Rational Irreligiosity via Cognitive Dissonance

Stephanie Dodd

May 13, 2022

> "The true realist, if he is not a believer, will invariably find within himself the strength and the ability not to believe in miracles either, and if a miracle stands before him as an incontrovertible fact, he will sooner disbelieve his senses than admit that fact."
>
> Fyodor Dostoevsky
> *The Brothers Karamazov*
> [3]

1 Introduction

This paper provides an argument for the economic rationality of irreligiosity. Arguments against the rationality of being religious typically rely on the lack of empirical evidence for the existence of the divine; this strain of logic as a disqualifier of religious behavior does not align with utility maximization under uncertainty in economics. Given this uncertainty exists, it is not necessarily obvious that the outcome of any utility maximization problem forces rational agents to be irreligious. Most models of religious behavior discuss decisions once the individual has chosen to be religious; few papers discuss problems in which agents choose over different levels of faith or religious participation.

Economic research of the rationality of religious participation was densely concentrated in the 1980s and 1990s. Much of this research discussed the allocation of resources to religious participation; a small portion of this work discusses the religion under the rational choice frame. However, the majority of the latter consists of commentary on the difficulty of such analysis or the flaws in extant work [9], [13]. Typically, these essays discuss the difficulty of assigning a well-defined utility function to religion or, relatedly, identifying probabilities people place on the existence of the spiritual.

The philosophical literature on the rationality of religion uniformly treats the decision to be religious as irrational and unscientific; theories of modernization typically

predict the death of religion as society progresses technologically—that the advancement of knowledge and reason will inevitably result in the abandonment of religion [16]. Work spanning the past four decades found that the shift away from religion did not occur (or at least occurred much more slowly than) as predicted by such theories [16], [7], [9]. The data are consistent with an economic rationality argument not typically utilized in the literature, in which it is rational for individuals to choose to be religious. Economic arguments that focused on utility maximization typically concluded that where there is nonzero probability of the existence of "God", a rational agent must choose to be religious.

The reasoning behind this rationality is captured in a motivating formulation for this paper: Pascal's Wager. This game frames religion as an expected utility maximization problem given in a 2×2 payoffs matrix and observes that it is always optimal for an individual to choose religious affiliation, which offers infinite gain in salvation, rather than risk eternal damnation by choosing atheism [15]. While Pascal's Wager has entered economic models of religion that determine an individual's choice to be religious [4], the arguments still seek to justify the choice to *be* religious, again operating under the assumption that the decision to be irreligious is not a utility-maximizing action. The present model does not centrally use Pascal's Wager as the choice set for an individual making a decision about religiosity, as this gives rise to poorly defined expectations over a discontinuous state space. Still, the essence of the Wager, which demands an investigation of an agent's manipulation of otherwise unknown probabilities, motivates the direction of this paper.

Under the frame established in Section 3 of the paper, we establish a faith threshold, such that when agents seek to maximize utility from religious variables, the choice to be irreligious is economically rational for agents that fall below this threshold. Under frame of Pascal's Wager, which uses discontinuous functions for afterlife utility and infinite potential payoffs, it is nearly trivial to observe that it is optimal for a rational agent to choose being religious over being irreligious [8], [14]. However, even in the case where the positive and negative returns to religion in the afterlife are taken as continuous and bounded functions, agents under rational expectations without cognitive dissonance are not economically justified in choosing atheism. We consider the choice between one religion and being irreligious, rather than allowing the agent in the model the option to choose over multiple religions.

Given that it is optimal for a utility-maximizing agent to choose to be religious, we present a model of rational irreligiosity via cognitive dissonance. This process is similar to that of the original cognitive dissonance model [1] and allows for the justification of

atheism under a utility maximization framework. This paper quantifies the threshold at which individuals choose to set their beliefs in the spiritual to zero.

2 Literature Review

2.1 Rational Atheism in the Economics of Religion

The seminal investigation of rationality in religion with uncertainty by Durkin & Greeley [4] discusses the existence of a religious utility function that individuals seek to maximize over a choice variable of faith. The particulars of this function[a], beyond a simple statement of its tendencies with faith, are not discussed. Pascal's Wager enters their investigation critically to structure the utility function. However, the paper does not explain why or how individuals could choose to be irreligious; it simply gives form to a utility function based on Pascal's Wager. The contradiction between the rationality of Pascal's Wager and the seemingly irrational choice to be irreligious remains unresolved in the text.

A response in the same journal to Durkin & Greeley first critiques the assumption that individuals can *choose* to believe, then notes that with uncertainty in payoffs and "stickiness" of religious beliefs, utility maximization is not well-defined, and decisions are not necessarily maximal. Specifically with reference to cognitive dissonance, Montgomery [14] claims that endogeneity of beliefs means that the rational choice in religion is not well-defined. This issue, however, is circumvented with the addition of various reasonable assumptions [5] (which are introduced in Sections 3 and 4), as it relies on the 2×2 game representation of Pascal's Wager.

The previously referenced response to Montgomery by Durkin & Greeley [5] highlights a critical aspect of religious faith, which forms the basis of this investigation of cogni-

[a] Durkin & Greeley present an expected utility maximization problem with two states: one in which the afterlife exists, and one in which it does not. The individual's income is a function of lifetime pecuniary wealth and afterlife quality; they face costs of maintaining faith, including an "insurance premium" for faith, the size of the local religious community, and the amount of faith generated by the participants of that community. The conclusions of their analysis of this utility problem give that the optimal level of faith is increasing in the probability of the afterlife, religious capital, and community faith level; it is decreasing in the size of afterlife income loss. The framing of the problem is the bulk of the paper; all subsequent content is an empirical investigation of how variables such as age, earnings, marriage, etc. affect variables such as mass attendance, perceived proximity to God, etc. This paper uses a number of costs that are similar to theirs.

tive dissonance: the crux of the religious salvation message is that individuals with a low level of faith should (and will) choose to be religious in spite of the uncertainty regarding the afterlife. That is, where faith represents a probability placed on the afterlife, the individual does not necessarily require that this probability is 1 to be religious. However, their subjective probability must be high enough to induce religiosity.

The second problematic aspect discussed by Montgomery in the response is the existence of priors that differ across individuals depending on upbringing. This is an important feature of religious involvement that we initially treat as exogenous. The results of the process of engaging cognitive dissonance depend on the individual's initial beliefs about the existence of the divine, as they determine the threshold at which individuals choose faith or set faith to zero. The endogeneity of beliefs is investigated in Section 5, after the initial solution to the basic decision problem with initially exogenous beliefs is presented.

