

Radicalization*

Jean-Paul Carvalho[†]

Michael Sacks[‡]

Abstract

Identity-based organizations coordinate forms of collective action that strengthen group identity (e.g. religion, race). Under what conditions can such an organization *radicalize* an identity group, mobilizing moderate individuals into a strict club with intensive participation in identity-based activities? We identify two paths to radicalization, based on cultural evolutionary theory: *Prestige-biased cultural transmission* occurs when active group members enjoy greater prestige and cultural influence within the group. *Niche construction* occurs when identity-based participation induces blanket discrimination against the identity group. By raising the group's identification and reducing its outside options, these forces produce a strict club which grows and becomes more extreme over time. Stigmatizing participation can backfire and spur radicalization under certain conditions. Fears of extremism can be self-perpetuating. Competition among identity-based organizations limits radicalization. We apply our model to Muslim communities in Europe and set out principles for preventing radicalization.

October 16, 2019

JEL Classification: D23, Z12

Keywords: extremism, cultural transmission, economics of religion, discrimination, club goods

*We are grateful for comments by Alberto Alesina, Ran Abramitzky, Samuel Bazzi, Eli Berman, Lisa Blaydes, Gary Cox, Martin Fiszbein, Larry Iannaccone, Saumitra Jha, Asim Khwaja, Timur Kuran, David Laitin, Mike McBride, Rachel McCleary, Alireza Naghavi, Nathan Nunn, Jared Rubin, Stergios Skaperdas, Peyton Young, as well as participants at the 2018 ASREC conference, 2017 ERINN Conference, 2017 AALIMS Conference, 2017 NES Conference 'Towards Effective and Equitable Development: the Role of Institutions and Diversity', 2016 ASREC conference, 2015 IMBS Workshop on the Economics of Religion and Culture, and seminar participants at Harvard, Stanford, Boston University and UC Riverside. Patrick Julius provided excellent research assistance. A previous version of this paper was circulated under the title "The Economics of Religious Communities: Social Integration, Discrimination and Radicalization".

[†]Department of Economics and Institute for Mathematical Behavioral Sciences, University of California, Irvine, 3151 Social Science Plaza, Irvine, CA 92697, jpcarv@uci.edu.

[‡]Department of Economics, University of Louisville, College of Business, Suite 144, Louisville, KY 40292, michael.sacks@louisville.edu.

1 Introduction

Since the end of the Cold War, attention has turned to the risk posed by radical non-state actors, both home-grown and foreign. It was once thought that different groups within a society would either (i) socially and economically integrate (Glazer & Moynihan 1963) or (ii) form a ‘liberal archipelago’, coexisting in harmony despite cultural distinction and social separation (Kukathas 2003). Today, a more pessimistic vision has emerged. Of particular concern is the rise of organizations that exploit latent, but powerful, forms of identification. These identity-based organizations promote a shared sense of belonging among members of an identity group (e.g. religious, racial group) and coordinate participation in rituals, ideology, and activism. Political radicalization has become a concern in the United States and Europe, where political movements are increasingly identity-based. More extreme examples include Islamist and jihadist groups around the world, white supremacist organizations, Hindu nationalist groups in India, and a large number of ethnic separatist groups in Chechnya, Spain, Sri Lanka and other countries.

Consider the various Islamist movements that arose in the 20th century, beginning with the Muslim Brotherhood in Egypt in 1928 and the Jamaat-i Islami of Pakistan in 1941. Their common aim was to reactivate and reshape Islamic identity after decades of colonial dominance and secularization. Sunni groups did so not through an established clergy, but through highly committed lay volunteers, who were set apart from the non-observant and accorded prestige. They built, in some places, a state within a state. For example:

In the 1940s, the Egyptian Muslim Brotherhoods developed a network of schools, clinics, and even small industries. They created student and professional sections, geared in fact toward modern professions (lawyers, engineers, doctors, teachers, bureaucrats), as well as worker sections in which they encouraged unionization.

(Roy 1994, p. 57)

Through this religious ecology they aimed to “resocialize” society from the bottom up along Islamic lines, while also seeking political power. Their success was evident in more visible forms of piety, including Islamic banking, the new veiling movement, and the rise of Islamic greetings (Kuran 2006, Carvalho 2013, Binzel & Carvalho 2017). From these Islamist organizations also came militant groups that sought political power through violent means. For example, Takfir wal-Hijra spun off from the Muslim Brotherhood in 1971, attempting to topple the Egyptian state through terrorist attacks and producing Ayman Al-Zawahiri, later second-in-charge of Al-Qaeda (Ibrahim 1980).

Today’s heightened concern about identity-based organizations owes partly to new communications technologies, which overcome geographical and legal barriers to radicalization and recruitment. ISIS used Twitter and Facebook to locate foreign recruits, followed by Skype video calls to screen and groom them. White supremacist groups in the United States radicalize and recruit through social media and online forums. Mob violence against Muslims in India has been coordinated through WhatsApp. Technology-driven radicalization is not new of course. Five centuries ago, the printing press, perhaps a more important innovation in communications technology, made possible the Reformation and the centuries of religious conflict that followed (Rubin 2017).

