The Evolution of Sectarianism

Ille, Sebastian

New College of the Humanities

18 January 2021

This is the accepted manuscript. Please refer to the published journal article in Communications in Nonlinear Science and Numerical Simulation https://doi.org/10.1016/j.cnsns.2021.105726. (http://www.sciencedirect.com/science/article/pii/S100757042100037X)

The Evolution of Sectarianism

Sebastian Ille

New College of the Humanities, London

Human cooperation for reasons other than self-interest has long intrigued social scientists leading to a substantial literature in economics. Its complement – sectarianism – has not received closer attention in economics despite its significant impact. Based on a dynamic model, the paper shows that sectarianism can be understood as the outcome of a repeated bargaining process in which sectarian affiliation evolves into a pure coordination signal that attributes economic and political benefits. It demonstrates that such sectarian social contracts co-evolve with the sects' degree of coerciveness and are self-reinforcing. Sectarian conflict may then not be a result of diverging religious ideologies but is shown to be caused by external manipulations of the signal (e.g. via identity politics), and internal political and economic grievances within a sect that spill over to the inter-sectarian level while adopting a sectarian appearance. Theoretical results are supported by empirical findings from the Middle East.

Keywords: Sectarianism, Cooperation, Evolutionary Game Theory, Agent-Based Modelling

1 Introduction

The evolution of cooperation has received much scientific attention during the past two decades and has brought together economists, anthropologist, and sociologists.[1] On the other hand, its antithesis - sectarianism - has not been studied in the same detail, and the evolutionary basis remains under-researched across the social sciences. It remains heavily debated by political scientists, while economists have mostly ignored sectarianism. Yet in August 2020, an explosion in Beirut demonstrated only to clearly the significant socio-economic impact it can have on a society. Furthermore, the Lebanese Civil War, the Yugoslav Wars, the Rwandan Genocide, and the Syrian Civil War are only but a few examples of sectarian-driven conflicts since the past

[1] The editors of Science ranked cooperation as one one of the top 25 questions facing scientists today (Pennisi, 2005).

century. Sectarian societies forgo opportunities of mutually advantageous exchange and cooperation between members of different religious affiliations while sectarian conflict is harmful both at a group and an individual level.[2]

Contrary to common perception, but supported by the literature (e.g. Barr, 2011; Makdisi, 2000; Salloukh et al., 2015), the paper argues that sectarianism and sectarian conflict may not be driven by religious ideologies, but can be seen as an evolutionary solution to a bargaining problem which distributes benefits according to sectarian affiliations. Sectarian conflict is then the result of a signalling mismatch or a contestation of the sectarian signal. To support this theory, the paper explores the factors leading to sectarianism and sectarian conflict, using analytical tools from evolutionary game theory and agent-based modelling. Going beyond existing literature, this paper demonstrates a first mathematical approach to the analysis of *sectarianisation* that systematically identifies and studies the determinants of sectarianism and sectarian conflict. Although the theoretical approach is based on a simple model of exchange which does not take account of the individual characteristics of sectarianisation in different regions, it is able to replicate the predominant and general characteristics of sectarianism and sectarian conflict. The model offers a clearer understanding of the underlying dynamics and demonstrates the co-evolution of seemingly ideological fronts along sectarian lines, and socio-economic or political benefits. It shows that religious affiliation may not be the cause of sectarianism but serves as its crucial descriptive characteristic. The evolutionary model provides strong support for the literature arguing that sectarianism is rooted in and reinforced by endogenous social and economic grievances and imposed inequalities linked to a group's cultural identity (for a broad overview of both the specific reasons and the general properties, see Hashemi and Postel, 2017 and for an empirical example, see Mitra and Ray, 2014). It further demonstrates that sectarian contracts are self-reinforcing and hence, the need for active policy measures.

2 Literature and Context

Sectarianism has been extensively studied by political scientists, and in recent years, especially in the Middle Eastern context. Recent in depth analyses of the region are provided by Hashemi and Postel (2017) and Wehrey (2017). A good overview of the historical context in the Middle East is provided in Barr (2011), while Salloukh et al. (2015), Salibi (1988), and Makdisi (2000) offer detailed analyses of the specific context of Lebanon. These authors are part of a modern interpretation of sectarianism that disagrees with other explanations (e.g. Nasr, 2006; Lewis, 1997) which argue on the basis of a *primordial* or *instrumentalist* perspective. The primordial, instrumentalist, and constructivist perspectives are not homogeneously defined in the literature and tend to conflate and intermix (for a critical study, see Tilley, 1997). Yet, in the context of sectarianism, the former primordial perspective is frequently based on a historical explanation postulating that religious conflicts are fuelled by instinctive impulses, innate social characteristics and ancient or *perceived* social givens (on the latter, see especially Geertz, 1973). Reinforced by tribal blood feuds for one and a half millennia, this has organically shaped and reinforced a cultural predisposition to violence. Sectarianism and sectarian conflict are deeply

[2] See, for example, the fatwas forbidding Sunnis to eat meat slaughtered by Shi'as or the marriage between both sects.

rooted in the social relations of sectarian societies defined by inevitable and lasting perpetual sectarian differences.

Instrumentalism (e.g. Fearon and Laitin, 2000; Lake and Donald, 1998; Taylor and Yapp, 1979), on the other hand, postulates that sectarian and ethnic conflict are externally created and reinforced by state and political actors who construct identities to advance their political and economic interests by emphasising on in-group similarities and out-group differences. This paper provides strong support for a mixed approach - the *constructivist* perspective - which recognizes the significance of pre-determined (and potentially imagined) differences in ethnic or religious identities, but also the changing character of these identities and perceives them as not inevitably driving a social system into conflict. Active manipulation by leaders, elites, and external actors play a critical role in the mobilisation process. Religious identities take on the role of markers of evolving socio-economic cleavages and demands.

Since the study of sectarianism is a natural extension of the study of cooperation, existing models of cooperation are therefore a fertile basis for developing analytical models of sectarianism. Sociobiology studies the evolution of social behaviour within species via natural selection (e.g. Wilson, 1975) and offers a wide plethora of models (Hamilton, 1975; Keller, 1999; Wade, 1978; Wilson, 1980). In social systems, fitness based replication is substituted by (biased) cultural transmission (see Cavalli-Sforza and Feldman, 1973; Dawkins, 1976; Popper, 1979 and especially, Boyd and Richerson, 1988, 2005a). However, models of cooperation frequently assume small populations defined by simple and recurrent interactions, where either detrimental actions against a group lead to a punishment of the specific individual perpetrator or behaviour that may not be beneficial for an individual is supposed to be beneficial for a group as a whole. These conditions do not generally hold in sectarian societies.

While economists have not studied sectarianism, some economists have examined ethnic conflict which is functionally analogous to sectarianism. However, literature discussing the theoretical underpinnings of intra-state ethnic conflict (e.g. Esteban and Ray, 2011) frequently assumes a pre-existing segregation and hostility between groups while relying on some version of the contest model. In these models, contest occurs only at the group level, often reduced to two uniform groups, i.e. rebels and government (other examples are Haavelmo, 1954; Hirshleifer, 1988, 1989; Garfinkel, 1990; Grossman, 1991; Skaperdas, 1992 and for a broader overview, see Garfinkel and Skaperdas, 2007; Blattman and Miguel, 2010), thus disregarding the co-evolution of conflict and ethnic divide. Other literature focuses on collective actions and free-riding (for further elaboration, see Tullock, 2005; Weinstein, 2005) as well as the principal motivators (on the grievance versus greed debate, see Collier and Hoeffler, 2004; Keen, 2012), and the role of individual incentives and social sanctions (for the impact of these effects on intra-state conflict, see Humphreys and Weinstein, 2008). Similarly, the multi-dimensionality of sectarianism requires that a proper study does not focus on religious or ethnic fault lines alone, but includes the socio-economic institutional framework. A number of scholars illustrate that ethnicity is exploited for economic and political competition (Bates, 1983; Fearon, 2005; Chandra, 2004). The feedback between political and economic benefits and sectarian conflict is therefore essential for understanding the evolutionary dynamics underlying sectarianism. This renders the paper more akin to the literature focusing on the co-evolution of institutions and culture (e.g. Belloc and Bowles, 2013; Bisin and Verdier, 2015; Bowles, 2004; Bowles and Gintis, 2011; Boyd and Richerson, 2005b; Levine and Modica, 2012).

Identity is another essential constituent of sectarianism. While identity economics (Akerlof and Kranton, 2010) fails to explain crucial characteristics of religious groups (for a discussion, see Ille, 2017), it forms a theoretical basis for studying the connections between group identification (e.g. perceived in-group similarity and out-group dissimilarity, see Basu, 2005) and the internalization of ideals and norms (Akerlof and Kranton, 2002; Bernheim, 1994, see also social identity theory, Goffman, 1959). Identity influences preferences and decisions, determines power relations and social interactions, and creates a sense of belonging (Horst et al., 2009). It explains characteristics of sectarianism, such as status effects (Veblen, 2009) and public actions (Granovetter, 1978), upholding norms and retaliating against norm violators (Fehr and Gächter, 2000), as well as ostracism of outsiders and favouritism of insiders (Sherif et al., 1961). Identity formation simultaneously feeds on and causes sectarianism leading to a recursive relationship.