The investigation most similar to this paper was conducted by Levy & Razin [12] and frames religious participation as a result of gains from social cooperation. The authors posit that in human relations, religion serves as a signal of cooperation. Interactions are presented as a random pairings Prisoner's Dilemma, and they characterize responses in beliefs and religious participation to shocks to society. They propose that individuals have two motivations to choose to be or remain religious: material reasons, in which the individual expects more cooperation from society as a whole if they are religious; and spiritual reasons, in which the individual expects they will cooperate more on the whole as a result of being religious. They find that positive shocks tend to push individuals away from religion; the effects of negative shocks are ambiguous.

The work by Levy & Razin [12] is similar to this investigation in that it is an attempt to explain the economic motivations behind the choice of religious participation. The frame they present, however, is substantially different from that which motivates this paper. Rather than focus on social gains from cooperation, this paper investigates the agent's evaluation of the perceived costs and gains from being religious. The dependence of beliefs on levels of religiosity within a community and societal levels of religious participation is empirically supported [4], [10].

2.2 Cognitive Dissonance

The original cognitive dissonance model proposed by Akerlof & Dickens [1] presented a two-period problem in which workers select into two industries: one that is safe and one that is hazardous. If the individual selects into the hazardous job, they will have the option to purchase safety equipment in the future; if they choose not to do so, they will perceive greater earnings from the job as they do not incur the cost of safety equipment. Given a certain relationship between the cost of fear about having an accident and the costs of safety equipment and accidents, workers in this hazardous industry will choose not to buy safety equipment and engage in cognitive dissonance so that they do not suffer utility losses from incurring the cost of fear.

The application of cognitive dissonance in that setting is similar in many ways to that of the present paper. In this model, agents set subjective probabilities on the existence of the afterlife, analogous to the probability the worker places on experiencing an accident in the workplace. However, this model diverges in the payoff structure from the original cognitive dissonance model. The afterlife benefits that the individual will receive from the being religious are positive and allowed to scale in this model. The decision to be religious is similar to the process of purchasing safety equipment in [1]. The cost of fear utilized in this framework was introduced in the original model, and the cost comparisons used to derive the faith threshold in this model are based on comparisons made over the costs and benefits of working in the hazardous industry in [1].

3 Characterizing the Agent Belief Problem

The first steps in characterizing the problems faced by agents in this model involves defining the values for utilities and probabilities that enter this model. As follows from the preceding section, analysis relating to the existence of god(s) and the afterlife require strong assumptions about the behavior of utility functions.

The original formulation of Pascal's Wager described a 2×2 space with four outcomes based on the individual's choice regarding religion and the true existence of god(s). Other formulations of the Wager give a 3×3 game in which individuals choose between atheism, "God" and "God' (God prime)" (that is, they have the option between atheism, the religion under consideration, and any other religion). This model is most similar to the original 2×2 formulation, in which we consider the simpler choice between atheism

and religiosity for a given religion. However, similar to the arguments presented in [8], the analysis is made more tractable and remains plausible when the benefits of salvation are not infinitely increasing.

We first introduce the structure of costs and beliefs involved in the utility calculation for an agent in this problem. Then, we propose assumptions over these costs to provide structure and tractability to the model. Beliefs in any god and the afterlife are referenced simply as "belief(s) in the afterlife."

3.1 Time Horizon

The framework for the agent's problem is similar to one in which they exist in a three-period world, in which they make the decision to become religious in the second period of the model. The first period is taken as the portion of the agent's life prior to making the choice to be religious. This might represent events such as reaching adulthood and making a choice about one's spirituality, or the aftermath of some positive or negative shock in an individual's life that drives them towards or away from spirituality [12]. The second period is the portion of the agent's life that follows the choice to be religious. This period captures the material gains an agent receives from choosing to be (or not to be) religious. The third period represents the afterlife—this period either exists or does not exist with some probability for all agents. The subjective probability they place on the occurrence of such a period is faith, the crucial choice variable of this model.

At the beginning of the second period, the individual enters with some level of faith and will choose based on that level of faith whether to be maintain that level of religiosity or become irreligious. The various costs of religion and levels of faith will determine the agent's decision. The first period is important in that it allows for the establishment of religious capital, which will be detailed further in Section 5. The two-period lifetime is designed to loosely represent the agent's youth and their transition to adulthood, at which point they make the decision to follow or abandon religious practice.

3.2 Beliefs

Either the afterlife exists, or it does not. That is, the true probability that the afterlife exists is either 0 or 1. However, agents vary their beliefs in the probability of the existence of the afterlife. Thus, an agent in the model will set their belief in the afterlife f to some

value in the range $[0, 1]$. As discussed in the previous part of this section, the individual enters the second phase of their life with a certain level of faith; if that level of faith is appropriately high with respect to the religion in question, they will maintain that level of faith. If it is too low for their religion, they become irreligious and set $f = 0$.

There are two obvious implications of this structure of beliefs. First, there is a narrow window for atheism—exactly $f = 0$. The choice of a zero level of faith depends on the individual's perceptions of costs and their level of faith in the first period. Following from this first observation is that, as noted by [5], the agent does not necessarily have to place subjective probability 1 on the existence of the afterlife. Religion simply requires the acceptance of the uncertainty of the probability via faith. However, below a certain level of faith, it is not economically rational for the agent to choose religion. Instead of incurring the negative costs associated with low but nonzero levels of faith, the agent engages in cognitive dissonance and sets $f = 0$.

3.3 Costs

We divide costs (c) and benefits (h) into two categories: *material* and *spiritual*. Material costs and benefits are incurred in the agent's lifetime and can be both pecuniary and nonpecuniary. Spiritual costs and benefits capture the nonpecuniary utility the individual gains from the various afterlife states as well as the fear associated with believing in an afterlife.

Lifetime pecuniary costs are represented as c_M. These costs are intended to capture faith-related pecuniary costs. The individual might also derive certain benefits from membership in the church [12], [2], which we denote h_M, and $h_M > 0$. We do not force these costs and benefits to scale by faith; conditional on having chosen religion, the material elements are realistically more dependent on wealth and similar factors than on an individual's chosen level of faith [2], [10].

The cost of religious capital accumulation, c_K, is comparable to an insurance cost or to the cost of purchasing safety equipment in [1]. An example of such costs is familiarization with spiritual doctrine; it requires effort to acquire knowledge about religion. Similarly, this captures barriers to entry or conversion for certain religions. This cost is also an entrant in determining the individual's choice to be religious or atheistic. Intuitively, individuals with higher religious exposure in the first period experience lower capital accumulation costs associated with increasing f in the second period [10]. c_K does not include

costs like tithes and offerings or any other pecuniary membership fees. In the first stage of analysis, we assume exogeneity of c_K—in Section 5, we allow c_K to be a variable of choice that determines f. When c_K is a choice variable, individuals gain the ability to make decisions regarding the accumulation of religious capital. Making investments in religious capital push the individual further from the atheism threshold, such that they are more likely to select a non-zero f in the second period of the model. The capital accumulation costs are generally nonpecuniary in nature.