The importance of identity to political economy has been established by Akerlof & Kranton (2000, 2010), McCleary & Barro (2006), Shayo (2009), Akerlof (2017), and others. This paper studies the conditions under which an identity-based organization can turn a moderate identity group into a strict club in which members are strongly attached to their identity and participate intensively in identity-based activities. We call this process *radicalization*. Identity-based activities include participation in (1) ideological production (e.g. narrative construction through traditional and social media), (2) rituals (e.g. communal worship, rallies), and (3) political activism (e.g. protests, political volunteering). Of course, rising group identification and participation is not necessarily pernicious. It could help produce collective action to correct identity-based inequality. It could also produce a ‘liberal archipelago’ with multiple experiments in living being conducted in a single society.¹ Nevertheless, radicalization presents risks in terms of political destabilization and identity-based conflict. These risks are not confined to any race or religion, but are a more general feature of identity, under certain conditions which we attempt to characterize.²

To understand the forces behind radicalization, we combine approaches from two literatures: a club model from the economics of religion (Iannaccone 1992) and intergenerational transmission of traits from cultural evolutionary theory (Cavalli-Sforza & Feldman 1981, Boyd & Richerson 1988, Bisin & Verdier 2000).³ Consider an identity group in which individuals have either high or low attachment to their identity. Group members choose whether to

¹Sen (2007) calls such a regime “plural monoculturalism” [p. 165].

²This process has parallels to the literature on social norm shifts, including Mackie (1996), Young (2015), and Hwang, Naidu & Bowles (2016).

³The religious club model has been further developed by Berman (2000), McBride (2008, 2010, 2015), Levy & Razin (2012), Aimone et al. (2013), Carvalho (2016, forthcoming) and Carvalho & Koyama (2016). An important distinction is made by Barro & McCleary (2003) between the effects of religious belief and participation. See Iyer (2016) for a review of the economics of religion literature and Carvalho (2019) in particular on religious clubs.

join a club that coordinates participation in identity-based activities. Greater participation today boosts the proportion of high-attachment types in the next cohort. The club is fully forward-looking, setting a minimum participation requirement (strictness) each period to maximize total participation over time. We analyze the conditions under which the club is willing and able to radicalize the group, taking it from low to high levels of identity attachment and participation. The short-run equilibrium structure is as follows: A rising share of high-attachment types fragments the identity group leading to the formation of a strict, exclusive club. Thenceforth, club participation scales superlinearly with the share of high-attachment types. Low development (e.g. real wages) and high blanket discrimination against identity group members (regardless of participation) raise club membership and participation. Stigmatizing only active participants in identity-based activities does the opposite. In the long run, however, radicalization does not occur in the baseline model. Eventually, either a moderate inclusive club forms or identity-based participation collapses completely.

Radicalization occurs only when two new mechanisms are added: prestige-biased cultural transmission and niche construction. Both are important mechanisms in cultural evolutionary theory (Henrich & Gil-White 2001, Odling-Smee et al. 2003), but have been largely ignored in economics.⁴ Prestige-biased transmission occurs in our model when active club participants enjoy greater prestige and cultural influence. Niche construction occurs in our model when blanket discrimination is endogenous, depending on total club participation. By raising strictness, the club induces blanket discrimination against identity group members, which reduces their outside options, enabling the club to further raise strictness. Thus the club can carve out a niche for the identity group. This niche construction strategy can be activated by rising fears of extremism, which then become self-perpetuating. In both cases, radicalization occurs through the formation of a small but extreme group of high-attachment types, who are used to raise the identity-attachment and reduce the outside options of subsequent cohorts. Over time, the initially small club expands and becomes more extreme and isolated.

Each mechanism relies on its own form of asymmetry. To cut off these two paths to radicalization, a shift in norms is required to ensure symmetry. Prestige bias can be eliminated by according active participants in identity-based activities no greater prestige and cul-

⁴Notable exceptions are Han, Hirshleifer & Walden (2018) on self-enhancing bias in transmission of investment return information and Carvalho & Koyama (2016, section 5.2) on niche construction by religious sects.

tural influence than inactive group members. Niche construction can be eliminated by a taboo against responding to rising participation by extreme group members with blanket discrimination against all members of a racial, religious or other identity group. More targeted stigmatization of active participants can work under some conditions. However, we also identify conditions in which it backfires and spurs radicalization. Finally, competition among identity-based organizations limits radicalization and more generally moderates participation in identity-based activities. We bring all of our results together in an application to contemporary Muslim communities in Europe.

In their study of Jewish emancipation, Carvalho & Koyama (2016) combine concepts from religious club and cultural transmission models (Section 5.1). Two important papers take this substantially further: Verdier & Zenou (2015) analyze cultural contests between competing (myopic) leaders pushing two different cultural traits, while Verdier & Zenou (2018) combine club goods production and cultural transmission, introducing new dynamic techniques to model forward-looking leaders.⁵ Hauk & Mueller (2015) examine how leaders can promote a cultural trait by tuning parameters in the standard Bisin-Verdier model, including the degree of cultural dislike. Carvalho (2016) examines a club model with competing identity-based organizations and an external socializing agent (e.g. media, state). Prummer & Siedlarek (2017) analyze the effect of a single leader who directly intervenes in cultural transmission to raise identification by group members. The leader is either myopic or optimizes steady-state identification. Hence, fully dynamic radicalization strategies are not analyzed. Chen, McBride & Short (2018) examine competing clubs with heterogeneous objectives, making a number of modeling innovations. Each club's strictness is fixed over time, however, which also rules out the dynamic radicalization strategies examined in this paper. None of these models study the radicalizing effects of prestige-biased cultural transmission and niche construction, or the moderating effects of competition among identity-based organizations.⁶

The remainder of the paper is structured as follows. Section 2 presents and analyzes the baseline model. Section 3 introduces the two mechanisms, prestige-biased transmission and niche construction, which enable radicalization. Section 4 examines the moderating effects of competition among identity-based organizations. Section 5 applies the model to the

⁵See also Bisin & Verdier (2017), who model the co-evolution of institutions and culture, as well as Carvalho (2013) and Carvalho & Koyama (2014), who examine how cultural transmission is shaped by identity choice and the education system, respectively.

