Rituals or Good Works:
Social Signalling in Religious Organizations

Gilat Levy and Ronny Razin, LSE[1]

Abstract: We develop a model of social signalling of religiosity and cooperative behaviour in religious organizations. The model embeds a ritual-based religious organization in which signalling arises through the use of costly rituals, and a discipline-based religious organization in which such signalling occurs through the monitoring of past behaviour. We use this framework to contrast -positively and normatively- these two forms of social signalling. We show that ritual-based religions, while using a costly and wasteful signal, also imply a higher level of coordination of behaviour in social interactions and a higher incidence of mutual cooperation. Our welfare analysis suggests that communities are more likely to support a switch to a discipline-based religion if strategic complementarities are high, and if there is sufficiently high level of public information about social behaviour. This accords with the success of Calvin's Reformation in Switzerland and France, a process characterized by the reduction of rituals along with the creation of institutions to monitor and publicise individuals' behaviour, such as the Consistory.

[1]We thank the editor and three anonymous referees for their comments. We also thank the ERC for financial support (grant no 210385).

1 Introduction

Religious beliefs typically place high value on pro-social behaviour through different theological systems.[2] Beliefs in rewards and punishments, whether in this life or the afterlife, are rife in many ancient and modern religions, and create an incentive to properly behave in a social context. Even Calvinistic beliefs which emphasise predestination, as Weber (1904) first recognized, may constitute an incentive for good works; this arises because an individual wishes to glean information about whether she will be salvaged and doing good deeds provides a positive signal about the individual's future.[3]

By enabling good behaviour, religious beliefs and religious organizations may also induce individuals to signal their ethics and religiosity to others. Adam Smith observes that religions tend to produce and distribute moral information about their members which allows traders to assess the risk involved in conducting business with them.[4] Weber (1906) writes of the social pressure in American Protestant communities, "*Unqualified integrity, evidenced by, for example, a system of fixed prices in retail trade...appears as the specific, indeed, really the only, form by which one can demonstrate his qualification as a Christian and therewith his moral legitimation for membership in the sect...admittance into the Baptist congregation was primarily of decisive importance...because of the on-going inquiries about moral and business conduct*". In this paper we compare two different mechanisms by which religious organisations may enable social signalling of religiosity and ethical behaviour.

A recent literature has focused on costly rituals as signals of religiosity. While religious rituals may perform several roles, their costly and public nature renders them suitable for signals of religious conviction. Iannaccone (1992, 1998) and Berman (2000) show how rituals allow religious groups to screen those who are less devout, and Levy and Razin (2012) show how costly and public rituals allow individuals to signal religiosity and hence good behaviour in social interactions.[5] While the above mechanism might be present in religions with a strong ritualistic emphasis, other religions may rely on observed behaviour instead, as the description by Weber (1906) above indicates. This theory is explored in Glaeser and Glendon (1998) who show how Protestant beliefs may lead individuals to signal their religiosity by taking actions that contribute to the common good. Arrunada (2010) refers to this as the Protestant "social ethic".

Religions which orchestrate behaviour in the social sphere using these two different systems

[2] This is the motivation behind several studies investigating the relation between religiosity and economic performance, such as Barro and McCleary (2003), Huber (2004) and Glaeser and Glenson (1998).

[3] On this Weber (1904) writes: "*The question: Am I one of the elect? must sooner or later have arisen for every believer*".

[4] See Anderson (1988).

[5] See also Chwe (2003).

of social signals might induce different distributions of social behaviour and economic outcomes. A community that relies on signalling by good behaviour might reap more benefits compared with one that uses a costly or a wasteful ritual. On the other hand, a costly ritual may do a better job at screening out individuals with low moral standards. Our aim in this paper is to compare these two mechanisms by focusing on their behavioural and normative implications.

This comparison is especially pertinent in the context of the Reformation of the Catholic church in the 16th century, and specifically that of Calvin in Geneva. In medieval times, the Catholic church had evolved to have an elaborate system of rent extraction and a heavy load of rituals.[6] In contrast, the reformers significantly reduced the number of rituals or religious sacraments an individual had to attend.[7] In fact, Barro and McCleary (2003) show that to this day Catholics participate in more church rituals than Protestants.

Moreover, Calvin's Reformation in Geneva has shifted the church's emphasis to discipline. In his second spell in the city, Calvin initiated the institution of the Consistory to monitor, discipline and publicise individuals' behaviour.[8] A great deal of its function was devoted to resolving civil disputes within families, between neighbours, and between business associates. Deviant behaviour was punished by public scolding, sometimes by Calvin himself. When other communities in Switzerland and France decided to adopt Calvin's religion, he insisted on the formation of local Consistories, which are better suited to monitor local behaviour. In fact, Arrunada (2010) shows that to this day, Protestants better monitor each other's conduct compared with Catholics.[9] Calvin's emphasis on discipline -religious and civil alike- is evident in his insistence that discipline is the third mark of a good Church (this was objected to by Lutherans) and is certainly a mark of his own reign in Geneva.[10] While fear of punishment itself may trigger discipline and good behaviour, punishments for deviant behaviour often consisted of either public scolding or of being denied access to communion; the key element of the punishment was therefore its public nature (we provide a more detailed discussion of the Consistory and these issues in Section 5.1).

Finally, our welfare comparison is motivated by the explicit choice faced by city-states in Switzerland between the highly ritualistic Catholic church and Calvin's Reformation. In such autonomous city-states, the choice of which religion to adopt was often decided by a vote in the city council. There are many political and economic factors behind the decision to adopt the Reformation, and our welfare analysis highlights a new dimension along which the two

[6] See Ekelund et al (1996, 2002).

[7] Calvin had rejected the seven sacraments of the Catholic church and accepted only two sacraments as valid (Baptism and the Lord's Supper).

[8] An institution of the same name existed before but dealt mainly with marriage law.

[9] McCleary (2007) also shows that Protestants tend to trust or place obligations on others as they do with family members.

[10] We discuss the differences between Calvin's Reformation and that of Luther in Section 6.2.

religions may be compared.[11]

We analyse a simple model which allows us to consider both types of social signalling. The model is based on the premise that religions moderate cooperative behaviour and thus possibly induce enhanced material utility, through a spiritual dimension and a signalling method.[12] We assume that a population of individuals is randomly matched into pairs to play a Prisoners' Dilemma (with strategic complementarities).[13] Religious beliefs, heterogeneous and privately known, consist of a perceived spiritual benefit from cooperative behaviour.

In a *ritual-based* religion, individuals can participate in costly and public rituals, and can condition their behaviour in the Prisoners' Dilemma game on whether their opponent participates as well. In a *discipline-based* religion, individuals' behaviour in an initial round of play is publicly observed. Individuals can then condition their subsequent behaviour on whether their opponents have behaved well in the past. The model allows for a *spiritual* as well as a *material* benefit from cooperative behaviour. A spiritual benefit arises in both religions due to religious beliefs. A material benefit arises due to successful social signalling which elicits more cooperation from others.

In the ritual-based religion, the cost of the signal determines the level of participation in rituals. We show that the Pareto dominant level of rituals induces an accurate signal, that is, all those who participate in rituals also cooperate with one another. In the discipline-based religion, the cost of signalling -i.e., the loss from cooperation- is endogenous, and depends on the share of those that cooperate in the first period. We show that this induces a noisy signal in equilibrium, in which a relatively large share of individuals cooperate in the first period. These large initial cooperation levels accord with Weber's (1906) observations of the "probation" period for new members in the North American sects that descended from the Calvinistic theology.[14]

We highlight a trade-off between the accuracy of the signal and its cost. The ritual-based religion allows for a costly but an accurate signal. The discipline-based religion on the other hand induces excessive signalling, i.e., some agents who initially cooperate, defect later on to

[11] Political factors such as the declined influence of the Roman church, or economic factors relating to urbanisation and the abuses of church power, are among the explanations for the Reformation. See Flick (1930).

[12] Wilson (2002) provides examples of the secular utility in the form of social order that religious institutions provide, from early Christianity offering a mini-welfare state in the Roman Empire (see also Stark 1996), through regulation of rice production in Indonesia, to modern US churches providing a social network to its members.

[13] Numerous papers have analysed social norms when social interaction is modelled by a Prisoner's Dilemma game. Greif (1989) studies how cooperation arises due to repeated interactions. A recent literature has analysed cooperation when players sustain different norms; see for example Dixit (2003), Tabellini (2008), and Andreoni and Samuelson (2006).

[14] Weber (1906) writes: "*And the Canonical limitation of the size of the unity, the congregation, to such dimensions that all members personally know one another and, therefore, can judge and supervise their "probation" reciprocally has always been a fundamental Baptist principle.*"

take advantage of others. This leads individuals to be more suspicious and less cooperative in the second period. We show that the implication of this is that the ritual-based religion can achieve higher levels of mutual cooperation as well as higher total coordination in behaviour.

We then consider average material welfare and identify two environments in which the above trade-off is resolved in favour of the discipline-based religion. We show that if strategic complementarities are sufficiently large, then both religions induce sufficiently similar and high levels of mutual cooperation, but the ritual-based religion is strictly costly and is thus dominated. The discipline-based religion also dominates if cooperation is more beneficial than coordination (so that even one-sided cooperation yields sufficiently large gains to society compared with mutual defection). In this case the accuracy and informativeness provided by the ritual-based religion is not valued enough, and moreover, the ritual cost must be high as individuals are keen to avoid mutual defection. If on the other hand coordination is sufficiently important, the ritual-based religion can dominate.

When we consider individual preferences, material and spiritual, we show that it is individuals with relatively weak beliefs that support a switch to a discipline-based religion. These are the individuals who enjoy the positive externalities that signalling by behaviour entails. Furthermore, if some individuals prefer to maintain the ritual-based religion, they must include individuals of intermediate beliefs. Such individuals value the accuracy of the signal provided by the ritual-based religion which allows them to change their behaviour in response to others' signalling. We also show that the support for a discipline-based religion increases with the availability of public information. This result is consistent with the experience of Calvin's Reformation in Geneva (see Section 5.1).

Our analysis brings to the fore a way to distinguish religious organizations according to the type of social signalling they generate. To be sure, both types of signalling mechanisms that we analyse may be used in any religion. However, we argue that in terms of its focus, Calvin's reformation can be seen as a shift of emphasis from rituals to discipline. In this sense our work is in the spirit of Botticini and Eckstein (2005, 2007) who consider the transformation of Judaism from a religion based on sacrifices in the Temple to a religion whose core is the reading of the Torah in synagogues, and Carvalho and Koyama (2011) who consider how religious restrictions change in response to growth. Complementary to our analysis is Glaeser and Glendon (1998) who compare the free will theology to the Weber's (1904) "Protestant work ethic" that induces individuals to focus on entrepreneurial actions that are more visible.[15] In contrast, we let both religions induce the same actions, and concentrate instead on the different social signalling methods and hence, in this sense, we follow Weber (1906) instead.

[15] Kantas and Stefanadix (2010) focus instead on the comparison between pride-based moral code (such as Protestantism) and guilt-based moral code (such as Catholicism) and show that the former leads to a more favourable attitude towards work.

Our second contribution is to provide a framework for the positive and normative analysis of these two signalling mechanisms. This framework is based on the ritual-based religion we have analysed in Levy and Razin (2012); in that paper we have analysed a more general model of the ritual-based religion, with a greater focus on religious beliefs, including a dynamic version which allows for belief updating. The current paper simplifies that model to embed an alternative signalling method in order to compare between the two. Given our discussion of the Reformation, religious organizations are a natural application to evaluate differences in signalling methods, but our model and results can be interpreted more generally; for example, the literature on signalling wealth or status has also considered different signalling mechanisms, either by conspicuous consumption or by a productive activity such as charity giving.[16]

Finally, our results can shed some light on recent empirical papers that have looked at the economic implications of the Catholic and the Protestant religions. Barro and McCleary (2003) show that economic growth responds positively to the extent of religious beliefs, notably those in hell and heaven, but negatively to church attendance. Our model shows that beliefs are indeed conducive for good economic outcomes and that costly and wasteful rituals are the main determinant behind the sometimes inferiority of the ritual-based religion. Guiso *et al* (2003, 2006) show that religious beliefs are associated with more trust and better economic attitudes and that these effects are larger for Protestants than for Catholics.[17] Becker and Woessmann (2009) suggest that literacy levels can explain the better economic outcomes of Protestant (mainly Lutheran) societies; literacy may be correlated with a higher level of public dissemination of information. Cantoni (2010) however finds that overall, the growth of Lutheran and Catholic cities is roughly the same.

We present the model of the two religions in Section 2. We analyse the equilibria in Section 3. Comparative -positive and normative- analysis is presented in Section 4. In Section 5 we discuss the Consistory in more detail and consider some supply side extensions. We discuss the potential link between theology and institutions and the comparison between Calvin and Luther in Section 6. All proofs are in the appendix.

2 A model of religious organizations

We present a model which embeds two religious organisations. Our aim is to make the two specifications as close as possible to one another so as to focus the comparison on the different signalling structures. We first present their common elements: the economic environment, religious beliefs and pay-offs.

[16] For examples see Konrad and Glazer (1996) and Pesendorfer (1995).

[17] La Porta *et al* (1997) show that countries with hierarchical religions perform comparatively worse on a wide range of outcomes, which accords with Putnam (1993) who suggests that such religions deter formation of trust.

2.1 Economic environment and religious beliefs

The social interaction. There are two periods of interaction; in each period, individuals are randomly paired to play a Prisoner's Dilemma (PD) game:

	C	D
C	d, d	$0, b$
D	$b, 0$	a, a

where a, b and d are bounded parameters, satisfying $b > d > a > 0$. We assume strategic complementarities in cooperation, i.e., that $d - b > -a$. This assumption implies that the relative benefit from cooperation is greater when the opponent cooperates. The assumption of strategic complementarities is often made in the literature that focuses on how cooperation stems from preferences or moral obligations.[18] Also, recent empirical evidence shows that religious affiliation affects levels of trust in society, which is potentially captured in the PD environment; but the model can easily be extended to other types of public good games in which the interaction is not necessarily pairwise and in which strategic complementarities typically play a role.[19]

We denote with σ the level of strategic complementarities, $\sigma \equiv \frac{a}{b-d} > 1$. Note that $\sigma > 1$ also implies that $2d > b$, so that mutual cooperation is the efficient outcome. For our welfare analysis, we distinguish between the case in which $b > 2a$, i.e., when one-sided cooperation is more socially efficient than coordination, and the case of $2a > b$, in which coordination of actions among agents is more important than cooperation.

