A Theory of Religious Organizations

Gilat Levy and Ronny Razin

This version: November 2007

Abstract:

In this paper we offer a new theory of religious organizations, which incorporates two important aspects of religion: individual beliefs about the causality between private actions and their consequences, and the social aspects of religious practices. We show how religious institutions arise endogenously, and characterize their features. Specifically, we find that members of the religious organization share similar beliefs and are more likely to cooperate with one another. Our theory predicts that religious organizations that are more demanding in their rituals attract less members, but that these members have more intense beliefs and behave in a more cohesive manner. We illustrate how religious conflicts may arise between religious organizations which differ in their intensity of rituals. Finally, we analyze the dynamics of religious organizations and religious beliefs after individuals experience personal shocks to their wellbeing.

1 Introduction

Religion is typically defined as a social organization that is based on a set of common individual beliefs and practices generally held by a group of people. The relationship between these two important elements of religion is not immediate; if religion is about beliefs in a supernatural entity, why should those beliefs also dictate how individuals interact with one another? If religions are best described as social organizations, why are they almost always associated with beliefs? In this paper we offer a new theory of religious organizations, which incorporates these two aspects of religion.

1Department of Economics, LSE. Email: g.levy1@lse.ac.uk; r.razin@lse.ac.uk. We thank Eddie Dekel, Erik Eyster, Wolfgang Pesendorfer and Torsten Persson for valuable comments, and Stephen Hansen for valuable research assistance. We thank the ESRC for financial support through grant number RES-000-22-1856. In addition we thank the EU for financial support: this project is part of the European Commission STREP project #506084: Polarisation and Conflict: Theory and evidence.

2See Boyer (2001) for a survey of definitions of religion.
We propose a simple model with the following ingredients: a social interaction game, (heterogenous) individual beliefs about the relation between social behaviour and utility shocks, and observable social activities (interpreted as rituals). In particular, we consider a society in which individuals are randomly paired to play a one-shot symmetric Prisoner’s Dilemma (PD) game. We assume that those individuals who choose to be religious must participate in some observable activity, which we interpret as religious rituals. Thus, agents can condition their behaviour in the PD game on their opponent’s affiliation.

A key assumption in the model is that following the strategic interaction, each agent will receive, in addition to the PD payoffs, either a negative utility shock or a positive utility shock. Agents hold different beliefs about the relation between their actions in the PD game and the probabilities of the shocks. While secular agents see no statistical relation between actions and shocks, we assume that religious individuals may look for "patterns", and believe that such a relation exists.\(^3\)

Examples of systems of religious beliefs that connect social behaviour and uncertain events are abundant in religious texts and preaching and are documented in the anthropological literature. In Buddhism, *Karma* and *Vikapa* represent actions and their results, in a cause-effect theory, both in an individual and in a social context. In Hindu, the term *Papa* refers to social actions that create negative (individual) karma. In Judaism, the biblical book of Job is considered problematic exactly because it questions the perceived relation between Job’s immaculate social behaviour and the punishments inflicted on him by God. Another representative example is taken from the seminal work of Evans-Pritchard (1956) about the Nuer Religion:

"...and in any argument about conduct the issue is always whether a person has conformed to the accepted norms of social life....the Nuer are of one voice in saying that sooner or later [and in one way or another] good will follow right conduct and ill will follow wrong conduct."\(^4\)

\(^3\)In many religions, Gods are viewed as entities which can be anywhere at anytime and have the knowledge of any individual’s private actions or thoughts (Boyer (2001)). Such all-knowing Gods are often believed to punish or reward agents according to their actions.

\(^4\)Note that while religious beliefs sometimes relate to rewards in the "afterlife", most religious writings involve discussions about both this life and afterlife rewards and punishments. One extreme example is that in Evans-Pritchard’s (1956) about the Nuer religion mentioned above: "Nuer avoid so far as
The above examples hint at a *particular* and *simple* relation between private actions and private utility consequences; a reward follows "pro-social" behavior by an individual whereas a punishment follows "anti-social" behavior. In this paper we adopt this simple form of religious beliefs, i.e., that utility shocks depend only on the individual’s behaviour.\(^5\) However, we do not take a stand on the particular nature of this relation; rather, we let religious beliefs be determined in equilibrium.

Religions play an active role in shaping individuals' beliefs; when individuals join the religious organization they usually adopt certain beliefs, and they typically change their view of the world when they leave the organization. In our discourse about conversions and deconversions we often use expressions like "seeing the light" or "loosing faith". To be sure, religious organizations invest time and effort in advocating certain kinds of messages while taking an active role in censoring others.\(^6\) In the model, we assume that an individual’s choice to become religious involves a change in his beliefs; as a secular, the individual sees no statistical relation between his actions and shocks to his well-being, while when religious he believes that such a relation exists. Although non-standard in the literature this is a natural assumption to make in the context of religious organizations.\(^7\)

We first analyze the basic model, in which we allow for only one religious organization to arise. Our equilibrium requirement is that each individual finds his affiliation and behaviour optimal, given his beliefs, and given others’ affiliation choices and behavior.\(^8\)

\(^5\)In general one can find examples of religious beliefs that condition consequences not only on an individual’s action but also on the behavior or the affiliation of his opponents. The analysis could be extended to consider these more complex systems of beliefs.

\(^6\)For a model of religion in which individuals choose their offsprings’ beliefs by censoring information see Benabou and Tirole (2006). For other papers in the literature which have focused on the process of the choice of beliefs in the context of cultural transmission see Bisin and Verdier (2001) and Tabellini (2007).

\(^7\)In Section 2 we discuss how our results are affected if we relax the assumption about the active role religions play in shaping beliefs.

\(^8\)We assume throughout that individuals realize that their beliefs will change when they change their affiliation. This "sophisticated" reasoning is not crucial for our results; our main results hold when we assume that agents are "naive" and do not realize that their beliefs will change as a result of a change...
We find that an organized religious group arises endogenously in the model. In particular we find that,

I. Behavior: Religious behaviour arises endogenously as a "pro-social" behaviour. Religious individuals are more cooperative towards fellow religious members than they are towards the seculars. Moreover, a religious agent obtains a higher level of cooperation from society as a whole than a secular agent does.

II. Beliefs: Religious beliefs are endogenously determined as the beliefs that "punishment" will follow non-cooperative behaviour. Any organized religion always includes and sometimes exclusively includes individuals who believe that they are more likely to receive a negative shock (be "punished") when they defect rather than when they cooperate.

III. Rituals: The model rationalizes costly religious rituals, whenever some agents in society are secular, rituals must be costly.

Intuitively, the model suggests two motivations to become (or to remain) religious. A "material" motivation to become religious arises in equilibrium, as a religious agent obtains a higher level of cooperation from society. A "spiritual" motivation to become (or remain) religious arises in the model through the individual beliefs about the consequences of behaviour. As religiosity induces more cooperative behaviour, agents who are relatively averse to defection (as they fear punishment), value being religious.

We analyze the effect of the intensity of religious rituals on the characteristics of religious organizations. We find that when utility shocks are large enough, there is a monotonic relation between the intensity of rituals, the beliefs and behavior of religious agents, as well as the size of the religion. In particular, religious groups that are more demanding in their rituals, are smaller, more cohesive, and composed of individuals who have more extreme beliefs. This prediction finds support in the empirical studies of Iannaccone (1992 and 1998), and in several experimental studies (see Orbell et al (1992) and Sosis and Ruffle (2003)).