⁶For example, Verdier & Zenou (2018) assume members of a particular group are homogeneous, so dynamic strategies in which leaders use extreme types to radicalize moderate types (e.g. prestige-based transmission) do not exist. In addition, there is no competition between leaders cultivating a particular cultural trait (only opposing cultural traits).

case of contemporary Muslim communities in Europe and sets out principles for preventing radicalization. Section 6 concludes.

2 The Baseline Model

Consider an identity group I , which is a continuum of agents with unit mass. Identity (e.g. race, religion) is determined at birth, as is attachment to one's identity.

The group I is embedded in an (unmodeled) society. While identity itself is not chosen, individuals choose whether or not to participate in an identity-based organization, known as a 'club', which caters to the identity group I .⁷ The club good, produced and consumed exclusively by club members, is a combination of collective identity-based activities, including rituals, ideology, political activism and other forms of collective action.

Each group member divides one unit of time between work/leisure outside the club and participation in the club's identity-based activities. The share of time devoted to club participation by i is x_i . The club sets the norms of participation by imposing a minimum participation requirement s : a group member is only a club member if she spends at least share s of her time on club participation. We begin with a single club and examine competition among identity-based organizations in Section 4.

Timing. Time is discrete and indexed by $t = 0, 1, 2, \dots$. Group members live for one period and each cohort t interacts as follows:

- DATE 0. I is partitioned into two subgroups, I_L with low attachment and I_H with high attachment to the identity. The share of high-attachment types is p^t , with initial $p^0 > 0$.
- DATE 1. The club announces a minimum participation requirement $s^t \in [0, 1]$.
- DATE 2. Observing s^t , individuals simultaneously choose a division of their time between outside activity and club participation.

Participation choices by L and H types are summarized by $x^t = (x_L^t, x_H^t)$, where x_ω^t is the (common) level of participation by ω types. Total club participation is $X^t = p^t x_H^t + (1 - p^t)x_L^t$.

- DATE 3. A club good is collectively produced and payoffs are received.

⁷Hence group members $i \in I$ need not be club members, but all club members are group members.

Organization Payoffs. Unlike group members, the club has an infinite horizon, maximizing total club participation over time. That is, at each period T , the club maximizes the discounted sum:

$$G^T = \sum_{t=T}^{\infty} \mu^{t-T} X(s^t, p^t). \quad (1)$$

where $\mu \in [0, 1)$ is a discount factor, subject to the cultural evolutionary dynamic:

$$p^{t+1} = X(s^t, p^t). \quad (2)$$

By (2), the distribution of identity attachment at $t + 1$ depends on identity-based participation in period t . This dynamic is broadly consistent with the cultural transmission models of Bisin & Verdier (2000, 2017), Carvalho (2016) and Chen, McBride & Short (2018).⁸ We shall generalize this specification in Section 3.1.

Our question is whether the club is willing and able to *radicalize* the identity group. By radicalization, we mean transitioning a group of mostly of low-attachment types with at most moderate participation in identity-based participation to one with high levels of attachment and participation through a sequence of strictness choices $\{s^t\}_{t=0}^{\infty}$.

Individual Payoffs. In doing so, the club is constrained by the agency of group members, who choose how much to participate, if at all.

As individuals live for one period and are self-regarding, we can drop time subscripts. The payoff function for a club member i is

$$u_i(x_i, x) = \underbrace{\pi(1 - x_i)}_{\text{outside activity}} + \underbrace{\theta_i X^{1/2}}_{\text{club participation}}, \quad (3)$$

where $\pi > 0$ is the payoff per unit of time spent outside of the community (e.g. wage) and θ_i is the value placed by i on the club good, which is larger for high-attachment types: $\theta_i = \theta_L$ for low-attachment types and $\theta_i = \theta_H$ for high-attachment types, with $\theta_H > \theta_L > 0$.

As agents are non-atomic, there is a severe free-rider problem in collective production of the club good. As in seminal work by Iannaccone (1992) on religious clubs, some intervention is required to mitigate this incentive problem. We assume the club imposes a minimum

⁸It is also consistent with religious capital models (Iannaccone 1990, McBride 2015) in which religious tastes are produced through an individual's participation, except here this occurs at the group level. McBride (2015) also studies population dynamics driven by endogenous fertility rather than cultural transmission.

participation requirement s , either through a leader or through coordination on norms by members themselves. Equivalently, s is a restriction on outside activity, which can be imposed indirectly by mandating various stigmatizing forms of dress, diet and sexual behavior which constrain social integration (see Iannaccone 1992, Aimone et al. 2013) or directly by monitoring time inputs (Carvalho 2016). We refer to s as the *strictness* of the club.

If $x_i \geq s$, individual i is deemed a member of the club and receives the payoff described by (3). The set of club members is denoted by $M \equiv \{i : x_i \geq s\}$. Because of the free-rider problem, $x_i = s$ for all club members. The identity group is *cohesive* if all group members are also members of the club, $M = I$. Otherwise, it is *fragmented*.

The payoff to outside activity by nonmembers of the club is $\Delta\pi > 0$, where Δ is the relative payoff vis-à-vis club members. If members of the club are discriminated against, Δ could be greater than one. If club members have access to a special production technology, Δ could be less than one.⁹ As nonmembers are excluded from participation and consumption of the club good, the payoff to all nonmembers is $\Delta\pi$.

We assume throughout that $\theta_H < 2\pi$.¹⁰ The structure of the game is common knowledge.

Equilibrium Definition. First consider individual participation choices before turning to the club's strictness choices. Let $x(s, p) = (x_L(s, p), x_H(s, p))$ be the participation strategies in state p given strictness s .

Coalitional deviations are central to the formation and fragmentation of clubs. Given increasing returns to group membership, we can restrict attention to the following set of coalitions: $\mathcal{C} = \{I_L, I_H, I\}$.