This literature includes two models that serve as a starting point for this research. The model by Choi and Bowles (2007) illustrates the conditions necessary for the evolution of ostracism and group conflict, as well as discriminatory (or parochial) altruism in the presence of unconditional altruism. The system's dynamics are determined by a biological selection process that operates simultaneously at an intra- and inter-group level. The model can be easily adapted to the given context by adjusting the migration rate, the replication process, and the frequency of interaction. While such an extension of Choi and Bowles (2007) can explain a number of properties of sectarian conflict (such as the co-evolution of parochial behaviour and conflict), the model's dynamics are fundamentally affected by group size. In larger group sizes, parochial and altruistic traits do not persist in the long-run. The model can therefore not explain sectarianism in larger societies. In contrast, the model in Axtell et al. (2001) is based on a Nash demand game, similar to the one used here, and provides an intuition for the occurrence of in-egalitarian and discriminatory institutions, while being unaffected by population and group size. However in this model, institutional change occurs by chance and institutions are generally efficient since individuals coordinate in most interactions. Instead of providing a strong rational of the underlying dynamics, individuals coordinate either on an egalitarian or in-egalitarian equilibrium out of an ergodic process and a coincidental sequence of non-best replies.[3] The model thus cannot explain the coercive nature of sectarianism that leaves potential for free-riding, as well as the occurrence of enduring sectarian conflict since agents only settle on pure strategy equilibria in the long-run. In contrast, the model presented in this paper takes account of the co-evolution of sectarianisation and the coercive measures taken by sectarian members. As a result, the evolution of a sectarian social contract is not driven by random actions, but by social imitation. The closed-form solution of the simplified model studies the dynamics defining a sect's coerciveness and provides conditions under which a society converges to a particular equilibrium, thereby eliminating social contracts that are implausible in the long-run. In addition, the simulations extend the initial model and show how institutional conditions can support the evolution of sectarianism and sectarian conflict.

In the absence of a uniform definition in the literature, I define sectarianism as a form of social contract that attributes socio-economic privileges and political rights based on the adherence to a specific religious sect. Sectarian conflict is then defined as a violent contestation

[3] With a small probability, an actor erroneously chooses a strategy other than the best response to her prior experience.

of this social contract, i.e. a persistent state of coordination failure in which sect members do not mutually agree on a social contract and aim to impose advantageous sharing norms based on coercive measures. Following the constructivist perspective, neither sectarian identities nor sectarian conflict are assumed to exist a priori, but are an emergent property determined by the history of endogenously established interactions and practice between individuals. I focus on the conditions necessary for the evolution of sectarianism and the mutual re-enforcement of its characteristics. Since sectarian affiliation only serves as a signal for the distribution of rights and privileges, I study under which condition a discriminatory marker defining an individual's position in a socio-economic contract emerges that is linked to sectarian affiliation. I further analyse the factors supporting an endogenous evolution of the coercive nature of sectarianism, which enforces abidance to a sectarian social contract, as well as the drivers of conflict in a sectarian society.

The following section elaborates the baseline model that will provide a general comprehension of the systems' dynamics based on a closed-form solution. This model is extended in section 4 to provide a fuller understanding of the underlying dynamics of sectarianisation and sectarian conflict by adding institutional drivers, such as a collective and individual history, identity politics, favouritism, and different forms of social learning to the model. Section 5 interprets the theoretical results, compares them with empirical findings and concludes.

3 The basic model with two sects

Assume that all members of the population are randomly matched in pairs to play a simplified Nash demand game with three possible strategies. The demand game is thereby representative of the determination of a general social contract in which individuals demand a share or set of political rights or socio-economic benefits. Each player l can choose a pure strategy $s_l \in S = \{L, M, H\}$, where L stands for demanding *low* equivalent to a share of $(1-\lambda)$ for $\lambda \in (0.5, 1)$, M for *medium* equivalent to a share of 50%, and H for *high* equivalent to a share of λ. Obtaining a larger share is in the interest of each individual, yet should the joint demands of both matched players exceed 100%, no social contract is formed and both will receive nothing. The payoffs are then as defined in matrix 1.

$$\begin{array}{c} \\ \mathbf{L} \\ \mathbf{M} \\ \mathbf{H} \end{array} \begin{array}{c} \begin{array}{ccc} \mathbf{L} & \mathbf{M} & \mathbf{H} \end{array} \\ \begin{pmatrix} 1-\lambda, 1-\lambda & 1-\lambda, 0.50 & \boldsymbol{1-\lambda, \lambda} \\ 0.50, 1-\lambda & \boldsymbol{0.50, 0.50} & 0,0 \\ \boldsymbol{\lambda, 1-\lambda} & 0,0 & 0,0 \end{pmatrix} \end{array} \tag{1}$$

The three Nash equilibria are shown in bold. A *socially efficient* outcome occurs if the joint net gains from exchange are equal to 1. All three Nash equilibria are socially efficient, but only (M, M) is egalitarian.[4]

[4]Note that the equilibria are identical under the assumption that the inegalitarian outcomes are socially inefficient (e.g. by defining the payoffs as λ for H and $(1-\lambda-\varepsilon)$ for L and a small $\varepsilon > 0$ in outcomes (L, H) and (H, L)). The reason for excluding this assumption is here to focus on social efficiency caused by coordination failures, instead of confounding both sources of inefficiency. The following results are unaffected by this assumption and also hold if a small dead-weight loss occurs at inegalitarian/sectarian social contracts.

Assume that two sufficiently large subpopulations, i.e. sects, each of size n exist and that only members of different sectarian affiliation are matched. Members of a sect are identified by an arbitrary marker. For a sect $i = \{1,2\}$, I define x_i as the frequency of L-players, y_i as the frequency of H players and $z_i = 1 - x_i - y_i$ as the frequency of M-players. Both players can only make one offer per matching. Since each player does not know what her counterpart will demand, she can only choose a strategy based on the likelihood of encountering a player choosing a particular pure strategy. In the absence of segmentation, the likelihoods equal the frequencies in the population.Since all frequencies add up to 1, any population distribution of sect i can be defined by the tuple (x_i, y_i). Assuming that interactions take place regularly over time and players update their strategy by comparing their own payoffs to the payoffs of others, the changes to the population frequencies can be modelled by the following replicator equations

$$\dot{x_i} = x_i(\pi_{xi} - \phi) \tag{2a}$$

$$\dot{y_i} = y_i(\pi_{yi} - \phi) \tag{2b}$$

where π_{xi} and π_{yi} define the expected payoff for choosing strategy L and H, respectively, and $\phi = x_i\pi_{xi} + y_i\pi_{yi} + z_i\pi_{zi}$ defines the average payoff in sect i. In other words, whenever a strategy does better (worse) than the average, its frequency increases (decreases).

Furthermore, players are assumed to choose whether to coerce other players into accepting their offer. Therefore, in addition to choosing between strategies low, medium, and high, they can be of one of two types: either coercive (indicated by superscript c in the following) or acquiescent (indicated by superscript a).[5] While an acquiescent type does not take any further actions if paired with another player and joint demands exceed 1, all coercive members join the coercive player of the same sect and collectively attempt to impose H on the other player. Thus, success of coercion depends on the abundance of coercive players in the respective sect. Let the type frequency of coercive members in sect $i = \{1,2\}$ interacting with a member of sect $j = \{1,2\}$, $i \neq j$ be defined by α_i and assume the probabilities of successful coercion is[6]

$$\frac{1}{2} + \frac{\alpha_i - \alpha_j}{2} \tag{3}$$

All probabilities extend over the entire unit interval, but include the possibility of a member in a sect with fewer other coercive members to decide the conflict in her favour. If the member succeeds, she receives her highest possible payoff of λ and receives a payoff of zero otherwise. Since the losing party receives a payoff of 0 instead of $(1-\lambda)$, any coercive action constitutes a socially inefficiency outcome.[7] Each time a coercive player loses, all her supporters pay a cost.

[5] In general, I will use *type* to refer to being coercive or acquiescent, whereas *strategy* refers to choosing to demand low, medium, or high.

[6] While the literature on conflict frequently employs a different functional form for contest success probabilities p_i, i.e. $p_i = \alpha_i/(\alpha_i + \alpha_j)$ (see Skaperdas, 1996), the functional form used here is common in the evolutionary literature (see e.g. Boyd and Richerson, 2002; Choi and Bowles, 2007) and the simulation results are independent of the form used. The former probabilities, however, generate significantly more complicated eigenvalues and render the analysis of the closed-form solution impracticable.

[7] One might argue that a player should only impose her original demand and not the maximum given by λ. However, to be consistent, the cost of coercion should also depend on the original demand. This will unnecessarily complicate the model without adding much to its explanatory power.