The spiritual costs are h_A, the unit cost of the afterlife (which can be positive or negative) and c_F, the unit cost of fear regarding the afterlife. $f \cdot h_A$ captures the utility an individual receives in the afterlife (zero if the individual chooses $f = 0$), and $f \cdot c_F$ captures the costs an individual receives from uncertainty—the fear of being wrong about the afterlife. At both ends of the range of faith, $c_F = 0$. That is, if an individual is certain that there is no afterlife, they do not experience costs due to fear; if the individual is certain that there is an afterlife, they do not experience costs due to fear. The costs of fear should be increasing towards the atheism threshold and decreasing away from it; the agent will experience this fear over the first period and will choose a level of faith in the second period to minimize this cost. This assumption regarding the cost of fear is formalized in Section 4.

Both the afterlife benefits and the costs of fear are assumed to be a single value for all individuals; the rewards in the afterlife scale (as is proposed above) by varying the level of faith the individual chooses, assuming they are above the atheism threshold.

3.4 Preliminary Assumptions

Assumption A: *For all $h_A > 0$ and h_M, $f \cdot h_A > h_M$. For all $h_A < 0$, $f \cdot h_A < h_M$.*

This assumption gives that the utility gained from spiritual rewards always exceeds the utility gained from material rewards[b]. Thus, material rewards alone are not sufficient to induce the agent to be religious. Note that we include all church-related benefits captured over the individuals lifetime, including goods such as community fellowship, member networking advantages, and quotidian routine, in the category of material goods because of their substitutability. The spiritual component is unique in that it does not have substitutes—where one might find community in book club, networking at poker night, and routine in daily exercise, one cannot acquire a high probability (or, if beliefs are set

[b]This follows an assumption proposed in [8].

appropriately, sure-gamble) afterlife end state anywhere else.

Assumption B: *Lifetime material utility h_M is bounded, continuous, weakly increasing, and greater than or equal to zero.*[c]

Individuals reach a saturation point in material benefits from religion, at which the accumulation of more material rewards from religious participation cease to contribute more to the utility that the individual experiences. We do not force this function to scale with f, as we expect this to be more associated with participation in religious activities rather than the pure level of faith. As will be discussed in Section 4, for example, trends for participation in religious activities are related to wealth. Thus, we allow such benefits to enter the agent's utility function, but do not expect it to scale critically with level of faith.

4 Irreligiosity & Cognitive Dissonance

In this section, we compute the atheism threshold based on the belief function formulation presented in the preceding section.

Assuming a certain level of faith f, an individual experiences total costs of religion:

$$fh_A - fc_F - c_K + h_M - c_M$$

Recall that the value of h_A can be positive or negative. However, as we will observe later, an individual should always choose $f = 0$ if they expect that their h_A will take on a negative value.

The function for faith and the threshold value of faith depends on the costs associated with religion and prior beliefs. There are a number of conditions under which the agent rationally chooses atheism; the parameters that alter the threshold level are the costs and prior beliefs. Again, we are assuming that individuals receive some endowment of religious capital, so that religious capital does not scale according to the level of f chosen. The goal of the agent is to maximize their total utility from pecuniary and nonpecuniary sources given their subjective probability assessment of the existence of an afterlife.

[c] As with Assumption A, this follows an assumption proposed in [8]. It is applied in [2].

4.1 Faith Threshold

We can generate the basic threshold for atheism considering the expected costs of being irreligious or religious where the true probability of the afterlife is unknown but binary. Thus, the threshold becomes simply a function of the costs associated with religion and is derived with the same rationale as used in [1]. This threshold requires that the costs of fear to the individual must exceed the costs of being religious if they are to choose religion. That is,

$$fc_F > c_K + c_M - h_M - fh_A$$

where the left hand side represents the total cost of fear, while the right hand side gives the net costs the individual faces in being religious. Rearranging the inequality, we find the atheism threshold:

$$\underline{f} = \frac{c_K + c_M - h_M}{c_F + h_A}$$

so that the individual will be religious for levels of faith that exceed the term on the right side of the inequality. Where there are higher costs associated with religion, there is a higher threshold of faith necessary for the individual to achieve for them to not choose atheism.

4.2 Assumptions Involving the Threshold

Assumption C: *Afterlife utility h_A is bounded, continuous, and weakly increasing.*

Afterlife utility h_A will be some function of the form:

$$h_A = \begin{cases} h_A > 0 & \text{if } f > \underline{f} \\ h_A \leq 0 & \text{if } f \leq \underline{f} \end{cases}$$

such that the agent receives negative afterlife utility if they are below the critical threshold and positive if they are above it. This captures the positive returns to being religious and negative returns to being irreligious if the afterlife does exist.

Assumption D: *Consider two levels of faith, f^* and f' placed symmetrically on opposite sides of $\underline{f}$ so that $|f^* - \underline{f}| = |f' - \underline{f}|$ and $f^* > \underline{f} > f'$. Then $f' \cdot h_A = -f^* \cdot h_A$. That is, payoffs are symmetric about the critical threshold*[d].

[d] These give the payoffs fh_A for any value of f, without invoking cognitive dissonance.

This frame of utility to be gained from religion is unique in that it allows for a continuous form of the afterlife function. Similar to the scaling costs of fear originally utilized in [1] (and as is applied to c_F in this model), h_A is assumed to scale with with f, such that higher levels of faith are associated with greater afterlife rewards. Note that while under the original cognitive dissonance model, agents set beliefs at exactly zero or at a specific threshold value, agents in this model can hold non-zero beliefs which can scale. Assumption D captures the variation across individuals in the degree of their beliefs for those who set their beliefs above the threshold for atheism. This is not implausible for many religions, where reincarnation or spiritual afterlives have gradations of utility [6]. In the context of Catholicism, for example, beliefs in Purgatory or different levels of Hell (or disbelief in such things) will generate different levels and distortions of beliefs. However, at the highest level of belief (in the Christian context, perhaps in an infinitely and equally rewarding Heaven), it is reasonable to impose an upper limit on gains from belief for tractability of analysis. Assumptions about continuity and boundedness of such functions are not unheard of in the literature (see [2] and [8], respectively).

