A Theory of Religion: Linking Individual Beliefs, Rituals, and Social Cohesion

Gilat Levy and Ronny Razin[1]

A preliminary version, September 2006

Abstract:

In this paper we offer a new theory of religion. Our theory relies on the assumption that individuals in society have different beliefs about the statistical relation between their actions in a strategic social interaction (specifically, a Prisoner's Dilemma) and uncertain events that affect their utility. This heterogeneity of beliefs allows for the endogenous emergence of religious institutions; A religious institution is defined by its members partaking in a costly and observable activity. We show that members of the religious organization share similar beliefs, tend to cooperate with one another, and face a larger degree of cooperation from society than do non-members. Our theory provides therefore a link between individual beliefs, social rituals, and social cohesion.

1 Introduction

Religion has fascinated scientists for hundreds of years. Anthropological theories of totem and taboos, sociological theories of functionalism, evolutionary theories of optimal adaptation or biological investigations of the "religious mind", are just notable examples of the many attempts to understand the origin of religion as well as its prominent existence in essentially every human society.

A fundamental feature of any theory of religion is, or should be, an explanation of the link between the two most observed aspects of religion, the individual aspect and the social aspect. The individual aspect is usually referred to as beliefs of the individual

[1]Department of Economics, LSE. Email: g.levy1@lse.ac.uk; r.razin@lse.ac.uk. We thank 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 the European Commission STREP project #506084: Polarisation and Conflict: Theory and evidence.

about the power or abilities of some supernatural entity. The social aspect refers to the observations that most religious rituals are conducted by the collective and not by individuals, and that most religions would prescribe how to behave in social interactions. The relation between these two observed attributes of religion is not immediate. If individuals believe in supernatural beings, why should they celebrate their beliefs in a social gathering? Why should these beliefs in supernatural entities also dictate how human beings should interact with one another? Our objective in this paper is to provide a theory which establishes this relation.

We analyze the following model of social interaction. Individuals are randomly paired to play a one-shot symmetric Prisoner's Dilemma (PD) game. We assume that the game exhibits strategic complementarities, i.e., that an agent's gain from cooperation is higher if his opponent cooperates as well.

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 the individuals. To capture this, we assume 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. We assume that agents hold different beliefs regarding the relation between their personal actions and the probabilities of the shocks mentioned above.

We define the following institutional setting as an *Organized Religion*: it is a group composed of individuals who partake in a costly and observable activity. When an organized religion exists in equilibrium, the strategy of an agent may depend on whether his opponent is affiliated with the religious organization or not. Given these strategies, agents choose whether to join the religion or not.

We distinguish between a *cohesive* and *non-cohesive* religious organizations. A cohesive religious organization is a group or sect whose members all behave in the same way towards one another (e.g., all defect or all cooperate). In a non-cohesive religion different members of the religion may behave differently when facing their fellow members. We show that in both types of religious organizations, a similar structure arises:

I. Behavior: A member of a religious organization faces a larger degree of cooperation

from the society overall than a non-member.[2] Moreover, religious organizations allow for a higher degree of internal cooperation (i.e., among its members) than in the rest of society. In particular, in the cohesive religion, all members cooperate internally.

II. Beliefs: In any equilibrium, religious members have similar beliefs. In any cohesive religion, all members of the religion believe that on average they are more likely to receive a negative shock (be 'punished') when they defect rather than cooperate. Any non-cohesive religion consists of these types as well; however, types who believe otherwise also become religious, in order to defect and take advantage of its members.

III. Rituals: The costly and observable activity that religious members engage in can be interpreted as social rituals. We show that cohesive and non-cohesive religious organizations differ in terms of the level of rituals and the size of their membership. Cohesive religions are generally smaller in size and more demanding in the level of rituals they impose on members.

Our theory is based on how individuals perceive the relation between their private actions and uncertain events. In particular, in the results stated above, it is the heterogeneity of these beliefs that allows for the emergence of the religious organizations. We further examine how the religious organization evolves over time, given that individuals update their beliefs based on their personal experiences.

We show that in our model the effect of a shock on the individual's beliefs depends on the action he was taking at the time of the shock. In particular, the 'direction' in which an individual's belief shifts is different when he cooperates to that in which he defects. We use this result to illustrate how the size of a religious organization decreases following a correlated positive utility shock for the whole population, and that it might increase following a correlated negative shock.

We then analyze the long-run ability of religious organizations to sustain themselves as individuals' beliefs evolve. We assume that the true environment exhibits no statistical

[2] This result accords with the view of Adam Smith, who argued that poor people had an incentive to join strict religious groups to signal their trustworthiness to others (not necessarily group members), with whom they wanted economic relationships (for a summary of Smith's view on religion, see Anderson (1988)).

relation between individual's actions and utility shocks. We characterize the types of religious organizations that can be sustained in the long run. We show that the only types of religions that survive are those which are 'redundant', i.e., they generate equilibria that may exist even without the existence of costly rituals. Moreover, our dynamic analysis lands support to the casual observations that surviving religious organizations typically involve beliefs about the "afterlife", which are immune to available information, and that they encourage "isolationism", inducing their members to interact only with other members.

Our paper is organized as follows. In the next section we discuss the related literature. Section 3 presents the model, whereas Section 4 provides the characteristics of endogenous organized religions. We analyze the dynamics of beliefs and of religious organizations in Section 5. In Section 6 we extend our analysis to investigate the question of tolerance between different religious organizations. We show that the relation between different sects is somewhat hierarchical, with the cohesive and thus more rigorous sect being less tolerant towards the non-cohesive and the more lenient sect. We also examine in this section the possible role of religious entrepreneurs. An appendix contains all proofs that are not in the text.

2 Related Literature

Our theory is aligned with functionalist theories of religion (see Durkheim (1912)); in our theory as well, religion allows for social cohesion, while social rituals are the mechanism providing the means to this end. However, as opposed to functionalist theories which do not explain why individuals participate in rituals, in our theory individuals are motivated by their self interest and find it optimal to participate in rituals.

