." *Public Library of Science Biology* 2:534–41.
- Schlesinger, George N. 1983. *Metaphysics*. Totowa, NJ: Barnes & Noble Books.
- Sethi, Rajiv, and E. Somanathan. 2001. "Preference Evolution and Reciprocity." *Journal of Economic Theory* 97:273–97.
- Smith, Adam. 1776. *The Wealth of Nations*, ed. Edwin Cannan, 1976. Chicago, IL: The University of Chicago Press.
- . 1790. *The Theory of Moral Sentiments*, ed. D. D. Raphael and A. L. Macfie, 1976. Oxford: Clarendon Press.
- Sober, Elliot, and David Sloan Wilson. 1998. *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge, MA: Harvard University Press.
- Southgate, Christopher. 2002. "God and Evolutionary Evil: Theodicy in the Light of Darwinism." *Zygon* 37:803–24.
- . 2008. *The Groaning of Creation: God, Evolution, and the Problem of Evil*. Louisville, KY: Westminster John Knox Press.
- Swinburne, Richard. 1978. "Natural Evil." *American Philosophical Quarterly* 15:295–301.
- Tooley, Michael. 1980. "Alvin Plantinga and the Argument from Evil." *Australasian Journal of Philosophy* 58:360–76.
- . 2010. "The Problem of Evil." *The Stanford Encyclopedia of Philosophy* (Spring 2010 Edition), ed. Edward N. Zalta. Stanford, CA: The Metaphysics Research Lab. <http://plato.stanford.edu/archives/spr2010/entries/evil/>.
- Trivers, Robert L. 1971. "The Evolution of Reciprocal Altruism." *Quarterly Review of Biology* 46:35–57.
- van Inwagen, Peter. 1991. "The Problem of Evil, The Problem of Air, and the Problem of Silence." *Philosophical Perspectives* 5:135–65.
- Wilson, Edward O. 2000. *Sociobiology: The New Synthesis, Twenty-Fifth Anniversary Edition*. Cambridge, MA: The Belknap Press of Harvard University Press.
- Wykstra, Stephen J. 1984. "The Humean Obstacle to Evidential Arguments from Suffering: On Avoiding the Evils of 'Appearance.'" *International Journal for Philosophy of Religion* 16:73–93.