Religious beliefs. To facilitate our analysis, we assume that in both religions, the theology instills beliefs that have similar effects on behaviour: We assume that each individual i believes that if he cooperates he is rewarded with a spiritual benefit $\gamma_i \geq 0$.[20]

Such beliefs can be interpreted under both the free will and predestination doctrines in Christian Theology. Under the free will theology of the Catholic Church, good deeds affect salvation. Different γ_i types may therefore have a different perception of the benefits from salvation or alternatively the probability that private actions will affect salvation. A richer model of Catholic Theology might allow for such rewards to be conditioned not only on good works but also on church obedience and participation in rituals; we discuss this in Section 6.1.

[18] See for example Tabellini (2008) and references therein.

[19] One can easily analyse the model under the alternative assumption of strategic substitutes; the model would then identify a need to signal one's willingness to defect, and in equilibrium, religious individuals would defect against each other. This goes against the empirical and experimental findings in the literature (see Sosis and Ruffle 2004 and Iannaccone 1998).

[20] The heterogeneity of types accords with evolutionary biology theories of the "religious mind". See Boyer (2002).

In contrast, the theology of predestination implies that salvation is independent of the individual's good works. For Calvin, it is by God's grace that an individual will be salvaged. But as Weber (1904) suggested, good works can become a mean to self signal one's membership in the *elect*. A self-signalling mechanism can induce individuals to cooperate (motivated by the behavioural prescriptions put forward through the stories on Christ in the scriptures) in order to assure themselves that they belong to the *elect*. In a companion working paper, Levy and Razin (2011), we derive such a model which provides a self-signalling interpretation for the reduced-form beliefs/types γ_i.[21]

Pay-offs and the distribution over types: In any period of play, the utility of an individual will be the sum of the material and the spiritual utility. For example, the relative payoff of cooperation vs. defection is $x + \gamma_i$, where $x \in \{d-b, -a\}$ depends on opponents' actions. It would be more interesting to concentrate on the strategic interaction of agents with types $\gamma_i \leq a$ (as otherwise an agent would have a strictly dominant strategy to cooperate). We then assume that with probability $1-\zeta < 1$, an individual's type is drawn uniformly from $[0, a]$. An individual with weak convictions, or "non-believer", would have $\gamma_i < b - d$ and thus a strictly dominant action to defect (although the possibility of signalling might change his behaviour). An individual with a conviction $\gamma_i \in [b-d, a]$, or a "believer", prefers to cooperate if his opponent cooperates for sure. More generally, his best response is to cooperate if the likelihood of facing cooperation is high enough whereas this likelihood decreases with γ_i. With the remaining probability $\zeta > 0$, the individual is a behavioural type who always cooperates. Similarly to the reputation literature, we assume that the fraction of behavioural types ζ is relatively small compared to the believers so that $\zeta < \bar{\zeta}$ and $a > \underline{a}$ for some $\bar{\zeta}, \underline{a} > 0$.[22]

A benchmark: Prior to introducing the possibility of signalling, we analyse the "autarky" benchmark of the two PD games, where no signalling arises. In other words, when individuals cannot condition their behaviour on any information:

Lemma 1: *In the unique Bayesian Nash equilibrium, there exists $\hat{\gamma} \in (b-d, a)$ such that all individuals i with $\gamma_i < \hat{\gamma}$ defect and all those with $\gamma_i > \hat{\gamma}$ cooperate in both periods of the PD games, where $\hat{\gamma}'(\zeta) < 0$, $\lim_{\zeta \to 0} \hat{\gamma} \to a$ and $\hat{\gamma}$ solves*

$$\frac{a - \hat{\gamma} + a\frac{\zeta}{1-\zeta}}{a + a\frac{\zeta}{1-\zeta}}(d-b) + \frac{\hat{\gamma}}{a + a\frac{\zeta}{1-\zeta}}(-a) + \hat{\gamma} = 0$$

[21] A recent literature has other, related, models of religious beliefs. Benabou and Tirole (2006, 2011) assume that agents differ in their beliefs with respect to how much hard work is rewarded, in this or in the afterlife, and actively choose to maintain such beliefs. In Scheve and Stasavage (2006) on the other hand, religious beliefs allow for a psychic benefit in bad times. Alaoui and Sandroni (2013) show an equivalence result between the utility functions of secular agents who have a moral obligation to accumulate wealth and those of religious agents who believe in Calvin's version of predestination.

[22] $\bar{\zeta}$ is derived from the proofs of Propositions 1 and 2.

When no signalling arises, individuals either cooperate or defect disregarding how others behave. As there is a share of behavioural types who always cooperate, some type with high enough reward from cooperation will cooperate as well, and in particular the type at the cutoff $\hat{\gamma}$ is indifferent between cooperating and defecting when all types above him are cooperating and all types below him are defecting. Note however that when ζ is small and specifically when $\zeta \to 0$, then $\hat{\gamma} \to a$ and no meaningful levels of cooperation can be sustained.

2.2 Signalling: rituals or discipline

Our model identifies a need for social signalling to increase coordination on cooperative outcomes. We now describe two signalling methods. One signalling method allows for information about behaviour in the PD game to be transmitted across the periods of play. Another signalling technology allows only for information about auxiliary activities -such as rituals- to be available for players in the PD game.

Signalling by discipline. In this specification we assume that in the second period, players have information on the first-period behaviour of their opponents. Players can use this information to make inferences about their opponents' religious types. We term a religion that has this (and only this) signalling technology a *discipline-based religion* (D). For now we assume that players are fully informed about the first period behaviour of opponents and in section 5.1 we consider lower levels of information.

Signalling by public rituals. In this specification we assume that individuals can participate in a ritual which is costly and observable. The cost of the ritual is denoted by $r \geq 0$. Players can then use the information on their opponents' choice of ritual participation to make inferences on their religious type. We term a religion that has this (and only this) signalling technology a *ritual-based religion* (R). To keep it simple, we assume that ritual participation can occur only once, prior to playing the two PD games.[23] For now we treat r as an exogeneous variable (we discuss supply side considerations in Section 5.2).

To recapitulate, the timing of the game is therefore as follows:

Period 0: In the R religion, individuals choose whether to pay r.

Period 1: Individuals are randomly matched with an opponent to play the PD game. In the R religion, before taking an action, they observe whether the opponent paid r.

[23] Allowing for just a single choice of ritual participation simplifies matters technically but note that the results will be qualitatively similar if individuals can pay r before any period of the PD game, as the value of r will adjust to reflect the value of signalling.

Period 2: Individuals are randomly matched with a new opponent to play the PD game. Before taking their action, in the R religion they observe whether the opponent had paid r, in the D religion they observe the opponent's Period 1 behaviour.

We assume no discounting; the payoff of an individual from the R religion is therefore her payoff from the PD games (material and spiritual), minus the cost of rituals if she chooses to participate, whereas the payoff of an individual in the D religion is her payoff in the two PD games.[24]

Remark 1: *The taxonomy of religious organizations.* For modelling purposes we have focused on two social signalling technologies: one by ritualistic participation and the other by discipline. Obviously this is a simplification as most religions or other social organizations will make use of both these technologies. Religious leaders in a ritual-based religion might also make use of institutions such as the Consistory and discipline-based religions might also have some ritualistic prescriptions. Still, organizations will differ in the focus they place on each signalling technology (as in the case of the Catholic and Protestant churches in the 16th century), and the analysis of each technology separately can shed light on such institutions.

Note that the two religions we analyze in the text can be constructed as extreme cases of a unified model, in which individuals can both pay r and observe first-period information. At one extreme, information on first-period behaviour is not available, and individuals can only use r to signal, which will mimick the R religion. At the other extreme, the value of r does not allow for meaningful signalling but information about first-period behaviour is available, which will mimick the D religion. In Section 5.3 and in Appendix B we analyse hybrid religions which involve both types of signalling technologies. We show that when both signals are used meaningfully, the gist of our results is maintained. We also discuss supply side considerations that might induce some religious leaders to choose one signalling institution rather than the other, and relate the choice of signalling methods to potential geographical and technological constraints, and to the different theologies.

Equilibria: We focus on Perfect Bayesian Equilibria in each religion. It is in general true that when endogenous population signalling games are considered, the benefit from signalling is not necessarily monotone in one's type, yielding equilibria with no generally defined charac-

[24] The choice of two interaction periods of the PD game is not an important assumption; it is the smallest number of periods that allows for signalling in the D religion, and for the sake of comparison we have the same number of interaction periods in each religion. As long as the information about past behaviour in the D religion is just about Period 1, all of our results could be easily extended to n periods of interaction. If this "probation period" is extended, the analysis becomes more complicated but qualitatively the results will have a similar nature.

teristics or with perverse forms of signalling.[25] In Levy and Razin (2012) we introduce a "belief activation" assumption which refines the set of social signalling equilibria to only monotone equilibria with potentially positive benefits from signalling behaviour. Such monotone equilibria imply that an individual who does not signal, i.e., does not participate in rituals in the R religion or defects in the first period in the D religion, will defect in all remaining PD games. For the sake of simplifying the model we will henceforth restrict attention to such monotone equilibria. We provide the details of the refinement in Appendix C.

3 Social signalling in the two religions

We start by describing the general features of equilibria in both models. We say that an agent *does not signal* if he does not pay r in the R religion, or defects in Period 1 in the D religion. In a monotone equilibrium, every agent who does not signal, indicates a clear intention to defect in the (remaining) PD games. This leaves us with only two types of equilibria, differing in terms of the intentions to cooperate of those who do signal.

In an equilibrium with *accurate signalling*, all agents who signal in the "signalling period" also cooperate. That is, in R, all agents who pay r also cooperate in the PD games in periods 1 and 2, and in D, all those who cooperate in period 1, also cooperate in period 2.

An equilibrium with *excessive signalling* implies that signalling is noisy, so that some agents who signal actually defect. In this case, non-believers also engage in signalling, in order to take advantage of their good reputation and defect while gaining cooperation from others. This is summarised in the Lemma below:

Lemma 2: *In both religions, there are only two types of monotone equilibria: (i) In an equilibrium with accurate signalling, there exists a cutoff* $\gamma^* \in [b-d, a)$*, such that all agents above* γ^* *and only agents above* γ^* *signal. In the remaining PD game(s), all below* γ^* *defect with all, and all in* $[\gamma^*, a]$ *cooperate with those who had signalled and defect against those who had not signalled. (ii) In an equilibrium with excessive signalling, there exist two cutoffs,* $\gamma_1 < b-d$ *and* $\gamma_2 \in (b-d, a)$*. All above (below)* γ_1 *signal (do not signal). In the remaining PD game(s), all below* γ_2 *defect against all, all those in* (γ_2, a) *cooperate with those who had signalled and defect against those who did not signal. Given* γ_1*,* γ_2 *is the unique solution to:*

$$\frac{(a-\gamma_2)+a\frac{\zeta}{1-\zeta}}{(a-\gamma_1)+a\frac{\zeta}{1-\zeta}}(d-b)+\frac{(\gamma_2-\gamma_1)}{(a-\gamma_1)+a\frac{\zeta}{1-\zeta}}(-a)+\gamma_2=0. \tag{1}$$

In the equilibria above, a spiritual as well as a material benefit arises for social signalling. A spiritual benefit arises as individuals who signal also tend to cooperate which provides

[25] For population games with signalling see Kranton (1996) or Athey *et al* (2010).

them with a spiritual benefit. A material benefit arises as agents in $[\gamma^*, a]$ in the accurate signal equilibrium, or in $[\gamma_2, a]$ in the excessive signalling equilibrium, change their behaviour favourably in response to an observation of an opponent who had signalled good intentions.

To see how the material benefit is determined in the excessive signalling equilibrium, note that, as in the autarky case, the cutoff type at γ_2 is indifferent between cooperating and defecting, and hence is determined according to the fixed point equation (1). In this equation, $\frac{(a-\gamma_2)+a\frac{\zeta}{1-\zeta}}{(a-\gamma_1)+a\frac{\zeta}{1-\zeta}}$ is the share of those who had signalled and will cooperate against those who had done so as well, and $\frac{(\gamma_2-\gamma_1)}{(a-\gamma_1)+a\frac{\zeta}{1-\zeta}}$ is the remaining share of those who had signalled but will defect. A unique solution arises with $\gamma_2 < a$ so a material benefit exists. Note that γ_2 is decreasing in γ_1, i.e., when signalling becomes more excessive, agents are more suspicious and less willing to cooperate later on (specifically, when $\gamma_1 = 0$, then $\gamma_2 = \hat{\gamma}$ identified in Lemma 1). Finally, the cutoffs γ_1 and γ^* will be determined according to the specific signalling method, which we analyse next.

3.1 Social signalling in the discipline-based religion

The next lemma characterizes the equilibria in the D religion.

Lemma 3: *There exists a unique equilibrium, which is characterized by excessive signalling.*

In the D religion the cost of signalling is endogenous and is determined by the measure of agents who cooperate in period 1. This, as we show in the appendix, implies that signalling must be excessive; all types above $b-d$ would rather pay the cost of cooperating, given the spiritual benefit they gain from cooperation and future cooperation with the types above them.

To determine γ_1, the type who is indifferent between cooperating and defecting in the first period, note that the following fixed point equation has to be satisfied:

$$\underbrace{(\zeta + (1-\zeta)(\tfrac{a-\gamma_1}{a}))(d-b) + (1-\zeta)\tfrac{\gamma_1}{a}(-a) + \gamma_1}_{\text{Period 1 Difference in Expected Payoff}} + \underbrace{(1-\zeta)(\tfrac{a-\gamma_2}{a})(b-a)}_{\text{Period 2 Difference in Expected Payoff}} = 0 \quad (2)$$

The second period difference in expected payoff between cooperating and defecting is composed of the benefit from changing the behaviour of other agents to be cooperative, while planning to defect. The first period difference is the endogenous cost of signalling by cooperation conditional on all above γ_1 cooperating. It is easy to show that given (1), γ_1 has a unique solution in $[0, b-d)$.

Endogenous signalling pins down a unique excessive signalling equilibrium with period 1's discipline effect of good behaviour, which induces even those below $b-d$ to cooperate. But this creates much lower cooperation levels in period 2 as believers are aware that some agents will defect and are thus less willing to cooperate themselves. The next example shows that

this might imply that the average level of cooperation across the two periods is relatively low. In Proposition 1 below we will generalize this result.