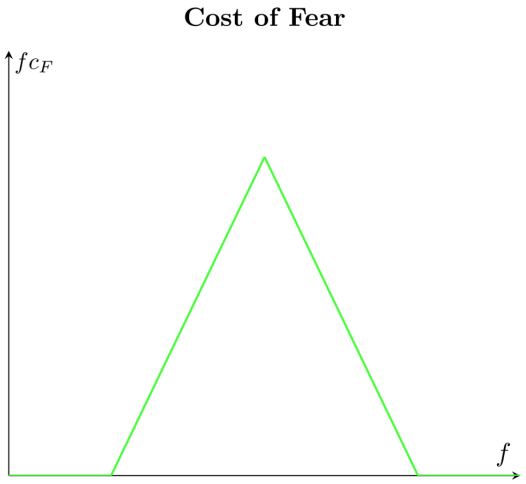
As discussed in the Section 2, cognitive dissonance in the economic context primarily deals with agents' choices over their beliefs. This assumes that agents have preferences over their beliefs and exert some level of control over the level of the beliefs that they set. This enters critically into the analysis of this paper: the investigation is primarily one of the process of setting beliefs. Atheists choose to set their beliefs in the afterlife to zero, while non-atheists are above a previously characterized threshold of perceived costs and benefits, such that they set their beliefs in the afterlife to some value greater than zero.

Following the notation in Assumption D and adding Assumption C means that $f^* h_A > f' h_A$. This assumption also means that if the individual's subjective probability assessment of the existence of the divine is too low, they will receive negative afterlife utility. This is not in contradiction with the assertion that individuals do not need to have $f = 1$ to achieve the positive-utility afterlife state—this merely requires that the individual chooses an f that is high enough to induce religiosity in pursuit of that afterlife state.

The cost of fear is taken to scale differently with faith. Consider the function that internalizes the scaling of fear costs with faith: $C_F(f) = f \cdot c_F$. The cost of fear is a unit cost; the behavior of C_F should reflect variation in the agent's level of certainty of the afterlife with different values of f[e].

[e]This cost does not scale monotonically across the [0,1] range, as it would in [1]. If an agent chooses $f = 0$, they should experience zero cost of fear. Similarly, if an agent chooses $f = 1$, they should experience

Assumption E: $\frac{dC_F}{df} > 0$ *if* $f \leq \underline{f}$, *and* $\frac{dC_F}{df} < 0$ *if* $f > \underline{f}$. *The cost of fear is increasing as* $f \to \underline{f}$ *and decreasing as* $f \to 0$ *and* $f \to 1$. *Furthermore,* $C_F(0) = 0 = C_F(1)$.



Assumption E gives that the costs of fear increase with the uncertainty that the individual experiences about the afterlife. If the agent places very low (or very high) probability on the existence of an afterlife, the costs of fear that they experience are lower than those faced by individuals closer to the critical threshold. The uncertainty experienced at $\underline{f}$ is greater than that experienced at values close to 0 or 1.

4.3 Cognitive Dissonance & Atheism

To find the point at which the agent engages in cognitive dissonance, we must consider the faith threshold in the context of the cost to the individual for making the wrong decision [1]. The cost to the individual for making the wrong choice in this case is the difference between the costs they would incur from being irreligious and the costs they would incur from participation. If the individual chooses not to believe in the afterlife, they experience a cost (utility loss of) $f \cdot h_A$. Considering a symmetric reflection of faith over the threshold, if the agent had chosen to be religious, they would have incurred a utility loss of $c_K + c_M - h_M - fh_A$. We did not write the cost of fear in this equations as it is symmetric across the threshold and will make identical contributions to these equations. The overall cost of making the wrong decision for the agent is then $fh_A - (c_K + c_M - h_M - fh_A)$.

zero cost of fear.

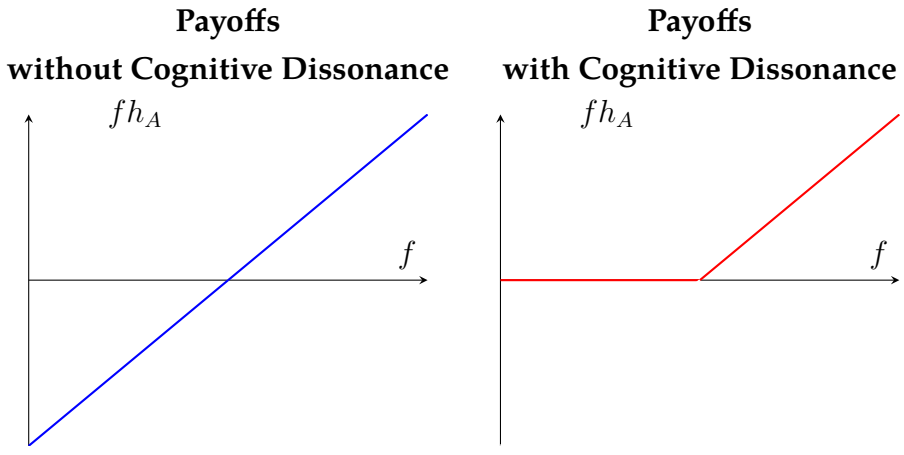
The behavior of f is interesting in the cases where individuals perceive large costs of accumulation of religious capital and pecuniary membership relative to lifetime material benefits. Recall that the level of faith is equivalent to the probability that the individual places on the existence of some afterlife. As discussed in the context of Assumption D, this means that when the individual places low enough probability on the existence of the divine, the individual will receive negative returns in the afterlife as they will not choose to be religious. Thus arises the conflict between rational expectations and irreligiosity: it is not economically rational for the agent to choose any level of f lower than the threshold derived above. If the decision of the agent to be irreligious is to be jusitifiably rational, this requires that they set their level of faith to zero. A comparison of the cost of making the wrong decision and the cost of being religious yields the following proposition:

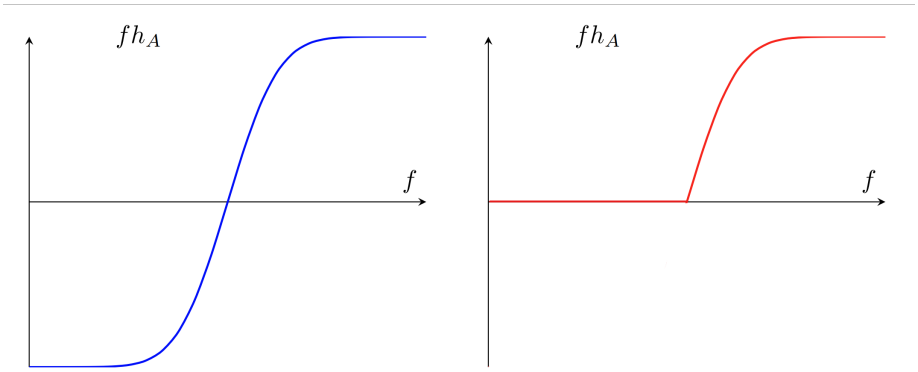
Proposition 1: *If the agent experiences costs such that their level of faith is below threshold utility and observes*

$$h_A < 2c_F,$$

then they will choose to set $f = 0$ when they make their decision in period 2. If the opposite is true, they will maintain their original level of faith with $f > \underline{f}$.