Recently, proponents of the evolutionary approach (see Wilson (2002), Dawkins (1993), Dennett (2006), Boyer (2001)) explain how religious beliefs and religious organizations might have been shaped by evolutionary forces. Under this view, religious beliefs, religious organization or religious ideas, have formed to enhance the fitness of their members (as in Wilson (2002)), are by-products of other successful evolutionary strategies (see Dennett (2006), Boyer (2001)), or maximize the fitness of parasite memes (the view

advocated by Dawkins). In our theory we also analyze the dynamic aspect of religious beliefs and organizations. In contrast to the evolutionary approach, institutions and beliefs evolve in our analysis not due to their superior fitness, but as a result of rational behavior on behalf of individuals. These individuals update their beliefs based on their personal experiences and subsequently choose optimally whether to join a religious organization or not.[3]

A related literature in social science, based on the assumption of rational behavior, was initiated by Iannaccone (1992), and followed by Berman (2000) and Iannaccone and Berman (forthcoming). In Iannaccone (1992), religion is modelled as a club good, in the sense that agents' religious satisfaction from praying for example, depends on the praying efforts exerted by others. Rigorous religious rituals evolve in this context in order to exclude free riders who will not exert enough effort. While this theory explains social cohesion in clubs it falls short of forming a convincing theory of religion, as religious beliefs are not accounted for in this approach.

In Starke and Finke (2000) and Berman and Iannaccone (forthcoming), the assumption is that individuals have a demand for supernatural goods and thus a market emerges for the supply of these goods. According to this theory, religion is inherently personal; individuals believe in the ability of a supernatural being to provide them with supernatural goods.[4] The social aspect, such as forming churches, arises in this theory as a by-product, as a collective of individuals is better equipped to insure quality in the supply side of religion. As in the club-good model, religious rituals serve to deter free riders (who would rely on others' efforts to screen good and bad priests). In contrast to these theories, in our theory, the demand for religious groups endogenously arises as a social need and individuals join a religious group as it alters both their behavior and the behavior of others towards them.

Finally, an interdisciplinary literature has analyzed religion from the perspective of the mind (see Boyer (2001)). According to this literature, the human mind has evolved

[3]For empirical support to the notion that religion is a result of the choices of rational individuals, see Stark, Iannaccone and Finke (1996) and Iannacone (1998).

[4]See also Benabou and Tirole (forthcoming) and Dixit and Grossman (1984) for a purely individualistic religion.

to contain some propensities to religious beliefs. Moreover, people might differ in these religious tendencies. These theories are in line with our assumption that there is an inherent heterogeneity of beliefs in society, where these beliefs might be interpreted as religious tendencies.

3 The Model

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

	C	D
C	d,d	c,b
D	b,c	a,a

where $b > d > a > c$. We further assume that the game exhibits strategic complementarities, i.e., that $d - b > c - a$.[5]

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, ε. There is some true underlying probability measure generating these shocks. The true probability measure may or may not be conditional on the action taken by the individual.

We assume that individuals differ in their beliefs about the relation between particular uncertain events and their personal actions. Such beliefs might depend only on the individual's personal actions, or on a more elaborate description of the course of play, for example, on both the individual's and his opponent's actions.

Here we assume that beliefs depend only on the individual's personal actions. We make this assumption for simplicity, as all our analysis can be carried through even if we allow beliefs to depend for example on the opponent's actions as well. Such an

[5] One could carry out the analysis with the alternative assumption of strategic substitutes. The type of religious institutions that we identify depend on the assumption of strategic complemantarities.

assumption is also reasonable as an individual has the best information about his actions and thus need not worry about mistakes in interpreting other players' actions. Moreover, from a cognitive point of view, this is the simplest possibility. We further discuss theses issues in Section 5, when we analyze the dynamics of religious beliefs.

Specifically, we assume that each agent believes that when he cooperates, he receives a negative shock with probability q_c^i (and a positive shock with probability $1-q_c^i$) that is distributed according to a full support density function $f^i(q_c^i)$, and that when he defects, he receives a negative shock with probability q_d^i (and a positive shock with probability $1-q_d^i$) that is distributed according to a full support continuous density function $f^i(q_d^i)$. At this point we need not make any assumption about whether agents believe that these probabilities are correlated with those of other agents.

The assumptions we have taken so far imply that the expected utility of an agent who cooperates is $x+\varepsilon(1-2E^i(q_c^i))$, for $x \in \{c,d\}$ (depending on his rival's action), and similarly, the expected utility of an agent who defects is $x+\varepsilon(1-2E^i(q_d^i))$, for $x \in \{a,b\}$.

Below, we will extend our analysis to discuss how beliefs evolve over time. For now we assume that these beliefs are fixed (and exogenously given). This will imply that without loss of generality we can focus on characterizing each agent in the population by the parameter $q^i = E^i(q_c^i) - E^i(q_d^i)$. Let q^i be distributed on $[-1,1]$ according to some continuous full support distribution function $F(.)$.

Given a distribution F, agents can be divided into three types. Those with high levels of q^i, larger than $\bar{q} = \frac{d-b}{2\varepsilon}$, will always defect. We denote their measure in the population by $\rho_d = 1 - F(\frac{d-b}{2\varepsilon})$. Those with low enough levels of q^i, lower than $\underline{q} = \frac{c-a}{2\varepsilon}$, will always cooperate. We denote their measure by $\rho_c = F(\frac{c-a}{2\varepsilon})$ and assume that $\frac{c-a}{2\varepsilon} > -1$. The best response of those with intermediate values of q^i is to cooperate if their opponent does, and defect otherwise. We denote the measure of these types by $\rho_{cd} = F(\frac{d-b}{2\varepsilon}) - F(\frac{c-a}{2\varepsilon})$.

Agents cannot identify the type they are matched with; the role that religious institutions will play in our model is to allow agents to do so. To abstract from a possible role of religious institutions as solving *coordination* problems, we assume that whenever it is somehow common knowledge among two matched agents that both have $q^i \leq \bar{q}$, then they cooperate in the game. This assumption also allows us to simplify the analysis by

reducing the possible configurations of religious institutions.

Absent religious institutions, in a Bayesian Nash equilibrium, each agent holds the true beliefs about the share of agents who cooperate in equilibrium, and best responds to these beliefs given his type. We then have the following benchmark result, which we will sometimes refer to as the state of "anarchy":

PROPOSITION 1 *(Anarchy): There exists an equilibrium. In any equilibrium there is a cutoff* $q^* \in (\underline{q}, \bar{q})$; *types below* q^* *cooperate and types above* q^* *defect.*

PROOF: In the anarchy state, agents cannot condition their actions on the type of the agent they are matched with, and thus can either always cooperate or always defect (there might be a measure zero of individuals who will be indifferent and that could possibly use a mixed strategy). Let ρ_{cd}^* denote the measure of the intermediate types who cooperate in equilibrium. Given that a measure of $\rho_c + \rho_{cd}^*$ cooperates, those who are willing to cooperate are all those with $q \leq \frac{(\rho_c+\rho_{cd}^*)(d-b)+(1-(\rho_c+\rho_{cd}^*))(c-a)}{2\varepsilon} \equiv q^*$. An equilibrium is therefore defined by

$$\rho_c + \rho_{cd}^* = F(\frac{(\rho_c + \rho_{cd}^*)(d-b) + (1-(\rho_c + \rho_{cd}^*))(c-a)}{2\varepsilon})$$

That an equilibrium exists with $q^* \in (\underline{q}, \bar{q})$ follows from the continuity of F and from the observations that $\rho_c + \rho_{cd}^{\max} = 1 - \rho_d > F(\frac{(1-\rho_d)(d-b)+(\rho_d)(c-a)}{2\varepsilon})$ and that $\rho_c < F(\frac{(\rho_c)(d-b)+(1-\rho_c)(c-a)}{2\varepsilon})$.■

Note that in this environment, the intermediate types would have liked to identify one another; this would allow them to both coordinate on mutual cooperation and to avoid being taken advantage by the types who always defect. We therefore proceed to analyze religious institutions as organizations which allow agents to identify one another.