Example 1: Assume that the PD pay-offs are given by:

	C	D
C	3, 3	0, 4
D	4, 0	2, 2

In the limit, when $\zeta \to 0$, the equilibrium conditions (1) and (2) imply that γ_1 is close to 1 and $\gamma_2 \simeq 1.5$. The distribution over outcomes is reported in the tables below. For example, in the first period, mutual cooperation arises when a type above γ_1 meets another type above γ_1. The share of such meetings in the population is $(\frac{a-\gamma_1}{a})^2 \simeq 0.25$. Note that potentially, all believers (above $b-d=1$) might cooperate (i.e., in 25% of the matches). This is indeed the case in Period 1, but cooperation is substantially lower in period 2:

period 1	C	D
C	25%	25%
D	25%	25%

period 2	C	D
C	6%	6%
D	6%	72%

average	C	D
C	15%	15%
D	15%	55%

3.2 Social signalling in the ritual-based religion

In the ritual-based religion, the cost of signalling is exogeneous and is determined by r.[26] The following Lemma characterizes the equilibria in the R religion:

Lemma 4: *For any $\gamma \in [0, a)$, there exists a ritual cost r so that all types above γ will participate in rituals.*

If $\gamma > b-d$, the equilibrium will be an accurate one, and otherwise it will be excessive. For equilibria with excessive signalling, the type at the cutoff γ_1 plans to defect against all. Paying r grants him the additional cooperation of all types in $[\gamma_2, a]$:[27]

$$r = 2(1-\zeta)(\frac{a-\gamma_2}{a})(b-a)$$

For any $\gamma_1 \in [0, b-d]$, γ_2 is as determined in (1), and we can then find the r that will support this equilibrium. Note that higher levels of rituals must give rise to a lower γ_2 and as a result a higher γ_1, and thus serve to improve the informativeness of the signal.

[26] For concreteness we assume that behavioural types, who always cooperate, also participate in rituals. This would also be the case if these types were rational. As the measure of behavioural types is small this assumption is without consequence to our results but simplifies the exposition.

[27] Note that the equilibrium is equivalent to an equilibrium in which a share $\frac{\gamma_2-\gamma_1}{a}$ of agents below γ_2 participate in rituals but defect, as in equilibrium all agents below γ_2 are indifferent between paying r or not. For concreteness we describe this equilibrium as one with a cutoff.

Consider now the accurate equilibrium, where for any cutoff $\gamma^* \geq b - d$, the relevant cost will satisfy:

$$r = 2\zeta(d - b + \gamma^*) + 2(1 - \zeta)(\frac{a - \gamma^*}{a})(d - a + \gamma^*)$$

The cost makes the cutoff type γ^* indifferent between paying r or not. Paying r has two effects on an agent's behaviour in the two periods of play. First, he cooperates against all behavioural types instead of defecting, which provides a relative reward of $d - b + \gamma^*$. Second, all types in $[\gamma^*, a]$ become cooperative, and he cooperates with them, which provides a relative reward of $d - a + \gamma^*$. Clearly for any cutoff $\gamma^* \in [b - d, a)$ we can find a cost level r which will support such an equilibrium.

Example 1 revisited: Consider again Example 1 and the accurate equilibrium with the largest participation, i.e., when $\gamma^* = b - d = 1$. In the limit, when $\zeta \to 0$, $r \simeq 2$, and the distribution of play is reported below. Note that all the potential for mutual cooperation is realized:

R	C	D
C	25%	0%
D	0%	75%

As the level of r is exogeneous, for all parameters, a continuum of equilibria exist in the ritual-based religion. To facilitate our comparison with the discipline-based religion, we focus on one particular equilibrium, the accurate signalling equilibrium with the largest participation as described in Example 1 above. This equilibrium is in the closure of both the accurate and the excessive signalling sets. We can also show:

Lemma 5: *The accurate equilibrium in which* $\gamma^* = b - d$ *and* $r^* = 2(1 - \zeta)(\frac{a-(b-d)}{a})(b - a)$ *is the unique Pareto dominant equilibrium whenever there exists some strictly positive measure of agents who do not participate in rituals.*[28]

To see why this equilibrium is Pareto dominant, consider for example the set of all excessive signalling equilibria. In these equilibria, the price is determined by the marginal type who is indifferent between paying r or not conditional on *defecting*, but for those who do participate, the gain from participation in rituals is conditional upon *cooperating*, which by strategic complementarities is higher. It is therefore worthwhile for them to pay a higher price for a less noisy signal and higher cooperation. On the other hand, those who never participate in rituals or never cooperate have the same utility across all equilibria (namely the utility of being identified and gaining cooperation only from behavioural types).

[28] In the limit when $\zeta \to 0$, indeed in all equilibria there is some strictly positive measure of individuals who do not participate in rituals.

The equilibrium above allows us to identify the trade-off between accuracy and cost; some of our results below extend to all values of r and the others will be robust to small deviations from r^*. Moreover, Pareto dominance may imply that religious leaders facing competition might wish to choose this level of rituals, and as we show in Section 5.2, in some environments, this equilibrium also maximises the revenue that can be extracted from agents from religious organization. Henceforth, in the R religion, we focus on the equilibrium described in Lemma 5.

4 Good works or rituals?

The above two organizations offer two different channels for individuals to publicly signal their religious convictions and their future behaviour. In this section we compare the equilibria of the two organizations in terms of the different distributions of behaviour they induce, and in terms of their normative implications.

We start by considering the positive implications of the two religions, where we compare behaviour in the two periods of the PD game in each religion. Next, we consider average material welfare, which also takes into consideration the cost of rituals. Average material welfare may be a relevant welfare criterion when one considers the long term survival of religious organizations.[29] There is also a substantial empirical literature looking at economic outcomes across countries with different religions and an analysis of average material welfare can possibly shed light on these different outcomes.[30] Finally, we look at individual preferences, which include both material and spiritual pay-offs. This normative analysis is more relevant when considering the political economy of religious reforms in communities. Indeed, the decision to switch alliances and adopt Calvin's Reformation in city states in Switzerland and communities in France was often taken by a vote in one or more City Councils. Calvin's eventual success to convince the Genevan council to tie the city to the Reformed Church hinged on the council members' approval.[31]

4.1 The distribution of behaviour

As the signal in the ritual-based religion is fully accurate, it leads to full coordination among players (abstracting from the behavioural types). When believers meet each other, they have both signalled and will thus cooperate with each other. In all other matches, which involve at least one non-believer who had not signalled, the players will coordinate on mutual defection.

[29] The evolutionary game theory literature often considers the survival of different preference types-measured by pure material pay-offs. Wilson (2002) advances the ideas of "group fitness" vis a vis individual fitness.

[30] See for example Barro and McCleary (2003), Guiso et al (2003, 2006) and Cantoni (2010).

[31] See Wilson (2002).

In contrast, in the discipline-based religion, signalling is excessive implying that some miscoordination will arise. In the next Proposition we show that the R religion leads to both higher coordination and higher mutual cooperation:

Proposition 1: *There is a higher level of mutual cooperation and a higher level of total coordination (mutual cooperation and mutual defection) in the ritual-based religion.*

The proof of Proposition 1 shows that in D, the additional cooperation in period 1, the "discipline" period, is overshadowed by the reduced cooperation in period 2. To see the intuition, consider the case when $\zeta \to 0$ which results in $\gamma_1 \to b - d$. In that case, signalling (in D) in period 1 via good behaviour is relatively accurate so only a few agents below $b - d$ cooperate. This small share, who plan to defect in period 2, induces those types just above $b - d$ to defect as well. But once more agents above $b - d$ are known to defect, others with slightly stronger beliefs will defect as well; this effect snowballs and keeps γ_2 bounded away from $b - d$, resulting in lower levels of overall mutual cooperation. The proof involves showing that this argument holds uniformly in the PD parameters a, b and d.

We now turn to consider the implications of Proposition 1 for the welfare comparison between the two religions.

4.2 Average material welfare

The result above had indicated that the ritual-based religion provides not only accurate signalling, but also more instances of the socially efficient outcome. We now show that this does not necessarily translate into higher social welfare.

In the next Proposition we focus on environments in which both religions offer similar behavioural prescriptions compared to the cost of rituals in the R religion, which implies that the latter is dominated. Specifically, when σ -the level of strategic complementarities, i.e., $\frac{a}{b-d}$- is large, there are relatively few non-believers compared with the population of believers. Intuitively, this implies that both religions easily generate mutual cooperation and so differences between the two in terms of behaviour will be small. The cost of rituals on the other hand might be substantial. This cost is given by

$$r^* = 2(1 - \zeta)(1 - \frac{1}{\sigma})(b - a)$$

and depends on the share of the believers, which is large for a large σ, and on $(b - a)$, which is the value that non-believers place on taking advantage of believers. Thus, when σ is large and $(b - a)$ not too small, the D religion will dominate. This is formalized in the following Proposition (which can be generalized to other values of r).

Proposition 2: *For any $\varepsilon > 0$, there exists a $\sigma^\varepsilon > 0$, such that for all $\sigma > \sigma^\varepsilon$, either (i) the discipline-based religion induces a strictly higher average material welfare compared with*

the ritual-based religion or (ii) the difference in average welfare between the two religions is smaller than ε.

We now analyse the case in which the difference in behaviour is not marginal, and the trade-off between the religions in terms of accuracy versus cost is more strongly manifested. We show that what matters for the resolution of this trade-off is the importance of coordination vis a vis cooperation. To see this, let us revisit first Example 1.

Example 1 revisited: Recall the PD, with $b = 4$, $d = 3$ and $a = 2$. Note that $2a = b$. R has more instances of mutual cooperation, so that 10% of the outcomes result in an average payoff of 3 instead of 2, a gain of 1 on 10% of outcomes per period. On the other hand, all agents above 1 -50% of the population- pay a cost of 2. Thus the relative cost of R is larger than its benefit, resulting in this religion being dominated.

When $2a$ is not sufficiently large compared with b, as in Example 1, two effects arise. First, coordination -which is what R is good at achieving- is not valued enough compared with miscoordination (an outcome that the D religion produces with a high probability). Second, the cost of ritual is quite high in equilibrium: As the benefit from mutual defection is too low, believers would agree to pay a high cost in order to change the behaviour of others towards them. Thus, accuracy is not valued enough and the cost is too high, which implies that the trade-off between accuracy and cost is resolved in favour of the D religion. We therefore have:

Proposition 3: *If coordination is not sufficiently important compared with cooperation (i.e., if $2a$ is not sufficiently large compared with b) then the discipline-based religion provides higher average material welfare.*

When on the other hand a is sufficiently high, the ritual-based religion can dominate:

Example 2: Consider the following PD game, in which, compared with Example 1, we have increased the value of a (together with d which must satisfy $d > a$):

	C	D
C	3.9,3.9	0,4
D	4,0	3.7,3.7

In R, in the limit when $\zeta \to 0$, $\gamma^* \simeq 0.1$ and $r^* \simeq 0.58$, which is paid by almost the whole population, and the distribution of play is close to:

	C	D
C	94%	0%
D	0%	6%

Average welfare is $2(0.94(3.9)+0.06(3.7))-(0.98)(0.58)=7.2$. In D, there is a large degree of cooperation in period 1 as $\gamma_1 \simeq 0.1$. But as even the small degree of non-believers defecting snowballs to substantially deter cooperation in period 2, $\gamma_2 \simeq 2.5$, so that only 10% of outcomes end in mutual cooperation:

period 1	C	D
C	94%	3%
D	3%	0%

period 2	C	D
C	10%	22%
D	22%	48%

average	C	D
C	52%	12%
D	12%	24%

Average welfare is $2(0.52(3.9)+0.24(3.7)+0.24(2)) \simeq 6.8$ and thus D is dominated; the cost of R is relatively low and in addition D creates a substantial level of one-sided cooperation which is sufficiently inferior, socially, compared to any other outcome.

4.3 Individual preferences

We now consider individual preferences. We find the following results. First, as signalling by good behaviour has positive externalities to the rest of society, this implies that non-believers prefer the discipline-based religion. Second, this analysis takes into account not only material but also spiritual utilities; the latter induces individuals to prefer the religion in which they cooperate more often. This, as long as strategic complementarities are strong enough, will imply that intermediate believers will support the ritual-based religion.

Let $U_J(\gamma)$ denote the (indirect) utility of an individual γ in the equilibrium in religion $J \in \{R, D\}$ and let $\Delta_{RD}(\gamma) = U_R(\gamma) - U_D(\gamma)$ denote the difference in utilities between the ritual-based religion and the discipline-based religion for a type γ.

Lemma 6: *There exists $\sigma' > 1$ such that: (i) If $\sigma \leq \sigma'$ then $\Delta_{RD}(\gamma)$ decreases with γ.(ii) If $\sigma > \sigma'$, $\Delta_{RD}(\gamma)$ increases with γ on $[b-d, \gamma_2]$ and decreases otherwise.*

Note that $\Delta_{RD}(\gamma)$ is composed of a material relative benefit and a spiritual relative benefit. The material relative benefit is fixed for all types that behave in the same way and elicit the same behaviour from others. On the other hand, the spiritual relative benefit is the difference in the probabilities with which one cooperates in the two religions, multiplied by the benefit from cooperation, γ. It is therefore positive and increasing in γ over an interval in which agents cooperate more often in R and negative and decreasing in γ otherwise.

When σ is sufficiently large (for example, when $\zeta \to 0$, we need $\sigma > \sigma' \to 2$) so that the share of believers is large, types in $[b-d, \gamma_2]$ cooperate on average more often in R; in R they cooperate vis a vis all believers (and thus with a relatively high probability), whereas in D they cooperate with all in the first period but with no one in the second period. Thus, the higher the γ in this region, the higher is the spiritual relative benefit from R. As the material relative benefit is fixed in this interval this implies that $\Delta_{RD}(\gamma)$ increases.

In all other regions $\Delta_{RD}(\gamma)$ always decreases with γ as all other types cooperate more often (at least weakly) in D. For example, types in $[0, b-d]$ never cooperate in R while some of them cooperate in D in the first period due to the discipline effect.[32] Using Lemma 6 we have:

Proposition 4: *There exists σ' such that: (i) If $\sigma \leq \sigma'$, all individuals prefer the discipline-based religion. (ii) If $\sigma > \sigma'$, there exists γ', γ'' with $b-d \leq \gamma' \leq \gamma_2 \leq \gamma'' \leq a$, so that only types in $[\gamma', \gamma'']$ prefer the ritual-based religion, and there exist parameters for which $\gamma' < \gamma''$.*

Consider first the types with weak beliefs, or non believers, who do not signal in any religion. In R all types (besides the behavioural ones) identify them as they do not participate in rituals. But in D many cooperate with them in the first period due to the discipline effect, which induces them to prefer D. This effect arises as D is based on a signal which provides positive externalities to others (note that this result holds for all r).