The computations for all propositions are located in the appendices. The equation above states that if the unit benefit of the afterlife is less than twice the unit cost of fear, and the individual observes that their faith is below the necessary threshold, they will choose to be irreligious and set $f = 0$. Visualizations of the change in an agent's afterlife payoffs when cognitive dissonance is applied as in Proposition 1 are presented below.





As is discussed in Section 3 and the preceding parts of Section 4, the utility that the individual will gain over their lifetime and in the afterlife, supposing it exists, depends on their level of faith. The inequality above gives the relationship for which the costs of making the wrong choice fail to exceed the costs of fear. For individuals entering the second period with levels of faith not high enough for the given religion, if the relationship between the cost of fear and afterlife benefits is as stated in Proposition 1, the individual will receive negative returns to religious participation and maintaining their previous level of faith, at which point they will engage in cognitive dissonance and set $f = 0$.

4.4 Wealth Differences at the Threshold

The inclusion of lifetime pecuniary and nonpecuniary costs also captures an interesting, empirically observable effect. As proposed theoretically by Iannaccone [10], wealthier religious individuals have the opportunity to offer monetary contributions to their religion. Critically, the condition derived in Proposition 1 gives that the costs that determine the individual's level of faith are the cost of fear and the utility the individual expects to receive in the afterlife. This is, of course, a result of the framing of the problem, where the other costs are independent of the level of faith.

Thus, it is worthwhile to make observations about the impacts of varying levels of h_M and c_M when the level of faith chosen by the two agents is identical. Different values for these costs will have different implications for the value of $\underline{f}$. Variable costs of participation and monetary value on time are common factors in models of religious participation—the human capital approach [10] contends (and finds some empirical evidence) that individ-

uals with a higher opportunity cost of time will select to participate in religion more via money-intensive practices, while those with a low opportunity cost of time will engage in time-intensive practices. Similarly, in modeling a household with uneven income earnings across two partners, [2] also find that the partner with a lower market wage[f] will spend more time participating in religious activities. The easiest frame to consider in which these costs and benefits would differ is across different wealth strata.

We consider two types of agents with different wealth; the agent with greater wealth is taken to place higher monetary value on time. Consistent with the predictions and empirical evidence provided in [2] and [10], this implies that the wealthier agent experiences higher pecuniary costs of participation c_M and lower pecuniary and nonpecuniary lifetime benefits h_M than the poorer agent. Suppose that we know that both agents have chosen to be religious (i.e. that Proposition 1 was satisfied such that neither agent has engaged in cognitive dissonance and set beliefs to $f = 0$). Then, we can compare across the different faith thresholds that they experience due to this wealth effect on participation costs. This comparison is contained in the following proposition:

Proposition 2: *Suppose there are two types of agents, poor (p) and rich (r), and suppose that rich agents have greater wealth and a higher opportunity cost of time than poor agents. Then,*

$$\frac{c_K + c_M^r - h_M^r}{c_F + h_A} > \frac{c_K + c_M^p - h_M^p}{c_F + h_A} \quad \text{so} \quad \underline{f}^r > \underline{f}^p$$

That is, rich agents have a higher faith threshold than poor agents.

Using the insights of the literature cited above, we find that individuals with greater wealth have a higher faith threshold than individuals with less wealth. Again, this reflects observable differences in participation in religious practice and monetary contributions to religion. One might thus expect to observe a higher c_M and a lower h_M for wealthy agents relative to poorer agents, as rich agents contribute more monetarily to their religion but receive fewer communal benefits given their lower participation. This is reflected in the faith threshold, where rich agents have a higher range of values of f over which they choose to be irreligious.

A similar logic might be applied to other sociodemographic groups. Differences across the form of religious participation will manifest in the material costs and benefits

[f] The empirical tests they run are limited in that the data used does not have disaggregated wages, so the effects of wage rates are not clear. However, they find strong non-labor income (i.e. wealth) effects on participation.

associated with religion. Similar empirical effects are observed in the aforementioned papers with income; one might contend that education has different implications for c_M and h_M, producing distinct faith thresholds in the same manner as wealth. Thus, under this frame we can establish that the faith threshold varies across agents of different types.

4.5 High Capital Costs

4.5.1 Switching Religions

While this paper is primarily looking at the individual's decision to either maintain their faith under their current religion or become irreligious, preliminary observations about switching religions can also be made using the faith threshold. High costs of accumulating religious capital are incurred by any agent that seeks to switch religions. This reflects observable trends in which individuals raised in religious households that remain religious typically remain within their religion or choose religions highly similar to their own [10]. One would expect the faith threshold to be high for an individual switching religions; this is reflected easily in the cost of accumulating religious capital.

This subsection remains a remark more than a formal investigation as there are further complications with the process of switching religions that make comparisons of the faith threshold difficult. A change in religion could yield ambiguous effects in the material costs and benefits side; furthermore, the h_A and c_F functions are likely to change between religions that are significantly different. These differences (especially across the spiritual costs and benefits side) require greater attention and more significant alterations of the framework. This is expanded upon briefly in Section 6.

4.5.2 Cults & Collectives

Another phenomenon to which this model of a faith threshold is applicable is cult behavior, which offers particularly high rewards h_M and costs for individuals c_M [11]. In lieu of the previous subsection, we can consider extremist wings of common religions in this frame and circumvent issues of different functions for afterlife payoff and the costs of fear. This section uses a similar process as the preceding section of the effects of wealth on the faith threshold. Note particularly in this context that cults fit this model where there are extremely high material costs in addition to higher costs of accumulating religious capital. Further exploration of endogeneity of faith with religious capital is given in

Section 5. Without introducing a greater role for religious capital, we can still make the unsurprising observation that, where the religion in consideration is a cult, even where the effects of c_M and h_M appear ambiguous, the faith threshold is still expected to be quite high as a result of the high cost of accumulating religious capital in a cult.