Our model, which is based on individual beliefs, allows for a new perspective for the study of the dynamic evolution of religious organizations. In particular, we examine the evolution of religious organizations when individuals update their beliefs based on their in affiliation.
their personal experiences. A key result is that the effect of a shock on an individual’s beliefs depends on his behavior. For example, following a positive shock, an individual who had cooperated will tend to have more religious beliefs (i.e., will believe that he is more likely to be punished when he defects) while an individual who defected will tend to have less religious beliefs. We use this result to show an asymmetry in the way religious organizations respond to common shocks. A correlated positive utility shock causes beliefs in society to polarize and induces deconversions and a reduction in the size of the religious organization. In contrast, a correlated negative shock moderates beliefs in society and promotes both conversions and deconversions, implying ambiguous effects to the size of religious organizations.

Finally, we extend the model to consider the coexistence of several religious organizations, and the (possibly conflictual) relationship among them. We find that the relationship between different religions is somewhat hierarchical; those with more demanding rituals and more internally cohesive behaviour may be uncooperative towards the less demanding and less cohesive ones, but not the other way around. This indicates that internal characteristics (e.g., behaviour towards one’s own group members) may be an important predictor of religious conflicts or intolerance among religious groups.

Previous theoretical models of religion that are related to our paper are those of Iannaccone (1992) and Berman and Iannaccone (2006). In an influential paper by Iannaccone (1992), whose model is later extended by Berman (2000), religion is a club good in which agents’ satisfaction from religious practices (e.g., praying), depends on the practice efforts exerted by others. Rigorous religious rituals are costly signals which allow to exclude free riders who will not exert enough effort. This theory, as ours, gives rise to a positive relation between intensity of rituals and cohesive religious behavior.

Our theory is complementary to Iannaccone (1992) in several ways. First, the club good feature of religion arises endogenously in our model, through the way religious participation affects individuals’ behaviour in the PD game. Second, our theory is

---

9Personal experiences are often mentioned as motivations for conversions or deconversions. For example, the 1966 Nobel laureate in Literature, S.Y. Agnon, became religious after he lost five years worth of writings in a fire. Individual accounts of Holocaust survivors on the other hand include cases of both conversion and deconversion.

10This effect is related to the one suggested in Piketty (1995).

11In other contexts, others have analyzed how costly and observable actions allow agents to discrim-
based on individual beliefs, an important feature of religions. This feature allows us to naturally ask and analyze questions about the evolution of religious beliefs and organizations. Finally, our theory illustrates how religious organizations may provide also "secular" or "material" benefits to their members, and not only the benefit from the consumption of religious goods.

In Berman and Iannaccone (2006), as well as in Stark (1996), the authors assume that individuals have a demand for supernatural goods and thus a market emerges for the supply of these goods. According to their theory religion is inherently personal; individuals believe in the ability of a supernatural being to provide them with supernatural goods. In addition, religion has social features because a collective of individuals is better equipped to insure quality in the supply side of religion.

Our paper is also related to recent literature in which agents choose their or their offsprings’ beliefs. In a related paper by Tabellini (2007), agents choose the scope of their children’s values, i.e., whether good conduct (cooperation in the PD game) should apply only to family and friends, or whether it should apply generally to all opponents. He shows that the equilibrium displays strategic complementarities between values and current behaviour. Similarly, Benabou and Tirole (2006) allow parents to consciously censor information from their offsprings, in order to influence their beliefs. This implies that beliefs that good fortunes arise to those who exert effort (possibly held by Protestants) can be maintained to sustain a political equilibrium in which taxes and redistribution are low, whereas beliefs that sheer luck, and not effort, has an important role in achieving high income, can be maintained to sustain an equilibrium with relatively high levels of redistribution.

Finally, our results may be useful in interpreting the findings in Barro and McCleary (2003). They find that economic growth responds positively to religious beliefs but negatively to religious participation.\footnote{In particular it is beliefs in heaven and hell that are found to have a positive effect on growth, and not just beliefs in the existence of God.} Our model implies a relation between religious beliefs about rewards and punishments, and cooperative behaviour, which may enhance economic growth. We also show that rituals are costly in equilibrium.

Our paper is organized as follows. In Section 2 we present the model, and analyze the
characteristics of an endogenous organized religion in Section 3. Section 4 discusses the coexistence of several religious organizations and religious conflicts. We extend the model to analyze dynamics of beliefs and of religious organizations in Section 5. In Section 6 we discuss some extensions and implications of the model. An appendix contains some omitted proofs.

2 The Model

In the model we consider societies in which some agents are secular and some are religious. Being religious involves participation in an observable action, and having certain beliefs. The agents are paired to play the PD game, and can condition their behaviour on their opponent’s affiliation. Our notion of equilibrium will involve optimal behaviour in the PD game and a stability condition on individuals’ affiliation choices, i.e., a condition requiring that no agent wishes to change his affiliation. We now explain the model in more detail.

2.1 The social interaction

Individuals are randomly paired to play a one-shot Prisoner’s Dilemma (PD) game:

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>d,d</td>
<td>c,b</td>
</tr>
<tr>
<td>D</td>
<td>b,c</td>
<td>a,a</td>
</tr>
</tbody>
</table>

where \( b > d > a > c \). As is standard in the literature, we assume strategic complements, i.e., that \( d - b > c - a \).

2.2 Social rituals

Being religious is associated with some observable action, with a cost \( r \) (\( r \) can be either positive or negative). All agents can identify who paid the cost. An agent that has paid the cost \( r \) is “religious”, otherwise, he is “secular”.

We interpret these observable actions as religious rituals. Such rituals can include investment in special clothes, attendance in religious sermons or any other activities which allow group members to identify and familiarize themselves with one another. This
implies that all agents can identify who is religious, and can condition their behaviour in the PD game on the affiliation of their opponent.\textsuperscript{13}

\subsection{Utility shocks}

One aspect of religion that is a crucial part of our theory, involves the interpretation of uncertain events and their relation to actions taken by individuals. To capture this, we assume that following the strategic interaction, each agent will receive, in addition to the payoffs of the PD game, either a negative utility shock, $-\varepsilon$, or a positive utility shock, $\varepsilon$.

Apart from participation in rituals, we differentiate between seculars and religious agents also according to their beliefs about the relationship between the shocks and their actions in the PD game, as we now explain.

\subsection{Individual beliefs}

We assume that religious agents believe that there is some pattern determining the shocks, and that this pattern depends on their actions in the PD game. Seculars, on the other hand, believe that there is no relation between actions taken in the PD game and utility shocks. Therefore, seculars view the social interaction as the "material" PD game.

In our model, switching affiliation involves therefore a change in beliefs. Specifically, each agent $i$ in the population has a type $(q_{c}^i, q_{d}^i)$, which represents his beliefs if he is religious. That is, a religious agent believes that when he cooperates, he receives the negative shock with probability $q_{c}^i \in [0, 1]$ and that when he defects, he receives the negative shock with probability $q_{d}^i \in [0, 1]$.\textsuperscript{14}

We assume that agents are risk neutral and thus the expected utility of a religious agent who cooperates is $x + \varepsilon(1 - 2q_{c}^i)$, for $x \in \{c, d\}$ (depending on his rival’s action),

\textsuperscript{13}In an alternative model, agents might also choose who to interact with, conditional on whether he had paid the cost or not. We discuss this possibility in Section 5.2.

\textsuperscript{14}An alternative, perhaps more standard, modelling approach is to assume that one’s affiliation does not affect her beliefs. In such a model each agent has beliefs of the form $(q_{c}^i, q_{d}^i) \in [0, 1]^2$. Equilibria will then determine the beliefs of both secular and religious agents (whereas in our model we fix secular beliefs). Equilibria similar to ours will arise in the alternative model (among others) and so the assumption of the choice of beliefs plays a role in refining the set of equilibria.
and similarly, the expected utility of a religious agent who defects is $x + \varepsilon(1 - 2q^i_d)$, for $x \in \{a, b\}$.