4 An organized religion

We define the following institutional setting as an *organized religion*: A group of agents pays some cost r, and the action of paying the cost is observable. In other words, *all* agents can identify who had paid the costs (and hence belongs to the religion) and who had not.

For the main part of the paper, we focus on the case in which there is only one such organized religion in society. In Section 6 we relax this assumption and consider the coexistence of several religions or sects.

When a religious organization exists, the strategy of an agent depends on whether his opponent is affiliated with the religious organization or not. In equilibrium, given these strategies, agents choose optimally whether to pay r or not, and given the affiliation choices, the strategies are best responses.

As all types below $\underline{q}$ (above $\bar{q}$) cooperate (defect) against all agents disregarding their own and their rival's affiliation, an organized religion will alter only the behavior of the intermediate types in $[\underline{q}, \bar{q}]$. Thus, necessary conditions for the existence of a religious organization is that the population includes types in $[\underline{q}, \bar{q}]$ and in $[\bar{q}, 1]$. Without types above $\bar{q}$, the population is presumably able to coordinate on the Pareto dominant equilibrium of full cooperation. But more importantly, without types in $[\underline{q}, \bar{q}]$, no agent will pay $r > 0$ as his affiliation will not alter others' behavior.

Note that there always exists an equilibrium with an organized religion that is redundant; in such an equilibrium, the behavior of agents is as in the "anarchy" benchmark and all agents pay the cost r, for r low enough.

PROPOSITION 2 *(Redundant religions): For any $r \leq \rho^*_{cd}(b-a)$, there exists an equilibrium with a religious organization in which the whole population pays r. The equilibrium replicates the anarchy equilibrium with a cutoff q^*.*[6]

PROOF: This equilibrium is supported by the off-equilibrium threat of the intermediate types in the region $(\underline{q}, q^*)$ to defect against an agent who does not pay the cost r. Whenever $r \leq \rho^*_{cd}(b-a)$, all agents in $[\tilde{q}, 1]$ for some $\tilde{q} \in (\underline{q}, q^*)$, who given the equilibrium threat will defect if they do not pay r, are willing to pay r. As $r < \rho^*_{cd}(d-c)$ by strategic complementarities, all agents in $[-1, \tilde{q}]$ who given the equilibrium threat will cooperate if they do not pay r, are willing to pay r.■

We now focus on equilibria with "meaningful" religious organizations, i.e., equilibria in which some agents in the population behave differently conditional on the affiliation

[6]Recall that $\rho^*_{cd} = F(q^*) - F(\underline{q})$.

of their opponent.

We will use the notation ρ_c^r to denote the measure of types below $\underline{q}$ who pay r ("religious" types), and ρ_c^s to denote the measure of those below $\underline{q}$ who do not pay r ("seculars"). We can similarly define the notation $\rho_{cd}^r, \rho_{cd}^s, \rho_d^r$ and ρ_d^s. We will use the term "internally" ("externally") to describe the behavior of agents towards agents with the same (different) affiliation.[7]

We first consider the case of a *cohesive religion*, i.e., a religious organization whose members all behave in the same way towards one another:

PROPOSITION 3. *(Cohesive religions): There exist equilibria with a meaningful cohesive religion. In any such equilibrium: (i) (behavior): All those who pay r cooperate internally, the degree of internal cooperation is higher among religious agents than among seculars, and a religious agent is facing a larger degree of cooperation in society overall than a secular agent does; (ii) (rituals): $r \in [\underline{r}, \bar{r}]$, $\underline{r} > 0$;(iii) (beliefs): A type q is in the religion if and only if $q \in [q', q'']$, where $q' \geq -1$ and $q'' \in [\underline{q}, \bar{q}]$.*

PROOF: see the appendix.

This simple result already highlights two important aspects of religion. First, it provides a rationale for religious rituals; one can interpret the costs r as investment in special clothes, attendance in religious sermons or any other costly activities which allow group members to identify and familiarize themselves with one another. The value of r cannot be too high, as those who cooperate would rather not pay the cost to identify themselves. But it also cannot be too low, as then outsiders might infiltrate the religion and defect against its members.

Second, and more importantly, it 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). When we allow for one cohesive religious group, it must be a group that *cooperates* internally and therefore includes individuals who believe that they will mostly be punished when they defect

[7] In a Bayesian Nash equilibrium, agents correctly conjecture the shares $\rho_c^r, \rho_c^s, \rho_d^r, \rho_d^s, \rho_{cd}^r$, and ρ_{cd}^s, how the types ρ_{cd}^s and ρ_{cd}^r behave internally and externally, and best respond to this conjecture.

against one another.

Finally, note that the members of the religious group have an incentive, if allowed this choice, to interact only internally and avoid the lack of cooperation from outsiders. Since the rituals themselves allow them to identify one another, they may choose - to some degree - to interact only with those who follow the same rituals. The model indicates therefore that isolationist attitudes may arise within religious groups. We will return to this possibility later on.

EXAMPLE 1: *The following religious organization exists for any distribution over beliefs F : An agent with type q pays r if and only if $q \in [-1, q'']$ for some $q'' > \underline{q}$; while all types below $\underline{q}$ always cooperate, all intermediate types in $[\underline{q}, q'']$ cooperate internally but defect against seculars. All secular agents defect disregarding the affiliation of their opponent. The level of rituals r satisfies $r \in (\rho_{cd}(b-a), \bar{r})$. When $r = \rho_{cd}(b-a)$, all intermediate types in the population are religious.*[8]

The religious organization characterized in Example 1 illustrates how religion generates more "targeted" cooperation, as intermediate religious types can tailor their behavior depending on the affiliation of their opponent. We will sometimes refer to this example when we provide additional characteristics of religious organizations.

We now consider a non-cohesive religion, in which some religious agents cooperate internally and some defect internally.