Definition 1. Consider the subgame following strictness choice s in state p . A deviation by individual/coalition $C \in \mathcal{C} \cup \{i\}$ from $x(s, p)$ is an alternative profile $x'(s, p)$ such that $x_i(s, p) = x'_i(s, p)$ if and only if $i \in C$. A coalitional deviation is profitable if

$$u_i(x(s, p)) < u_i(x'(s, p)) \tag{4}$$

for all $i \in C$. The deviation is coalitionally stable if there are no further profitable deviations from $x'(s, p)$ by any coalition $C' \in \mathcal{C} \cup \{i\}$.

⁹See, for example, Bernstein (1992) on diamond trading networks among Orthodox Jews.

¹⁰When $\theta_H \geq 2\pi$, we have to consider uninteresting boundary conditions, since the club can always demand the maximum possible time contribution ($s = 1$) by high types and get it.

Definition 2. The profile $x^*(s, p)$ implements a coalition-proof equilibrium of the subgame induced by the club's choice of s in state p if no coalitionally stable deviation exists.

Definition 3. A identity equilibrium (IE) of the entire game consists of (i) individual participation strategies that form a coalition-proof equilibrium $x^*(s, p)$ for each $(s, p) \in [0, 1]^2$ and (ii) a function $s^*(p)$ that in each state p sets strictness to maximize the club's payoff (1).

Coalition-proofness rules out *coordination failures* based on self-fulfilling expectations that nobody will join and contribute to the club good. It cannot, however, select a Nash equilibrium with efficient contributions to the club good (a free rider problem), as no such Nash equilibrium exists in the absence of the minimum participation requirement s .¹¹ Since individual group members live for one period, restricting the club to Markov strategies in Definition 3 is without loss of generality until we introduce competitive clubs in Section 4.

2.1 Short-Run Equilibrium

We now characterize the stage-game equilibrium in each state p , before studying the long-run evolution of p . We call this the short-run identity equilibrium (SRIE). In addition to p , the short-run equilibrium structure depends on the outside options available to group members.

Define $\underline{\Delta} \equiv 1 + \left(\frac{\theta_L}{2\pi}\right)^2$ and $\bar{\Delta}(p) \equiv 1 + p \left(\frac{\theta_H}{2\pi}\right)^2$. Lemma A1 in the Appendix shows that when $\Delta > \underline{\Delta}$, $x_L(s, p) = 0$ for all $(s, p) \in [0, 1]^2$, i.e. zero participation. When $\Delta > \max\{\underline{\Delta}, \bar{\Delta}(p)\}$, there is zero participation for all types regardless of $s \in [0, 1]$. Thus an attractive outside option, due to a large (common) payoff from outside activity π and/or large relative payoff from outside activity to nonmembers Δ , makes it impossible to induce participation in identity-based activities.

Second, even if the club could attract all group members at some level of strictness, it may not wish to do so. Instead, the club may maximize total participation by setting a high level of strictness, *screening* out L types and extracting larger contributions from high-attachment H types. Whereas screening of low-attachment types is simply assumed in standard religious club models, screening here emerges endogenously under certain conditions.

To analyze screening, we turn to the incentives of the club. We refer to a club which all community members join, $M^*(s, p) = I$, as *inclusive*. A club which only high-attachment

¹¹Unlike the coalition-proof Nash equilibrium of Bernheim et al. (1987), which is applied recursively, we need only apply the notion of coalitional stability once in our setting.

types join, $M^*(s) = I_H$, is *exclusive*. As the club wishes to maximize participation, he will set strictness as high as possible. For an inclusive club, that means setting strictness up to $s = 1$ or the maximum strictness that satisfies the low type's participation constraint IR_L :

$$\underline{s} = \begin{cases} 1 & \text{if } \Delta \leq \frac{\theta_L}{\pi} \\ \left(\frac{\theta_L}{2\pi} + \sqrt{\left(\frac{\theta_L}{2\pi}\right)^2 - (\Delta - 1)} \right)^2 & \text{if } \Delta \in \left(\frac{\theta_L}{\pi}, \underline{\Delta} \right]. \end{cases} \quad (5)$$

For an exclusive club, this means up to $s = 1$ or the maximum strictness that satisfies the high type's participation constraint IR_H :

$$\bar{s} = \begin{cases} 1 & \text{if } \Delta \leq \frac{\theta_H \sqrt{p}}{\pi} \\ \left(\sqrt{p} \frac{\theta_H}{2\pi} + \sqrt{p \left(\frac{\theta_H}{2\pi}\right)^2 - (\Delta - 1)} \right)^2 & \text{if } \Delta \in \left(\frac{\theta_H \sqrt{p}}{\pi}, \bar{\Delta} \right]. \end{cases} \quad (6)$$

If outside options are poor, in particular if $\Delta \leq (\theta_L/\pi)$, then the club can demand that individuals devote all of their time to identity-based participation ($s = 1$) and still induce all group members to participate. For higher values $\Delta \in (\theta_L/\pi, \underline{\Delta}]$, it may be possible to achieve cohesion, but not at full strictness $s = 1$. In this case, an inclusive club forms at the strictness level that makes IR_L bind, i.e. L types are indifferent between joining the club and not joining. The same logic applies to the optimal strictness of an exclusive club \bar{s} . The main difference is that \bar{s} is increasing in the share of high-attachment types p . Given increasing returns to membership, the larger an exclusive club, the more productive it is. Hence the larger is p (the size of an exclusive club), the further the club can push its members in terms of demands on their time. Note that $\bar{s} > \underline{s}$ if and only if $p > (\theta_L/\theta_H)^2$.

Having determined \underline{s} and \bar{s} , we can precisely define the types of short-run equilibria that can arise.

Definition 4. A cohesive equilibrium is an SRIE in which $M^*(s^*) = I$, $s^* = \underline{s}$ and $x_i^* = \underline{s}$ for all $i \in N$.