If σ is too low, then $\Delta_{RD}(\gamma)$ decreases for all γ, which by the above implies that all individuals support D. This is the case for the parameters of Example 1. When σ is sufficiently large, by Lemma 6, some individuals may prefer R. Moreover, the type at γ_2 will be its strongest supporter. As he cooperates more in R, he will have a higher spiritual payoff there. But also his material payoff might be higher, as the accurate signal in R allows him to better protect himself against defectors compared to the signal in D. This is the case for the parameters of Example 2:

Examples 1 and 2 revisited: In Example 1, $\sigma = \lim_{\zeta \to 0} \sigma' = 2$, which implies that all individuals prefer D. In Example 2 on the other hand strategic complementarities are very large, with $\sigma = 37$. Computing individual utilities, we find that all types above $\tilde{\gamma} \simeq 0.4 \in [b-d, \gamma_2]$ prefer R, which constitutes 88% of the population.

One may wonder how individual preferences interact with average material welfare. For example, based on their material and spiritual welfare and a simple majority rule, would individuals choose environments which also yield higher material welfare for their community? In Examples 1 and 2 this was the case. Example 3 considers parameters (namely a and d) which are between Examples 1 and 2 and shows that this can fail.

Example 3: Consider the PD game with $b = 4$, $d = 3.8$ and $a = 3.2$. In this case, when $\zeta \to 0$, we have $r^* \simeq 1.5$. The average distribution of play in the two religions is:

Ritual	C	D
C	88%	0%
D	0%	12%

Discipline	C	D
C	46%	12%
D	12%	30%

[32] Similarly, types in $[\gamma_2, a]$ cooperate more in D: in D they cooperate with all in the first period and with all types above $\gamma_1 < b-d$ in the second period. In R they cooperate with all above $b-d$.

and we find that average material utility is higher in the D religion. However, in this case, $\gamma' = 1$ and $\gamma'' = a$ so that all individuals above $\gamma = 1$ which represent 66% of the population would prefer R. Any voting rule or political process which will give voice to such a supermajority will create some stickiness towards the less socially efficient ritual-based religion.

Remark 2: Our model abstracts away from comparisons related to specific religious beliefs. One consideration when comparing across religions might be how individuals forecast changes in their beliefs. The analysis above is suitable for the case in which individuals believe that the relative strength of their religious beliefs, if society switches between religions, will remain the same. That is, what is important is that their relative ordering in society remains the same. This accords with the evolutionary biology idea of a "religious mind" or a "religious gene" which is distributed in society and can adapt to different religious systems.[33]

5 Supply side considerations

In this Section we discuss possible extensions of the model, focusing on the supply side of religious organizations.

5.1 The Consistory and the value of information

We now provide more specific details about the Consistory and then extend the model to consider comparative statics on the technology of public monitoring. Our discussion below is based on Kingdom (1992), Dommen and Bratt (2007), McGrath (1990) and Wilson (2002).

In 1541, upon his return to Geneva from exile in Strasbourg, Calvin had become convinced of the need for a disciplined and well-ordered church. In his letter to the city Council, Calvin writes: "*If you desire to have me for your pastor correct the disorder of your lives...I cannot possibly live in a place so grossly immoral...of what use is dead faith without good works?? Re-establish there pure discipline*".[34]

Calvin drew up the structure of his well-ordered church in the *Ecclesiastical Ordinances* (1541). The most distinctive and controversial aspect of this organization was the Consistory. It was formed in 1542, "*their office is to have oversight of the life of everyone...there were to be twelve of them, chosen from the members of the three councils, to keep an eye on everybody*".[35]

The main objection to this body by the city council was because it feared that the line between ecclesiastical and civil matters would be crossed. Indeed, a great deal of its function was devoted to resolving disputes within families, neighbours, and among business associates.

[33] See Boyer (2002).

[34] Cited in Beza (1996, rep.)

[35] *Ecclesiastical Ordinances* (1541), in Gilbert (1998).

Robert Kingdom who analyses the registrars of the Consistory, writes: "*A number of times businessmen were called in and questioned about complicated deals involving loans of money..and those found guilty of usury were subject to harsh penalties in an effort to form ethical business practice...At the end the consistory was extremely successful in achieving discipline".* The Consistory's normal cases ended with either an admonition or a remonstrance, a kind of public scolding delivered by one of the ministers, usually Calvin himself. Some of the cases ended with excommunication, which denied access to one of the four annual communion services in Geneva.

In Calvin's attempt to spread his influence into France, he supplied pastors that were trained in Geneva, but insisted that local churches elect local Consistories. The consistorial structure was made obligatory by the Venerable Company of Pastors in 1557; In 1562, the number of local consistories in France had risen to 1785. The fact that elders and deacons were to be provided locally indicates that the consistory had an important role in monitoring, gathering and disseminating information, an activity best done by locals. Thus, although Calvin was striving for strong control of the church over individuals' and pastors' daily life, his most important institution was a local, decentralized, one.

The discussion above illustrates that the decision to invest in a Consistory, or to create a culture of monitoring, may be a concrete choice by religious leaders. To be sure, such an investment may be costly, and the effectiveness of the Consistory may also depend on other exogeneous conditions such as urbanization or literacy levels.

To look at the effect of public information let the parameter $\pi \in [0,1]$ measure the probability that first period information is observed. Thus, with probability π the PD game in period 2 proceeds with information about Period 1, whereas with probability $1-\pi$ the game proceeds with no information which takes us back to the equilibrium in the autarky benchmark of Lemma 1. The result below establishes that the higher is π, the more signalling arises in the first period (as the value of signalling is higher), implying excessive signalling for sufficiently high π. We also show how the individual support for D is affected by π:

Proposition 5: *(i) For any π, the equilibrium is unique and there exists a π' such that the equilibrium is characterized by excessive signalling for $\pi > \pi'$ (with $d\gamma_1/d\pi < 0$) and by accurate signalling for $\pi < \pi'$ (with $d\gamma^*/d\pi < 0$, $\gamma^*(\pi)_{\pi\to 0} \to \hat{\gamma}$); (ii) For any π, there always exist a set of agents with weak beliefs who prefer the D religion to R. (iii) When π is small enough, there exist parameters for which an increase in π decreases the support for D; (iv) A strictly larger measure of agents prefer D to R when π increases from being sufficiently small to being sufficiently large.*

Note that all agents who defect in D always prefer a lower π as then more miscoordination arises. Moreover, agents who defect both in D and in R would prefer D to R for any π, as in

R they do not benefit at all from signalling while in D they always face some miscoordination.

We also show that an increase in information does not necessarily increase the support in the population for D, and in particular that when π is sufficiently small, an increase in π may imply that more agents actually prefer R. To see why, note that when $\pi \to 0$, we have that $\gamma^* \to \hat{\gamma}$. But a small increase in π implies that the utility from D for all agents below γ^* decreases; these agents always defect in D and a higher π implies that accurate signalling arises more often (where others can screen them out). If ζ is small enough and hence $\gamma^*(\pi)_{\pi \to 0}$ sufficiently large, a large share of agents becomes worse off in D with more information. However, globally, when π increases sufficiently, the support for D increases. This arises as believers and in particular those with strong enough beliefs will switch to prefer D once enough cooperation and signalling levels are guaranteed.

The Proposition indicates that religious beliefs alone may not be sufficient to convince individuals to adopt the Calvinistic religion and that a sufficient level of public monitoring needs to be provided for societies to switch to D. Note that Calvin was successful in Geneva only in his second spell in the city, when he initiated the Consistory, which supports the above result. It is also interesting to note that in Strasbourg, Zurich and Basel, the city councils did not give the church the power over excommunication and no Consistory was created despite attempts of Reformers such as Martin Bucer; the result above may also shed light on why these other attempts of the Reformation which had similar theological systems, reduced the role of rituals, but created no consistories, had initially failed.

5.2 The choice of rituals

We have looked above at the possibility of religious leaders designing the mechanism of monitoring behaviour; naturally, how one models the cost of such a monitoring institution will affect the choice of these leaders, while the benefits, at least at the time of the Reformation, could be captured by the degree of participation or the success of shifting a society from a ritual-based to a discipline-based organization.

The choice of religious leaders in the ritual-based religion may be more straightforward. Such leaders need to determine the level of rituals. They may maximize participation, or revenues from the religion, if some of the cost of rituals can be extracted as actual rent. We now show that in some environments, the equilibrium with accurate signalling and largest participation will also be chosen by a religious leader who maximises $r(\zeta + (1-\zeta)(1 - \frac{\gamma(r)}{a}))$, where $\gamma(r)$ is either γ^* in an accurate signalling equilibrium or γ_1 in an excessive signalling equilibrium:

Proposition 6: *(i) In the set of excessive equilibria, the higher is γ_1 the higher are the revenues from the ritual-based religion; (ii) In the set of accurate equilibria, when σ is low enough, the lower is γ^* the higher are the revenues from the ritual-based religion.*

Together, (i) and (ii) imply that the cost of rituals characterized in Lemma 5 may be chosen by religious leaders who maximize revenues. To see the intuition, consider first the set of accurate equilibria. When σ is low enough, when r increases, it is also the case that γ^* decreases; a higher fee implies then that more types need to change their behaviour. Thus decreasing γ^* increases both the demand for rituals and its price, which implies that the religious leader will choose the lowest possible γ^*, i.e., $\gamma^* = b - d$. In the set of excessive signalling, even though whenever r increases the demand for rituals also decreases, it is also the case that r has to be substantially lower to attract the non-believers, and such elasticity implies that revenues are maximized when r is highest in this set.

Remark 3: Note that we have assumed that ritual cost is equal for all. It is possible to analyse a model and maintain the general results when the cost of rituals has an individual component depending on γ, as perhaps some individuals find the rituals intrinsically beneficial. We have also assumed that the rituals are a deadweight loss to society. Suppose instead that the ritual cost represents some monetary or charity element that can be redistributed back to some of the members of the community, if the church chooses to do so. This effectively will lower the cost of participation for the members on the receiving end, taking us back to the case of heterogeneous cost which can be analysed as discussed above.

5.3 Emphasis on rituals or discipline

So far, we have focused on the extreme cases of religious organizations that either provide signalling by rituals or by monitoring of behaviour. Most religions will probably use a mixture of both signals. In the appendix we formalise and analyse equilibria in a model with such a hybrid religion; we show that the gist of our results is maintained and thus it is sufficient to focus on the extreme cases.

We now discuss how different factors might imply that some religious organisations will put more emphasis on particular types of signalling technologies. First, the choice of the signalling mechanism might be related to the evolution of the religious market. In a ritual-based religion, religious leaders may be able to extract some rent or appropriate some portion of the cost of rituals. When they have monopoly power, they might therefore prefer to stick to social coordination via rituals. In the appendix we revisit Examples 1 and 2 and show that when hybrid religions exist, a religion that fully focuses on rituals generates the highest revenues to the church.

It is also reasonable from the point of view of reformers to abolish rituals and focus on discipline; one of the main explanations for the success of the Reformation is the high costs of Catholic rituals (for example see Ekelund *et al* 2002). The process of urbanization increased professionalism which implied that in terms of opportunity costs, rituals as well as rent extrac-

tion by the church became more costly. The reformers had to break away with old practices and specifically those that were costliest to the believers.

Second, geographical considerations and the constraints on information transmission may also affect the organisation of religious life. For example, subtle information on past behaviour might be harder to exchange when considering trade between villages, cities or countries, whereas rituals with physical attributes such as clothes, language, and participation in sermons, may be more easily transmitted across distant locations. With more dense population in cities, monitoring of discipline may become easier. The urban population is also more literate and hence may make dissemination of subtle information more viable. Finally, advancement in the technology of transmitting information, such as the advent of printing, might affect the decision of religious leaders to shift emphasis to a more discipline-based organisation.

Below we also discuss how the theology may be related to religious institutions in general and the signalling mechanisms in particular; this implies that religions will focus on some form of signalling which fits best their theology instead of choosing a mixture.

6 Discussion and conclusion

Below we discuss the possible role of Theology in our model, as well as its link to the social signalling mechanism. In addition, as a way of motivating our focus on the Calvinistic Reformation, we discuss its differences compared with Luther's Reformation and provide some evidence of life in Geneva pre-Calvin in light of our welfare results. We conclude by discussing similar phenomena in other religions, and the comparative advantage of religion in sustaining cooperation vis a vis other mechanisms.

6.1 Theologies and institutions

To facilitate our analysis we have abstracted from differences in theologies and assumed that both religions motivate good works in the same way. But different theologies might affect individuals' beliefs through different psychological channels and might not be easily compared. For example, the self-signalling interpretation by Weber (1904) of Calvinistic beliefs might induce incentives for good works but such a mechanism is rather indirect. Also, Luther and Calvin encouraged their supporters to go back to the scriptures and to read the Bible by themselves (enabled by the advent of printing and higher levels of literacy), and this may lead to weaker or stronger beliefs than when one participates in rituals conducted by priests, depending both on the individuals and the priests in question. On the other hand, mechanisms such as forgiveness and indulgences (the system of exchange between money and redemption) that have evolved in the Catholic church might erode the connection between beliefs and good works to some degree. It would be interesting to analyse these more nuanced systems of beliefs.

We note though that the differences between these two theologies might be consistent with the differences in the institutional structure (although the causality between institutions and theology is not obvious). Specifically, in the Catholic church, good works alone do not suffice; according to Thomas Aquinas, three are required for salvation: direct reliance on the church and its sacraments, the free turning of the will to God and away from sin, and the remission of the guilt incurred by sin by priestly absolution.[36] In medieval times, this had evolved into a heavy load of public rituals and an impressive system of rent extraction. More generally, this theology easily lands itself to a hierarchical structure in which priests have to certify which actions provide rewards and can possibly deliver forgiveness. In the absence of free will, such a role of the Church's hierarchy, which is reinforced by rituals, is reduced.

6.2 Calvin vs. Luther

We have focused on Calvin's Reformation and not on Martin Luther's. Luther differs from Calvin both in terms of his theology and in terms of his general attitude towards the relation between Church and morality.