Given the winning probabilities, a player of sect i loses an encounter with probability

$$\eta_i = \frac{1+\alpha_j-\alpha_i}{2} \tag{4}$$

The share of encounters at which joint demands exceed 1 and a member is coercive in sect i is

$$\zeta_i = \alpha_i \left(z_i y_j + y_i z_j + y_i y_j\right) \tag{5}$$

Further assume that each unsuccessful encounter results in a loss of $\sigma_i > 0$. The cost is shared among all coercive sect members given by $\alpha_i n$. The expected cost of being coercive is then

$$\gamma_i = n\zeta_i\eta_i \frac{1}{\alpha_i n}\sigma_i = \frac{(1+\alpha_j-\alpha_i)\left((1-x_i)y_j + y_i z_j\right)\sigma_i}{2} \tag{6}$$

The expected payoffs for the six strategy-type combinations for sect i is given by

$$\pi^a_{xi} = 1-\lambda \tag{7a}$$
$$\pi^a_{yi} = \lambda\ x_j \tag{7b}$$
$$\pi^a_{zi} = 0.50(x_j+z_j) \tag{7c}$$
$$\pi^c_{xi} = 1-\lambda-\gamma_i \tag{7d}$$
$$\pi^c_{yi} = \lambda\left(x_j + (y_j+z_j)\frac{1+\alpha_i-\alpha_j}{2}\right) - \gamma_i \tag{7e}$$
$$\pi^c_{zi} = 0.50(x_j+z_j) + \lambda y_j \frac{1+\alpha_i-\alpha_j}{2} - \gamma_i \tag{7f}$$

Consequently, being coercive and choosing strategy L is strictly dominated. Since expected benefit of coercion has to outweigh cost, the surplus of a coercive type is

$$\pi^c_{zi} - \pi^a_{zi} = (1-\eta_i)\lambda y_j - \gamma_i \qquad \text{for an } M \text{ player} \tag{8a}$$
$$\pi^c_{yi} - \pi^a_{yi} = (1-\eta_i)\lambda(1-x_j) - \gamma_i \qquad \text{for an } H \text{ player} \tag{8b}$$

Both types benefit from an increase in the share λ obtained by successfully enforcing a demand of high, and suffer from an increase in the additional cost of an unsuccessful coercion. In a state of randomized strategy and type frequencies, $\pi^c_{zi} > \pi^a_{zi}$ if $\lambda > \sigma_i$ and $\pi^c_{yi} > \pi^a_{yi}$ if $\lambda > 0.5\sigma_i$.

The dynamic system is then defined by 3 equations for each sect i

$$\dot{x}_i = x_i\left[\left((1-\alpha_i)\pi^a_{xi} + \alpha_i\pi^c_{xi}\right) - \bar{\pi}_i\right] \tag{9a}$$
$$\dot{y}_i = y_i\left[\left((1-\alpha_i)\pi^a_{yi} + \alpha_i\pi^c_{yi}\right) - \bar{\pi}_i\right] \tag{9b}$$
$$\dot{\alpha}_i = \alpha_i(1-\alpha_i)(\bar{\pi}^c_i - \bar{\pi}^a_i) \tag{9c}$$

with $\bar{\pi}_i = x_i\left[(1-\alpha_i)\pi^a_{xi} + \alpha_i\pi^c_{xi}\right] + y_i\left[(1-\alpha_i)\pi^a_{yi} + \alpha_i\pi^c_{yi}\right] + z_i\left[(1-\alpha_i)\pi^a_{zi} + \alpha_i\pi^c_{zi}\right]$, and for the type frequencies $\bar{\pi}^c_i = x_i\pi^c_{xi} + y_i\pi^c_{yi} + z_i\pi^c_{zi}$, and $\bar{\pi}^a_i = x_i\pi^a_{xi} + y_i\pi^a_{yi} + z_i\pi^a_{zi}$. The equilibria of this dynamic system are given by any distribution at which no change in the population frequencies occurs, i.e. at $\dot{x}_1 = \dot{x}_2 = \dot{y}_1 = \dot{y}_2 = \dot{\alpha}_1 = \dot{\alpha}_2 = 0$. However, the long-run evolving properties

of this model are of greater interest. Individuals are prone to commit errors and to rarely but recurrently choose a strategy that is not their best response. Evolutionary stable or asymptotically stable states are unaffected by this kind of idiosyncratic play, as long as the latter occurs sufficiently infrequently. The stability of an equilibrium is then determined by the eigenvalues of the Jacobian of the six-dimensional system at this fixed point.

However, solving for $\dot{x}_1 = \dot{x}_2 = \dot{y}_1 = \dot{y}_2 = 0$ defines six *strategy* equilibria, in which joint demands do not exceed 1. Consequently, coercion is never enforced, and coercive and acquiescent members obtain the same payoff. Since $\bar{\pi}_i^c = \bar{\pi}_i^a$, equation 9c is zero for any (α_1, α_2) across the unit interval. Any such *type* equilibrium is neutrally stable at the strategy equilibria. To address this problem, I will first study the stability of the strategy equilibria for any $\alpha_1, \alpha_2 \in [0,1]$, and thereafter, the stability of the type equilibria in the vicinity of the stable strategy equilibria.

The strategy equilibrium analysis shows that the three potentially evolutionary/asymptotically stable states are identical to those three pure strategy Nash equilibria of the simpler Nash demand game without types. However, the eigenvalues define the following additional conditions[8]

$\{x_i = 1, y_i = 0, x_j = 0, y_j = 1\}$ only asymptotically stable if

$$\alpha_i < \hat{\alpha}_i^s = \frac{1}{2}\left(-(1-\alpha_j) + \sqrt{(1-\alpha_j)^2 + 8\frac{1-\lambda}{\lambda}}\right) \tag{10a}$$

$\{x_1 = 0, y_1 = 0, x_2 = 0, y_2 = 0\}$ only asymptotically stable if for both $i, j = (1,2), i \neq j$

$$\alpha_i < \hat{\alpha}_i^e = \frac{1}{2}\left(-(1-\alpha_j) + \sqrt{(1-\alpha_j)^2 + \frac{4}{\lambda}}\right) \tag{10b}$$

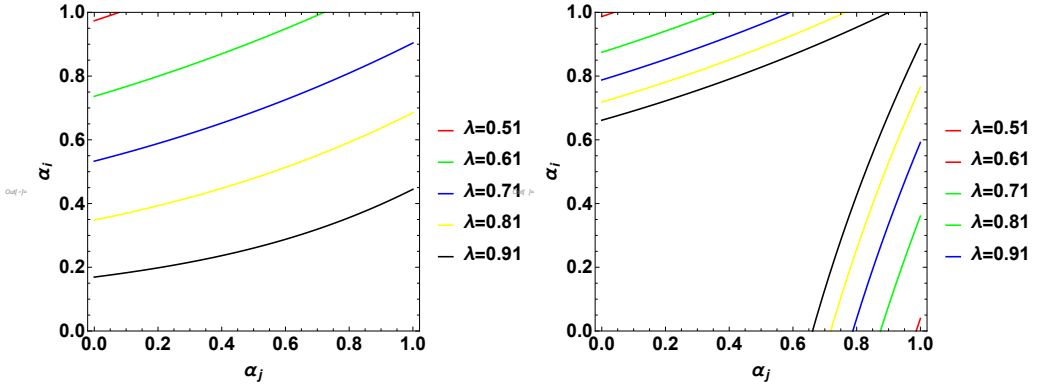
In the equilibrium of 10a, members discriminate according to sectarian affiliation, while the equilibrium of 10b is non-discriminatory. I will therefore refer to the former as a sectarian contract, and to the latter as an egalitarian contract. It holds that

$$\frac{\partial \hat{\alpha}_i^s}{\partial \lambda} = -2\left(\lambda^2\sqrt{(1-\alpha_j)^2 + 8\frac{1-\lambda}{\lambda}}\right)^{-1} < 0 \tag{11a}$$

$$\frac{\partial \hat{\alpha}_i^e}{\partial \lambda} = -\left(\lambda^2\sqrt{(1-\alpha_j)^2 + \frac{4}{\lambda}}\right)^{-1} < 0 \tag{11b}$$

Figure 1 provides an illustration for the shape of $\hat{\alpha}_i^e$ and $\hat{\alpha}_i^s$ given various parameter values for λ. All equilibria are stable at $\alpha_1 = \alpha_2 = 0$. While the egalitarian contract remains stable at $\alpha_1 = \alpha_2 = 1$, the sectarian contracts are only stable at this type distribution if $\lambda < 2/3$. Sectarian contracts, on the other hand, are stable at $\alpha_j = 1$ and $\alpha_i = 0$, while at the egalitarian contract, it must hold that $\alpha_j < (4\lambda/(\lambda+4))^{-1/2} - \frac{1}{2} < 1$ for $\alpha_i = 0$. An increase of the unilateral gains λ of a sectarian contract consolidates this type of social contract if the benefiting sect is strongly coercive and the other sect acquiesces. However, it promotes the evolution of an egalitarian contract if both sects are strongly coercive.

[8] Note that we can write the conditions as $\alpha_i < (1-\lambda)/((1-\eta_i)\lambda)$ and $\alpha_i < (2(1-\eta_i)\lambda)^{-1}$, respectively, which shows that the thresholds are functions of the winning probability and the gains from H.