It will be sufficient, as will become apparent later on, to characterize the types in the population by the parameter $q^i = q_c^i - q_d^i$, where the higher is $q^i$, the less a religious agent is worried about defection. Note that the best response of a religious agent with $q^i \geq \bar{q} = \frac{d-b}{2c} < 0$ is to defect in the PD game, whereas the best response of a religious agent with $q^i \leq \hat{q} = \frac{c-a}{2c}$, is to cooperate. We assume that $\frac{c-a}{2c} > -1$, so that such types exist in the population. By strategic complementarities, $\bar{q} < \hat{q}$. The best response of those with $q^i \in (\bar{q}, \hat{q})$, henceforth “intermediates”, is to cooperate if their opponent does, and defect otherwise.

Let the types $q^i$ in the population be distributed on $[-1, 1]$ according to some continuous full support distribution function $F(.)$. We assume that $F(.)$ is common knowledge but that individuals do not observe the belief $q^i$ of their religious opponent $i$.\(^{15}\)

Note that the only source of heterogeneity in the model is individuals’ types.\(^{16}\) Assuming such heterogeneity will allow us to endogenize which beliefs are shared by religious individuals. Such heterogeneity may arise when a religious preacher advocates some particular belief, but that agents have different propensities to adopt it.\(^{17}\) Or, it might be that these different beliefs are the result of some process of cultural transmission.

### 2.5 Actions, timing and equilibrium

For the main part of the paper, we focus on the case in which there is only one organized religion in society. For any $(r, F)$, we look at a configuration of affiliation and PD strategies choices, and check whether it is stable, i.e., that no individual wishes to change either her affiliation, or her behaviour in the PD game, given others’ affiliations and behaviour and given her beliefs.

\(^{15}\)We maintain the assumption that there is no restriction imposed on the personal beliefs of agents given their knowledge of the distribution of beliefs in society at large. This is motivated either by an assumption of non-common priors or by assuming that agents believe that there is no statistical relation between their parameters, $q^i_c$ and $q^i_d$, and others’ parameters.

\(^{16}\)A possibly interesting extension is to incorporate additional heterogeneity in the model, e.g., by assuming that individuals differ in their costs or benefits from the rituals.

\(^{17}\)This interpretation is in line with recent literature in evolutionary theory (see Boyer (2001)). According to this literature, about the "religious mind", the human mind has evolved to contain some propensities to religious beliefs and people might differ in these religious tendencies.
Our equilibrium definition consists therefore of two main conditions. First, agents must optimally choose how to play in the social interaction. As paying $r$ is observable, the strategy of an agent in the PD game may depend on whether his (randomly matched) opponent is religious or not. Note that seculars defect disregarding the affiliation of their opponent, as this is a dominant action in the PD game.

Second, agents’ affiliation choices have to be optimal. That is, religious agents must prefer to stay religious, while secular agents must prefer to stay secular, given others’ behaviour and affiliations. This equilibrium requirement is more subtle, as we need to determine how individuals evaluate counterfactual affiliation choices. We make two assumptions. First, we assume that an agent evaluates both affiliations given his current beliefs (i.e., $q^i \in [-1,1]$ for religious, and the belief that there is no relation between shocks and actions for seculars). Second, we assume that individuals anticipate that their beliefs will change once they switch their affiliation.

The first assumption is made for tractability. It also accords with the "partial empathy" approach, which involves parents and offsprings. Under this approach (see for example Tabellini (2007), Bisin and Verdier (2001) and Benabou and Tirole (2006)), adopted to our context, parents’ beliefs are fixed, but they choose which beliefs to transmit to their children - religious or secular. They evaluate the future welfare of their children given their own beliefs, hence empathy is only partial.

Relaxing the second assumption will not change the qualitative nature of our results. If individuals are naive and do not anticipate that their beliefs (and as a result their behaviour) will change once they switch affiliation, then a subset of the equilibria analyzed here will arise. We will explain in the course of the analysis which equilibrium does not survive the naivety of agents.

Equilibria in the model satisfy therefore the following conditions:

1. *(Optimal behaviour in the social interaction):* Given the affiliation choices, and individual beliefs, the strategies in the PD game are best responses.

2. *(Optimal affiliation):* Given the strategies in the PD game, and their current beliefs: (i) a religious agent $i$ prefers to behave according to $q^i$ and to pay $r$, than not to pay $r$ and defect. (ii) a secular agent $i$ prefers to defect and not to pay $r$, than to pay $r$ and behave according to $q^i$. 

10
3. *(Maximizing religious participation):* There is no secular agent, such that if he becomes religious, strictly prefers to stay religious.

The third requirement is motivated by a surplus maximizing incentive of religious organizations. The assumption is that if there exists an individual that may be converted, the religious organization is able to do so. As will be clear in what follows, this refinement does not change the qualitative structure of religious organizations.

### 3 Religious organizations and patterns of cooperation

We are interested in the patterns of behavior created by the introduction of organized religions. To this end we first define some variables that will help us formalise our results. Let \( \gamma_{RR} \) be the probability that a randomly chosen religious agent cooperates when he meets another religious agent. Similarly, let \( \gamma_{RS} \) denote the probability that a randomly chosen religious agent cooperates when he meets a secular agent. These probabilities are determined in equilibrium given the share and the strategies of the types who join the religion. We denote by \( \rho \) the share of religious individuals in society, where \( 1 - \rho \) is the share of seculars. Finally, the net cooperation towards religious agents (that is, relative to the cooperation towards seculars), \( \gamma^{net} \), is given by

\[
\gamma^{net} = \rho(\gamma_{RR} - \gamma_{RS}).
\]

The following Proposition characterizes all the equilibria in the game:

**Proposition 1** For any \( r \leq \bar{r} \), there exists an equilibrium with a religious organization. Any equilibrium satisfies: (i) Behavior: \( \gamma_{RR} \geq \gamma_{RS} > 0 \), \( \gamma^{net} \geq 0 \). (ii) Rituals: whenever there are seculars, \( r \geq \bar{r} > 0 \). (iii) Beliefs: \( \exists q' \leq q < 0 \), such that all agents with types \( q^i \leq q' \) are religious.

The model yields precise predictions about the intensity of religious rituals, the form of religious beliefs, and the behaviour of religious individuals, which are jointly determined in equilibrium. First, the model provides a rationale for *costly* religious rituals; one can interpret the costs \( r \) as investment in special clothes or attendance in religious sermons, or other activities which have some opportunity cost. The value of \( r \) cannot be too high so as to outweigh the benefits of entering the religion. But it also cannot be too low; whenever some agents remain secular, the cost of rituals has to be strictly positive, as
otherwise seculars might infiltrate the religion and "free ride" by defecting against its members.

Second, the result ties between the individual aspects of religion (beliefs about the causality between private actions and uncertain events), and the observed social aspects of religion (i.e., rituals and social cohesion). In particular, those who participate in rituals are also agents who are more cooperative towards their fellow religious agents and, to a lesser degree, towards seculars. Moreover, religious agents tend to share similar beliefs. Specifically, an endogenous “religious belief” arises, that defection in the social interaction leads to “punishment”.

Note that for each agent, the benefit of being religious is a function of how many agents choose to become religious and how they behave, and is thus determined endogenously in equilibrium. We now illustrate how the benefit of being religious is determined in more detail.

3.1 Characterization of religious organizations

3.1.1 Behaviour

We first explore how agents behave in the PD game depending on the affiliation of their opponents. Secular agents will defect irrespective of whom they meet, as for them this is a dominant action. This implies that all religious agents with beliefs above $q$ will defect against seculars.