In terms of theology, while Calvin advocates justification by the grace of God, Luther focuses on justification by faith: *"It is faith in Christ which makes him live in me and move in me and act in me...faith receives Christ's good works; love performs good works for the neighbours"* (cited in Green 1964). Luther offered individuals personal certitude of salvation already in this life, provided only that they have faith. These beliefs reduce the anxiety about salvation and as a result, good works become less important (McGrath 1990).

More generally, Luther permitted religion to be identified with neither ethics nor social justice as religion transcended both. An interesting illustration of this is Luther's response to the Peasants' Revolt in 1525: Luther firmly resisted the slightest diminution of religion and criticized the peasants' characterization of their demands for social justice as being Christian demands (Ozment 1980).

In terms of the institutional structure of the Church, Luther has created no institutions, let alone the Consistory, and discipline was not considered an issue for the Church. In 1530, in the Confession of Augsburg, Luterhans insisted that there are only two marks of a true church: the church is the assembly of saints in which the gospel is taught purely and the sacraments are administered rightly. In particular, there is no requirement of good behaviour, which Calvinists considered as the third mark of a good Church.[37]

[36] Thomas Aquinas, summa theologia. p.39.

[37] It is also worth mentioning that similar arguments formed the debate in England between Calvinist Puritans and Erastian Anglicans and that the debate on this goes on even today. In the formula adopted by the churches of the Reformation in US in 1997 it was decided that there should be no mention of the mark of discipline.

6.3 Geneva before Calvin

Our analysis indicates that the discipline-based religion is more likely to be adopted when monitoring and information dissemination institutions arise (see Proposition 5). We have argued above that this could explain why Calvin was only successful in his second tenure in Geneva, when he created the Consistory. Also, urbanization and increased literacy levels which is correlated with urban professionalism, have both enabled the dissemination of more subtle information and have made Calvinism more attractive in line with our results.

Propositions 3 and 4 indicate that discipline-based religions are more likely to be successful when strategic complementarities are large enough and when cooperation is more important than coordination. While an empirical analysis that tackles this question is beyond the scope of our analysis, it is clear that cooperation was much in demand in Geneva in the pre-Calvin times. McGarth (1990) describes the city as being in decay, exhibiting reduced morality, and in need in some reformation, not only religious but social, political and economic. Wilson (2002) writes more specifically: "*By Calvin's own account, his primary challenge was to unite the fractious city of Geneva in to an effective corporate unit...Geneva also had an infrastructure and the burden of supporting it was probably greater for the average citizen than now...a massive wall around the city had to be maintained, the swiss mercenaries that protected Geneva from the duchy of savoy had to be paid, a plague hospital had to be build, charity had to be given to the poor, an educational system had to be built, the list of public goods goes on and on. The temptation to avoid the burden must have been great, not to speak of subverting the entire system.*"

6.4 Rituals vs. discipline in other religions

While our leading example in the paper is that of the Calvin's Reformation other religious organisations share similar attributes. In this Section we discuss a few examples.

The Orthodox church and the Old Believers

With some similarity to the Reformation, in 1666 the Old Believers separated from the Orthodox church after the reforms of Patriarch Nikon which introduced a number of ritual and textual revisions with the aim of achieving uniformity between Russian and Greek Orthodox practices. The Old believers have been compared to Calvin's reformation in Geneva and to its successors the Quakers and Baptist sects, both in terms organisation and economic achievements.[38] Several papers analyse their contribution to the rise of private industrial enterprise in early nineteenth-century Russia and in particular their role in the success of the Moscow textile industry (see Blackwell 1965 and Raskov 2012). Old believer communities still

[38] See Levintova (2007) and Vorontsova and Filatov (2000) for example.

exist today. The following is an example of the public scrutiny of behaviour in a community of Old Believers in Oregon, USA:[39] "*In recent years, the nastayatyel has been increasingly called on to administer punishments and other forms of discipline to miscreant young people. These punishments usually consist of a public announcement of the individual's sins to the congregation at the end of the service, whereupon the transgressors may be compelled to perform several prostrations before the congregation, or some other act of contrition and penance*".

Islam and Ismailies

Islam is a religion with a large number of costly (many of them observable) rituals including daily prayers, fasting in the month of the Ramadhan, and the *hajj* to Mecca. The Ismaili branch of Shia Islam however, has fewer rituals[40] and moreover, around the 17th century, the Ismailis have created social structures similar to Calvinist sects. Specifically, two institutions were created: the *Jamat*, an assembly in the council of all the adult males, and the *Jamat Khana*, a council hall or guildhall of the community. There is a large literature in Sociology comparing these institutions and the Ismaili sect to Calvinist ones. Clarke (1976) describes the practice of these institutions in Britain as mixing the secular and the religious: "*The role of the Jamat is multiple..it is well nigh impossible for an ismaili to make a meaningful distinction between the religious and the social. The jamat is regarded as a community centre which serves as a variety of purposes..a conciliar system of organization and administration is now operative which is similar to that of Calvinism...*".

Similarly, Bocock (1971), analysing the large Ismaili community in Tansania, describes features as in Calvinism such as the abundance of local leaders, and the role of the religion in monitoring secular affairs: "*Unlike the pope, the Imam is not surrounded by a large structure with full time priests and bishops staffing it, but can choose local leaders himself...A local community of ismailies govern their own secular affairs and meet for worship every day..this is similar to some of the organizational forms among calvinists, or calvinist influenced groups such as congregationalists, quakers, baptists*".

Finally, Goldthorpe (1996) stresses the discipline element: "*Both [calvinist and ismaili] religions were notably congregational in their form of organization, and the local community of Ismailis who met daily in the Jamat Khan, the Ismaili mosque, constituted also a community exercising moral control over its members, and providing them - when in good standing - with mutual aid and credit networks".*

[39] see http://pages.uoregon.edu/sshoemak/325/texts/old_believers.htm

[40] For example it has less daily prayers and its mosque is actually closed at noon time.

6.5 Other institutions supporting cooperation

We have advocated a view of religious organizations as enabling cooperation through a system of beliefs which induce preferences for cooperation, and an organizational structure that allows for the signalling of these preferences. There is a vast literature in Economics analysing how cooperation can arise in situations such as the Prisoner's Dilemma and our analysis is complementary to this literature.

One large literature concerns how cooperation might arise when groups are involved in repeated interactions. The scope for cooperation in these environments is constrained by discounting of future pay-offs and by the size of the community. In addition, Greif (1989) shows that even within the repeated game literature religious organizations have a role; specifically, they allow to create coalitions which are not too large and which are conducive to both information transmission and enforcement of punishments. Our results are complementary to these. We focus instead on environments in which cooperation per se is less of an issue, as some individuals are motivated to cooperate by their religious beliefs, and show how religious organisations use this starting point to induce higher levels of cooperation.

A second type of literature assumes that agents have preferences for cooperation (e.g., Tabellini 2008) and analyses the conditions for their evolution. Greif and Tabellini (2010) focus on the evolution of cooperation towards kin vs indiscriminate cooperation. In their dichotomy, clans have strong morality and rely on repeated games while cities have looser morality but rely on institutions. Our analysis is related to this as we also combine preferences for cooperation in an institutional set up, albeit both preferences and institutions are provided by the religion.

Finally, it may be argued that when state capacity is developed, institutions and legal courts are sufficient to promote cooperation as they allow enforcement of contracts. This might imply that these institutions will crowd out religion in its role of facilitating cooperation. The literature on contract theory has however also noted that there are intrinsic problems in enforceability of contracts even in the presence of symmetric information; e.g., imperfect verifiability or the inability to describe complicated scenarios. Equilibrium behaviour as in our model is naturally immune to such problems.[41]

[41] It is also worthwhile to mention that the theories of the role of religion as enabling signalling and hence cooperation, by both Adam Smith and Max Weber, were written at times and places in which state capacity was already sufficiently developed.

7 Appendix

7.1 Appendix A: Proofs

Proof of Lemma 1: As agents can cooperate or defect without any other conditions, the equilibrium must be a cutoff equilibrium as described in the Lemma. To show that such an equilibrium exists, note that if we set $\hat{\gamma} = a$ then the LHS is positive for any $\zeta > 0$, while if we set $\hat{\gamma} = 0$, then the LHS is negative, and thus there will be a cutoff $\hat{\gamma}$. Moreover, the LHS is monotone in $\hat{\gamma}$ which will imply uniqueness.■

Proof of Lemma 2: By monotonicity of beliefs, if a cutoff type signals at γ, all above with $\gamma' > \gamma$ will wish to signal as well, at least weakly. This implies that the only monotone equilibria are as described in the Lemma. To see that (1) has a unique solution, note that it implies,

$$\gamma_2 = a\frac{(b-d-\gamma_1) + a\frac{\zeta}{1-\zeta}\frac{(b-d)}{a}}{(b-d-\gamma_1) + a\frac{\zeta}{1-\zeta}}.$$

As $\frac{\partial(\gamma_2)}{\partial\zeta} < 0$, then $\lim_{\zeta\to 1}\gamma_2 = b-d < \gamma_2 < a = \lim_{\zeta\to 0}\gamma_2$.■

Proof of Lemma 3: We will now consider existence and uniqueness for the excessive signalling equilibrium in D. The fixed point equation for γ_1 is:

$$(*)\ ((1-\zeta)\frac{(a-\gamma_1)}{a} + \zeta)(d-b) + (1-\zeta)\frac{\gamma_1}{a}(-a) + \gamma_1 + ((1-\zeta)\frac{(a-\gamma_2)}{a})(b-a) = 0$$

where γ_2 is given by

$$(**)\ \ \gamma_2 = a\frac{(b-d-\gamma_1) + a\frac{\zeta}{1-\zeta}\frac{(b-d)}{a}}{(b-d-\gamma_1) + a\frac{\zeta}{1-\zeta}}$$

Note that from the above γ_2 is monotonically decreasing in γ_1.

Suppose that $\gamma_1 = 0$. Then

$$\gamma_{2|_{\gamma_1=0}} = a\frac{(b-d) + a\frac{\zeta}{1-\zeta}\frac{(b-d)}{a}}{(b-d) + a\frac{\zeta}{1-\zeta}}$$

If $((1-\zeta)a+\zeta)(d-b) + ((1-\zeta)(a-\gamma_2(0)) + \zeta)(b-a) \geq 0$ then there exists an equilibrium with $\gamma_1 = 0$. If $((1-\zeta)a+\zeta)(d-b) + ((1-\zeta)(a-\gamma_2(0)) + \zeta)(b-a) < 0$ then there exists an equilibrium with $0 < \gamma_1 \leq b-d$. To see this, note that at $\gamma_1 = 0$, the LHS of (*) is negative. On the other hand at $\gamma_1 = b-d$ from (*) and (**) we have

$$\gamma_{2|_{\gamma_1=b-d}} = b-d$$

and the LHS of (*) becomes,

$$\begin{aligned}
& ((1-\zeta)\frac{(a-\gamma_1)}{a}+\zeta)(d-b)+(1-\zeta)\frac{\gamma_1}{a}(-a)+\gamma_1+((1-\zeta)\frac{(a-\gamma_2)}{a}+\zeta)(b-a) \\
= \; & ((1-\zeta)\frac{(a-b+d)}{a}+\zeta)(d-b)-(1-\zeta)(b-d)+(b-d)+((1-\zeta)\frac{(a-b+d)}{a}+\zeta)(b-a) \\
= \; & \frac{1}{a}(1-\zeta)((a-b+d)(d-a)+\zeta(b-a)>0
\end{aligned}$$

So a value of γ_1 satisfying (*) exists and is the solution to the two equations. To see the uniqueness of a solution note that using (*) and (**) we get,

$$((1-\zeta)\frac{(a-\gamma_1)}{a}+\zeta)(d-b)+(1-\zeta)\frac{\gamma_1}{a}(-a)+\gamma_1+((1-\zeta)\frac{(a-\gamma_2(\gamma_1))}{a}+\zeta)(b-a)=0$$

Note that this expression is monotone in γ_1,

$$\begin{aligned}
\frac{\partial LHS}{\partial \gamma_1} &= \\
(1-\zeta)(b-d)+\zeta-(1-\zeta)(b-a)\frac{\partial \gamma_2(\gamma_1)}{\partial \gamma_1} &> 0,
\end{aligned}$$

which insures uniqueness. We now show that there is no accurate equilibrium. The fixed point equation for γ^* is:

$$\begin{aligned}
& (***)\ ((1-\zeta)\frac{(a-\gamma^*)}{a}+\zeta)(d-b)+(1-\zeta)\frac{\gamma^*}{a}(-a)+\gamma^*+(((1-\zeta)\frac{(a-\gamma^*)}{a}+\zeta)d \\
& +(1-\zeta)\frac{\gamma^*}{a}a+((1-\zeta)\frac{(a-\gamma^*)}{a}+\zeta)\gamma^*)-(\zeta b+(1-\zeta)a) \\
= \; & 0
\end{aligned}$$

At $\gamma^*=b-d$ the LHS becomes,

$$(1-\zeta)(a-b+d)(\frac{d}{a}-1)>0$$

At $\gamma^*=a$ the LHS becomes,

$$2\zeta(d-b+a)>0$$

Therefore if the derivative with respect to γ^* is monotone we will not have such an equilibrium; the derivative of the *lhs* of (***) after some manipulation is

$$-(1-\zeta)(\frac{2d-b+2\gamma^*-a}{a})+1+\zeta$$

Note that this expression is decreasing in γ^*. Therefore it is either first positive and then negative or always negative. In either case an equilibrium does not exist.■

Proof of Lemma 4: The equations for the equilibrium cost of rituals are provided in the text from which it is clear to see that for any γ, a ritual cost can support this γ as a signalling cutoff in the first period.[42]■

[42]Note that when r is small enough, it is possible to support an equilibria in which all participate in rituals (i.e., $\gamma_1=0$). This equilibrium relies on out of equilibrium beliefs that all in (γ_2,a) will defect against the one who does not participate in rituals.

Proof of Lemma 5: Consider an equilibrium with excessive signalling which is characterized by,

$$r = 2(1-\zeta)(\frac{a-\gamma_2}{a})(b-a)$$

Note that higher levels of rituals must give rise to a lower γ_2 and as a result a higher γ_1, and thus serve to improve the content of the signal.