(a) Threshold value $\hat{\alpha}_i^s$ of sectarian contract (b) Threshold value $\hat{\alpha}_i^e$ of egalitarian contract

Figure 1. Threshold values $\hat{\alpha}_i^e$ and $\hat{\alpha}_i^s$ given various parameter values for λ.

Given equation 9c, the interior stationary values for α_i for $i, j = \{1,2\}$ and $i \neq j$ are given by the isolines at $\dot{\alpha}_i = 0$ which are

$$\widetilde{\alpha}_i = \frac{(1+\alpha_j)\sigma_i - (1-\alpha_j)\lambda}{\lambda + \sigma_i} \tag{12}$$

Since for $\alpha_i > \widetilde{\alpha}_i$ ($\alpha_i < \widetilde{\alpha}_i$) the share of coercive members in sect i will increase (decrease), the isolines in 12 determine the equilibria's basins of attraction.[9] The isolines are independent of the strategies, and hence, the direction (but not the speed) of the dynamics are unaffected by the strategy distributions of both sects. The stability of the type equilibria can thus be studied independently of the values of x_1, x_2, y_1, and y_2. The nulls of equation 9c and the eigenvalues of the Jacobian return two unconditionally asymptotically stable equilibria, defined by

$$\{\alpha_1 = 0, \alpha_2 = 1\} \text{ and } \{\alpha_1 = 1, \alpha_2 = 0\} \tag{13}$$

two conditionally asymptotically stable equilibria

$$\{\alpha_1 = 0, \alpha_2 = 0\} \text{ only asymptotically stable if } \lambda < \sigma_1, \sigma_2 \tag{14a}$$
$$\{\alpha_1 = 1, \alpha_2 = 1\} \text{ only asymptotically stable if } \lambda > \sigma_1, \sigma_2 \tag{14b}$$

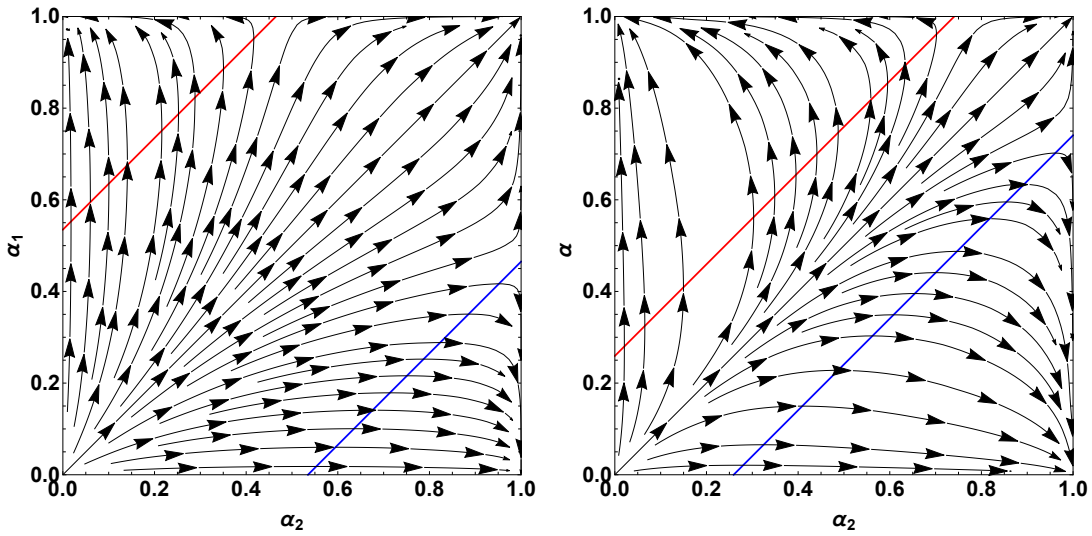
as well as two unstable interior equilibria.

We have

$$\frac{\partial \widetilde{\alpha}_i}{\partial \lambda} = -\frac{2\sigma_i}{(\lambda+\sigma_i)^2} \qquad < 0 \tag{15a}$$
$$\frac{\partial \widetilde{\alpha}_i}{\partial \sigma_i} = \frac{2\lambda}{(\lambda+\sigma_i)^2} \qquad > 0 \tag{15b}$$

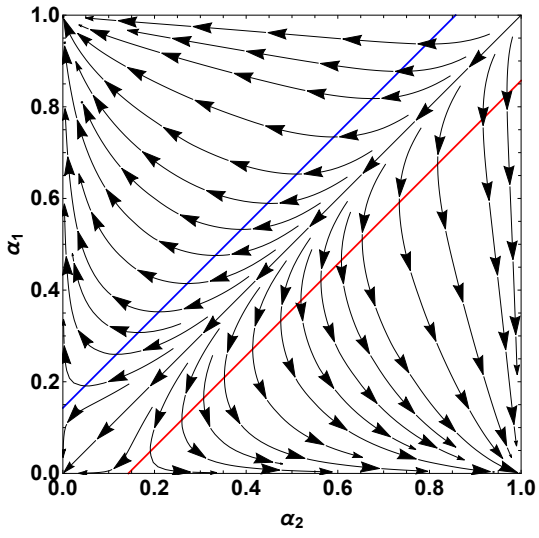
[9] An equilibrium's basin of attraction defines the set of states which dynamically evolve to this equilibrium. In the presence of infrequent idiosyncratic and erroneous individual choices, the basin of attraction therefore is indicative of the likelihood with which this equilibrium occurs.

implying that an increase in the cost of coercion for sect i moves $\widetilde{\alpha}_i$ towards the $\{a_i = 1, a_j = 0\}$ state, thereby decreasing its basin of attraction. An increase in the benefits of a sectarian contract, given by λ has the inverse effect. Thus, the basin of attraction of the $\{a_1 = 1, a_2 = 1\}$ state is largest at $\sigma_1, \sigma_2 \approx 0$ and $\lambda \approx 1$, i.e. at very small costs of coercion and large unilateral gains at the sectarian contract. In the inverse scenario at $\sigma_1, \sigma_2 \gg 0$ and $\lambda \approx 0.5$, the basin of attraction of $\{a_1 = 0, a_2 = 0\}$ is maximal. In the absence of coercive members, the interaction follows the classical Nash demand game and the basin of attraction of the egalitarian contract increases in λ. If the costs of coercion of both groups are not identical the likelihood of the sectarian equilibrium increases with the difference in costs and relative to the potential gain λ.

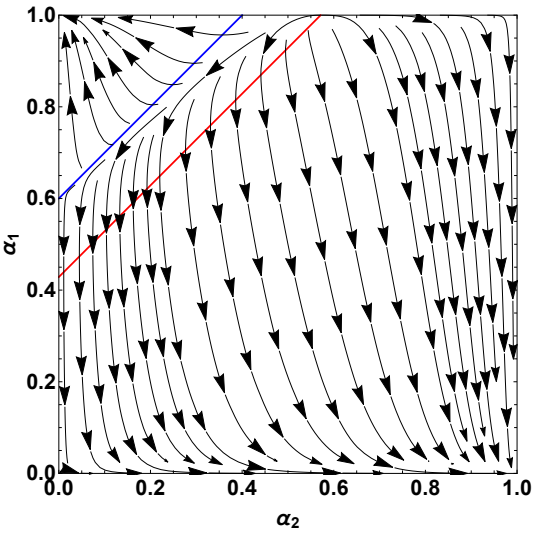


(a) $\sigma_1 = 0.3, \sigma_2 = 0.3, \lambda = 0.99$

(b) $\sigma_1 = 0.3, \sigma_2 = 0.3, \lambda = 0.51$



(c) $\sigma_1 = 1, \sigma_2 = 1, \lambda = 0.75$



(d) $\sigma_1 = 3, \sigma_2 = 0.3, \lambda = 0.75$

Figure 2. Dynamics of α_1 and α_2 for different values of σ_1, σ_2 and λ.

Figure 2 illustrates different scenarios and demonstrates the impact of σ_1, σ_2, and λ on the dynamics of α_1 and α_2. Conditions 10a, 10b, 13, 14a and 14b define the system's long-run equilibria. For example, assuming that a sufficiently large population is initially entirely randomly distributed implies that it is initially situated at the centre of the mixed-strategy/type simplex, i.e. at $x_1 = x_2 = y_1 = y_2 = 1/3$ and $\alpha_1 = \alpha_2 = 1/2$. If $\lambda > \sigma_1, \sigma_2$, dynamics will push the population towards $\alpha_1 = \alpha_2 = 1$ and $z_1 = z_2 = 1$. Low costs of coercion and a high prospective gain from a sectarian contract will lead to a balance of power and hence, an egalitarian contract in which all individuals are coercive. The coercive character of this state renders the equilibrium more stable against individual idiosyncratic choices and thus decreases the likelihood of a sectarian contract (due to the violation of condition 10a). I denote such an equilibrium a consociational contract. On the other hand, if $\sigma_j > \lambda > \sigma_i$, equation 12 shows that such an initially random distribution will converge to a sectarian contract, defined by $y_i = x_j = 1$, $\alpha_i = 1$, and $\alpha_j = 0$. The conditions do not directly define the long-run strategy composition of the population if $\lambda < \sigma_1, \sigma_2$. However, since $\alpha_1 = \alpha_2 = 0$ in the long-run, the interactions will revert back to the standard Nash demand game. If the convergence of the types occurs sufficiently rapidly, the population will evolve towards an egalitarian contract $z_1 = z_2 = 1$ since the centre of the unit simplex is within the basin of attraction of the latter asymptotically stable state.