PROPOSITION 4 *(Non-cohesive religions): There exist equilibria with a meaningful non-cohesive religion. In any such equilibrium: (i) (behavior): Some of those who pay r cooperate internally and some defect internally, and a religious agent is facing a larger degree of cooperation in society overall than a secular agent does; (ii) (rituals): $r \in [\underline{r}', \bar{r}']$, $\underline{r}' > 0$;(iii) (beliefs): A type q is in the religion if $q \in [-1, q'']$, for some $q'' \in (\underline{q}, \bar{q}]$, whereas some types above $\bar{q}$ also belong to the religion.*

PROOF: see the appendix.

As opposed to a cohesive religion, a non-cohesive religion is infiltrated by opportunists

[8] We characterize the equilibrium conditions in the Appendix.

who enjoy the cooperation of the group members and take advantage of them. Using Example 2 below, we will explore whether such a religion is also more lenient in its religious rituals compared with a cohesive religion, and is therefore perhaps larger.

EXAMPLE 2: *The following religious organization exists for any distribution over beliefs F : An agent with type q pays r if $q \in [-1, q'']$ for some $q' > \underline{q}$, and some agents with $q > \bar{q}$ also pay r; all types below $\underline{q}$ always cooperate, all religious intermediate types in $[\underline{q}, q']$ cooperate internally and defect against seculars, whereas religious types above $\bar{q}$ and all secular agents always defect. The level of rituals r satisfies* $r = x(b-a)$, for $x \in [\rho^*_{cd}, \rho_{cd}]$.

Note that in the religious organizations characterized in Example 2, a larger value of r is associated with a larger share of intermediate religious types who cooperate internally. The reason is that the opportunistic types who infiltrate the religion in order to benefit from its cooperative members need to be compensated for the rise in r.

We now focus on the religious institutions specified in Examples 1 and 2 to gain additional insights about the characteristics of religious organizations.

The relation between the size of a religious organization and its level of rituals: Taken together, Proposition 2 and Example 1 and 2 imply that for any $r \leq \bar{r}$, there exists an equilibrium with a religious organization and that globally, the size of the religion is monotonic in r. For all $r \leq \rho^*_{cd}(b-a)$, the religion is redundant, as it comprises the whole population and is identical to the anarchic situation. When $r \in [\rho^*_{cd}(b-a), \rho_{cd}(b-a)]$, the religion is non-cohesive but exclusive and hence smaller. When r increases from its minimum value $\rho^*_{cd}(b-a)$ to its maximum value $\rho_{cd}(b-a)$, the overall size of the religion decreases. Specifically, when $r = \rho^*_{cd}(b-a)$ the religion comprises the whole population, whereas when $r = \rho_{cd}(b-a)$, there is no such redundant religion, so that some agents remain secular. Moreover, when r increases, the religion becomes more cohesive in the sense that the share of types who cooperate internally increases.

When $r \in [\rho_{cd}(b-a), \bar{r}]$, the unique type of religion is even smaller and the religion is cohesive, i.e., all religious types cooperate. When $r = \rho_{cd}(b-a)$ all intermediate types

are religious and by Proposition 3, the religion obtains its maximum size. Thus, when r increases from $\rho_{cd}(b-a)$ to $\bar{r}$, as less intermediate types are willing to pay its cost, the size of the religion decreases.

Religion and welfare: Are religious organizations improving the welfare of the population? Obviously, the redundant religious organizations which mimic the anarchy equilibrium but "force" its members to pay some costs, reduce the welfare of the whole population compared with "anarchy".

To conduct a welfare analysis for the meaningful religious organizations, there are two ways to proceed. We can look at the welfare gains or losses of each individual from his point of view, that is, given his particular beliefs q^i. Alternatively, we can assess the welfare of individuals from some "objective" point of view, by considering the true parameters determining the distribution of the shocks. To fix ideas, assume that the true parameters satisfy $q_c^i = q_d^i$ (so that $q^i = 0$) for any i.

Given Examples 1 and 2, it is clear that religious organizations decrease the welfare of all seculars above q^*, compared with anarchy.[9] The reason is that these types defect against all agents both in the anarchic state and in the presence of religion (to which they are not affiliated). However, when a religion exists, no agent above $\underline{q}$ cooperates against them, as opposed to the anarchic state.

For other agents, the welfare analysis is ambiguous. The agents in the religion always benefit from more "targeted" cooperation, as opposed to the anarchic state, as they can better identify those who would defect against them. However, they have to pay the cost r. In some cases, for example when q^* is very close to $\underline{q}$, the religion induces a sufficiently higher degree of cooperation to outweigh its cost so that the welfare of all religious agents is higher. In other cases, however, in particular, when q^* is close to $\bar{q}$, it is possible that the high cost needed to sustain the religion outweigh its benefits, so that *all* religious agents are worse off compared with the anarchic state.

We have identified in this section two types of meaningful endogenous religious organizations, the cohesive and the non-cohesive one. In Section 6 we investigate the

[9] As there may be multiple equilibria in the anarchic state, one can look at $q^*_{\sup} = \sup q^*$.

coexistence of such different religions (or different sects within one religion) and how they behave towards one another. But first we turn to some dynamic analysis.

5 Dynamics of religious beliefs and organizations

In this section we consider the dynamics of religious beliefs following the shocks experienced by the population. We first ask how the size of the religious organization changes in response to a correlated one-off shock experienced by the whole population. We then investigate the long run consequences of individuals updating their beliefs given their experiences.

To model the dynamic process, we fix the initial beliefs in the population (specifically, $f^i(q_c^i)$ and $f^i(q_d^i)$), and let each individual update his beliefs following his course of play and the shock he receives. Before the next period, each such individual is replaced by a new member of the population who inherits his updated beliefs.[10] Moreover, we assume that each individual updates his beliefs only given his action in the game and the positive or negative shock he experiences. If individuals believe that their shocks are independent, these are indeed the only relevant observations. More generally, even if there is no such independence, these might be the only available observations for each individual.[11]

5.1 Religion in "good times" and in "bad times"

We first consider the effect of a correlated shock (negative or positive) on religious organizations. For example, a positive correlated shock can represent an unexpected economic boom, whereas a negative correlated shock may represent a natural disaster.

[10] Our dynamic process can therefore be intepreted as a process of transmission of beliefs from one generation to another. We choose this modelling strategy as in our anlaysis we do not focus on dynamics arising from a repeated game between the same individulas. See Greif (1993) for the relation between ethnic or religious groups and repeated interaction.