Let us compare this equilibrium to one in which $r^* = 2(1-\zeta)(\frac{a-b+d}{a})(b-a)$, i.e., an accurate equilibrium with $\gamma^* = b-d$. Remember that $r^* > r$. Note that all those that did not pay r, or that defect (are below γ_2), are indifferent among these $r's$ as their net utility when they either pay or not pay for rituals is $\zeta b + (1-\zeta)a$ per period. On the other hand, all agents above γ_2, have a higher utility when r increases to r^*. To see why, note that γ_2 decreases to γ^*, and thus they receive, for the interval of change $[\gamma^*, \gamma_2]$, a relative benefit of $(1-\zeta)\frac{d-(\gamma_2-(b-d))}{a} > 0$ per period in the equilibrium with r^*, where $(1-\zeta)\frac{d-0}{a}$ represent the difference in material pay-off from the PD game whereas $(1-\zeta)\frac{(\gamma_2-(b-d))}{a}$ represents the increase in the payment from r to r^* (per period). Finally, these types have another increase in their utility as when γ_1 increases, they defect against more agents who defect against them which provides them a higher utility according to their beliefs. Therefore, all excessive signalling equilibria are Pareto dominated by the equilibrium with r^*.

Now let us now look at r which sustains accurate signalling, i.e., $\gamma' \in (b-d, a)$. Those who do not pay r gain the same utility $\zeta b + (1-\zeta)a$ per period, whereas types $\gamma > \gamma_1$ who pay r, get $2(\zeta d + (1-\zeta)\frac{a-\gamma'}{a}d + (1-\zeta)\frac{\gamma'}{a}a + (\zeta + (1-\zeta)\frac{a-\gamma'}{a})\gamma) - r$, for

$$r = 2(\zeta(d-b+\gamma') + (1-\zeta)(\frac{a-\gamma'}{a})(d-a+\gamma'))$$

The utility of agents above γ' as a function of γ is

$$\zeta d + (1-\zeta)\frac{a-\gamma'}{a}d + (1-\zeta)\frac{\gamma'}{a}a + (\zeta + (1-\zeta)\frac{a-\gamma'}{a})\gamma - \zeta(d-b+\gamma') - (1-\zeta)(\frac{a-\gamma'}{a})(d-a+\gamma')$$

per period. The derivative w.r.t γ' for some type γ is

$$\begin{aligned}(1-\zeta)(1-\frac{d+\gamma}{a}) - \zeta + (1-\zeta)(\frac{d-a+\gamma'}{a}) - (1-\zeta)(\frac{a-\gamma'}{a}) &< 0 \Leftrightarrow \\ (1-\zeta)(\frac{2\gamma'-\gamma}{a}) - 1 &< 0\end{aligned}$$

which is satisfied as $\gamma' < \gamma < a$. Therefore, all these types prefer a religion with a lower cutoff and again any such equilibrium will be Pareto dominated by the equilibrium with r^*.■

Proof of Proposition 1: We start with the following helpful Lemma.

Lemma A1 (i) *There is more mutual cooperation in R iff*

$$\frac{(\frac{\gamma_2}{a} - \frac{1}{\sigma})(2 - \frac{\gamma_2}{a} - \frac{1}{\sigma})}{(\frac{1}{\sigma} - \frac{\gamma_1}{a})(2 - \frac{\gamma_1}{a} - \frac{1}{\sigma})} > 1$$

(ii) If

$$\frac{\gamma_1+\gamma_2}{2a} > \frac{1}{\sigma}$$

then there is more total coordination in the R religion.

Proof of Lemma A1:

(i) Mutual cooperation in the R religion is $2(1-\frac{1}{\sigma})^2$ where it is $(1-\frac{\gamma_1}{a})^2+(1-\frac{\gamma_2}{a})^2$ in the D religion.

$(1-\frac{1}{\sigma})^2 > \frac{(1-\frac{\gamma_1}{a})^2+(1-\frac{\gamma_2}{a})^2}{2} \Leftrightarrow (\frac{1}{\sigma}-\frac{\gamma_1}{a})(\frac{\gamma_1}{a}+\frac{1}{\sigma}-2) > (\frac{\gamma_2}{a}-\frac{1}{\sigma})(\frac{\gamma_2}{a}+\frac{1}{\sigma}-2) \Leftrightarrow \frac{(\frac{\gamma_2}{a}-\frac{1}{\sigma})(2-\frac{\gamma_2}{a}-\frac{1}{\sigma})}{(\frac{1}{\sigma}-\frac{\gamma_1}{a})(2-\frac{\gamma_1}{a}-\frac{1}{\sigma})} > 1.$

(ii) Total miscoordination in the R religion is given by $2\zeta\frac{1}{\sigma}$. In the D religion miscoordination is larger than $\zeta(\frac{\gamma_1+\gamma_2}{a})$.□

We now prove the Proposition. In particular we prove that the statement is true when ζ is small enough, uniformly for the parameters of the model. We therefore consider a convergent sequence of parameters $\{a_n, b_n, d_n\}_{n=1}^{\infty}$ and a sequence $\{\zeta_m\}_{m=1}^{\infty}$ such that $\lim_{m\to\infty}\zeta_m = 0$. Let $\sigma_n = \frac{a_n}{b_n-d_n}$. By Lemma 2, for any m and n, there is a unique equilibrium, $(\gamma_1^{n,m}, \gamma_2^{n,m})$. Equilibrium equations are,

$$\begin{aligned}((1-\zeta_m)(1-\frac{\gamma_1^{n,m}}{a_n})+\zeta_m)(d_n-b_n)+\zeta_m\gamma_1^{n,m}+((1-\zeta_m)(1-\frac{\gamma_2^{n,m}}{a_n})(b_n-a_n) &= 0\\ \frac{(1-\frac{\gamma_1^{n,m}}{(b_n-d_n)})+\frac{\zeta_m}{1-\zeta_m}}{(1-\frac{\gamma_1^{n,m}}{(b_n-d_n)})+\sigma_n\frac{\zeta_m}{1-\zeta_m}} &= \frac{\gamma_2^{n,m}}{a_n}\end{aligned}$$

There are two cases to consider.

Case 1: Suppose that $\sigma_n \to_{n\to\infty} \infty$. The second equilibrium equation can be written as,

$$\frac{(\frac{\gamma_2^{n,m}}{a_n}-\frac{1}{\sigma_n})}{(\frac{1}{\sigma_n}-\frac{\gamma_1^{n,m}}{a_n})} = \frac{(\sigma_n-1)}{(1-\frac{\gamma_1^{n,m}}{(b_n-d_n)})+\sigma_n\frac{\zeta_m}{1-\zeta_m}}$$

Taking the double limit, first with respect to n and then with respect to m, we get,

$$\lim_{m\to\infty}\lim_{n\to\infty}\frac{(\frac{\gamma_2^{n,m}}{a_n}-\frac{1}{\sigma_n})}{(\frac{1}{\sigma_n}-\frac{\gamma_1^{n,m}}{a_n})} = \lim_{m\to\infty}\lim_{n\to\infty}\frac{(\sigma_n-1)}{(1-\frac{\gamma_1^{n,m}}{(b_n-d_n)})+\sigma_n\frac{\zeta_m}{1-\zeta_m}} = \lim_{m\to\infty}\frac{1-\zeta_m}{\zeta_m} = \infty$$

This implies that $\frac{\gamma_2^{n,m}+\gamma_1^{n,m}}{2a_n} > \frac{1}{\sigma_n}$ for a low enough ζ.

Note also that,

$$\frac{(2-\frac{\gamma_2^{n,m}}{a_n}-\frac{1}{\sigma_n})}{(2-\frac{\gamma_2^{n,m}}{a_n}-\frac{1}{\sigma_n})} > \frac{1-\frac{1}{\bar{\sigma}}}{2} > 0.$$

This implies that for any $\sigma^* > 1$ there exists a $\zeta^* > 0$ such that for any $\sigma > \sigma^*$ and $\zeta < \zeta^*$,

$$\frac{(\frac{\gamma_2}{a}-\frac{1}{\sigma})(2-\frac{\gamma_2}{a}-\frac{1}{\sigma})}{(\frac{1}{\sigma}-\frac{\gamma_1}{a})(2-\frac{\gamma_1}{a}-\frac{1}{\sigma})} > 1$$

and so there is more mutual cooperation in the R religion.

Case 2: Suppose that $\sigma_n \rightarrow_{n\rightarrow\infty} \sigma \leq \sigma^*$. First note that as a, b and d are bounded and as $a > \underline{a}$ there exists a $\mu > 0$ such that $\frac{b_n - a_n}{b_n - d_n} < \mu$.

Case 2(i): Suppose that, $(1 - \frac{\gamma_1^{n,m}}{(b_n - d_n)}) \rightarrow_{n\rightarrow\infty} 0$. In this case we get,

$$\frac{(\frac{\gamma_2^{n,m}}{a_n} - \frac{1}{\sigma_n})}{(\frac{1}{\sigma_n} - \frac{\gamma_1^{n,m}}{a_n})} \rightarrow_{n\rightarrow\infty} \frac{(\sigma_n - 1)}{\sigma_n \frac{\zeta_m}{1-\zeta_m}}$$

But this means that

$$\lim_{m\rightarrow\infty} \lim_{n\rightarrow\infty} \frac{(\frac{\gamma_2^{n,m}}{a_n} - \frac{1}{\sigma_n})}{(\frac{1}{\sigma_n} - \frac{\gamma_1^{n,m}}{a_n})} = \infty$$

and as before $\frac{\gamma_2^{n,m} + \gamma_1^{n,m}}{2a_n} > \frac{1}{\sigma_n}$ for a low enough ζ as well as

$$\frac{(2 - \frac{\gamma_2^{n,m}}{a_n} - \frac{1}{\sigma_n})}{(2 - \frac{\gamma_2^{n,m}}{a_n} - \frac{1}{\sigma_n})} > \frac{1 - \frac{1}{\bar{\sigma}}}{2} > 0$$

so that we have that

$$\lim_{m\rightarrow\infty} \lim_{n\rightarrow\infty} \frac{(\frac{\gamma_2^{n,m}}{a_n} - \frac{1}{\sigma_n})}{(\frac{1}{\sigma_n} - \frac{\gamma_1^{n,m}}{a_n})} \frac{(2 - \frac{\gamma_2^{n,m}}{a_n} - \frac{1}{\sigma_n})}{(2 - \frac{\gamma_2^{n,m}}{a_n} - \frac{1}{\sigma_n})} = \infty$$

This implies that for any $\sigma < \sigma^*$ there exists a $\zeta' > 0$ such that for any $\zeta < \zeta'$,

$$\frac{(\frac{\gamma_2}{a} - \frac{1}{\sigma})(2 - \frac{\gamma_2}{a} - \frac{1}{\sigma})}{(\frac{1}{\sigma} - \frac{\gamma_1}{a})(2 - \frac{\gamma_1}{a} - \frac{1}{\sigma})} > 1$$

and so there is more mutual cooperation in the R religion.

Case 2(ii): Suppose that $(1 - \frac{\gamma_1^{n,m}}{(b_n - d_n)}) \rightarrow_{n\rightarrow\infty} \eta > 0$. Using the second equation we get,

$$\frac{\gamma_2^{n,m}}{a_n} = \frac{(1 - \frac{\gamma_1^{n,m}}{(b_n - d_n)}) + \frac{\zeta_m}{1-\zeta_m}}{(1 - \frac{\gamma_1^{n,m}}{(b_n - d_n)}) + \sigma_n \frac{\zeta_m}{1-\zeta_m}} \rightarrow_{n\rightarrow\infty} \frac{\eta + \frac{\zeta_m}{1-\zeta_m}}{\eta + \sigma \frac{\zeta_m}{1-\zeta_m}}$$

But looking at the first equation as $n \rightarrow \infty$:

$$-((1 - \zeta_m)(\frac{\sigma + \eta - 1}{\sigma}) + \zeta_m) + \zeta_m(1 - \eta) + \zeta_m(\frac{\sigma - 1}{\eta + \sigma \frac{\zeta_m}{1-\zeta_m}}) \frac{(b_n - a_n)}{(b_n - d_n)} = 0 \qquad (3)$$

Where we have substituted the following equations,

$$\begin{aligned}
\lim_{n\rightarrow\infty} (1 - \frac{\gamma_1^{n,m}}{a_n}) &= \lim_{n\rightarrow\infty} (1 - \frac{\gamma_1^{n,m}}{(b_n - d_n)\sigma_n}) = (\frac{\sigma + \eta - 1}{\sigma}) \\
\lim_{n\rightarrow\infty} (\frac{\gamma_1^{n,m}}{(b_n - d_n)}) &= 1 - \eta \\
\lim_{n\rightarrow\infty} (1 - \frac{\gamma_2^{n,m}}{a_n}) &= (\frac{\zeta_m}{1 - \zeta_m})(\frac{\sigma - 1}{\eta + \sigma \frac{\zeta_m}{1-\zeta_m}})
\end{aligned}$$

But note that in (3) for high enough m this equation cannot hold as it is negative. Therefore this case cannot arise for large enough m.

To conclude the proof of this part we choose $\bar{\zeta} < \min\{\zeta', \zeta^*\}$.■

Proof of Proposition 2: Suppose that $\sigma_n \to_{n\to\infty} \infty$. By the second equilibrium equation, we have that

$$\lim_{n\to\infty} \frac{\gamma_2^{n,m}}{a_n} = \lim_{n\to\infty} \frac{(1 - \frac{\gamma_1^{n,m}}{(b_n - d_n)}) + \frac{\zeta_m}{1-\zeta_m}}{(1 - \frac{\gamma_1^{n,m}}{(b_n - d_n)}) + \sigma_n \frac{\zeta_m}{1-\zeta_m}} = 0.$$

As $\sigma_n \to_{n\to\infty} \infty$ and $\gamma_1^{n,m} < (b_n - d_n)$ (by Lemma 1) we have that

$$\lim_{n\to\infty} \frac{\gamma_1^{n,m}}{a_n} = 0$$

Thus, in both religions, for high enough n, cooperation is almost full. However, $r_{n,m} = 2((1-\zeta_m)(1 - \frac{b_n - d_n}{a_n})(b_n - a_n)) > 0$ when $\frac{b-d}{a} \nrightarrow 1$. Therefore, either the D religion is strictly preferred, for high enough n, or they converge to yield the same average welfare.■

Proof of Proposition 3: In R, material welfare of all types below $b - d$ is $\zeta b + (1-\zeta)a$ per period, whereas the material welfare of all types above $b - d$ per period is $\zeta d + (1-\zeta)(1 - \frac{b-d}{a})(d - b + a) + (1-\zeta)\frac{b-d}{a}a \lesssim a$ by strategic complementarities, for a small enough ζ. On the other hand, in D, social welfare for all is some combination of a, $b/2$ and d. Thus if a is not sufficiently larger than $b/2$, D dominates.■

Proofs of Lemma 5 and Proposition 4: Let $\Delta_{RD}^{[\gamma_i,\gamma_j]}$ denote the difference in expected utility of types in the interval $[\gamma_i, \gamma_j]$ from R vs. D. We consider below average utility per period of play i.e. in R the utility in the one-period PD game minus $r/2$, and in D average utility across the two periods. Consider first all types in $[\gamma_2, a]$.