In general terms, this implies that a sectarian society is likely to evolve if only one sect has a higher gain from such a contract than the cost of coercion.[10] If both sects are initially defined by similar characteristics, the pervasive social contract will be egalitarian or consociational depending on how effortlessly norm adherence can be enforced by few coercive members. Once a population settles into a contract, only exogenous changes to the ability and the cost of coercion in conjunction with idiosyncratic actions can destabilize the equilibrium. Since the equilibria are asymptotically/evolutionarily stable, any social contract is self-reinforcing in the absence of such exogenous change.

4 Extensions

Although the simple model offers an explanation for some of the dynamics we observe in sectarian societies, it ignores several important drivers of sectarianisation. It can therefore explain sectarianism only on the basis of differences in the cost of coercion and does not demonstrate reasons for sectarian conflict, since any population will always converge to a coordination equilibrium. In its current form, the model takes account of the limited cognitive abilities of individuals, since individual choices are based on social learning. However, the model ignores the historical component of sectarianisation and thus, the impact of past experience and a collective memory on individual decision-making (Hashemi and Postel, 2017). In addition, the model neglects the role of identity politics, differences in social learning, sectarian favouritism, and intra-sectarian socio-economic grievances. This section will therefore extend the baseline model and analyse the impact of factors other than the cost of coercion.

Akin to the literature using fictitious play with limited memory, assume that a member l of sect i has a memory of size m and recollects the last history of interactions with sect members of sect j. Thus, each time l makes a new experience with a member of i, she forgets the oldest

[10] Note that this also holds for the theoretical case in which costs equal gains for both sects.

interaction and adds the latest interaction to her memory. Member l does not know with whom she will be matched nor the strategy that will be played against her, but only that her counterpart is of the opposite sect. She assumes the likelihood with which the latter chooses a strategy is defined by the empirical frequency distribution of past play with the other sect. She then chooses a strategy that maximizes her expected payoff.[11] For example, if an individual observes that members of the other sect sufficiently frequently demand L, her best response in the current period is to choose H.

If both strategies exceed a joint demand of one, and coercion is successfully attempted by one sect, the winning member recollects that the losing sect played L and correspondingly, the losing members recollects that the winning sect played H. If no contest occurs or no winner is determined (i.e. the coercion has been unsuccessful), both parties retain the other's original offer. In this type of fictitious play, a player takes account of the type distribution in a limited rational manner and not directly in the cognitively demanding way that a recollection of the share of supporting coercive members for each period would require, but by remembering which strategy was successful at the end of the interaction.

I further assume that members (denoted as *learner*) regularly update their types by comparing their success over the past n periods to another member of the same sect (denoted as *teachers*) and, for simplicity, I assume $n = m$.[12] In addition, in some of these social learning encounters, the learner does not only adopt the type of the teacher, but also the latter's recollection of past interactions. Depending on the relative frequency of this event with respect to the length of memory m, individuals weigh higher either collective or individual memory. In the most extreme case in which social imitation and an adoptions of a teacher's past history occur in each period, an individual bases her choice exclusively on the teacher's memory. As long as members are randomly matched, sect members will then base their choice on a common collective memory after some time. If a player never embraces her teacher's past experience, choice is exclusively based on individual experience.

Separating the mechanism defining an individual's choice of her strategy from that of her type entails a form of multi-level selection. The strategy choice is based on an individual's interaction with the other sect (i.e. it is subject to cross-sectarian selection), while the type is defined based an encounters with and in comparison to members of the same sect (i.e. it is subject to intra-sectarian selection). At the same time, I also take account of the interplay between individual experience and collective memory. To test the model's robustness, I study the dynamics with respect to four different common replicators. In each period, a share $\tau \in (0, 0.5)$ of members of sect i either

1. adopt the type of the best performing sect member in that period and with probability $\zeta \in [0,1)$, the latter's past memory (*Variant 1*).

2. are randomly paired with a member of the other sect, while one adopts the role of learner k, the other of teacher l. If the teacher performs strictly better, the learner adopts the teacher's

[11] Given the history of the m last interactions as defined by pure strategy-tuple $s_{lj} = \big(s_j(t), s_j(t-1), \ldots, s_j(t-m+1)\big)$, let $\eta_l^t(s_j) : S_j \to \mathbb{N}$ be the frequency that l observed a member of sect j playing strategy s_j in the past m interactions. Given the player's payoff function $\pi_l : S \to \mathbb{R}$, player l assumes the likelihood of this strategy in period t is then $\mu_l(s_j) = \eta_l^t(s_j)/m$ and chooses $s_l \in \arg\max_{s_l \in S_l} \pi_l(s_l, \mu_l)$.

[12] This assumption has no impact on simulation results as long as n is of similar scale to m.

type with probability $\theta(\pi_k, \pi_l) \in (0,1)$ proportional to the payoff difference $\pi_k - \pi_j$ and in this case, the latter's past memory with probability ζ (*Variant 2*).[13]

3. are randomly paired, and type and past memory are imitated as in Variant 2, but the teacher is not randomly drawn from the entire sect, but from a predefined peer-group within the sect i (*Variant 3*).

4. are randomly paired, and type and past memory are imitated as in Variant 2, but adoption can even occur if the payoff difference is negative, such that $\theta(\pi_k, \pi_l) \in (0, 0.5)$ ($\theta(\pi_k, \pi_l) \in (0.5, 1)$) if learner k performs worse (better) than teacher l (*Variant 4*).[14]

These variants cover a broad spectrum of social learning replicators and are listed according to their degree of selection pressure against the less efficient type. A less successful type will not be imitated in variant 1 and has the highest likelihood of adoption if members follow variant 4. To incorporates the superior ability of politically and economically more successful individuals to impose their conditions, loss is defined as $\sigma_i = \kappa\left(\bar{\pi}^h / \pi_k^h\right)$, with strictly positive constant κ and $\bar{\pi}^h$ being the population average after history h and π_k^h being the individual's average payoff.

4.1 Simulation Results

All simulations show that the convergence time and likelihoods of the equilibria are independent of groups size. Similarly, the size of the peer group relative to the entire sect's population has an insignificant impact on the likelihood of the final equilibrium. To understand the dynamics of an unbiased system starting from a state in which no social contract has yet been determined nor that is affected by external influence, each individual is initiated with a memory of past play of length m drawn from a random distribution of all three strategies.

Table 1 shows the general results for each model. Since the likelihoods of each of the three social contracts differ only marginally between variant 2 to 4, the table illustrates the joint results for these replication variants. Table 2 provides details on the marginal impact of the variants and parameters. Table 1 further demonstrates that likelihoods are critically dependent on whether λ is below a critical value $\bar{\lambda}$. Sectarian conflict occurs whenever the likelihoods of the egalitarian and the sectarian contracts sum to less than 100 percent.

In the *balanced* model, half of all members of each sect are initiated as coercive and sects are identical in the first period of each simulation run, i.e. the population is initially at the centre of the unit simplex in Figure 2. Table 1 shows that if social learning is defined by variant 2 to 4 and the potential gain λ is less than 0.9, the population converges to the egalitarian contract in almost each simulation run. Extensive simulations show that this result is robust to changes in the length of memories m and n, the share of members open to social learning τ, and the relative importance of the collective memory ζ. For $\lambda \geq 0.9$, a prolonged sectarian conflict occurs in approximately half of all simulation runs, while the population converges to an egalitarian contract in the remaining half. If, in addition, $\zeta = 1$ a sectarian conflict occurs in approximately three fourth of all simulation runs. Thus, very high potential gains from a sectarian contract combined

[13] Note that this form of replication is closest to the one modelled by the replicator dynamics in section 3, see Weibull, 1997

[14] See Boyd and Richerson, 1992, for the underlying replicator dynamics.

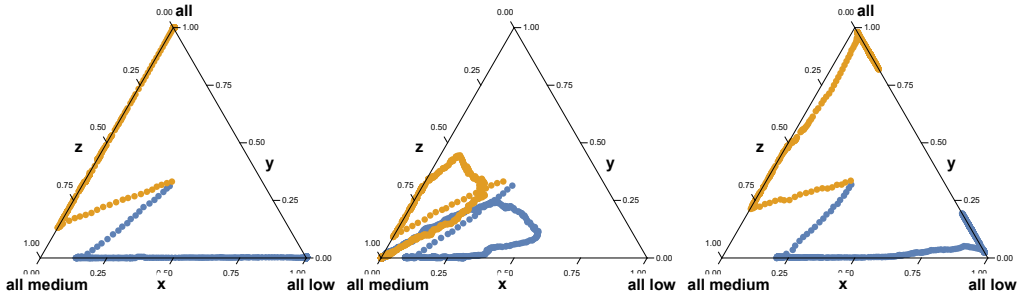
Table 1. Simulation Results: Likelihoods of Social Contracts in Percent

		Variant 2 - 4				Variant 1			
λ		balanced	unbalanced	Identity	Intern	balanced	unbalanced	Identity	Intern
$< \bar{\lambda}$	egal.	100	40	0	99	70	50	0	95
	sect.	0	54	76	0	20	50	81	2
$\geq \bar{\lambda}$	egal.	50	9	0	0	23	9	0	5
	sect.	0	57	2	1	30	76	32	0

Note: For balanced and unbalanced, $\bar{\lambda} = 0.9$, and for Identity and Intern, $\bar{\lambda} = 0.8$.

with a persuasive collective memory are the main drivers of a sectarian conflict if members interact based on common social learning and sects are defined by roughly equal characteristics.