We now consider those religious agents that would defect against their fellow religious agents. A religious individual will defect against another religious agent, who cooperates with probability $\gamma_{RR}$, if:

$$\gamma_{RR} b + (1 - \gamma_{RR}) a + 2 \varepsilon (1 - q^i_d) \geq \gamma_{RR} d + (1 - \gamma_{RR}) c + 2 \varepsilon (1 - q^i_c)$$

or when $q^i \geq q_{RR}$, for

$$q_{RR} \equiv \frac{1}{2 \varepsilon} ( (d - b) \gamma_{RR} + (c - a) (1 - \gamma_{RR}) ) = \gamma_{RR} \tilde{q} + (1 - \gamma_{RR}) q$$

(1)

In equilibrium, given $\gamma_{RR}$, individuals can be distinguished into three types. Religious agents in $[-1, q]$ cooperate against any other agent (we refer to these as A types). Those in $(q, q_{RR})$, when religious, cooperate with fellow religious agents but defect against seculars (B types). Finally, those in $(q_{RR}, 1]$, defect irrespective of the affiliation of their
opponent (C types). Thus, even though we cannot assess at this point which agents will become religious, it is clear that religious agents are (weakly) more cooperative towards other religious agents than towards seculars:

**Claim 1** In any equilibrium, $\gamma_{RR} \geq \gamma_{RS}$, implying that $\gamma^{net} \geq 0$.

Note that while the share of A types in the population, $F(q)$, is fixed, the share of B and C types will be endogenously determined (together with $\gamma_{RR}$).

### 3.1.2 The relative benefit of being religious

We can now determine the relative benefit of being religious, i.e. the expected utility difference between being religious and being a secular, disregarding the cost $r$. Such a calculation depends on the affiliation of the individual (as individuals use their current beliefs to evaluate their affiliation choices). We first characterize the relative benefit of being religious as evaluated by religious agents, i.e., given some $q^i$. Note that for such agents, there are two reasons for preferring to stay religious. A "material" motivation arises as a religious agent obtains a higher level of cooperation from society (by Claim 1). A "spiritual" motivation arises when religious agents realize that they cooperate more often when they are religious, both because of the beliefs instilled by the religion, but also because others cooperate against them more often. If they become secular on the other hand, they will defect - a scenario they may wish to avoid given their current beliefs $q^i$.

**Claim 2** For a religious agent, the relative benefit of being religious is monotonically decreasing and continuous in $q^i$.

Proof: For an A type the relative benefit for being religious is given by:

$$
\rho(\gamma_{RR}^d + (1 - \gamma_{RR})c) + (1 - \rho)c + 2\varepsilon(1 - q^i_d)
- \rho(\gamma_{RS}^b + (1 - \gamma_{RS})a) + (1 - \rho)a - 2\varepsilon(1 - q^i_d)
= R^A(\rho, \gamma_{RR}, \gamma_{RS}) - 2\varepsilon q^i,
$$

where $R^J(\rho, \gamma_{RR}, \gamma_{RS}), J = A, B, C$, represents "material" payoffs from the PD game, while the second term above, $-2\varepsilon q^i$, represents the "spiritual" payoff (recall that $q^i < 0$ for A and B types).
For a B type the relative benefit for being religious is given by:

\[
\rho(\gamma_{RR} d + (1 - \gamma_{RR}) c) + (1 - \rho)a + 2\varepsilon\rho(1 - q_d^i) + 2\varepsilon(1 - \rho)(1 - q_d^i) \\
- \rho(\gamma_{RS} b + (1 - \gamma_{RS}) a) - (1 - \rho)a - 2\varepsilon(1 - q_d^i) \\
= R^B(\rho, \gamma_{RR}, \gamma_{RS}) - 2\varepsilon\rho q^i.
\]

Note that the material payoff changes for these agents, conditional on their type \(J\), as their belief dictates a different behaviour in the game. Finally for a C type this relative benefit is:

\[
\rho(\gamma_{RR} b + (1 - \gamma_{RR}) a) + (1 - \rho)a + 2\varepsilon(1 - q_d^i) \\
- \rho(\gamma_{RS} b + (1 - \gamma_{RS}) a) - (1 - \rho)a - 2\varepsilon(1 - q_d^i) \\
= R^C(\rho, \gamma_{RR}, \gamma_{RS}) = \gamma^{net}(b - a) \geq 0.
\]

Moreover, one can show that \(R^A(\rho, \gamma_{RR}, \gamma_{RS}) - 2\varepsilon q = R^B(\rho, \gamma_{RR}, \gamma_{RS}) - 2\varepsilon\rho q\) and that \(R^B(\rho, \gamma_{RR}, \gamma_{RS}) - 2\varepsilon\rho q_{RR} = \gamma^{net}(b - a)\), hence the benefit is continuous in \(q^i\). □

Figure 1 plots the relative benefits of being religious for religious types, given a particular behavioral configuration \(\rho, \gamma_{RR}, \gamma_{RS}\):

![Figure 1: The relative benefit for a religious agent from being religious versus secular.](image-url)
Intuitively, agents below $q_{RR}$ cooperate more often when they are religious than when they are secular. But the higher is $q^i$, the lower is the benefit from cooperation, and so the "spiritual" payoff decreases. In terms of the material payoff, it is easy to see that

$$R^C_0(\rho, \gamma_{RR}, \gamma_{RS}) > R^C_1(\rho, \gamma_{RR}, \gamma_{RS}) > R^C_2(\rho, \gamma_{RR}, \gamma_{RS}).$$

When secular, all these types of agents behave in the same way. When they are religious though, they defect more often if their type is higher. But given $(\rho, \gamma_{RR}, \gamma_{RS})$, defection always yields a higher material payoff than cooperation.

Note that the expressions for the material payoffs represent the relative benefit of being religious as viewed by the seculars, as these agents consider only the material payoff when evaluating affiliation choices. For seculars, the relative benefit from being religious is therefore (weakly) monotonically increasing in their type $q^i$. Intuitively, the higher is their type, the less they cooperate when they are religious, which is beneficial from the point of view of secular beliefs. Moreover, the relative benefit from being religious is equal for a secular C type, and for a religious C type.

The above two observations, together with condition (iii) of the equilibrium definition, allow us to proceed in the following way. First, we can check, given $r, \rho, \gamma_{RS}$ and $\gamma_{RR}$, which religious agents indeed prefer to be religious. By the monotonicity property in Claim 2, this will be the case for all agents below some $q'$. If $q'$ is high enough such that a religious C type agent strictly prefers to be religious, it must be that all society is religious in equilibrium. If on the other hand $q'$ is low enough so that a religious C type prefers to be secular, then also a secular C type prefers to stay secular, and as a result, also secular A and B types prefer to stay secular. Thus, all secular agents with $q \geq q'$ will prefer to be secular, insuring that no agents cycles from one affiliation to another.

By the above, to characterize the religion in equilibrium, it will be sufficient to focus on religious types, and to characterize $q'$ given $r$, using Figure 1. Note that as $\rho, \gamma_{RR}, \gamma_{RS}$ change, so do the graphs depicted in the Figure (their levels, slopes, and $q_{RR}$). Given $r$, the equilibria will therefore be determined by a fixed point argument.

### 3.1.3 Fully cooperative religion

In the first family of equilibria we analyze, a "fully cooperative" religion arises, whose members are only A types, as depicted below:
Let \( r_1 = 2\varepsilon F(q)(\bar{q} - q) \) and \( r_2 = 2\varepsilon(\bar{q} + 1) \). Note that \( \min\{r_1, r_2\} > 0 \). We now show that these equilibria hold for any \( r \in [\min\{r_1, r_2\}, \max\{r_1, r_2\}] \).