Definition 5. An exclusive equilibrium is an SRIE in which $M^*(s^*) = I_H$, $s^* = \bar{s}$ and $x_i^* = \bar{s}$ for $i \in N_H$ and zero for $i \in N_L$.

First, note that a highly liberal club never forms: $s^* > 0$ whenever the club is nonempty. The short-run equilibrium structure is depicted in Figure 1 as a function of the proportion of

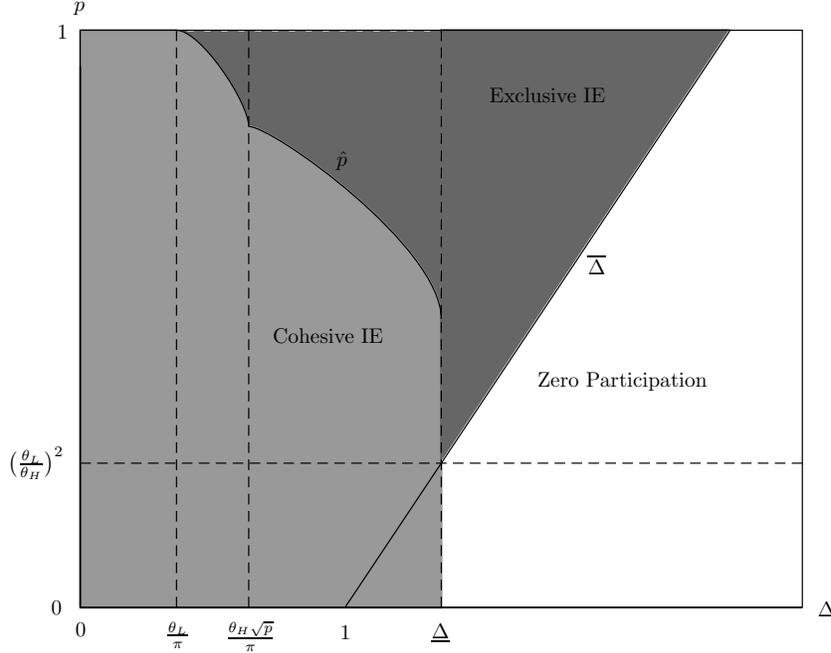


Figure 1: Short-run equilibrium structure

high-commitment types p and the outside option determined by π and Δ . (See Proposition A1 in the Appendix for a formal characterization and proof.) To understand the structure, consider $\Delta \leq \underline{\Delta}$.¹² In this case, the identity group is cohesive if the proportion of H types is sufficiently low, $p \leq \hat{p}$. If $p > \hat{p}$, the group fragments with L types not participating and H types forming a stricter club. The intuition is as follows. H types value the club good more highly than L types. To induce L types to join, strictness must be set relatively low $s \leq \underline{s}$. Alternatively, the club could raise strictness and elicit larger contributions from the mass p of H types. An inclusive club's equilibrium output is \underline{s} , while an exclusive club's equilibrium output is $p\bar{s}$. Hence an exclusive club forms when the proportion of H types is greater than the relative strictness of an inclusive club: $p > \underline{s}/\bar{s} \equiv \hat{p}$. Thus, an identity group can fragment when it has a large share of members with high attachment to the identity.

2.1.1 Identity Attachment & Participation

Figure 2 plots total religious participation as a function of the share of high-attachment types p in the interesting case $\Delta \in (\theta_L/\pi, \underline{\Delta}]$. When p is low, an inclusive equilibrium is in place. All group members choose club participation \underline{s} , so total participation is independent of p . As the proportion of high-attachment types grows large ($p \geq \hat{p}$), the group fractures and an

¹²This condition is satisfied for a relative payoff from outside activity of $\Delta = 1$ (the standard assumption in the religious clubs literature) and for a range of Δ greater than one.

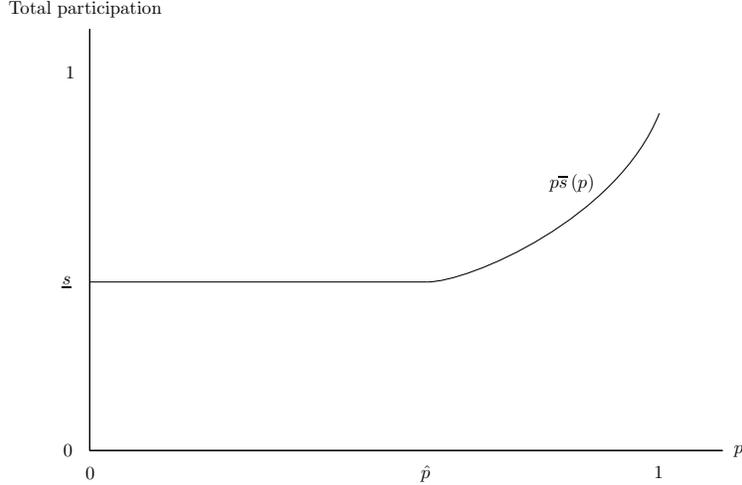


Figure 2: Total participation as a function of the share of high-attachment types, when $\Delta \in (\theta_L/\pi, \underline{\Delta}]$.

exclusive club is formed. Both the size of this club and its participation intensity $\bar{s}(p)$ are increasing in p . Hence the club benefits from a larger share of high-attachment types if and only if p is sufficiently large. Thenceforth, identity-based participation scales superlinearly with p .

2.1.2 Discrimination & Participation

The identity group in our model is embedded in a larger society. We will now examine how society’s attitudes toward group members affects their participation in identity-based activities. By discrimination we mean actions taken by outgroup members that lower the payoff to outside activity by members of the identity group I . This includes social discrimination which negatively affects social interactions with outgroup members and labor market discrimination which lowers the expected wage faced by community member.