$\Delta_{RD}^{[\gamma_2,a]} = \zeta(d+\gamma) + (1-\zeta)(1 - \frac{b-d}{a})(d - b + a + \gamma) + (1-\zeta)(\frac{b-d}{a})a$

$-\frac{1}{2}[(\zeta + (1-\zeta)(1 - \frac{\gamma_1}{a}))d + \gamma + (\zeta + (1-\zeta)(1 - \frac{\gamma_2}{a}))d + (\zeta + (1-\zeta)(1 - \frac{\gamma_1}{a}))\gamma + 2(1-\zeta)\frac{\gamma_1}{a}a] =$

$(1-\zeta)(1 - \frac{b-d}{a})(d - b + a) + (1-\zeta)(\frac{b-d-\gamma_1}{a})a - [((1-\zeta)(1 - \frac{\frac{\gamma_1+\gamma_2}{2}}{a}))d] + \gamma[(1-\zeta)(\frac{\gamma_1}{2a} - \frac{b-d}{a})]$

Hence $\Delta_{RD}^{[\gamma_2,a]}$ is decreasing in γ in this region. This is true as $\gamma_1 < b - d$.

Consider now types in $(0, \gamma_1)$.

$\Delta_{RD}^{(0,\gamma_1)} = \zeta b + (1-\zeta)a - \frac{1}{2}[(\zeta + (1-\zeta)(1 - \frac{\gamma_1}{a})b + (1-\zeta)(\frac{\gamma_1}{a})a + \zeta b + (1-\zeta))a] < 0$ for all ζ.

Consider types in $(\gamma_1, b - d)$: they have the same utility in R as the types below but a higher utility in D from their own point of view. Hence $\Delta_{RD}^{[\gamma_1,b-d]}$ must be lower and decreasing. Specifically:

$\Delta_{RD}^{[\gamma_1,b-d]} = \zeta b + (1-\zeta)a -$

$\frac{1}{2}[(\zeta + (1-\zeta)(1 - \frac{\gamma_1}{a}))d + \gamma + (\zeta + (1-\zeta)(1 - \frac{\gamma_2}{a}))b + (1-\zeta)(\frac{\gamma_2}{a})a] < 0$ for all ζ.

Consider now types in $(b - d, \gamma_2)$.

Note that type $b-d$ is indifferent between paying for the ritual or not and hence is utility from R is $\zeta b+(1-\zeta)a$. For all types above $b-d$, the utility from R will be $\zeta b+(1-\zeta)a+(\gamma-(b-d))(\zeta+(1-\zeta)(1-\frac{b-d}{a}))$, as their type affects their spiritual utility in the order of the probability by which they are cooperating. Their utility from D differs only in the spiritual payoff that accrue in the first period with probability one. Thus we have,

$\Delta_{RD}^{[b-d,\gamma_2]}=\Delta_{RD}^{[b-d,\gamma_2]}|_{\gamma=b-d}+(\gamma-(b-d))(\frac{1}{2}\zeta+(1-\zeta)(\frac{a-2(b-d)}{2a}))$.

This may be increasing or decreasing, depending on the sign of $\frac{1}{2}\zeta+(1-\zeta)(\frac{a-2(b-d)}{2a})$. If it is decreasing, then the highest Δ_{RD} is for the type at 0 and it is negative. So, if $\sigma<\sigma'(\zeta)\rightarrow_{\zeta\rightarrow 0}2$, then all prefer R, which proves (i). If it is increasing, so $\sigma>\sigma'(\zeta)$, then the highest Δ_{RD}, if positive, is for the type at γ_2. Thus two cutoffs $\gamma'>b-d$ and $\gamma''\in[\gamma',a]$ arise so that all supporters of R are in $[\gamma',\gamma'']$, with $\gamma_2\in[\gamma',\gamma'']$.■

Proof of Proposition 5: (i) We first analyze equilibrium existence for all π. For the excessive signalling equilibrium,

$$(*)\ ((1-\zeta)\frac{(a-\gamma_1)}{a}+\zeta)(d-b)+(1-\zeta)\frac{\gamma_1}{a}(-a)+\gamma_1+\pi((1-\zeta)\frac{(a-\gamma_2)}{a})(b-a)=0$$

We can repeat the analysis in Lemma 3 to show that for a π large enough, such an equilibrium would still hold. Note that the higher is π, the higher are the gains from signalling which implies that γ_1 must decrease and admits the lowest value for $\pi=1$.

For small enough π, we might now have a negative lhs for all $\gamma_1\leq b-d$; thus due to the monotonicity of the lhs in γ_1, the equilibrium must have accurate signalling. The equilibrium condition for the accurate signalling equilibrium is:

$$\begin{aligned}&((1-\zeta)\frac{(a-\gamma^*)}{a}+\zeta)(d-b)+(1-\zeta)\frac{\gamma^*}{a}(-a)+\gamma^*+\pi((((1-\zeta)\frac{(a-\gamma^*)}{a}+\zeta)d\\&+(1-\zeta)\frac{\gamma^*}{a}a+((1-\zeta)\frac{(a-\gamma^*)}{a}+\zeta)\gamma^*)-(\zeta b+(1-\zeta)a))\\=\ &0.\end{aligned}$$

For a small enough π this now can be an equilibrium and in particular in the limit when $\pi\rightarrow 0$, we have that $\gamma^*\rightarrow\hat{\gamma}$.

Due to the linearity in π, it is easy to see that there exists π' such that for all $\pi>\pi'$ there exists an equilibrium with excessive signalling, and moreover that $\gamma_1'(\pi)<0$. For $\pi<\pi'$ there exists an equilibrium with accurate signalling, again with $\gamma^{*\prime}(\pi)<0$.

(ii) In the accurate equilibrium, all agents below $b-d$ must prefer D always to R, as in R they get a while in D they get some b in the first period. This is the same in the excessive signalling equilibrium for all agents below γ_1. This establishes (ii).

(iii) Explained in the text.

(iv) We will now show that globaly, all agents prefer a high enough π to a low enough π. We will show this result for extreme $\pi's$, specifically $\pi = 0$ and $\pi = 1$. By continuity and given (i), this holds for high enough and low enough $\pi's$.

As the utility under $\pi = 0$ for individuals below $b - d$ is $(\zeta + (1-\zeta)(1-\frac{\hat{\gamma}}{a}))b + (1-\zeta)\frac{\hat{\gamma}}{a}a$ per period whereas in R their utility is $\zeta b + (1-\zeta)a$ per period, they prefer D for all ζ. On the other hand, for a small enough ζ, individuals above $b-d$, have a utility which converges to a per period under $\pi = 0$ and a utility which is strictly higher than a per period in R (they can always guarantee a if they defect and thus in any equilibrium their material and spiritual utility must be greater than a). More specifically, consider individuals in $\gamma \in [b-d, \hat{\gamma}]$. Their utility difference between rituals and $\pi = 0$ is increasing in γ and is the lowest for the type at $b-d$ (as he is gaining the least from spiritual payoff) for which, per period, it is $(\frac{\hat{\gamma}-a}{a})(b-a) < 0$. Thus, there must be a cutoff in $[b-d, \hat{\gamma}]$, above which, for small ζ (as then $\hat{\gamma} \to a$), all types prefer R and below which they prefer $\pi = 0$. For all ζ though, this cutoff is strictly above $b-d$.

It is easy to show that all types at least weakly prefer $\pi = 1$ to $\pi = 0$ in D. Thus, weakly, the set of supporters must increase. We now show that indeed there are types that switch their preferences. Consider the type at $b-d$ and his difference in utility between R and $\pi = 1$. From the proof of Proposition 4 we know that this type strictly prefers $\pi = 1$ for all ζ. On the other hand, when ζ is small enough, the utility of this type from $\pi = 0$ approaches his utility from R. By continuity, there exist a type $\gamma > b-d$ but close enough that switches to prefer D when π increases.■

Proposition 6: (i) For all equilibria with excessive signalling,

$$r = 2(1-\frac{\gamma_2}{a})(1-\zeta)(b-a)$$

Note that $dr = -d\gamma_2 \frac{2(b-a)(1-\zeta)}{a}$. To maximize $r(\zeta + (1-\zeta)(1-\frac{\gamma_1}{a}))$, the foc is $dr(\zeta + (1-\zeta)(1-\frac{\gamma_1}{a})) - d\gamma_1 r\frac{(1-\zeta)}{a} = -d\gamma_2\frac{2(b-a)(1-\zeta)}{a} - d\gamma_1 r\frac{(1-\zeta)}{a}$. We therefore care about the sign of $-d\gamma_2 2(b-a) - d\gamma_1 r$.

But according to (1), $\gamma_2 = \frac{a(b-d-\gamma_1)+a\frac{\zeta}{1-\zeta}(b-d)}{b-d-\gamma_1+a\frac{\zeta}{1-\zeta}}$, and then we have $d\gamma_2 = d\gamma_1 a\frac{\zeta}{1-\zeta}\frac{b-d-a}{((b-d-\gamma_1)+a\frac{\zeta}{1-\zeta})^2}$. Thus we need to check the sign of $\frac{a}{b-d-\gamma_1+a\frac{\zeta}{1-\zeta}} - (1-\zeta)$ but $\frac{a-(1-\zeta)(b-d-\gamma_1)-a\zeta}{b-d-\gamma_1+a\frac{\zeta}{1-\zeta}} > 0$ iff $b-d-\gamma_1 > 0$ which is indeed the case and hence this expression is positive. We therefore have revenues increasing in γ_1.

(ii) Now consider accurate equilibria, where the expression for the ritual cost is $r = 2(\zeta(d-b+\gamma^*) + (1-\zeta)(\frac{a-\gamma^*}{a})(d-a+\gamma^*))$. The revenues again are $r(\zeta + (1-\zeta)\frac{a-\gamma^*}{a})$ and the foc is $dr(\zeta+(1-\zeta)\frac{a-\gamma^*}{a}) - d\gamma^* r\frac{(1-\zeta)}{a}$. We then have $dr = d\gamma^* 2(\zeta+(1-\zeta)(\frac{-d+2a-2\gamma^*}{a}))$ so we need to check the sign of $(\zeta+(1-\zeta)(\frac{-d+2a-2\gamma^*}{a}))(\zeta+(1-\zeta)\frac{a-\gamma^*}{a})2 - 2(\zeta(d-b+\gamma^*)+(1-\zeta)(\frac{a-\gamma^*}{a})(d-a+\gamma^*))\frac{(1-\zeta)}{a}$, which for a small ζ is $(\frac{-d+2a-2\gamma^*}{a})\frac{a-\gamma^*}{a} - (\frac{a-\gamma^*}{a})(d-a+\gamma^*))\frac{1}{a}$. We then need to

check the sign of $-d+2a-2\gamma^*-d+a-\gamma^*=3a-2d-3\gamma^*$. Note that for this to be negative for all γ^* we need to check at $\gamma^*=b-d$. We then have $3a-2d-3b+3d=3a+d-3b$ which is negative when σ is not too large.■

7.2 Appendix B: A hybrid religion

We now consider the game assuming that both rituals and information about behaviour in the first PD game are fully available. It will be simpler and more sensible to consider a slight change in timing, where the game would be as follows: in Period 0 agents are randomly matched to play the PD game. In Period 1 they can pay r, and in Period 2 they are matched again to play the PD game now with possibly two signals available about their opponent.[43] As usual, we are considering monotone equilibria, so that (i) those who do not signal at all defect for sure in any remaining PD game(s); (ii) those that pay r are a connected interval and those that cooperate are a connected interval and one of these intervals is nested in the other. Thus society in equilibrium is divided into $[0,\gamma_1],[\gamma_1,\gamma_2],[\gamma_2,a]$, where $[\gamma_2,a]$ uses both signals, $[\gamma_1,\gamma_2]$ uses one signal and $[0,\gamma_1]$ does not signal.

We now consider the behaviour in Period 2 following the signalling behaviour. Note that it cannot be that all in $[\gamma_1,\gamma_2]$ defect against $[\gamma_2,a]$ as then by monotonicity they would also defect against each other. But also $[\gamma_2,a]$ must defect against them; hence no one changes their action in response to the signal used by $[\gamma_1,\gamma_2]$ which implies that their signalling is a waste, a contradiction. Also it cannot be that all agents in $[\gamma_1,\gamma_2]$ fully cooperate against $[\gamma_2,a]$ as then again by monotonicity all agents in $[\gamma_2,a]$ will cooperate against its own which implies that their signalling is a waste, a contradiction. However, it must be that all agents in $[\gamma_2,a]$ fully cooperate against each other. If not, then it must be that $\gamma_2<b-d$ which implies that all agents in $[\gamma_1,\gamma_2]$ will defect against agents in $[\gamma_2,a]$ which we showed above cannot be.

Given the above, it must be that: (i) $[0,\gamma_1]$ defect against all; (ii) $[\gamma_1,\gamma_2]$ partially cooperates with $[\gamma_2,a]$ and at least partially defecting against itself; (iii) $[\gamma_2,a]$ fully cooperates against itself and at least partially cooperates against $[\gamma_1,\gamma_2]$. The simplest form of equilibrium will include agents in $[\gamma_1,\gamma_2]$ defecting against each other and agents in $[\gamma_2,a]$ fully cooperating against $[\gamma_1,\gamma_2]$. Finally let us set the type at γ_2 indifferent between cooperating or not.[44] Let $\gamma_3\in[\gamma_1,\gamma_2]$ be the lowest type that cooperates against $[\gamma_2,a]$. Note that it must be that $\gamma_3=b-d$.

We now analyze the two possible cases: either the group at $[\gamma_1,\gamma_2]$ cooperates at period 0, or the group at $[\gamma_1,\gamma_2]$ pays r in period 1.