In these types of equilibria, \( 1 = \gamma_{RR} = \gamma_{RS} > 0 \). Thus, \( q_{RR} = \bar{q} \). What we need to determine is the marginal A type, who is indifferent between being religious and being secular, given his beliefs \( q' \). We therefore need to solve:

\[
R^A(\rho, \gamma_{RR}, \gamma_{RS}) - 2\varepsilon q' = r \iff \quad F(q')(d - b) + (1 - F(q'))(c - a) - 2\varepsilon q' = r \iff \\
F(q')\bar{q} + (1 - F(q'))\bar{q} - q' = \frac{r}{2\varepsilon}
\]

To see why a solution exists to the above fixed point equation in \( q' \), consider some \( r \) in the range described above. If \( q' = -1 \), then, as \( F(q') = 0 \), we have that the left-hand-side of the equation is \( \bar{q} + 1 \) which is greater (smaller) than the right-hand-side when \( r_1 < r_2 \) (\( r_1 > r_2 \)). If on the other hand, \( q' = \bar{q} \), then the left-hand-side becomes \( F(\bar{q})(\bar{q} - \bar{q}) \), and it is smaller (greater) than the right-hand-side when \( r_1 < r_2 \) (\( r_1 > r_2 \)). By continuity, a fixed point exists.

In these equilibria, there is no "material" payoff from being religious, as no agent is changing his behaviour conditional on the affiliation of his opponent. Still, agents are religious because of the "spiritual" payoff. Being religious allows them to "better
behave", as once they become secular, they will defect, which they find unappealing given their low level of \( q' \). The cost of rituals must be positive to ensure that only a selected group of agents are religious.

### 3.1.4 Selective cooperation equilibrium

In the second type of equilibrium, we have a religion whose members cooperate selectively. The members of the religious organization are all A types, and some B types:

![Figure 3](image)

Let \( r_1 = (F(q) - F(q))(b - a) \) and \( r_2 = 2\varepsilon F(q)(\bar{q} - q) \). Note that \( \min\{r_1, r_2\} > 0 \). We now show that these equilibria hold for any \( r \in [\min\{r_1, r_2\}, \max\{r_1, r_2\}] \).

In this equilibrium, \( 1 = \gamma_{RR} > \gamma_{RS} > 0 \). Again, as \( \gamma_{RR} = 1 \), we have that \( q_{RR} = \bar{q} \). We need to determine which of the B types is the marginal type, i.e., to solve:

\[
R^B(\rho, \gamma_{RR}, \gamma_{RS}) - 2\varepsilon \rho q' = r \Leftrightarrow \\
F(q')(d - a) - F(q)(b - a) - 2\varepsilon F(q')q' = r
\]

The existence of a fixed point, for the relevant range of \( r \), follows from the same argument as described in the previous section with \( q \) and \( \bar{q} \) as the end points.

In these equilibria, intermediate religious types change their behaviour depending on whom they meet. They cooperate with fellow religious agents but defect against seculars \( (\gamma_{net} > 0) \). All religious agents prefer to stay religious given the "material" benefit, as
well as the "spiritual" benefit associated with the religion. The cost of rituals is positive, as otherwise, C types will become religious to enjoy a higher level of cooperation.

3.1.5 Free riders equilibrium

In this family of equilibria, all A and B types, and a measure $\alpha$ of C types are religious. Such an equilibrium exists for any $r \in [r^*, (F(\bar{q}) - F(q))(b - a)]$.

![Figure 4](image)

In this type of equilibrium, $1 > \gamma_{RR} > \gamma_{RS} > 0$ (and $\gamma^{net} > 0$). The level of internal cooperation in the religion is reduced, as C types, who always defect, become religious in order to take advantage of the other members of the organization and enjoy its "material" benefit. In response to the introduction of religious C types, there are less B types in society (although all of them are religious) so that $q_{RR} < \bar{q}$.

To find the equilibrium, we need to determine $q_{RR}$, which pins down, given $\alpha$, the size of the religion and the set of types which cooperate in the religion. The equilibrium is a solution to the fixed point equation in (1), where $\gamma_{RR} = \frac{F(q_{RR})}{F(q_{RR}) + \alpha}$. Moreover, as the relative benefit of C types from being religious does not depend on $q^i$, they must be indifferent so that some of them, but not all, are religious. Thus, given the solution to
\( q_{RR} \), we can fix \( r \) so that \( C \) types are indifferent. For consistency, we let \( q_{RR} = q' \):

\[
q' = \frac{F(q')}{F(q')} \bar{q} + \left( 1 - \frac{F(q')}{F(q')} \right) q,
\]

\[
r = \gamma_{RR} (b - a) = (F(q') - F(q))(b - a)
\]

When \( \alpha = 0 \), \( q' \) is set at \( \bar{q} \), and is the limit of the selective cooperation equilibria, when there are no "free riders". When \( \alpha \) is higher, and the religion has a larger share of free riders, \( q' \) decreases. To determine the lower bound \( r^* \) we can find the limit of these equilibria when the whole population is religious (so that \( \rho = 1 \)). Specifically, let \( \alpha = 1 - F(q') \). We then have \( r^* = (F(q^*) - F(q))(b - a) \) where \( q^* \) is the solution to

\[
q' = F(q') \bar{q} + (1 - F(q')) q.
\]

Thus, globally, the size of the religion decreases with \( r \), together with its level of internal cooperation, \( \gamma_{RR} \), which decreases from 1 (when \( \alpha = 0 \) and \( r \) is the highest possible in the free rider equilibria) to \( F(q^*) \) when the whole population is religious and \( r = r^* \). Note that \( r^* > 0 \) as for any full support distribution function \( F \), \( q^* > \bar{q} \).

### 3.1.6 Society with no seculars

Finally, for any \( r \leq r^* \), there exists an equilibrium in which the whole society is religious. This equilibrium is as the limit equilibrium described above, which holds for \( r^* \). In this equilibrium, all agents below \( q^* \) cooperate, and all agents above \( q^* \) defect. If an agent deviates and becomes secular, then a share \( F(q^*) - F(q) \) of the population (the \( B \) types) will defect against him. Obviously, this equilibrium can be sustained for any \( r \leq r^* \), including negative values of \( r \). We can therefore conclude that whenever there are seculars, \( r \) is positive and is bounded away from zero, but that when there are no seculars, then religious rituals can be directly beneficial to one’s utility.

**Remark 1: Equilibria with naive agents:** Note that when agents are naive, all the equilibria above hold, besides the first type of equilibria with a fully cooperative religion. The key observation is that even with naive agents, the two reasons for being (or staying) religious, the "material" benefit and the "spiritual" benefit are still relevant. In particular, given their current religious beliefs, intermediate agents still realize that when they become secular they will cooperate less often, even if they maintain their current \( q^i \). The reason is that being secular, they will face less cooperation from society, which will alter their best response. This reasoning cannot arise in the fully cooperative
religion.

We have characterized all possible equilibria, for any $r$ and $F$. This completes the proof of Proposition 1. In the next section we use this equilibrium characterization to investigate the relation between the intensity of rituals and the size and nature of religious organizations.

### 3.2 Intensity of rituals and the size of religious organizations

The religious organizations we have identified in the previous section exist for any full support distribution $F$. We now show that for any $F$, if the shocks are large enough, these equilibria induce, globally, a monotone relation between $r$, religious behaviour, and religious beliefs.

**Proposition 2** For any full support $F$, there exists an $\varepsilon$ such that for any $\varepsilon > \varepsilon$, globally, when $r$ increases: $\rho$ decreases, $\gamma_{RR}$ and $\gamma_{RS}$ increase, and $q'$ decreases.