5 Endogenizing Faith with Religious Capital Accumulation

Previously we assumed that faith was determined exogenously; now, we relax this assumption to allow individuals to participate in behaviors that increase their faith. This section allows for the entrance of human capital theories of religion [10] into the cognitive dissonance model, by which engaging in religious practices affects the agent's choice to be religious [5]. In earlier sections we assumed that individuals are endowed with some level of faith and religious capital (associated with c_K). Now, we allow for individuals to make choices in the first period that alter their second-period costs of religious capital accumulation. The agent's level of faith is now made to scale with religious capital; decreasing the cost of the accumulation of religious capital in the second period can then increase the individual's level of faith.

First, we redefine faith to make it a function of religious capital. The cost of acquiring religious capital is similar to the cost of fear in its intangibility. This involves practices such as reading spiritual doctrine, meditation or prayer, and other practices that yield benefits in terms of faith. The other material costs and benefits describe factors that one might derive indirectly *from* one's spirituality, while religious capital costs relate directly to the level of one's spirituality. We now assume that c_K is the costs incurred by choice of the individual. c_K will still enter the faith threshold has before, such that an individual with levels of faith too low would have to incur c_K too high for religiosity to be optimal.

In the original problem, the time period division was not highly important to the agent's decision process save as a means of temporally structuring the choice to be religious. To endogenize faith, we allow for the individual to accumulate religious capital in the first period of the model, such that their second-period costs of religious capital and threshold value for faith are lower than assumed throughout Section 4.

In the first-period of the model, the individual has the option to engage in the accumulation of religious capital. This might realistically occur where parents invest in their children's religious upbringing (focusing on time-intensive ways such as religious services

and related activities). We first specify intertemporal relationships for faith and religious capital of the form $f_t(f_{t-1}, c_{K,t-1})$ and for religious capital of the form $c_{K,t} = k(\delta)c_{K,t-1}$.

Assumption F:

(*a*) $f_t(f_{t-1}, c_{K,t-1}) = f_{t-1} \cdot g(c_{K,t-1}) \in [0,1]$ *is semi-continuous and weakly increasing.* $g \in \mathcal{C}^1[a,b]$ *and* g' *is strictly increasing on* $[a,b]$. $\frac{\partial f_t}{\partial f_{t-1}} \geq 0$. *With respect to religious capital,* $\frac{\partial f_t}{\partial c_{K,t-1}} \geq 0$ *and* $\frac{\partial^2 f_t}{\partial c_{K,t-1}^2} \leq 0$.

(*b*) $c_{K,t} = k(\delta)c_{K,t-1} \geq 0$ *and* $0 < \delta(c_{K,t-1}) \leq 1$ *with* $\frac{\partial \delta}{\partial c_{K,t-1}} \geq 0$.

The first part of this assumption gives us that faith is increasing in the costs of religious capital that the agent chooses to incur, but returns to incurring more religious capital diminish as c_K grows large. The convexity or concavity of the relationship between f_t and f_{t-1} does not need to specified as it could take on any form and as f_{t-1} has the same range as f_t.

The second part of this assumption gives us that incurring religious capital in period $t-1$ reduces the cost of religious capital for the agent in period t. $c_{K,t}$ is a function of some scaling parameter δ and capital in the preceding period. The value of δ depends on the value of $c_{K,t-1}$, so that larger investments in religious capital in period $t-1$ are associated with larger reductions in the cost of religious capital in period t. Thus, by investing in $c_{K,1}$, the individual increases their level of faith and scales down the level of religious capital investments that they will need to incur in the future.

5.1 General Form

Suppose the individual is initially born or instilled with some exogenous predisposition for religiosity. We call this innate level of faith `f`. The individual will choose c_K for the first period depending on their expectations for religiosity and utility overall. We let the intertemporal faith function and intertemporal religious capital function be represented as

$$f_t(f_{t-1}, c_{K,t-1}) = f_{t-1} \cdot g(c_{K,t-1})$$

and

$$c_{K,t}(\delta, c_{K,t-1}) = k(\delta) \cdot (c_{K,t-1})$$

respectively. The agent only sets their levels of faith via investments in religious capital. Overall, the agent observes utility:

$$U(c_{K,1}) = \mathtt{f}g(c_{K,1})(h_A - c_F) - k(\delta)c_{K,1} + \sum_{t}^{2} h_{M,t} - c_{M,t}$$

Differentiating and solving for the optimal level of c_K as a function of the individual's initial level of faith yields the following proposition:

Proposition 3: *The optimal level of investment in religious capital as a function of the agent's initial level of faith is*

$$c_K = (g')^{-1}\left(\frac{k(\delta)}{\mathtt{f}(h_A - c_F)}\right)$$

Thus, the optimal level of capital investment is increasing in $k(\delta)$ and decreasing in $\mathtt{f}(h_A - c_F)$.

When this expression is substituted into the equation for $\underline{f}$, we find the expression for the endogenous faith threshold, $\underline{\underline{f}}$, where faith depends on the original level of faith via the capital accumulation term:

$$\underline{\underline{f}} > \frac{(g')^{-1}\left(\frac{k(\delta)}{\mathtt{f}(h_A - c_F)}\right) + c_M - h_M}{c_F + h_A}$$

Thus, we find that the optimal level of investment in religious capital is, unsurprisingly, increasing in the $k(\delta)$, as $k(\delta)$ is structured such that larger values of δ yield lower values of $c_{K,2}$. As the initial level of faith increases, the agent requires less of a capital investment to move them beyond the faith threshold. As the difference between afterlife rewards and the cost of fear increases, the individual requires less in religious capital investment to achieve an optimal level of faith; this mechanism is similar to that of $\mathtt{f}$.

5.2 Specific Case

Consider the case where the agent with innate faith level $\mathtt{f}$ has intertemporal faith relationship $f_t = f_{t-1}\left(1 - \frac{1}{c_{K,t-1}}\right)$ and intertemporal religious capital accumulation function $c_{K,t} = \frac{1}{\delta}c_{K,t-1}$. Upon making the decision to be religious or irreligious at the beginning of the second period, the individual observes that they will receive utility $h_{M,t} - c_{M,t} + f_t(f_{t-1}, c_{K,t-1})h_A - f_t(f_{t-1}, c_{K,t-1})c_F = h_{M,1} - c_{M,1} + f_0(1 - \frac{1}{c_{K,1}})(h_A - c_F)$. They incur additional capital costs in the second period of $c_{K,2}$, pecuniary religious costs $c_{M,2}$, and

benefit of $h_{M,2}$. Overall they observe utility:

$$U(c_{K,1}) = \mathtt{f}\Big(1 - \frac{1}{c_{K,1}}\Big)(h_A - c_F) - \frac{1}{\delta}c_{K,1} + \sum_{t}^{2} h_{M,t} - c_{M,t}$$