[11] One can possibly allow individuals to observe other individual's course of action and shocks either within their group, or those of the whole population. These alternative assumptions will change the path of play that we analyze and the speed of long run convergence of beliefs. The result reported in Proposition 7 still holds in these cases.

In these situations, as agents re-evaluate their beliefs, it is interesting to examine the effect of such updating process on the characteristics of religious organizations.

An important feature of our model is that in response to a shock, individuals update their beliefs conditional on the action they have played in the game. For example, an individual who was cooperating and experienced a positive shock, believes that on average q_c^i is lower but does not change his beliefs on q_d^i, whereas an individual who was defecting and experienced a positive shock, believes that on average q_d^i is lower and does not change his beliefs on q_c^i. Recall that $q^i = E^i(q_c^i) - E^i(q_d^i)$. In the analysis below we look at the equilibria analyzed in Example 1. We then have:

LEMMA 1 *(Belief updating): An individual's belief is updated depending on both the sign of the shock and the action he played. Specifically: (i) A negative shock has a moderating effect on beliefs; agents who cooperated (relatively low q^i) have a higher q^i and agents who defected (relatively high q^i) have a lower q^i.(ii) A positive shock has a polarizing effect on beliefs; agents who cooperated (relatively low q^i) have a lower q^i and agents who defected (relatively high q^i) have a higher q^i.*

PROOF: 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 $q_c' \geq q_c''$:

$$\frac{Pr(q_c'|-\varepsilon)}{Pr(q_c'|\varepsilon)} = \frac{\frac{f(q_c')q_c'}{\int f(q_c)q_c dq_c}}{\frac{f(q_c')(1-q_c')}{\int f(q_c)(1-q_c)dq_c}} \geq \frac{\frac{f(q_c'')q_c''}{\int f(q_c)q_c dq_c}}{\frac{f(q_c'')(1-q_c'')}{\int f(q_c)(1-q_c)dq_c}} = \frac{Pr(q_c''|-\varepsilon)}{Pr(q_c''|\varepsilon)}.$$

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

Given Lemma 1, we can characterize the effect of a correlated shock on the size of the religious organization. To this end we fix some r and an equilibrium, as in Example 1. We analyze the effect of the correlated shock on the characteristics of this equilibrium. We define the *impressionability* of an individual as the magnitude of the shift in his q^i after observing a shock.[12] To do this exercise we need to avoid the usual problems that arise when there are multiple equilibria. The following Proposition relates to the

[12]Impressionability depends on the exact form of the prior density functions, $f^i(q_c^i)$ and $f^i(q_d^i)$.

equilibria characterized in Example 1 in cases in which they are unique both before and after the shock (see footnote 14 below).

PROPOSITION 5 *(The size of religion and a correlated shock): Following a positive shock, the size of the religion decreases. (ii) Following a negative shock, the size of the religion either increases or decreases, as a function of both its original size and individual impressionability. Controlling for impressionability, the size of a religion will increase (decrease) if the religion is small (large) enough.*

PROOF: see the appendix.

To see how a positive shock reduces the size of the organized religion, note that no secular agent joins the religion following a shock. All secular agents have defected, and their belief that this is the right course of action has only strengthened following a positive shock to their utility. Similarly, religious agents who have cooperated have strengthened their beliefs. However, some religious agents who were matched with outsiders have defected; their updated beliefs indicate that defecting is a relatively favorable action and they do not find it worthwhile anymore to be affiliated with the religion. Thus, some religious agents switch to be secular and the size of the religion decreases.

A negative shock, on the other hand, induces both religious agents to switch to be secular (those who have cooperated), and secular agents to become affiliated (those who have defected). The total relative size of the incoming and outgoing flow to the religion depends therefore on whether religious agents cooperate or defect (which in turn depends on the relative size of the religion), and whether agents impressionability is high enough to change their q^i beyond the cutoff that defines the religion.

5.2 Religion in the long-run

We now turn to consider the long-run relationship between beliefs and religious organizations. In Section 4 we have analyzed how religious organizations are shaped given a distribution of individuals' beliefs. The previous section has established how religious organizations determine the path through which individuals update their beliefs; updating depends on the actions taken by the individuals which are determined by the

religious organization. This reciprocal relationship between beliefs and religious organization suggests that we look for a steady state to determine what kinds of religious organizations and what kind of beliefs can sustain in the long-run.

Our results in Section 4 imply that religious organizations are generally characterized by being comprised of types with $q^i \leq \bar{q} < 0$. As individuals use Bayesian updating given their experiences, it is interesting to ask whether religious organizations can be sustained in the long run, even when the true distribution of shocks is "secular", i.e., when the true distribution is such that $q^i = 0$ for each i (or, $q_c^i = q_d^i$). Such distribution exhibits no (statistical) relation between actions and shocks. In what follows, we therefore assume that the true distribution over shocks satisfies $q_c^i = q_d^i = \lambda$ for all i.

We now define our notion of a steady state equilibrium and distribution of beliefs. Recall that the beliefs of each individual are characterized by $f^i(q_c^i)$ and $f^i(q_d^i)$. We denote the distribution over such beliefs in the population by G.We say that G is *consistent* with the truth (q_c, q_d) if for any i, (q_c, q_d) is in the support of $f^i(q_c^i)$ and $f^i(q_d^i)$ respectively.

DEFINITION 1 *A distribution G and an equilibrium given G is a steady state iff: (i) G is consistent with the truth, (ii) G and the equilibrium do not change after individuals update their beliefs.*

PROPOSITION 6 *(i) For some λ, there exists a steady state with a redundant religion. (ii) For any λ, there doesn't exist a steady state with a meaningful religion.*

PROOF: (i) Consider some G and a redundant religion equilibrium with some cutoff q^*. We now check whether G can be generated by the individual densities $f^i(q_c^i)$ and $f^i(q_d^i)$,for which the truth is in their support. As all those with $q^i \geq q^*$ defect, they must believe that $q_d^i = \lambda$ with probability one. Let these types have $E^i(q_c^i) \in (\lambda + q^*, 1)$. They therefore satisfy $q^i \in [q^*, 1 - \lambda]$. Similarly, as those with $q^i \leq q^*$ put probability one on $q_c^i = \lambda$, we can let these types have $E^i(q_d^i) \in (\lambda - q^*, 1)$, so that these types satisfy $q^i \in [\lambda - 1, q^*]$.

Note that for the above, we need that $\lambda + q^* \geq 0$ and $\lambda - q^* \geq 0$ implying that we need $\lambda \geqslant -q^*$ (recall that $q^* < 0$). Therefore, $\lambda > -\bar{q}$. Moreover, to insure that an equilibrium with a redundant religion exists, we need that $\lambda - 1 < \underline{q}$. Thus, for any λ

satisfying:

$$-\bar{q} < \lambda < 1 + \underline{q}$$

we can find some G which generates a distribution F over q^i with support on $[\lambda - 1, 1 - \lambda]$ as constructed above, for which there exists a steady state with a redundant religion.