When the level of the shock is large enough relative to the payoffs in the PD game, the types of equilibria described above are segregated in $r$. For relatively high levels of $r$, only fully cooperative religions can exist. For intermediate levels of $r$, only religions with selective cooperation can exist. These equilibria are characterized by larger religions, and less cooperation with seculars. For low levels of $r$, only religions with free riders can exist, where these are characterized by even larger religions and less internal cooperation. Finally, for even lower levels of $r$, the unique form of equilibria is society with no seculars.

To see the intuition for the result, recall the two motivations for being religious. The "material" motivation arises as a religious agent obtains a higher level of cooperation from society. This implies that the cost of religion should increase when more intermediate types (who cooperate selectively) are religious. On the other hand, the "spiritual" motivation arises when agents who become religious change their behaviour and become more cooperative themselves. This implies that the cost of religion should decrease when more intermediate types are religious, as the marginal type is less averse to defection. When the shocks are large enough, the "individual belief" motivation becomes more important, inducing a negative monotonic relationship between the cost of religion and its size.

Our prediction about the relationship between intensity of rituals and the level of
cooperation finds some support in recent empirical and experimental studies. Sosis and Ruffle (2007) conduct an experiment involving religious and secular individuals in Israeli Jewish Kibbutzim. They find that religious males are more cooperative compared with religious females (who participate in less rituals) and with seculars, and that the more religious males attend synagogue, the more cooperative they are. Secular rituals (such as eating in the common dining room) are found to have a weaker effect on cooperation than do religious rituals.

Orbell et al (1992) find, also in an experimental setting, that religious agents do not show higher levels of cooperation than seculars, but that those who attend church more frequently show higher levels of cooperation than others; they conclude that rituals and not religious beliefs promote cooperation. In another study, using a self-reported survey in the US, Iannaccone (1992) finds that the stricter is the church in its demands, the higher is the level of contributions within the church. Finally, Iannaccone (1998) reports some stylized facts from the US about the relationship between rituals, cooperation, and beliefs. Within religions, more conservative denominations have members who contribute proportionally more income, attend more services, and have more orthodox doctrinal beliefs.

4 Many religions and religious conflicts

In this section we ask whether several religious organizations can coexist in one society, and whether such coexistence is likely to induce tolerance or conflicts between the different organizations. In this section we view the non-cooperative behaviour in the PD game as a form of conflict; in Section 6 we discuss other types of conflicts, such as direct hostile actions vis-a-vis other groups.

It is easy to see that there always exist equilibria with several different religious organizations. For example, for any equilibrium as described above, a new equilibrium is sustained if we split the religious organization into two or more "churches", maintain the original cost $r$ for all groups, and assume that members of any group cooperate with all religious individuals. But, there may also exist equilibria with several churches such that some who cooperate internally defect against members of the other. We now investigate these equilibria.

As usual, in our model, the interesting agents to consider are the intermediate types
who may alter their behavior conditional upon their opponent’s affiliation. We say that
Religion 1 defects (cooperates) against Religion 2, if some (all) intermediate Religion
1 agents who cooperate internally in Religion 1, defect (cooperate) against members
of Religion 2. The next result illustrates that the way religions behave towards one
another depends on their own internal level of cooperation (denoted by \( \gamma_{R_i^1R_j^1} \) for religions
\( i, j \in \{1, 2, \ldots, n\} \)):

**Proposition 3** (i) There exists no equilibrium in which a religion \( i \) with \( \gamma_{R_i^1R_j^1} \leq 1 \)
defects against a religion \( j \) with \( \gamma_{R_i^1R_j^1} = 1 \) while religion \( j \) cooperates against religion
\( i \); (ii) There exists an equilibrium with two religions in which religion \( i \) with \( \gamma_{R_i^1R_j^1} = 1 \)
defects against religion \( j \) with \( \gamma_{R_i^1R_j^1} < 1 \) but religion \( j \) cooperates against religion \( i \). In
this equilibrium, \( r_i > r_j \), i.e., religion \( i \) has a higher intensity of rituals.

Proposition 3 indicates that the relationship between different religions is somewhat
hierarchical; those with more demanding rituals and more internally cohesive behaviour
(i.e., religions whose members always cooperate with one another) may be less tolerant
towards the less demanding and less cohesive ones, but not the other way around. This
prediction of our model also suggests a possible direction for empirical analysis.

The proof of the first part of the Proposition is straightforward. If members of religion
\( j \) all cooperate with one another, then it must be composed of individuals with beliefs
below \( \bar{q} \). Moreover, as all its intermediate individuals cooperate against religion \( i \), then
it must be that all members of \( j \) cooperate against \( i \) members. Thus, the best response
of intermediate types in religion \( i \) is to cooperate against members of \( j \).

To prove the second part, we construct an equilibrium with two religious organizations,
Religion 1 and Religion 2, satisfying the characteristics described in the Proposition (the
exact construction is in the appendix). In this equilibrium, the distributions of types, \( F^i \)
for \( i = 1, 2 \), is the uniform distribution, for both religious organizations. The equilibrium
is depicted in Figure 5, which described the affiliation (and strategies) of the religious
agents, given their \( q^i \). In the Figure, \( R^1 \) and \( R^2 \) denote the affiliation to Religion 1 and
Religion 2 respectively, and the strategy \( xyz \) for \( x, y, z \in \{c, d\} \) denotes the behaviour
of the agents towards Religion 1, Religion 2, and the seculars respectively:
Figure 5: Two religions with one-sided tolerance

In this construction, Religion 1 is similar to the one described in Figure 4, i.e., it also contains “free riders” who wish to exploit the relatively high level of internal cooperation and defect against their own fellow group members. Some agents with \( q^i \leq q \) belong however to Religion 2, which also includes intermediate agents with a relatively high level of \( q^i \).

Religion 2 is characterized by full internal cooperation, but all its intermediate agents defect against both seculars and members of Religion 1. On the other hand, all intermediate agents of Religion 1 who cooperate internally also cooperate against Religion 2. The seculars defect against all agents. Thus, a Religion 2 agent gains a higher level of cooperation from society overall, than does a Religion 1 agent. To sustain this in equilibrium, Religion 2 demands a higher intensity of rituals than Religion 1 \((r_2 > r_1 > 0)\).

Note that in this equilibrium, the sect or the religion which is considered as more extreme in terms of religious rituals is actually composed of types with relatively high levels of \( q^i \) or in other words, types that according to our distinction, have less “religious beliefs”. Moreover, these agents -and not the more “fundamentalist” religious agents with very low level of \( q^i \)- are the ones which are hostile to the rest of society in general and to the other religion in particular.

5 The evolution of beliefs

We now analyze how individuals update their beliefs following the shocks they experience. In order to be able to discuss the evolution of beliefs, assume that each individual does not know the value of \( q^i_c \) and \( q^i_d \) but instead his type is such that \( q^i_c \) and \( q^i_d \) are taken (independently) from a full support density functions \( f^i(q^i_c) \) and \( f^i(q^i_d) \) respectively. The definition of \( q^i \) is now slightly altered to \( q^i = E^i(q^i_c) - E^i(q^i_d) \). We therefore fix the initial latent types in the population (specifically, \( f^i(q^i_c) \) and \( f^i(q^i_d) \)), and let each individual update his type following his course of play, and the shock he receives.\(^{18}\) We then find:

\(^{18}\)Note that we are assuming that individuals do not learn from others’ experiences. This is consistent with our earlier assumption about the heterogeneity of beliefs about the parameters, \( q^i_c \) and \( q^i_d \), and the

23
Lemma 1 (i) Following a negative shock, an agent who cooperated will increase his $q^i$ and an agent who defected will decrease his $q^i$. (ii) Following a positive shock, an agent who cooperated will decrease his $q^i$ and an agent who defected will increase his $q^i$.