Following Bisin et al. (2011), we distinguish between two types of discrimination:¹³ ‘Blanket discrimination’ is directed toward all group members regardless of their club membership. ‘Stigma’ depends on a group member’s actions and is directed only toward active club members. Recall that π is the common component of the payoff to outside activity. Let $\pi = w/(1 + \delta)$, where w is economic development (e.g. real wage) and δ is the level of blanket discrimination against group members. In contrast, Δ is the relative payoff to outside activity for those who do not participate in the club. Thus Δ is a measure of the stigma faced by active club members.

¹³See also Fouka (2017) for an empirical investigation.

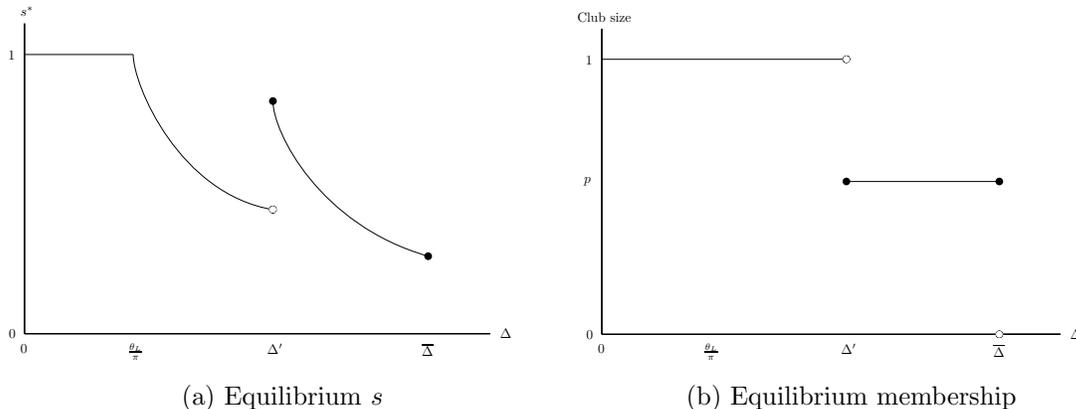


Figure 3: The effect of stigma

The two forms of discrimination have different effects on club participation. We know (from Proposition A1) that total participation is $X^* = 1$ for $\Delta \leq \theta_L/\pi$ and $X^* = 0$ for $\Delta > \max\{\underline{\Delta}, \bar{\Delta}\}$, $X^* = 1$. Hence consider the intermediate range:

Proposition 1. *Suppose $\Delta \in (\theta_L/\pi, \max\{\underline{\Delta}, \bar{\Delta}\})$. Total participation X^* is increasing in blanket discrimination δ and decreasing in stigma Δ , and strictly so whenever $s^* < 1$.*

Hence, in the short run, blanket discrimination against group members promotes identity-based participation, while stigmatizing active participation inhibits it.

Even in the short run, the relationship is more complex than this suggests. Total participation depends on both equilibrium club membership and strictness. For p sufficiently large, there is a non-monotonic relationship between equilibrium strictness and both δ and Δ .¹⁴ Let us focus on the effect of stigma Δ depicted in Figure 3. The opposite effect of blanket discrimination is depicted in Figure 4.

Starting from a low level of Δ , a rise in Δ makes it less attractive to join the club, prompting the club to lower strictness s^* in order to keep L types in the club. This is depicted in panel (a) by the graph up to point Δ' . Once Δ is sufficiently high, however, it is not worthwhile to

¹⁴Proof available upon request. A similar pattern of non-monotonicity in δ emerges through a different mechanism in a religious club model by Carvalho & Koyama (2016). (They only analyze δ and make the standard assumption that $\Delta = 1$.) In their model, individuals can make both time and money contributions to a club, club goods are rival, and the club does not impose a minimum contribution requirement, but rather chooses a tax on outside activity. The fact that the same result emerges under substantially different assumptions suggests there is something robust about the pattern.

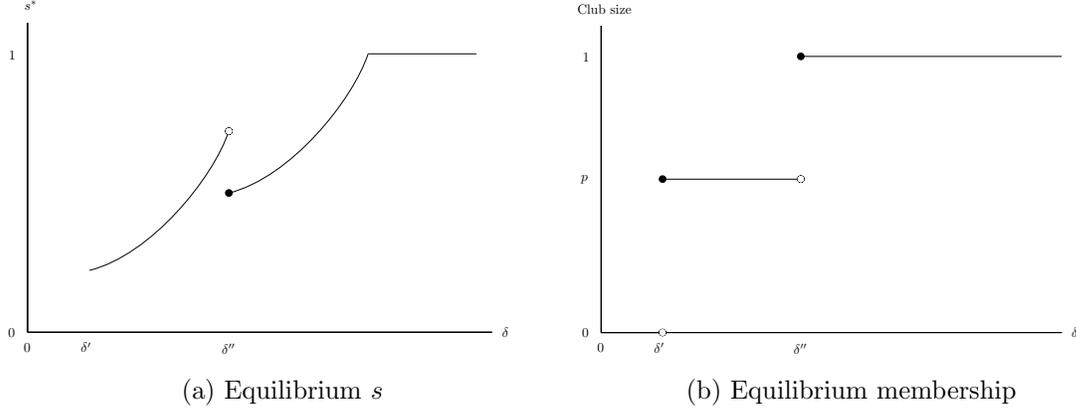


Figure 4: The effect of blanket discrimination ($\Delta > 1$)

liberalize any further. Instead the club benefits from raising strictness, inducing L types to exit and catering exclusively to H types. The discrete jump in strictness at point Δ' is shown in panel (a), while the exit of low commitment types is shown in panel (b).¹⁵ Therefore, stigmatization club membership reduces club participation, but can fragment the identity group leading to the formation of a small but extreme club.