As in the general case, we differentiate U and solve for the optimal level of $c_{K,1}$, which yields:

$$c_{K,1} = \sqrt{\delta \mathtt{f}(h_A - c_F)}$$

so that the threshold for faith becomes:

$$\underline{f} > \frac{\sqrt{\delta \mathtt{f}(h_A - c_F)} + c_M - h_M}{c_F + h_A}$$

6 Conclusion

In this paper, we have identified a threshold for religious participation that depends on various material and spiritual costs incurred over the agent's lifetime and afterlife periods. This first required a characterization of the associated costs and benefits and the relationship between these costs and the level of faith an individual chooses to hold. Where an individual is guaranteed to receive severe negative afterlife penalties for choosing to be irreligious (i.e. have faith of insufficient magnitude) when the afterlife does exist, the agent can rationally choose to be irreligious by engaging in cognitive dissonance, such that they expect to receive nothing in the afterlife rather than the large negative cost. The conditions for this threshold are presented in Proposition 1 and its associated calculations. Proposition 2 described the faith threshold in a comparative setting, which is applicable to a variety of observable empirical trends in religious participation data. Finally, we derive Proposition 3, which characterizes a capital accumulation function, where the level of faith at which the individual makes their decision regarding religion depends on their earlier choices to accumulate religious capital. With this capital accumulation function, we derive a faith threshold that depends on the costs outlined in Section 3 and the individual's assumed innate level of faith.

The first portion of this paper is dedicated to creating a framework for considering problems of religion that allows for the consideration of economically rational atheism. Models of religiosity that investigate the individual's choice to be religious typically utilize variants of Pascal's Wager to characterize the problem, in which it is rarely (if ever)

possible to justify atheism due to the formulation of the decision problem. The paper closest in studying the choice to participate focuses on religion as a cooperation signal in a game theoretic frame, in which critical intangible costs, such as those discussed in Section 3, are absent from the model.

We discuss three potential limitations and extensions of this paper. First, the model is weak in its ability to explain the choice of an atheistic individual to become religious. While this differs from the aim of the paper, which primarily seeks to justify the decision for the individual to become atheistic from some initial level of $f > 0$; nonetheless, the inclusion of some mechanism by which irreligious individuals can be explained to become religious would serve the model. For example, one might expand the model to include such behaviors by involving shocks (events perceived as miracles or tragedies), which could drive an individual towards or away from religion. Where such shocks would enter is not immediately clear—these might affect the h_M terms for a religious individual, or perhaps alter the structure of c_F; on the other hand, these shocks might constitute an entirely separate term to determine faith, which would then allow for atheists (with $f = 0$ initially) to become religious.

A second extension of this model would involve the addition of risk preferences and discounting. This is an element that would enter in the utility problem, but this would require further consideration of the cost of fear and of afterlife benefits, which might differ prior to entering the utility formulation. Discounting would be an interesting addition to model with a larger number of time periods and uncertainty over the timing of the final period (i.e. the agent's death and beginning of the afterlife period, such as in [2]), perhaps where the individual faces on multiple occasions a choice to be religious or irreligious over the lifetime periods.

Finally, the application of the model to a religious competition framework is an appealing and challenging extension. One obvious obstacle is the difficulty of comparing across afterlife costs and lifetime nonpecuniary benefits of religion. While it might be easier to consider the costs of accumulating capital or pecuniary religious contributions, comparing across the severity of afterlife punishments, costs of fear, and the lifetime gains across different religions is no simple task. This model benefits in the formulation of h_A from considering an agent's choice to remain within or abandon a given religion. Comparing across related religions, such as across different denominations of Protestantism, is easier than across religions with highly different roots, such as Shintoism and Catholicism. Even the comparison between Catholicism and Protestantism is difficult in that the

afterlife functions are quite dissimilar.

Overall, the application of such a threshold for religious participation can be applied to a variety of theoretical and empirical questions. Improvements on the threshold formulation, such as those listed in the preceding paragraphs, would strengthen and diversify its applicability. Nonetheless, work in the economics of religion can utilize such a threshold that gives conditions under which agents can choose to be irreligious, even where religions that threaten large, negative costs for such a decision.

Appendix A Glossary of Terms

f	individual's chosen level of faith
$\underline{f}$	faith threshold
$\underline{\underline{f}}$	endogenous faith threshold
$\mathtt{f}$	individual's innate level of faith
c_M	lifetime pecuniary costs of religion
h_M	lifetime pecuniary and nonpecuniary benefits of religion
c_K	cost of accumulating religious capital
h_A	afterlife benefits
c_F	cost of fear regarding the afterlife
c^i, h^i	cost or benefit to poor $(i = p)$ or rich $(i = r)$ agent
$\delta(c_K)$	religious capital discounting term
$k(\delta)$	function on δ to scale capital accumulation
$g(c_K)$	function characterizing relationship between capital and faith level

Appendix B Proposition 1 Algebra

On the left-hand side:

$$\frac{h_A(c_K + c_M - h_M)}{c_F + h_A} - (c_K + c_M - h_M - h_A)$$
$$\frac{h_A(c_K + c_M - h_M)}{c_F + h_A} - (c_K + c_M - h_M) + \frac{h_A(c_K + c_M - h_M)}{c_F + h_A}$$
$$\frac{2h_A(c_K + c_M - h_M)}{c_F + h_A} - (c_K + c_M - h_M)$$
$$\frac{h_A(c_K + c_M - h_M)}{c_F + h_A} - \frac{(c_F + h_A)(c_K + c_M - h_M)}{c_F + h_A}$$

$$\frac{h_A - c_F}{h_A + c_F}(c_K + c_M - h_M)$$

Rearranging around the inequality:

$$\frac{h_A - c_F}{h_A + c_F}(c_K + c_M - h_M) < \frac{c_F(c_K + c_M - h_M)}{c_F + h_A}$$
$$(h_A - c_F)(c_K + c_M - h_M) < c_F(c_K + c_M - h_M)$$
$$h_A < 2c_F$$

Appendix C Proposition 2 Algebra

Relative to the original threshold, we are observing changes in the material cost terms for each agent type.

$$\frac{c_K + c_M^r - h_M^r}{c_F + h_A} > \frac{c_K + c_M^p - h_M^p}{c_F + h_A}$$
$$c_M^r - h_M^r > c_M^p - h_M^p$$

Here we simply have that because $c_M^r > c_M^p$ and $h_M^r < h_M^p$, the inequality above holds. All other costs and benefits experienced by the agents are the same, as they have chosen the same level of faith. Thus, the threshold for rich agents is higher than the threshold for poor agents.