In response to a shock, individuals update their beliefs conditional on the action they have played in the game. An individual who has been cooperating and experienced a positive shock, believes that on average $q^i_c$ is lower but does not change his beliefs on $q^i_d$, whereas an individual who has been defecting and experienced a positive shock, believes that on average $q^i_d$ is lower and does not change his beliefs on $q^i_c$. As $q^i = E^i(q^i_c) - E^i(q^i_d)$, we have the result of the Lemma above.

To see the implication of the above result for the evolution of religious organizations, consider the selective cooperation equilibria as described in Figure 3. Those equilibria are characterized by a cutoff $q'$ below which individuals are religious, where some religious individuals (those with $q^i < q$) always cooperate, and some religious individuals (with $q^i \in [q, q']$) defect when meeting seculars and cooperate when meeting a fellow religious member. All those who remain secular defect, while their (latent) type is higher than the cutoff $q'$.

Suppose now that all individuals in society experience a positive shock (e.g., an economic boom). By Lemma 1, all individuals with relatively low $q^i$, who were cooperating, will decrease their $q^i$ even further, and all those with relatively high $q^i$, who were defecting, will increase their $q^i$ even further. Thus, a correlated positive shock will tend to polarize beliefs in society. On the other hand, a negative shock (e.g., a natural disaster) will have the opposite effect and will tend to moderate beliefs.

Moreover, following a correlated positive shock, some religious individuals with intermediate $q^i$, who happened to meet seculars and were therefore defecting, will actually leave their faith. Therefore, such a shock will not only polarize beliefs but will also decrease the size of the religion. We illustrate this below, where the arrows in the figure mark the direction of the change in beliefs, given the action of the individual, and the bold arrow is the relevant one in terms of (de)conversion:

common knowledge about the distribution of (heterogenous) beliefs in society.
On the other hand, the effect of a correlated negative shock on religious participation is less clear cut. The shock tends to moderate beliefs; we will therefore observe seculars who convert but also religious individuals who were cooperating, and hence will leave the faith:

Which effect dominates will generally depend on the exact shapes of \( f^i(q^i_c) \) and \( f^i(q^i_d) \), but holding these fixed, the size of the religion will increase following a negative shock if the share of the religion is small enough, whereas it will decrease if this share is large. To see why, note that the size of the religion has no effect on the flow into the religion but has a negative impact on the flow out of the religion: the larger is the religion, the more likely it is that intermediate religious types encounter other religious individuals and hence more of them end up cooperating. As the negative shock indicates the inferiority of cooperation, more of these individuals will leave the faith. On the other hand, seculars defect against all individuals so that the size of the religion does not affect the rate of religious conversion.

### 5.1 The survival of religious organizations

The discussion above has illustrated how religious organizations affect the distribution of beliefs; the updating of beliefs depends on the actions taken by individuals, and these actions are determined endogenously in equilibrium, by the religious organization. On the other hand, our analysis in Section 3 has focused on how the distribution of beliefs shapes the form of religious organizations. This reciprocal relationship between beliefs and religious organizations suggests that it would be interesting to analyze a
dynamic model in which beliefs and organizations evolve over time, and are possibly
determined together in a steady state equilibrium. In such a long-run analysis, it will be
important to specify the true distribution generating the shocks to individuals’ utilities,
as individuals are allowed to learn and update their beliefs (and the religious organization
adapts accordingly).

One question that our model allows to analyze is whether religious organizations can
be sustained in the long run, even if the true distribution exhibits no statistical relation
between actions and shocks. From the analysis above it is clear that in some cases,
individuals who are always cooperating may fool themselves about the relative benefit
of defecting, and may stick to their beliefs forever as they do not learn anything about
defection. However, in the selective cooperation equilibria, or in the free riders equilibria
(see Figures 3 and 4), intermediate types tend to experiment, i.e., they sometimes coop-
erate and sometimes defect. They will then converge to learn information which is close
to the truth; no such intermediate types will exist in the long run. Therefore, the model
predicts that religious organizations that are sustained in the long run will be one of
two types. Either the equilibria with the fully cooperative religions in which only types
below $q$ are religious and they always cooperate (Figure 2), or equilibria in which all
individuals in society are religious, which may be characterized by very low intensity of
rituals. In both these types of equilibria, individuals either always cooperate, or always
defect, disregarding the affiliation of their opponent.

Clearly, if the true distribution exhibits no statistical relation between actions and
shocks, then religious organizations could be sustained only if beliefs are immune to avail-
able information. Thus, our model can indicate why some religions focus on specification
of rewards and punishments in the “afterlife”, and hence, why religious organizations
may be stronger institutions than other social organizations.

5.2 Religious segregation

In our analysis we assume that the players are randomly matched disregarding their affil-
iation. Another modelling possibility is to allow individuals to condition their matching
decisions on their own affiliation or on their opponent’s one. In particular, it might be
that affiliation with a religious organization also allows the individual to interact solely
or perhaps mainly with members of that organization. In such a model, individuals who
belong to the same organization also choose to congregate in the same neighborhood, and thus limit interaction with non-affiliated members.

The benefit of religious congregation can be two-fold. First, religious agents will meet only agents who cooperate and will avoid being taken advantage of. Second, such isolation might allow them to avoid information that is unfavorable for the prevalence of their beliefs and hence for the existence of the religion in the long-run. For example, religious agents may not learn about the relative benefit of defecting if they are not exposed to the part of the population that does so. Indeed, many religious organizations tend to segregate their members from non-members either physically or in other ways that contain the amount of information about the experiences of non-members.

6 Discussion

In this section we discuss some additional implications and extensions of the model.

Religion and welfare: One explanation for the existence of religious organizations is that such organizations improve the fitness of either the individual or of the group as a whole compared with the fitness of seculars (see Wilson (2002)). We can assess the welfare of individuals in the model according to the standard Prisoner’s Dilemma matrix. In that case, a religious organization could potentially be beneficial to society as a whole. To see this note first that by changing the beliefs of individuals, the religious organization introduces some cooperation, whereas when there are only seculars, all individuals defect. In particular, this implies that all secular agents are better off when a religious organization exists.

Religious individuals however, do not always benefit from the introduction of religion. First, it is not clear that they benefit from a higher degree of cooperation in society. While type B agents employ selective cooperation and cooperate only when matched with fellow religious agents, type A agents cooperate also with seculars, who therefore take advantage of them. Second, whenever seculars exist, religious individuals must pay \( r > 0 \) in order to be affiliated with the religion. In some cases, the material payoff even from selective cooperation may be outweighed by the cost \( r \) of entering the religion.

Religious entrepreneurs: Our framework can be useful to account for the supply side of religion. To do so, we can consider religious entrepreneurs who will accommodate
the demand for religion. We can think of the entrepreneur as initiating a religion by choosing the cost \( r \) and the distribution of types \( F \). Such an entrepreneur will maximize his revenues, which are possibly a function of \( r \) (a portion of which he might be able to extract) and the level of the demand for religion that \( r \) and \( F \) generate. We can then analyze which religious organization is chosen by such an entrepreneur.

Similarly, one can analyze the effects of competition in the supply side of the market for religion. If we incorporate religious entrepreneurs in the model, it might allow us to derive more precise predictions about cross-religion relations. A reasonable conjecture is that in the presence of competition, each entrepreneur will initiate a religious organization that will defect against others, as this will increase the relative advantage of belonging to his group.