Why should we care about the formation of such a club? Berman (2009) shows that strict, close-knit religious clubs are most at risk of transitioning from providing religious club goods into less benign forms of collective action. At the extreme end, terrorist organizations make extreme demands of members, including suicide attacks. They are also highly sensitive to defection. Due to their strictness and social isolation, small radical religious groups screen out all but the most committed types and elicit extreme contributions by members (Iannaccone 1992). This gives them an advantage in violent forms of collective action which require intensive participation and safeguards against defection and infiltration, as proposed by Berman & Laitin (2008) and Berman (2009). We shall explore the dynamic radicalizing effects of such exclusive clubs in Section 3.

2.2 Dynamics

Let us now examine the evolution of p^t , reintroducing time notation.

Proposition 2. *In every IE:*

- (i) $\Delta \leq \theta_L/\pi$: *The identity group is cohesive, $s^t(p^t) = 1$, $p^t = 1$ for all t .*

¹⁵Zero participation is unavoidable above $\bar{\Delta}$ (Lemma 1, Appendix).

- (ii) $\Delta \in (\theta_L/\pi, \underline{\Delta}]$: p^t is non-increasing. For $p^0 < (\Delta\pi/\theta_H)^2$, p^t is strictly decreasing until a finite time T at which an inclusive club forms and $p^t = \underline{s}$ for all $t \geq T$.
- (iii) $\Delta \in (\underline{\Delta}, \overline{\Delta}(p^0)]$: p^t is non-increasing. For $p^0 < (\Delta\pi/\theta_H)^2$, p^t is strictly decreasing until a finite time T at which zero participation occurs and $p^t = 0$ for all $t \geq T$.
- (iv) $\Delta > \max\{\underline{\Delta}, \overline{\Delta}(p^0)\}$: Zero participation occurs and $p^t = 0$ for all $t \geq 1$.

Thus radicalization does not occur in the baseline model. Under no condition does p^t increase over time. In case (i), there is always full participation and only high-attachment types from period 1 onward. In case (iv), there is always zero participation and no high-attachment types from period 1 onward. Consider the intermediate cases (ii) and (iii). Unless the share of high-attachment types is large to begin with, p^t falls over time, either stabilizing with the formation of a moderate inclusive club or at zero participation. Hence something else is required for the club to be willing and able to radicalize group members, transitioning it from low to high attachment/participation.

3 Two Paths to Radicalization

The baseline model combines cultural transmission and club formation, but does not produce radicalization. In this section, we introduce two mechanisms from the cultural evolution literature that have been largely ignored in economics: (i) prestige-biased cultural transmission and (ii) niche construction. When added to the baseline model, these mechanisms enable the club to radicalize the group, raising attachment and participation over time.

3.1 Prestige-Biased Cultural Transmission

Let us generalize the dynamic in equation (2) to admit a form of bias in cultural transmission. Recall that group-average participation is $X(s^t, p^t) = p^t x_H(s^t, p^t) + (1 - p^t)x_L(s^t, p^t)$. While this is calculated over the identity entire group I , club-average participation, which we shall denote by $\tilde{X}(s^t, p^t)$, depends only on the participation of active club members $i \in M^*(s^t, p^t)$. Specifically, club-average participation $\tilde{X}(s^t, p^t)$ differs from the group average $X(s^t, p^t)$ only when the club is exclusive, in which case it equals $x_H(s^t, p^t)$.

In the new cultural dynamic, let the proportion of high-attachment types be a convex combination of group-average and club-average participation levels:

$$p^{t+1} = (1 - \alpha)X(s^t, p^t) + \alpha\tilde{X}(s^t, p^t), \quad (7)$$

where $\alpha \in [0, 1]$. The baseline dynamic, (2), is a special case of (7) with $\alpha = 0$.

When $\alpha > 0$, active club members have a disproportionate influence on cultural transmission. This cultural power could derive from greater prestige and visibility in the identity group, a classic form of transmission bias in the cultural evolution literature (Boyd & Richerson 1988, Henrich & Gil-White 2001). For example, Chudek et al. (2012) show that children are more likely to copy adult role models to whom bystanders in an experiment pay more attention. In the religious context, it is clear how such prestige bias might work. Acts such as prayer and fasting, proselytizing, and community service give participating members greater visibility and status. As a consequence, they might exert greater influence in the religious socialization process. Hence we refer to α as the degree of *prestige bias* in cultural transmission.

We shall now see that when $\alpha > 0$, the club can radicalize the identity group by screening out low-attachment types and using the active high-attachment participants in the club to raise identification among all group members over time. We focus on the intermediate case $\Delta \in (\theta_L/\pi, \underline{\Delta}]$, in which the club must choose between inclusivity and exclusivity:

Proposition 3. *Consider the case of prestige-biased cultural transmission, $\alpha > 0$.*

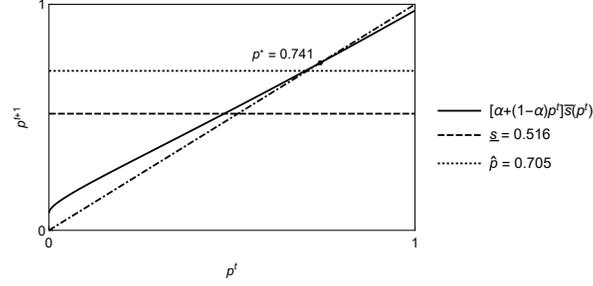
Suppose $\Delta \in (\theta_L/\pi, \underline{\Delta}]$.

In every IE, if α , θ_H and μ are sufficiently large, there exists a finite time T such that an exclusive club forms in all $t \geq T$ and $p^t > \underline{s}$ for all $t > T$.