In the the *unbalanced* model, both sects are not initially equally coercive, but one sect is initially composed of only 1 percent coercive members while the other sect consists of 99 percent coercive members, i.e. the population is initially at the top left or bottom right corner of the unit simplex in Figure 2. The chance of an egalitarian contract is now lower, while roughly half of the simulation runs settle into a sectarian contract. The initially highly coercive sect demands *H* and the acquiescent sect *L*. This is consistent with the results of section 3. In the remaining simulation runs, the population remains in a state of sectarian conflict. For $\lambda \geq 0.9$, the chance of a sectarian conflict strongly increases while roughly half of all runs converge to a sectarian contract. If a sectarian conflict occurs, all members are universally coercive. Figure 3 exemplifies the evolution of the strategy composition of both sects for three different runs. The position of a sect is defined by the frequency distribution of the three strategies across the members' memory of past play. Since in the initial state, all members draw a random sample from a strategy distribution of the three strategies, the sects are positioned at the centre of the unit simplex at the beginning of each simulation run (i.e. $x = y = z = 1/3$). Figures 3a to 3c show examples for the three possible scenario and the corresponding path each sect takes across $1,000$ simulation periods.



(a) Convergence to sectarian contract (b) Convergence to egalitarian contract (c) No convergence to an equilibrium

Figure 3. Example trajectories of both sects within the strategy unit simplex for 1,000 periods

Model 1a in Table 2 provides a more detailed analysis. It illustrates the impact of all parameters that significantly affect the likelihood of the egalitarian contract for replication variant 2 to 4.[15] The dummy variable *unbalanced* separates between the above models. Dummy *Clientelism* indicates whether a learner also receives the teacher's accumulated payoffs from past interactions. Clientelism should be interpreted as sectarian favouritism, i.e. obtaining social rights and status by following one's peers, or benefiting from legislation or economic ties that favour one group over the other.

All parameters, except for the slightly positive effect of a longer memory, decrease the chance of an egalitarian contract. An increase in memory length implies a longer reverberation of the initial randomly generated memory and biases the dynamics towards an egalitarian contract. All other parameters are positively correlated with selection pressure. Higher potential gains from a sectarian contract, a stronger adherence to a collective memory, defined by ζ, a higher likelihood of being matched in each period, defined by τ, and sectarian favouritism reduces the likelihood of an egalitarian social contract. Similarly, replication variants 2 and 3 show an adverse effect since both variants imply a higher selection pressure against an inferior type compared to variant 4, in which low performing members are imitated with positive probability.

Results change, if replication follows variant 1 and details are shown in Model 1b in Table 2. A stronger influence of collective memory and more frequent matching benefit the egalitarian contract. This contrasts with the other replication variants. If members are likely to imitate singular dominant members, an increased exposure to these members and a stronger reliance on collective memory increases the probability of an egalitarian social contract, contrary to the other variants. As shown on the right of Table 1, both the initial type distribution and the gains from demanding high affect dynamics. In both models, the egalitarian contract is significantly less likely and the sectarian contract more probable if $\lambda \geq 0.9$. A balanced initial distribution is more prone than the unbalanced type distribution to settle in a sectarian conflict under variant 1 replication. In addition, while an increase in λ exerts limited influence on the likelihood of a sectarian contract, it strongly shifts the odds against an egalitarian contract in favour of a sectarian conflict under any replication variant.

4.2 Identity, Conflict, and Spill-overs

Sectarian identities are subject to changing institutional conditions and discursive practices (Hashemi and Postel, 2017) and are thus constructed by a mixed socio-economic, political and religious heritage. These endogenously evolving sectarian identities are implicitly included in the model. The memory of past interactions, on which individuals base their choices, defines the role and status that each individual attributes to members of a specific sect, i.e. the latter's social identity. At the same time, an individual's offer and thus, the self-concept of her position in the prevalent social contract, i.e. her personal identity, is defined in relation to this perception. Both the sectarian contract and the egalitarian contract are each defined as a pure strategy equilibrium in which mis-coordination is entirely absent. Consequently, under these social contracts, each individual's personal identity coincides with her social identity rendering interactions socially

[15] Note that if the sectarian contract is the dependent variable instead, the signs for each parameter are inverted while the absolute values are approximately identical. This holds both for model 1a and 1b.

efficient. Therefore, the model provides an explanation for the evolution of sectarian social contracts that assign socio-economic benefits to a religious marker, but, at this stage, it still fails to explain social conflict beyond being a rare random event if potential gains from a sectarian contract are moderate.

Sectarian conflict occurs when discursive practices do not strictly follow an undirected evolutionary process but are channelled by elites, and clientelist politics shape incentive structures that prevent alternative non-sectarian social contracts. Elites accentuate differences to shape the distribution of rights and economic privileges in their favour. This external exertion of influence is simulated in model *Identity* (see Table 1) by assuming that individuals do not initially base their choices on a history of randomly drawn strategies from the strategy set (L,M,H), as in the previous simulations, but on a history that is defined by a random sample taken from the truncated strategy set (L,H), i.e. exogenous politics stressed extreme positions and have completely eliminated the memory of an egalitarian offer M. In this case, the population is not only unable to coordinate on an egalitarian contract, but fails to converge to a sectarian contract with a high likelihood for $\lambda \geq 0.8$.[16] Model 2a and 2b in Table 2 illustrate the impact of the independent variables on the likelihood of a sectarian conflict for variants 2 to 4 and variant 1, respectively. Increased selection pressure (both in regards to frequency of interactions and replication variant) promotes sectarian conflict in model 2a. In model 2b, the share of matches per period decreases the likelihood of a conflict, since members are matched with only the most successful member, thus increasing coordination.

In addition, intra-sectarian inequalities based on exploitative socio-economic contracts are not confined to sectarian boundaries, but can spill over and cause conflict at the larger cross-sectarian level, while taking on a sectarian veneer. To test the model's predictions, I split each sect into two sub-groups in model *Intern* (see Table 1). I assume that individuals from different sub-groups but the same sect are able to identify the other's affiliation to a subgroup, whereas individuals matched with a member of the other sect cannot discriminate beyond sectarian affiliation. At the initial stage, a sequence of strategies drawn randomly from the entire strategy set define each individual's memory, equivalent to model 1. One sect is defined by one subgroup being 99 percent coercive, and the other subgroup being 99 percent acquiescent, while the other sect has an equal share of coercive and acquiescent members in both subgroups. Thus while at the aggregate, both groups seem identical (having the same total number of coercive and acquiescent members), at the meso-level, one sect is defined by heterogeneous sub-groups and the other by two homogeneous sub-groups.

Simulations show no significant difference between the four sub-groups and two sects. It therefore is of no importance where the inequality originated, and the exploitative internal contract can spill over to the other sect. Table 1 shows the impact of the potential gains λ from a sectarian contract. At $\lambda < 0.8$, the population almost always converges to an egalitarian norm but a cross-sectarian conflict occurs for $\lambda \geq 0.8$. Model 3a and 3b in Table 2 shows which variables promote cross-sectarian conflict for variants 2 to 4 and variant 1, respectively. All variables, except memory length, positively contribute to a sectarian conflict. Model 3b offers further details. Similar to model 1, the dependence of a collective memory and the frequency of

[16] Note that in the baseline model, strategy M is strictly dominated at $\lambda < 0.75$ if all acquiesce. Since λ is increased in increments of 0.1, this explains the shift in regimes at $\lambda = 0.8$.

social learning have an inverse effect compared to model 3a.

If the in-egalitarian intra-sectarian contract spills over to the other sect, conflict arises between both sects with a likelihood of 30 percent under variants 2 to 4 and 70 percent under variant 1. Similarly, in a sectarian conflict, at least one sect remains internally in-egalitarian with a likelihood of 37 percent under variants 2 to 4 and of 69 percent under variant 1, illustrating the interaction between intra-sectarian inegalitarian contracts and cross-sectarian conflict. Model 4a and 4b in Table 2 show the regression results for this state as the dependent variable for variants 2 to 4 and variant 1, respectively. Comparing the signs between model 3 and 4 shows that the parameters have a similar impact in both models while collective memory and the frequency of interaction differ in their impact based on the variant of social learning.