Appendix D Proposition 3 Calculation

Proposition 3: *The optimal level of investment in religious capital as a function of the agent's initial level of faith is*

$$c_K = (g')^{-1}\left(\frac{k(\delta)}{\mathtt{f}(h_A - c_F)}\right)$$

Thus, the optimal level of capital investment is increasing in $k(\delta)$ and decreasing in $\mathtt{f}(h_A - c_F)$.

Computation. We differentiate the utility function and solve for the optimal level of $c_{K,1}$:

$$
\begin{aligned}
U(c_{K,1}) &= \mathtt{f} g(c_{K,1})(h_A - c_F) - k(\delta)c_{K,1} + \sum_{t}^{2} h_{M,t} - c_{M,t} \\
\frac{\partial U}{\partial c_{K,1}} &= \mathtt{f} \cdot (h_A - c_F) \cdot g'(c_{K,1}) - k(\delta) = 0 \\
g'(c_{K,1}) &= \frac{k(\delta)}{\mathtt{f} \cdot (h_A - c_F)} \\
c_{K,1} &= (g')^{-1}\left(\frac{k(\delta)}{\mathtt{f} \cdot (h_A - c_F)}\right)
\end{aligned}
$$

Proof. From Assumption F(a), we have that $g \in \mathcal{C}^1[a,b]$ with $a < b$, and g' is strictly increasing on $[a,b]$. We want to show that $(g')^{-1}$ is strictly increasing on $[a,b]$.

LEMMA: *Because g' strictly increasing on $[a,b]$, it has inverse function $(g')^{-1}$ exists.*

Proof. This simply follows from the fact that because g' is strictly increasing, it is *one-to-one*. Thus, $(g')^{-1}$ exists. □

LEMMA: *Because g' is strictly increasing on $[a,b]$, $(g')^{-1}$ is also strictly increasing.*

Proof. Suppose for g' as described above for $x < y$. Let $a = (g')^{-1}(x)$ and $b = (g')^{-1}(y)$; then, we have that $g'(a) = x$ and $g'(b) = y$. Suppose now that $b \leq a$. Then, $g'(b) \leq g'(a) \implies y \leq x$, as g' is strictly increasing. This is a contradiction. We must have that $a < b \implies (g')^{-1}(x) < (g')^{-1}(y)\ \forall\ x, y$ s.t. $x < y$. Thus, $(g')^{-1}$ is also a strictly increasing function. □

Via the two lemmas given above, we have that the function $(g')^{-1}$ exists and is strictly increasing over the interval $[a,b]$. Thus, we have that c_K is increasing in $k(\delta)$ and decreasing in $\mathtt{f}(h_A - c_F)$. □

Example

We differentiate U and solve for the optimal level of $c_{K,1}$:

$$
\begin{aligned}
U(c_{K,1}) &= \mathtt{f}\Big(1 - \frac{1}{c_{K,1}}\Big)(h_A - c_F) - \frac{1}{\delta}c_{K,1} + \sum_{t}^{2} h_{M,t} - c_{M,t} \\
\frac{\partial U(c_{K,1})}{\partial c_{K,1}} &= -\frac{\mathtt{f}}{c_{K,1}}(h_A - c_F) - \frac{1}{\delta} = 0 \\
\frac{1}{\delta} &= \frac{\mathtt{f}}{c_{K,1}^2}(h_A - c_F) \\
c_{K,1}^2 &= \delta \mathtt{f}(h_A - c_F)
\end{aligned}
$$

So the optimal level of capital investment is:

$$
c_{K,1} = \sqrt{\delta \mathtt{f}(h_A - c_F)}
$$

Plugging this expression into the faith threshold, we find the optimal level of faith as a function of the initial level of faith, the parameter δ, and the other costs:

$$
\underline{f} > \frac{\sqrt{\delta \mathtt{f}(h_A - c_F)} + c_M - h_M}{c_F + h_A}
$$

References

[1] George A Akerlof and William T Dickens. "The economic consequences of cognitive dissonance". In: *The American Economic Review* 72.3 (1982), pp. 307–319.

[2] Corry Azzi and Ronald Ehrenberg. "Household allocation of time and church attendance". In: *Journal of Political Economy* 83.1 (1975), pp. 27–56.

[3] Fyodor Dostoyevsky. *The Brothers Karamazov*. Penguin UK, 2003.

[4] John T Durkin Jr and Andrew M Greeley. "A model of religious choice under uncertainty: on responding rationally to the nonrational". In: *Rationality and Society* 3.2 (1991), pp. 178–196.

[5] John T Durkin Jr and Andrew M Greeley. "Response to Montgomery". In: *Rationality and Society* 4.1 (1992), pp. 121–122.

[6] Robert B Ekelund Jr, Robert F Hébert, and Robert D Tollison. "An economic analysis of the protestant reformation". In: *Journal of Political Economy* 110.3 (2002), pp. 646–671.

[7] George Gallup Jr. "Religion in america". In: *The Annals of the American Academy of Political and Social Science* 480.1 (1985), pp. 167–174.

[8] Alan Hájek. "Waging war on Pascal's wager". In: *The Philosophical Review* 112.1 (2003), pp. 27–56.

[9] Laurence R Iannaccone. "Introduction to the Economics of Religion". In: *Journal of Economic Literature* 36.3 (1998), pp. 1465–1495.

[10] Laurence R Iannaccone. "Religious practice: A human capital approach". In: *Journal for the Scientific Study of Religion* (1990), pp. 297–314.

[11] Laurence R Iannaccone. "Sacrifice and stigma: Reducing free-riding in cults, communes, and other collectives". In: *Journal of Political Economy* 100.2 (1992), pp. 271–291.

[12] Gilat Levy and Ronny Razin. "Religious beliefs, religious participation, and cooperation". In: *American Economic Journal: Microeconomics* 4.3 (2012), pp. 121–51.

[13] James D Montgomery. "Contemplations on the economic approach to religious behavior". In: *The American Economic Review* 86.2 (1996), pp. 443–447.

[14] James D Montgomery. "Pascal's wager and the limits of rational choice: A comment on Durkin and Greeley". In: *Rationality and Society* 4.1 (1992), pp. 117–121.

[15] Blaise Pascal. *pensées*. Dezobry et E. Magdeleine, 1852.

[16] Rodney Stark, Laurence R Iannaccone, and Roger Finke. "Religion, science, and rationality". In: *The American Economic Review* 86.2 (1996), pp. 433–437.