(ii) In a meaningful religion, some agents sometimes cooperate and sometimes defect. These agents must converge to know the truth with probability one and will therefore converge to have $q^i = 0$. This is a contradiction as such agents will always defect.□

In a redundant religion, the intermediate types do not experiment with different actions which allows such religion to survive. In a meaningful religion, which allows for sufficient experimentation, the intermediate types disappear, and so does the possibility for a meaningful religious organization. Note however that in cases close to the extreme case of $q_c^i = q_d^i = 0$, in the long run, there are still agents in the population who are cooperating (even though this is not the right action). Thus, if a religious organization can also allow its members to isolate themselves from the rest of the population, it might arise in such environments. For example, suppose that r is the cost of moving to a particular neighborhood, and agents can interact only with agents who live in this neighborhood. All agents whose long-run beliefs are below $\bar{q}$ would agree to pay some positive r for such a possibility. This will not be the case when the true distribution over shocks is close to the other extreme case of $q_c^i = q_d^i = 1$. In such a case, all types end up defecting.

Given the above result, there might be several ways to change the model in order to enable the survival of meaningful religious organizations in the long run, even if the true distribution over shocks provides no reason for such an institution to exist. First, a religious organization might be sustained if beliefs are somehow immune to available information. In other words, our model can possibly explain why many religions are based on events relating to the "afterlife". When actions are related to punishments and rewards which are not easily observable, such beliefs are possible to sustain in the long run.

Second, religious organizations that involve more elaborate beliefs about the relation between the shocks and the outcomes of the PD game might survive in the long run. Recall that the intermediate types, who are experimenting with different actions, learn everything precisely because we have assumed that they condition their beliefs only on their private actions. On the other hand, were they to hold more complex beliefs that depend for example on their opponent's action, they might not learn everything as not all outcomes of the PD game arise in equilibrium. In that case, a religious organization may be sustained in the long term. This argument may provide some justification for religions to provide a code of behavior which specifies how to act depending on whether the opponent is himself cooperative or not.

6 Discussion and extensions

Our basic model can be extended in several ways. Here we consider two main extensions. First, we examine the coexistence of several religious organizations or sects. Second, we consider the "supply" side of religious organizations, by allowing religious entrepreneurs to optimally choose the characteristics of their organization in order to maximize their profits.

6.1 Multiple sects and religious tolerance

In this section we ask whether several religious organizations can coexist in one society, and what may be the differences between such organizations. In reality we observe religious groups that differ across various dimensions. One dimension is the identity of the one (or many) who may inflict the punishments or rewards that individuals experience. Another dimension is the specifics of the religious practices, while more subtle differences may relate to the intensity of the rituals (as is the case for example between Reform and Orthodox Jews, or between Protestant and Catholic Christians). Our model only allows us to distinguish between different religious groups according to the intensity of their rituals r and according to their behavior, both internally and towards other affiliations.

Note that there always exist in the model equilibria with two different religious organizations. For example, consider an equilibrium with a cohesive religion as in Example

1; an equilibrium is sustained if we split the religion into two groups or "sects", maintain the original cost r for both groups, and assume that members of one group cooperate with the other group. Similarly, one can split a non-cohesive religion as in Example 2 into a cohesive and a non-cohesive sects, such that anyone who cooperates internally also cooperates with the members of the other sect, and maintain an equilibrium.[13] But, there may also exist equilibria with one cohesive and one non-cohesive sects such that some who cooperate internally defect against members of the other sect.

Thus, given the coexistence of multiple sects, the key question is how do members of one sect or religious organization behave towards members of the other religion. Obviously, agents below $\underline{q}$ or above $\bar{q}$ cooperate or defect respectively, disregarding their opponent's identity. It is, as above, the affiliated intermediate types who may alter their behavior towards agents of other affiliations. We say that a sect or a religion is "tolerant" ("intolerant") towards another sect if the (some) members of the sect who cooperate internally, cooperate (defect) when matched with members of the other sect. The next result illustrates that whether sects are tolerant towards one another depends on their relative level of cohesiveness:

PROPOSITION 7 *(i) Suppose that several religious organizations coexist in equilibrium. Whenever there is non-tolerance between a cohesive and a non-cohesive organization, then the cohesive one must be non-tolerant. (ii) There exists an equilibrium with one tolerant non-cohesive religion and one intolerant cohesive religion.*

PROOF: (i) Consider a cohesive religion and a non-cohesive religion. If the cohesive religion is tolerant, so that all its intermediate types cooperate with the non-cohesive religion, then it must be that all intermediate types affiliated with the non-cohesive religion will cooperate with members of the cohesive religion, since it is their best response. Thus, to sustain intolerance between non-cohesive and cohesive religions, it must be that the cohesive one is intolerant. (ii) See the appendix.■

In the appendix we prove the existence of an equilibrium with a tolerant non-cohesive

[13] The way to do so without imposing additional equilibrium conditions is to have, in the non-cohesive sect, the measure ρ^r_{cd} of the original religion and only types below $\underline{q}$ in the cohesive one.

religion and an intolerant cohesive religion. By (i) however, the opposite equilibrium configuration cannot be sustained. Thus, the relation between different sects is somewhat hierarchical, with the more rigorous sect being less tolerant towards more lenient sects.

6.2 Religious entrepreneurs

We next discuss the possibility of adding to the model the supply side of religion. In particular, we can analyze how religious entrepreneurs, who possibly enjoy some portion of the costs r that are born by the members of the religion, accommodate the demand for religion.

The term 'religious entrepreneurs' should not be taken at face value. It is a methodological device that allows to formally capture the social rents generated by religion and to refine the set of equilibria analyzed in Section 4. The question of who or what forces generate these institutions, be they individuals, groups or supernatural beings, may depend on the historical or social context.

Consider then a single religious entrepreneur (a monopoly). Let $R(r)$ denote the demand for a religious organization, as characterized by Examples 1 and 2.[14] In the first stage of the game, the entrepreneur initiates one religious group by choosing r, to maximize his revenues represented by $rR(r)$. Following the entrepreneur's announcement, agents choose whether to pay r or not, and subsequently play the PD game.

As the demand $R(r)$ depends on the characteristics of $F(.)$, we can find examples in which the entrepreneur prefers to choose a cohesive religion, which is small but raises relatively large revenues per member, and examples in which the entrepreneur chooses a non-cohesive religion, or a redundant religion, which is more encompassing but raises lower revenues per member.