5 Interpretation and Conclusion

In contrast to the instrumentalist and primordial perspective, the model supports the perception of sectarian identities as non-perpetual, but constructed by an endogenous process of past interactions (for a detailed example, see Makdisi, 2000). Sectarianism therefore constitutes an emergent property in which sectarian affiliation is not the cause of sectarianisation, but rather serves as a marker or descriptive characteristic that assigns a role to each sect member in a social contract. While the instrumentalist argument stresses that economic competition and political opportunities are the principal incentives of political leaders to mobilize identities, the model further shows that these components also directly constitute elements essential to the endogenous evolution of sectarianism. The model demonstrates a co-evolution of political and economic influence, and sectarian institutions; a fact which is observable in countries like Bahrain, Syria, and Lebanon. Consequently, the paper argues that sectarianism is not necessarily motivated by differences in religious ideologies and doctrines, but by political and economic grievances. The self-stabilizing character of the sectarian equilibrium provides an intuition for the self-reinforcing nature of sectarianism and stresses the need for active policy interventions. As Hashemi and Postel (2017, pg. 21) write: "Despite the constructed character, sectarianisation has the ominous potential to become a self-fulfilling prophecy. Putting the sectarian genie back in the bottle is unlikely to be easy." Further and contrary to the primordial perspective, a natural social evolution of two groups sharing equal characteristics is unlikely to generate a sectarian contract. The model identifies two principal cases in which sectarian contracts can develop: if (1) an antecedent imbalance in the ability for coercive actions exists, or if (2) sect members strongly identify with pre-eminent members.

In the first case, the model illustrates that ease and frequency of coercion, as well as sectarian favouritism leading to political and economic imbalances foster the evolution of a sectarian state. For example, the *Kalashnikov culture* and trade in contraband and narcotics have significantly contributed to sectarianism and sectarian violence in Pakistan and Afghanistan (Nasr, 2017). Consistent with the model's predictions, while a tremendous number of militiae eventually consolidated a consociational contract in Lebanon, power imbalances in Syria and Iraq supported the co-evolution of an increasingly coercive nature of individual interactions and a sectarian social contract. The latter contract is more likely to evolve if individuals rely more on a collectively shared memory than their own past experience, and are more inclined to adopt

rituals and forms of engagements from supposedly more successful peers. This explains the incentives behind state actors and elites with an interest in an increased sectarianisation to invoke the history and norms shared by sect members.

In the second case, however, adopting actions performed by the pre-eminent members, as well as endorsing their past memories support the evolution of an egalitarian and non-sectarian society, while the transfer of benefits from these members to their followers benefits a sectarian contract. This might explain the *'Do as I say, not as I do'* mentality of elites, such as the Saudi regime acting as a protector of the Shia minority (Al-Rasheed, 2017), or the collaboration of the Al Khalifa family with Shia elites in Bahrain (Matthiesen, 2017), and the assignment of rights and positions by sectarian elites, as is frequently found in countries like Lebanon.

The model further demonstrates that sectarian conflict is not a naturally evolving property following a non-biased evolutionary process. The literature argues that a fragile state, class dynamics, and geopolitical rivalries form the principal impetus of sectarian conflict (Hashemi and Postel, 2017). Correspondingly, the model demonstrates that socio-economic deprivations and inequalities, as well as power asymmetries between sub-sect actors can spill-over to other groups while taking on a sectarian appearance leading to a state-wide sectarian conflict. This explains the transformation of non-sectarian uprisings, which initially represented a cross-section of society, into conflicts along seemingly sectarian boundaries, like in Syria, Yemen, and Bahrain in 2011. Similarly, the model also shows that the contestation of existing power relation and externally reinforced rivalries over socio-political dominance can lead to a disintegration of a social contract and thus, to sustained sectarian conflict. Exogenous mobilisation based on economic, social or political differences along sectarian identities can lead to persistent rivalries between sects and a struggle over their position in a social contract (see for example the recent politics employed by the Saudi regime in Al-Rasheed, 2017). This also shows why attempts to devitalize critical movements and to deflect state criticism by strengthening other religio-political groups, as has been done by Sadat and Zia ul-Haq in the 1970s, has led to an increase in sectarian violence. What then appears to be a narcissism of small differences is a conflict not motivated by minor ideological disparities, but by a divergent perception of the role and rights assigned to the sectarian marker.

The model shows that the prospect of significant socio-economic benefits for a sect under a sectarian contract is an essential driver of sectarian conflict. The correlation between potential benefits and conflict is not simply linear, but the likelihood of a sectarian conflict changes considerably if prospective gains cross a certain threshold. In addition, the model also demonstrates the ambiguous role of collective memory (Ross, 2013; Páez and Liu, 2015; McGrattan and Hopkins, 2016). While collective memory can help avoid sectarian conflict and support peace building, it can severely destabilize any form of social contract in the presence of identity politics stressing sectarian differences and socio-economic inequalities within sects. This has been empirically shown in Salloukh (2019), which illustrates how a selective deployment of war memories is used as a tool against geopolitical and domestic threats while inciting sectarian tensions. In illustrating the self-reinforcing character of sectarian conflict, the model demonstrates the need for external intervention. It stresses the need to thwart politics which emphasize sectarian differences and entitlements, and the important role of a balance of power among the various religio-political groups, the reduction of socio-economic inequalities, thus the need for contestation, and lastly, the individual gains from a sectarian contract in order to reinstate peace.

6 Acknowledgements

I would like to extend my deepest gratitude to Bassel Salloukh and Raymond Hinnebusch not only for their detailed feedback but also for their encouragement and interest in this research. I am also particularly grateful to Jurgen Brauer and Chuck Anderton for their comments and support. Furthermore, I would like to extend my sincere thanks to the attendees of the Annual Meeting of the American Economic Association, the St Catharine's Political Economy Seminar Series at Cambridge University as well as the two anonymous peer-reviewers for their remarks and feedback.

References

Akerlof, G. A. and R. E. Kranton (2002). Identity and schooling: Some lessons for the economics of education. *Journal of Economic Literature 40*(4), 1167–1201.

Akerlof, G. A. and R. E. Kranton (2010). *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being*. Princeton University Press.

Al-Rasheed, M. (2017). Sectarianism as counter-revolution - saudi responses to the arab spring. In N. Hahsemi and D. Postel (Eds.), *Sectarianization - Mapping the New Politics of the Middle East*, Chapter 8, pp. 143–158. Hurst & Company.

Axtell, R., J. M. Epstein, and H. P. Young (2001). The emergence of classes in a multi-agent bargaining model. In S. N. Durlauf and H. P. Young (Eds.), *Social Dynamics*, Chapter 7, pp. 191–211. MIT Press.

Barr, J. (2011). *A Line in the Sand*. Simon & Schuster.

Basu, K. (2005). Racial conflict and the malignancy of identity. *Journal of Economic Inequality 3*, 221 – 241.

Bates, R. H. (1983). Modernization, ethnic competition, and the rationality of politics incontemporary africa. In D. Rothchild and V. A. Olunsorola (Eds.), *State versus Ethnic Claims: African Policy Dilemmas*, pp. 153–171. Westview Press.

Belloc, M. and S. Bowles (2013). The persistence of inferior cultural-institutional conventions. *American Economic Review 103*(3), 93–98.

Bernheim, D. (1994). A theory of conformity. *Journal of Political Economy 102*(5), 841–877.

Bisin, A. and T. Verdier (2015). On the joint evolution of culture and institutions.

Blattman, C. and E. Miguel (2010). Civil war. *Journal of Economic Literature 48*(1), 3–57.

Bowles, S. (2004). *Microeconomics – Behavior, Institutions, and Evolution*. Princeton University Press.

Bowles, S. and H. Gintis (2011). *A Cooperative Species - Human Reciprocity and Its Evolution*. Princeton University Press.

Boyd, R. and P. J. Richerson (1988). *Culture and the Evolutionary Process*. The University of Chicago Press.

Boyd, R. and P. J. Richerson (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and Sociobiology 13*, 171–195.

Boyd, R. and P. J. Richerson (2002). Group beneficial norms can spread rapidly in a structured population. *Journal of Theoretical Biology 215*, 287–296.

Boyd, R. and P. J. Richerson (2005a). *The Origin and Evolution of Cultures*. Oxford University Press.

Boyd, R. and P. J. Richerson (2005b). Solving the puzzle of human cooperation. In S. C. Levinson and P. Jaisson (Eds.), *Evolution and Culture*, pp. 105–132. MIT Press.

Cavalli-Sforza, L. L. and M. W. Feldman (1973). Cultural versus biological inheritance: Phenotype transmission from parents to children. *American Journal of Human Genetics 25*, 618–637.

Chandra, K. (2004). *Why Ethnic Parties Succeed: Patronage and Ethnic Headcounts in India*. Cambridge University Press.

Choi, J.-K. and S. Bowles (2007). The coevolution of parochial altruism and war. *Science 318*, 636–640.

Collier, P. and A. Hoeffler (2004). Greed and grievance in civil war. *Oxford Economic Papers 56*, 563–595.

Dawkins, R. (1976). *The Selfish Gene*. Oxford University Press.

Esteban, J. and D. Ray (2011). A model of ethnic conflict. *Journal of the European Economic Association 9*(3), 496–521.