[43] This is simpler as otherwise the game in Period 1 will be both a signalling period and a period in which individuals respond to the previous signal r. This timing also captures more the essence of how to accomodate two signals.

[44] If other types of equilibria exist they will admit a similar structure.

Case 1: types in $[\gamma_1, \gamma_2]$ signal by cooperation in Period 0:

The equilibrium conditions are:

$$\begin{aligned}(\zeta + (1-\zeta)(\tfrac{a-\gamma_1}{a}))(d-b) + (1-\zeta)\tfrac{\gamma_1}{a}(-a) + \gamma_1 + (1-\zeta)(\tfrac{a-\gamma_2}{a})(b-a) &= 0 \\ \frac{(\gamma_2-\gamma_3)+a\frac{\zeta}{1-\zeta}}{(\gamma_2-\gamma_1)+a\frac{\zeta}{1-\zeta}}(d-b) + \frac{(\gamma_3-\gamma_1)}{(\gamma_2-\gamma_1)+a\frac{\zeta}{1-\zeta}}(-a) + \gamma_2 &= 0 \\ (1-\zeta)(\frac{(\gamma_2-\gamma_3)}{a}d + (\gamma_2-a)\frac{(\gamma_2-\gamma_1)}{a}) &= r\end{aligned}$$

For this case it is easy to show the following:

(i) For some parameters, as in Example 1, the revenues of the church are higher in the R religion than in the hybrid religion.

To create a fair comparison, we consider r in the R religion that applies for one signalling period, as here. We then have to show that:

$$(\frac{(\gamma_2-\gamma_3)}{a}d + (\gamma_2-a)\frac{(\gamma_2-\gamma_1)}{a})\frac{a-\gamma_2}{a} < \frac{a-\gamma_3}{a}(b-a)\frac{a-\gamma_3}{a}$$

Note that if for example $b > a + \frac{d}{2}$ (as in example 1) the above holds.

(ii) This equilibrium does not exist in Example 2.

Case 2: types in $[\gamma_1, \gamma_2]$ signal by paying r in Period 0:

In this case the equilibrium equations are:

$$\begin{aligned}(\zeta + (1-\zeta)(\tfrac{a-\gamma_2}{a}))(d-b) + (1-\zeta)\tfrac{\gamma_2}{a}(-a) + \gamma_2 + (1-\zeta)(\frac{(\gamma_2-\gamma_3)}{a}d + (\gamma_2-a)\frac{(\gamma_2-\gamma_1)}{a}) &= 0 \\ \frac{(\gamma_2-\gamma_3)+a\frac{\zeta}{1-\zeta}}{(\gamma_2-\gamma_1)+a\frac{\zeta}{1-\zeta}}(d-b) + \frac{(\gamma_3-\gamma_1)}{(\gamma_2-\gamma_1)+a\frac{\zeta}{1-\zeta}}(-a) + \gamma_2 &= 0 \\ (1-\zeta)(\tfrac{a-\gamma_2}{a})(b-a) &= r\end{aligned}$$

Again, as above, we can show:

(ii) In Example 1, the revenues of the church are higher in the R religion than in the hybrid religion.

(iii) This equilibrium does not exist in Example 2.

(iii) If we are limited to have $r \geq \frac{r^*}{2}$ (where r^* is the level of r in the Pareto efficient equilibrium) then this equilibrium cannot arise (to see why note that here $\gamma_2 > b - d$ and hence it must be that $r < r^*$).

As a final comment note that the gist of our results is maintained. Specifically, the second type of equilibrium often does not hold, which implies that paying r is much more likely to be an accurate signal whereas signalling by discipline is more likely to be an inaccurate signal. When comparing the hybrid religion to the "pure" religions it is the case that the R religion provides more coordination, whereas the D religion saves on cost, as in our main analysis.

7.3 Appendix C: "Belief Activation" refinement

Ritual-based religion: Suppose that individuals are not endowed with religious beliefs, but that they gain such beliefs only if they participate in rituals, i.e., pay the cost r. More specifically, individuals have "latent" types in $[0, a]$ and this type will be activated when they pay r but not activated otherwise. Many religious organizations play an active role in shaping beliefs and invest time and effort in advocating certain kinds of messages while censoring others. This assumption is therefore reasonable when considering religious organizations and it specifically fits the Catholic religion where the rewards from good works were also conditioned on participation in rituals. The role of rituals in such an alternative model is therefore two-fold: to endow individuals with beliefs favouring cooperation and to serve as a public signal.

Since this model involves a choice of beliefs, we need to add to the equilibrium concept a stability condition. Namely, an individual of type γ who in equilibrium had paid r and has activated her beliefs, will, given her current beliefs γ and other equilibrium behaviour, prefer to do so than not pay r and defect against all. Similarly an individual who had not paid r and had not activated her beliefs, prefers to do so than to acquire beliefs and sometimes cooperate, given her current beliefs (e.g., $\gamma = 0$) and equilibrium behaviour of others. For more on this stability notion and the robustness of the results to other stability notions, see Levy and Razin (2012).

The assumption on belief activation implies that whoever does not participate in rituals, has no beliefs in favour of cooperation, and will therefore defect. Together with the stability notion above, it implies that as in our model, equilibria can only be as described in Lemma 2 with exactly the same equilibrium conditions specified in the text, and that such equilibria indeed exist.

Discipline-based religion: as above, suppose that individuals have "latent" types in $[0, a]$ and that they have to choose to activate these beliefs prior to the two period PD game. As both Calvin and Luther called for believers to return back to the scriptures and read the bible themselves, suppose that it is costless and private to activate beliefs, and as a tie-breaking rule, that if individuals forecast that their utility from activating beliefs and not activating beliefs is the same, then they do not activate their beliefs. Again, the equilibrium will demand that individuals who did not activate their beliefs, given all other equilibrium behaviour, will be happy with this decision and vice verse.

With this assumption one can show that there will be no individual who defects in the first period and cooperates later on. Thus only monotone equilibria can arise as described in Lemma 2. Moreover, these equilibria indeed exist in this alternative model as it is optimal for all those who defect not to acquire beliefs (and hence defect from that point onwards).

References:

Alaoui, L. and A. Sandroni, (2013), "Predestination and the Protestant ethic," Economics Working Papers 1350, Department of Economics and Business, Universitat Pompeu Fabra.

Athey, S. Calvanoy E. and S. Jha (2010), "A Theory of Community Formation and Social Hierarchy", mimeo.

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

Andreoni, J. and L. Samuelson (2006), "Building rational cooperation", Journal of Economic Theory, vol. 127(1), pages 117-154.

Arrunada, B. (2010), "Protestants and Catholics: Similar Work Ethic, Different Social Ethic," Economic Journal 120 (547), 890–918.

Barro, R. and R. McCleary (2003), "Religion and Economic Growth across Countries", American Sociological Review, vol. 68(5), pp. 760-81.

Becker, S. O. and L. Woessmann, "Was Weber Wrong? A Human Capital Theory of Protestant Economic History," Quarterly Journal of Economics, May 2009, 124 (2), 531–596.

Benabou, R. and J. Tirole (2011), "Identity, Morals and Taboos: Beliefs as Assets", discussion paper no. 6123, CEPR.

Benabou, R. and J. Tirole (2006), "Belief in a Just World and Redistributive Politics", Quarterly Journal of Economics, vol. 121(2), pp. 699-746.Benabou, R. and J. Tirole (2004), "Willpower and Personal Rules", Journal of Political Economy, 112(4), PP. 848-886.

Berman, E., (2000), "Sect, Subsidy and Sacrifice: An Economist's View of Ultra-Orthodox Jews", Quarterly Journal of Economics, vol. 115(3), pp. 905-953.

Beza, T. (1996), *The Life of John Calvin*, Reprinted, Back Home Industries.

Blackwell, W. L., (1965), "The Old Believers and the Rise of Private Industrial Enterprise in Early Nineteenth-Century Moscow", *Slavic Review*, Vol. 24, No. 3, pp. 407-424.

Bocock, R. J., (1971), "The Ismailis in Tanzania: A Weberian Analysis", *The British Journal of Sociology*, Vol. 22, No. 4, pp. 365-380.

Bottocini, M. and Z. Eckstein (2005), "Jewish Occupational Selection: Education, Restrictions, or Minorities?", Journal of Economic History.

Botticini, M. and Z. Eckstein (2007), "From Farmers to Merchants, Conversions and Diaspora: Human Capital and Jewish History", Journal of the European Economic Association 5, no. 5, pp. 885-926.

Boyer, P. (2001), *Religion Explained*, Basic Books, New York, NY.

Cantoni, D. (2010), "The Economic Effects of the Protestant Reformation: Testing the Weber Hypothesis in the German Lands", mimeo, Universitat Pompeu Fabra.

Carvalho, J.P and M. Koyama, "Development and Religious Polarization: The Emergence of Reform and Ultra-Orthodox Judaism", mimeo.

Chwe, M.S. (2003), *Rational Ritual: Culture, Coordination, and Common Knowledge*, Princeton: Princeton University Press 2003.

Clarke, P. B., (1976), "The Ismailis: A Study of Community", *The British Journal of Sociology*, Vol. 27, No. 4, pp. 484-494.

Dixit, A. (2003), "Trade expansion and contract enforcement", Journal of Political Economy, Vol. 111, pp. 1293-1317.

Dommen E. and J.D. Bratt (2007), ed., *John Calvin rediscovered : the impact of his social and economic thought*, WJK.

Ekelund, R., Herbert R. and R. Tollison (2002), "An Economic Analysis of the Protestant Reformation", Journal of Political Economy, Vol. 110.

Ekelund, R., R. Herbert, R. Tollison, G. Anderson and A. Davidson (1996), *Sacred trust : the medieval church as an economic firm*, Oxford University Press.

Flick, A.C (1930, reprinted 1967), *The decline of the medieval church*, Bert Franklin Publishing, New York.

Gilbert, W. (1998), *Renaissance and Reformation*, Electronic edition, Kansas University.

Glaeser E. and S. Glendon (1998), "Incentives and Predestination", Economic inquiry, vol 36.

Goldthorpe, J. E., The Sociology of Post-Colonial Societies

Economic Disparity, Cultural Diversity and Development, Cambridge University Press, 1996.

Green, V.H.H (1964), *Luther and the Reformation*, Putnam's, New York.

Greif, Avner (1989), "Reputation and Coalitions in Medieval Trade: Evidence on the Maghribi Traders", The Journal of Economic History, vol. 49(04), pp. 857-882.

Greif, Avner & Tabellini, Guido, 2012. "The Clan and the City: Sustaining Cooperation in China and Europe," CEPR Discussion Papers 9072, C.E.P.R. Discussion Papers.

Guiso, L. P. Sapienza, and L. Zingales (2006), "Does Culture Affect Economic Outcomes?" Journal of Economic Perspectives, 20, 23-48.

Guiso, L. P. Sapienza, and L. Zingales (2003), "People's opium? Religion and economic attitudes," Journal of Monetary Economics 50, 225–282.

Huber, J.D. (2005), "Religious belief, religious participation, and social policy attitudes across countries", working paper, Columbia University.

Iannaccone, L. (1992), "Sacrifice and stigma: reducing free-riding in cults, communes and other collectives", Journal of Political Economy, vol. 100(2), pp. 271-291.

Iannaccone, L. (1998), "Introduction to the Economics of Religion", Journal of Economic Literature, vol. 36(3), pp. 1465-1495.

Kanatas G., and C. Stefanadis (2010), "Culture, Institutions, and Appropriation", SSRN working paper.

Kingdom, D. (1992), "Calvin and the Family: The Work of the Consistory in Geneva", in

Calvin's *work in Geneva*, ed. R. Gamble, Graland.

Kranton, R. (1996), "Reciprocal Exchange: A Self-Sustaining System", The American Economic Review Vol. 86 (4), pp. 830-851.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., and Vishny, R. W. (1997), "Trust in large organizations", American Economic Review, vol. 87(2), pp. 333-38.

Levintova, E. (2007), "Does History Repeat Itself? Public Discourse of the Contemporary Russian Old Believer Elite", *The Slavonic and East European Review*, Vol. 85, No. 4, pp. 753-779.

Levy, G. and R. Razin (2011), "Calvin's Reformation in Geneva: Self and Social Signalling", mimeo, LSE.

Levy, G. and R. Razin (2012), "Religious Beliefs, Religious Participation and Cooperation", American Economic Journals: Micoreconomics.

McCleary, R. M. (2007) "Salvation, damnation and economic incentives", Journal of Contemporary Religion, vol. 22(1), pp. 49-74.

MacCulloch, D. (2010), *A History of Christianity*, Penguin Books.

McGrath, A. (1990), *A life of John Calvin*, Oxford University Press.

Ozment, S. E. (1980), *The age of reform, 1250-1550 : an intellectual and religious history of late medieval and Reformation Europe*, Yale University Press.

Pesendorfer, W. (1995), "Design Innovations and Fashion Cycles", American Economic Review, vol. 85, pp. 771–792.

Potter, G. R. and M. Greengrass (1983), John Calvin, St. Martin's Press, New York.

Putnam, R. (1993), *Making Democracy Work: Civic Traditions in Modern Italy*, Princeton, New Jersey: Princeton University Press.

Raskov, D. E., "Norms in the Economic Evolution: Old Believers in the Russian Nineteenth-century Textile Industry", working paper, St. Petersburg State University.

Scheve, K. and D. Stasavage (2006), "Religion and Preferences for Social Insurance", Quarterly Journal of Political Science, vol.1, pp.255-286.

Sosis, R. & Ruffle, B. (2004) Ideology, Religion, and the Evolution of Cooperation: Field Tests on Israeli Kibbutzim. Research in Economic Anthropology 23, 89-117.

Stark, R. (1996), *The Rise of Christianity*, Princeton University Press, Princeton.

Tabellini, G. (2008), "The Scope of Cooperation: Values and Incentives", The Quarterly Journal of Economics, MIT Press, vol. 123(3), pages 905-950.

Vorontsova L. and S. Filatov, (2000), "Paradoxes of the Old Believer Movement" *Religion, State and Society,* Vol. 28, Iss. 1.

Weber, M. (1904), *The Protestant Ethic and the Spirit of Capitalism*, New York: Scribners (translated by T. Parsons, 1930).

Weber, M (1906), ""Churches" and "Sects" in Northern America: An Ecclesiastical Sociopo-

litical Sketch", trans. Colin Loader. *Sociological Theory* 3: 7-13. (1986 translation).