Assume however that the religious organization allows for "isolationism", i.e., it allows

[14] The demand $R(r)$ is well defined when for any r there exists a unique religious organization, which is the case if F satisfies some conditions (for example, in the cases of redundant and cohesive religions, there exists a unique equilibrium for any r whenever F is concave over $[\underline{q}, \bar{q}]$). Alternatively, one can assume that the entrepeneur specifies r and the behavioural aspects of the religion, i.e., announces a particular equilibrium among the ones characterized by Examples 1 and 2, and the population coordinates on this equilibrium. In that case the demand is well defined although it is not only a function of r.

its members to use the social rituals not only to identify whom they play against, but also to avoid interaction with non-members. If the religious organization can take this form, then it is easy to show for any $F(.)$, the optimal religion chosen by the entrepreneur is actually a cohesive religion.

The possibility of isolationism implies therefore a smaller and a cohesive religion. To see the intuition, note that with isolationism, an entrepreneur who considers a non-cohesive religion is limited in his ability to increase the demand for religion by reducing its price. This arises as when a secular agent considers joining the non-cohesive religion following a decrease in the price, he is aware that many other secular agents might join as well. This deters him from joining the religion as when affiliated, he is bound to interact with these agents. On the other hand, without isolationism, such agent ignores how many other secular agents join the religion as he will interact with them whether he is affiliated or not. This difference in the elasticity of demand results in the non-cohesive religion being less attractive for the entrepreneur when isolationism is possible.

We now briefly discuss several extensions relating to the introduction of religious entrepreneurs, which can offer some directions for future research.

Competition between entrepreneurs: A reasonable conjecture is that a competition between religious entrepreneurs leads to religions with intolerant attitudes towards each other. In the competition for members, an intolerant religious group that otherwise shares its attributes with the tolerant religious group, may be an attractive choice for the entrepreneur. Joining the intolerant religion is *de facto* a membership in both religions as such a member is guaranteed the cooperation of the other religion. On the other hand, membership in the tolerant religion does not guarantee that. Consequently, individuals will tend to join the intolerant religion, resulting in higher revenues for the "intolerant" entrepreneur.[15]

Political power of religious entrepreneurs: Religious leaders can sometimes translate their rents into political power or into political success for the group they are heading.

[15] When we model comptition between entrepeneurs, we have to consider some methodological problems related to multiplicty of equilbria and coordination between individulas.

In our model, 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 appropriate. Alternatively, the entrepreneur can demand that group members exert effort towards a political goal. For example, the group can demonstrate for or against a particular law. The members might be asked to participate in a demonstration over a position that is not necessarily aligned with their preferences. But individuals agree to do so as such costly participation is part of the general costs r and is thus crucial for them in order to be identified by others in the group. The power of the leader arises therefore from his ability to supply the need of the individuals to coordinate on a particular institution. Our model can possibly be extended in future research to analyze the scope of such power.

References

[1] G. M. Anderson (1988), "Mr. Smith and the Preachers: The Economics of Religion in the Wealth of Nations", The Journal of Political Economy, Vol. 96, No. 5. (Oct., 1988), pp. 1066-1088.

[2] Benabou, R. and J. Tirole, "Belief in a Just World and Redistributive Politics," Quarterly Journal of Economics, forthcoming.

[3] Berman, E. , (2000) A Sect, Subsidy and Sacrifice: An Economist's View of Ultra-Orthodox Jews,Quarterly Journal of Economics.

[4] Berman and Iannaccone, "Religious Extremism: the good, the bad and the deadly,", Public Choice (forthcoming).

[5] Boyer, P. (2001), Religion Explained, Basic Books, NY, NY.

[6] Dawkins, R. (1993) "Viruses of the Mind", In Bo Dalbom, ed., *Dennett and His Critics.* Oxford: Blackwell, pp. 13-27.

[7] Dennett, D. C. (2006), *Breaking the Spell: Religion as a Natural Phenomenon,* Penguin, Allen Lane.

[8] Dixit A.; G. Grossman (1984) "Directly Unproductive Prophet-Seeking Activities", The American Economic Review, Vol. 74, No. 5. (Dec., 1984), pp. 1087-1088.

[9] Durkheim, E. (1912), The Elementary Forms of Religion, reprinted in A Reader in the Anthropology of Religion, Michael Lamback (ed.), Blackwell publishing, 2005.

[10] Greif, A. (1993), Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition The American Economic Review > Vol. 83, No. 3 pp. 525-548.

[11] Iannaccone, L. (1992), "sacrifice and stigma: reducing Free-riding in cults, communes and other collectives", Journal of Political Economy, Vol 100, No. 2, pp. 271-291

[12] Stark, R., L. R. Iannaccone, R. Finke (1996) "Linkages Between Economics and Religion, Religion, Science, and Rationality", The American Economic Review, Vol. 86, No. 2, Papers and Proceedings, pp. 433-437.

[13] Wilson, D.S. (2002), Darwin's Cathedral, Chicago University Press.

Appendix

Lemma 2 below will be useful for the proofs of Propositions 3 and 4.

Lemma 2 *In all equilibria with one religious organizations, either there exist types below $\underline{q}$ who are religious, or the religious organization is composed of an interval of intermediate types $[q^1, q^2]$, for $q^1 > \underline{q}, q^2 < \bar{q}$, who all cooperate internally and defect against seculars.*

Proof: to be completed.

Proof of Proposition 3: Let $\hat{\rho}_{cd}$ be the net share of intermediate types who cooperate against religious and defect against seculars (disregarding their own affiliation). If $\hat{\rho}_{cd} \leq 0$, then no one pays for the religion. Thus, it has to be that $\hat{\rho}_{cd} > 0$.

Suppose now that all agents in the religion defect against one another. In this case, the religion must then include only agents with types above $\underline{q}$. If there are some types above $\bar{q}$ who are in the religion, then $r \leq (b-a)\hat{\rho}_{cd}$. As there are no types below $\underline{q}$, it must be that $r \geq (d-c)\hat{\rho}_{cd}$. By strategic complementarities, this cannot be satisfied and hence there cannot be any types above $\bar{q}$ in the religion. Thus, all religious types can cooperate internally, a contradiction. Thus, the religion must be composed only of agents with types below $\bar{q}$, as each agent must cooperate internally with some probability. But then agents will cooperate internally with probability one.