Fearon, J. D. (2005). Why ethnic politics and "pork" tend to go together.

Fearon, J. D. and D. D. Laitin (2000). Ethnicity, insurgence and civil war. *American Political Science Review 97*(1), 75–90.

Fehr, E. and S. Gächter (2000). Cooperation and punishment in public goods experiments. *The American Economic Review 90*(4), 980–994.

Garfinkel, M. R. (1990). Arming as a strategic investment in a cooperative equilibrium. *American Economic Review 80*(1), 50–68.

Garfinkel, M. R. and S. Skaperdas (2007). Economicsofconflict: An overvie. In S. Todd and K. Hartle (Eds.), *Handbook of Defense Economics*, Volume 2. Elsevier, North-Holland.

Geertz, C. (1973). Thick description: Toward an interpretive theory of culture. In C. Geertz (Ed.), *The Interpretation of Cultures: Selected Essays*. Basic Books.

Goffman, E. (1959). *The Presentation of Self in Everyday Life*. Doubleday.

Granovetter, M. (1978). Threshold models of collective behavior. *The American Journal of Sociology 83*(6), 1420–1443.

Grossman, H. I. (1991). A general equilibrium model of insurrections. *The American Economic Review 81*(4), 912–921.

Haavelmo, T. (1954). *A Study in the Theory of Economic Evolution*. North-Holland.

Hamilton, W. D. (1975). Innate social aptitudes of man: An approach from evolutionary genetics. In R. Fox (Ed.), *Biosocial Anthropology*, Chapter 9, pp. 133–155.

Hashemi, N. and D. Postel (2017). *Sectarianization - Mapping the New Politics of the Middle East*. Hurst & Company.

Hirshleifer, J. (1988). The analytics of continuing conflic. *Synthese 76*(2), 201–233.

Hirshleifer, J. (1989). Conflictand rent-seeking success functions: Ratio vs. difference models of relative success. *Public Choice 63*(2), 101–112.

Horst, U., A. Kirman, and M. Teschl (2009). Changing identity: The emergence of social groups.

Humphreys, M. and J. Weinstein (2008). Who fights? the determinants of participation in civil war. *American Journal of Political Science 52*(2), 436–455.

Ille, S. (2017). Towards better economic models of social behaviour? identity economics. *Studies in Ethnicity and Nationalism 17*(1), 5–24.

Keen, D. (2012). Greed and grievance in civil war. *International Affairs 88*(4), 757–777.

Keller, L. (1999). *Levels of Selection in Evolution*. Princeton University Press.

Lake, D. A. and R. Donald (1998). *The International Spread of Ethnic Conflict Fear, Diffusion, and Escalation*. Princeton University Press.

Levine, D. K. and S. Modica (2012). Conflict and the evolution of societies.

Lewis, B. (1997). *The Middle East: A Brief History of the Last 2, 000 Years* (Pocket Books: Reprint edition ed.). Touchstone Book.

Makdisi, U. (2000). *The Culture of Sectarianism*. University of California Press.

Matthiesen, T. (2017). Sectarianization as securitization - identity politics and counter-revolution in bahrain. In N. Hahsemi and D. Postel (Eds.), *Sectarianization - Mapping the New Politics of the Middle East*, Chapter 8, pp. 143–158. Hurst & Company.

McGrattan, C. and S. Hopkins (2016). Memory in post-conflict societies: From contention to integration? *Ethnopolitics 16*, 488 – 499.

Mitra, A. and D. Ray (2014). Implications of an economic theory of conflict: Hindu-muslim violence in india. *Journal of Political Economy 122*(4), 719 – 765.

Nasr, V. (2006). *The Shia Revival: How Conflicts within Islam Will Shape the Future*. W. W. Norton.

Nasr, V. (2017). International politics, domestic imperatives and identity mobilization - sectarianism in pakistan 1979-1998. In N. Hahsemi and D. Postel (Eds.), *Sectarianization - Mapping the New Politics of the Middle East*, Chapter 8, pp. 143–158. Hurst & Company.

Páez, D. and J. H. Liu (2015). *The collective remembering of conflict and its role in fueling an ethos of conflict in society.*, Volume 27, pp. 61 – 72. Springer.

Pennisi, E. (2005). How did cooperative behavior evolve, in what we don't know. *Science 309*(5731), 75–93.

Popper, K. (1979). *Objective Knowledge: An Evolutionary Approach*. Clarendon Press.

Ross, M. H. (2013). *The politics of memory and peacebuilding*, Chapter 7. Routledge.

Salibi, K. (1988). *A House of Many Mansions - The History of Lebanon Reconsidered*. Reprint, I. B. Tauris, 2014.

Salloukh, B. F. (2019). War memory, confessional imaginaries,and political contestation inpost-war lebanon. *Middle East Critique 28*, 341 – 359.

Salloukh, B. F., R. Barakat, J. S. Al-Habbal, L. W. Khattab, and S. Mikaelian (2015). *The Politics of Sectarianism in Postwar Lebanon*. Pluto Press.

Sherif, M., O. J. Harvey, B. J. White, W. R. Hood, and C. W. Sherif (1954/1961). *Intergroup conflict and cooperation. The Robbers Cave experiment* (Classics in the History of Psychology ed.). York University.

Skaperdas, S. (1992). Cooperation, conflict, and power in the absence of property rights. *American Economic Review 82*(4), 720–739.

Skaperdas, S. (1996). Contest success functions. *Economic Theory 7*, 283–290.

Taylor, D. and M. Yapp (1979). *Political Identity in South Asia*. Curzon Press.

Tilley, V. (1997). The term of the debate: Untangling language about ethnicity and ethnic movements. *Ethnic and Racial Studies 20*(3), 497–522.

Tullock, G. (2005). *The Selected Works of Gordon Tullock*, Volume 8. Liberty Fund.

Veblen, T. (2009). *Theory of the Leisure Class: An Economic Study in the Evolution of Institutions* (Original work published 1899, Oxford World's Classics ed.). Oxford University Press.

Wade, M. J. (1978). A critical review of models of group selection. *Quarterly Review of Biology 53*, 101–114.

Wehrey, F. (2017). *Beyond Sunni and Shia: The Roots of Sectarianism in a Changing Middle East*. Oxford University Press.

Weibull, J. W. (1997). *Evolutionary Game Theory*. MIT Press paperback edition.

Weinstein, J. M. (2005). Resources and the information problem in rebel recruitment. *Journal of Conflict Resolution 49*(4), 598–624.

Wilson, D. S. (1980). *The Natural Selection of Populations and Communities*. Benjamin / Cummings.

Wilson, E. O. (1975). *Sociobiology: The New Synthesis*. Belknap Press.

Table 2. Simulation Results: Probit Regression showing Marginal Effects

Model	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)
Dependent Variable	Egalitarian Contract	Egalitarian Contract	Sectarian Conflict	Sectarian Conflict	Cross-sect. Conflict	Cross-sect. Conflict	Intern. Inegal. Cross-sect. Confl.	Intern. Inegal. Cross-sect. Confl.
Unbalanced	-.7705	-.2831						
	(.00254)	(.00592)						
Play High Gain λ	-.0369	-0.271	.0821	.0323	.0468	.0434	.0064	.0284
	(.00024)	(.0034)	(.00068)	(.00042)	(.00046)	(.00058)	(.00011)	(.00041)
Collective Memory ζ	-.4042	.3546	.8818	.2065	.3687	-.2372	.1233	-.2884
	(.00585)	(.00924)	(.01254)	(.01211)	(.00716)	(.01397)	(.00262)	(.01094)
Share of Matches τ	-.2583	.2362	.6968	-.6095	.2973	-.4449	.1431	-.2297
	(.00560)	(.00874)	(.01158)	(.01251)	(.00685)	(.01400)	(.00278)	(.01101)
Memory Length m	.0096	.0111	.0110	-.0026	-.0035	-.0108	-.0029	-.0102
	(.00013)	(.00020)	(.00022)	(.00026)	(.00013)	(.00031)	(.00005)	(.00024)
Clientelism	-.1035	-.0920	.1308	.0886	.1238	.2679	.0170	.1310
	(.00387)	(.00599)	(.00611)	(.00835)	(.00430)	(.00901)	(.00161)	(.00726)
Replication:								
Variant 2	-.1467		.2859		.0200		-.0277	
	(.00500)		(.00618)		(.00514)		(.00167)	
Variant 3	-.1436		.2595		.0210		-.0297	
	(.00497)		(.00637)		(.00519)		(.00166)	
N	115,200	38,400	57,600	19,200	57,600	19,200	57,600	19,200
Pseudo R^2	0.5307	0.3176	0.7572	0.3397	0.6776	0.4497	0.3293	0.3732

Note: Robust standard errors in parentheses. Unbalanced, Clientelism and Type of Replication indicate discrete change of dummy variable from 0 to 1, all results statistically significant at $p < 0.01$. Reference group for the type of replication dummies is defined by variant 4. Clientelism is a dummy variable indicating that past payoff is acquired from a *teacher* during social learning. Logistic regression generates very similar marginal effects.