*Religious leaders and political power:* The above idea of religious entrepreneurship suggests a theory of the political power of religious leaders; the entrepreneur chooses the level of rituals/costs, but presumably he also chooses the content of these costs. One component of these costs can be pure monetary contributions which the entrepreneur can expropriate. Alternatively, the entrepreneur can demand that group members exert effort towards a political goal. As long as this effort (such as participation in a demonstration) is costly and observable, individuals will adhere to these demands.

*Religious conflicts:* When discussing religious relations, we have focused in this section on a particular notion of religious tolerance, that is, how religious agents treat one another in their daily economic interactions. Another notion of tolerance or intolerance is one in which one religion takes explicit hostile actions against a rival one. This can be naturally incorporated in our model when we account for the possibility of religious entrepreneurs or religious leaders. Specifically, the religious entrepreneur can also determine that the costly and observable actions will include hostile activities towards rival sects (i.e., in addition to their actions in the PD game). This is beneficial to the religious leader in the context of competition between religious entrepreneurs, as it will increase the relative advantage of belonging to his group rather than others. According to this idea, competition between groups might lead to more religious intolerance.
References


Appendix

Proof of Proposition 2: If

\[(F(\bar{q}) - F(q))(b - a) < 2\varepsilon F(q)(\bar{q} - q) < 2\varepsilon(q + 1),\]

then equilibria in which only type A are religious exist only for \(r \in [2\varepsilon F(q)(\bar{q} - q), 2\varepsilon(q + 1)]\), equilibria in which all A and some B types are religious exist only for \(r \in [(F(\bar{q}) - F(q))(b - a), 2\varepsilon F(q)(\bar{q} - q)]\), equilibria in which all A, all B and some C types are religious exist only when \(r \in [(F(q^*) - F(\bar{q}))(b - a), (F(\bar{q}) - F(q))(b - a)]\) and finally for all \(r \leq (F(q^*) - F(q))(b - a)\) there exist only equilibria in the whole society is religious. Thus, as \(r\) increases beyond the next cutoff, \(\gamma_{RS}\) and \(\gamma_{RR}\) (weakly) increase, \(\rho\) decreases and \(q^*\) decreases.

We now show why (2) holds when \(\varepsilon\) is high enough. Consider first the left-hand-side inequality:

\[
(F(\bar{q}) - F(q))(b - a) < 2\varepsilon F(q)(\bar{q} - q) \iff \\
(F(\bar{q}) - F(q))(b - a) < F(q)(d - c - (b - a)) \iff \\
\frac{F(\bar{q})}{F(q)} < \frac{d - c}{b - a}
\]

By strategic complements, \(\frac{d - c}{b - a} > 1\). On the other hand, the left-hand-side approaches 1 when \(\varepsilon\) increases. Consider now the right-hand-side inequality:

\[
2\varepsilon F(q)(\bar{q} - q) < 2\varepsilon(q + 1) \iff \\
F(q)(\bar{q} - q) < q + 1
\]

When \(\varepsilon\) increases the left-hand-side approaches zero whereas the right-hand-side is positive and bounded away from zero.\(\blacksquare\)

Proof of Proposition 3(ii): We now construct an equilibrium with the following characteristics:

\[1 = \gamma_{R^2R^2} > \gamma_{R^2R^1} = \gamma_{R^2S} > \gamma_{R^1R^1} = \gamma_{R^1R^2} > \gamma_{R^1S} > 0.\]
In particular, in this equilibrium, all \( R^1 \) agents which cooperate in \( R^1 \) also cooperate with agents in \( R^2 \), whereas the opposite is not true, as agents in \( R^2 \) enjoy full internal cooperation but some of them defect against members of \( R^1 \).

In this case, type A individuals, have \( q^i \leq q \), and will cooperate with all, when in \( R^1 \) or \( R^2 \). Type B individuals, with \( q^i \in (q, q_{R^1 R^1}) \), will, when in \( R^1 \) or \( R^2 \), defect against seculars, and cooperate otherwise. Type C individuals, in \( (q_{R^1 R^1}, q_{R^2 R^1}) \), will, when in \( R^1 \) or \( R^2 \), cooperate only with \( R^2 \) members, and defect otherwise. Type D individuals, in \( (q_{R^2 R^1}, \bar{q}) \), when in \( R^2 \), will cooperate only with \( R^2 \) members, and will defect otherwise. Finally, type E individuals, with \( q^i > \bar{q} \), will defect against all. Note that for all these types, the benefit from joining \( R^2 \) is always greater than the benefit from joining \( R^1 \), as they will enjoy a higher level of cooperation from society overall.

We set \( r_1 = (F(q_{R^1 R^1}) - F(\bar{q}))(b - a) \) and \( r_2 = (F(q_{R^1 R^1}) - F(\bar{q}))(b - a) + (F(q_{R^2 R^1}) - F(q_{R^1 R^1}))(d - c) > r_1 \). This implies that A, B, and C types agents strictly prefer to be religious than to be secular, but are indifferent with regard to which religion to join, and D and E types are indifferent between joining \( R^1 \) and being secular, and strictly prefer it to joining \( R^2 \). We can therefore consider an equilibrium in which \( R^1 \) includes a share \( \alpha \) of the A types, all B types, and a share \( \phi \) of D and E types, and \( R^2 \) includes the remaining share \( F(q) - \alpha \) of the A types, and all C types. The remaining agents are secular. As in (1), the equations for the cutoff points are:

\[
q_{R^1 R^1} = \gamma_{R^1 R^1} \bar{q} + (1 - \gamma_{R^1 R^1})q; \quad q_{R^2 R^1} = \gamma_{R^2 R^1} \bar{q} + (1 - \gamma_{R^2 R^1})q \rightarrow
\]

\[
q_{R^1 R^1} = \frac{F(q_{R^1 R^1}) - F(\bar{q}) + \alpha}{F(q_{R^1 R^1}) - F(\bar{q}) + \alpha + \phi}(\bar{q} - q) + q
\]

\[
q_{R^2 R^1} = \frac{F(\bar{q}) - \alpha}{F(q_{R^2 R^1}) - F(q_{R^1 R^1}) + F(\bar{q}) - \alpha}(\bar{q} - q) + q
\]

Note that to solve for an equilibrium, we simply need to find \( \alpha \) and \( \phi \) such that the solution for the above fixed points equations (note that we first solve for \( q_{R^1 R^1} \) and then for \( q_{R^2 R^1} \), will satisfy the following conditions:

\[
1 > \gamma_{R^2 R^1} > \gamma_{R^1 R^1} > \gamma_{R^1 S} > 0
\]

\[
0 < \alpha < F(\bar{q}); \quad 0 < \phi < 1 - F(q_{R^2 R^1})
\]

32
Consider the uniform distribution on \([-1,1]\), i.e., \(F(q) = \frac{1}{2}(1 + q)\). Let \(\bar{q} = -0.25\) and let \(q = -0.5\). Finally, let \(\alpha = 0.2\) and let \(\phi = 0.3\). The solution is is \(q_{R1R3} = -0.38446, q_{R2R3} = -0.33393\) and it satisfies all the conditions set above. ■

**Proof of Lemma 1:** Suppose that an individual cooperated in the game (the analysis for an individual who defected is analogous). Note that such an individual will only update his belief about \(q_c\). His updated beliefs satisfy, for any \(q'_c \geq q''_c\):

\[
\frac{Pr(q'_c | \varepsilon)}{Pr(q'_c | \varepsilon)} = \frac{\int f(q'_c)q'_c dq_c}{\int f(q'_c)(1-q'_c) dq_c} \geq \frac{\int f(q''_c)q''_c dq_c}{\int f(q''_c)(1-q''_c) dq_c} = \frac{Pr(q''_c | \varepsilon)}{Pr(q''_c | \varepsilon)}.
\]

The MLRP therefore holds, implying the result reported in Lemma 1. ■