By the proof of Lemma 2, if the religion includes also types below $\underline{q}$, then the religion must be comprised of a connected interval. If not, again by Lemma 2, it is comprised of a connected interval of intermediate types.■

Characterizing the equilibrium in example 1: In this equilibrium, seculars defect against seculars. As all seculars have $q > \underline{q}$, this can hold in equilibrium. In order for types below $\underline{q}$ to be willing to pay for the religion, and for types above $\bar{q}$ to prefer not to pay, we need that

$$\rho^r_{cd}(b-a) \leq r \leq \rho^r_{cd}(d-c) \tag{1}$$

Consider now the intermediate religious types. If they become secular, they defect against seculars. If they still cooperate against religious, then the condition $r \leq \rho^r_{cd}(d-c)$

insures that they pay for the religion. Consider the case in which they defect against religious agents once they become seculars. They do so if $q^i \geq \hat{q}(\rho^r_{cd})$, for

$$\hat{q}(\rho^r_{cd}) = \frac{1}{2\varepsilon}[\frac{\rho^r_{cd}}{\rho^r_{cd}+\rho_c}(c-a) + \frac{\rho_c}{\rho^r_{cd}+\rho_c}(d-b)]$$

and they are willing to pay for the religion if $q^i \leq q'(\rho^r_{cd}, r)$ where:

$$q'(\rho^r_{cd}, r) = \frac{1}{2\varepsilon}[\frac{\rho^r_{cd}}{\rho^r_{cd}+\rho_c}(d-a) + \frac{\rho_c}{\rho^r_{cd}+\rho_c}(d-b) - \frac{r}{(\rho_c+\rho^r_{cd})}]$$

By (1), $\hat{q}(\rho^r_{cd}) \leq q'(\rho^r_{cd}, r) \leq \bar{q}$ for any r and ρ^r_{cd}. This implies that at least some of the intermediate agents who will defect against religious agents if they become secular, are willing to pay r, and that moreover, all agents with $q^i \leq q'(\rho^r_{cd}, r)$ are religious and all with $q^i > q'(\rho^r_{cd}, r)$ are secular. The equilibrium conditions are therefore (1) and:

$$\rho_c + \rho^r_{cd} = F(\frac{1}{2\varepsilon}[\frac{\rho^r_{cd}}{\rho^r_{cd}+\rho_c}(d-a) + \frac{\rho_c}{\rho^r_{cd}+\rho_c}(d-b) - \frac{r}{(\rho_c+\rho^r_{cd})}]) \tag{2}$$

To see that an equilibrium exists, plug $r = x\rho^r_{cd}$ for some $x \in [b-a, d-c]$ in (2). We then have:

$$\rho_c + \rho^r_{cd} = F(\frac{1}{2\varepsilon}[\frac{\rho^r_{cd}}{\rho^r_{cd}+\rho_c}(d-a-x) + \frac{\rho_c}{\rho^r_{cd}+\rho_c}(d-b)]). \tag{3}$$

At $\rho^{\max}_{cd}$ ($\rho_{cd} = 0$) the lhs of (3) is larger (smaller) than the rhs. Thus, by continuity, a fixed point exists. As $d-a-x < d-b$, the rhs of (3) decreases in ρ^r_{cd}, so that: (i) for a fixed x there exists a unique fixed point ρ^r_{cd}; (ii) $\frac{d\rho^r_{cd}}{dx} < 0$. The largest religion is therefore attained when $x = b-a$ (this religion has $\rho^r_{cd} = \rho^{\max}_{cd}$), and the smallest religion is attained when $x = d-c$, in which case $q'(\rho^r_{cd}, r) = \frac{1}{2\varepsilon}[\frac{\rho^r_{cd}}{\rho^r_{cd}+\rho_c}(c-a) + \frac{\rho_c}{\rho^r_{cd}+\rho_c}(d-b)]$.■

Proof of Proposition 4: to be completed.

Characterizing the equilibrium in example 2: The conditions are:

$$r = \rho^r_{cd}(b-a); \tag{4}$$

$$\rho_c + \rho^r_{cd} = F(\frac{1}{2\varepsilon}[\frac{\rho^r_d}{\rho^r_{cd}+\rho_c+\rho^r_d}(c-a) + \frac{\rho_c+\rho^r_{cd}}{\rho^r_{cd}+\rho_c+\rho^r_d}(d-b)]). \tag{5}$$

Given some r, (4) implies that types above q are indifferent between joining the religion or not (as they only influence the behavior of ρ^r_{cd} by joining the religion). This condition

also insures that all other religious members as characterized in the example are willing to pay this cost. Thus, r determines ρ_{cd}^r. Condition (5) determines therefore the level of ρ_d^r which is consistent with a share of $\rho_c + \rho_{cd}^r$ willing to internally cooperate. Plugging (4) in (5), one can see that an equilibrium exists (and is unique) for some range of r, as the right-hand-side of (5) is decreasing in ρ_d^r whereas the left-hand-side is fixed.

PROOF OF PROPOSITION 5: (i) Consider a correlated positive shock, and an equilibrium with a cohesive religion as described in Example 1, with a cutoff at $q'' \in (\underline{q}, \bar{q})$. Given a distribution over beliefs F, the equilibrium satisfies (2), which can be written as:

$$F(q'') = F(\frac{1}{2\varepsilon}[\frac{F(q'') - F(\underline{q})}{F(q'')}(d-a) + \frac{F(\underline{q})}{F(q'')}(d-b) - \frac{r}{F(q'')}]) \Leftrightarrow \tag{6}$$
$$q'' = \frac{1}{2\varepsilon}[\frac{F(q'') - F(\underline{q})}{F(q'')}(d-a) + \frac{F(\underline{q})}{F(q'')}(d-b) - \frac{r}{F(q'')}].$$

Moreover, for any F, when $q'' = \underline{q}$,the rhs is greater than the lhs of (6).

All those below $\underline{q}$ were cooperating and by Lemma 1 will end up with an even lower q. All those above q'' were defecting and will end up with a higher q. A fraction of the types in $(\underline{q}, q'')$ will end up with lower q and some with higher q. Thus, the new distribution over beliefs $F'()$ satisfies $F'(q'') < F(q'')$. This implies that the rhs of (6) decreases, and thus the fixed point equation has a solution at a lower cutoff, decreasing the size of the religion.

(ii) Consider now a negative correlated shock and an equilibrium as described above. All those above q'' will decrease their q. Types in $(\underline{q}, q'')$ who interact with religious will increase their q and those who interact with seculars will decrease it. The resulting value of $f'(q'')$ will depend on the net flow of those crossing q'', which in turn depends on the original density function around q'', the relative size of the religion (which determines how many individuals in $(\underline{q}, q'')$ defect and how many cooperate), and the impressionability of voters around the cutoff.■

PROOF OF PROPOSITION 7(ii): to be completed.