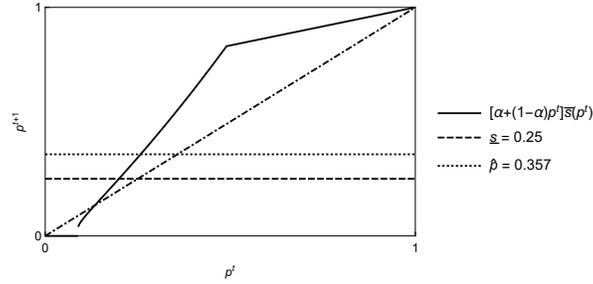
Otherwise, the result is the same as for $\alpha = 0$ in Proposition 2.

In the baseline case $\alpha = 0$, we know the share of high-attachment types p^t falls until either an inclusive club is formed at moderate strictness or the club is empty. With a large enough degree of prestige-bias $\alpha > 0$, identity attachment of H types θ_H , and club discount factor μ , an exclusive club forms and is stabilized. In addition, the limiting share of high-attachment types p^t is greater than the share generated by an inclusive club, \underline{s} . We remark that such radicalization can be generated even by a myopic club ($\mu = 0$) when the initial share of high attachment types p^0 is sufficiently large.

To illustrate how this radicalization strategy works with a forward-looking club, see Figure 5. In both panels, if the share of high-attachment types starts out below \hat{p} , and $\alpha = 0$,



(a) $\alpha = 0.9$, $\Delta = 0.915$, $\theta_L = 3$ and $\theta_H = 4.5$.
 $\bar{s}(p^*) = 0.76$.



(b) $\alpha = 0.667$, $\Delta = 1.05$, $\theta_L = 3$ and $\theta_H = 7.5$.

Figure 5: Representative equilibrium paths for $\alpha > 0$ and $\pi = 5$.

an inclusive club is formed in every period at moderate strictness \underline{s} . The share of high-attachment types is constant: $p^t = \underline{s}$ for $t \geq 1$. With prestige bias $\alpha > 0$, however, a forward-looking club might screen out low-attachment types. While this lowers club participation in the short run, the active (high-attachment) club members raise the identification of the group p^t over time and thereby increase the membership rate and participation intensity of club members. When α and θ_H are sufficiently large, p^t rises far enough to make an exclusive club produce higher per-period participation X^t than an inclusive club. In Panel (a) of Figure 5, the dynamic converges to a state in which approximately 74% of group members are club members, who spend 76% of their time on club participation. Panel (b) is a more extreme example in which the exclusive club eventually consumes the entire identity group, inducing group members to spend 100% of their time on club participation. A sufficiently forward-looking club (high μ) is willing to incur the short-run loss in participation from screening out low-attachment types to eventually produce a large and strict exclusive club. Thus, an identity group can be radicalized when there is prestige-biased cultural transmission.

3.2 Niche Construction

To examine the niche construction path to radicalization, we make blanket discrimination δ^t endogenous and study its co-evolution with p^t . Again, we allow for prestige-biased transmission: $\alpha \in [0, 1]$.

Let the initial level of blanket discrimination against group members be δ^0 . For comparison, we assume δ^0 equals the exogenous δ of previous sections.¹⁶ The question is whether the club can further radicalize the group by inducing rising blanket discrimination δ^t over time. Let $\underline{s}(\delta^0)$ be the strictness of an inclusive club at the initial level of blanket discrimination. We assume:

$$\delta^{t+1} = \delta^0 + \gamma \max\{p^{t+1} - \underline{s}(\delta^0), 0\}, \quad (8)$$

where p^{t+1} is the cultural dynamic (7) and $\gamma \geq 0$.¹⁷ We interpret γ as society's *propensity for blanket discrimination*. Note that $\gamma = 0$ corresponds to the special case of exogenous δ examined in previous sections. When $\gamma > 0$, blanket discrimination δ^t is a function of identity-based participation in a manner similar to p^t and, as with p^t , the club can shape blanket discrimination through its choice of strictness.

To further understand the discrimination dynamic (8), first note that p^{t+1} is a convex combination of group-average and club-average participation in period t . Beginning in state δ^0 , if an inclusive club forms, then $p^1 = \underline{s}(\delta^0)$. Hence $\delta^1 = \delta^0$, as in the exogenous discrimination case. Suppose instead that the club can raise strictness and participation above this level, by screening out low-attachment types. When $\gamma > 0$, society responds to greater participation by high-attachment types with blanket discrimination against all group members. In this case, $\delta^1 > \delta^0$, and the club is able to further raise strictness at $t = 1$. Thus a second, complementary path to radicalization is produced.

Again we focus on the intermediate Δ case in which the club must choose between inclusivity and exclusivity:

Proposition 4. *Suppose $\Delta \in (\theta_L/\pi, \underline{\Delta}(\delta^0)]$.*

As in Proposition 3 ($\gamma = 0$): In every IE, if α , θ_H , and μ are sufficiently large, there exists a finite time T at which an exclusive club forms. In addition, $p^t > \underline{s}(\delta^0)$ for all $t > T$. Otherwise, the result is the same as for $\alpha = 0$ and $\gamma = 0$ in Proposition 2.

¹⁶The original assumption that $\theta_H < 2\pi = 2w/(1 + \delta)$ is simply replaced by $\theta_H < 2w/(1 + \delta^0)$.

¹⁷We could assume $\delta^{t+1} = \delta^0 + \gamma(p^{t+1} - \underline{s}(\delta^0))$ without changing any of the propositions.

Denote by $\Omega(\gamma)$ the set of parameters (α, θ_H, μ) for which there exists a finite time T such that $p^t > \underline{s}(\delta^0)$ for all $t > T$. A higher propensity for blanket discrimination, γ , enlarges this set:

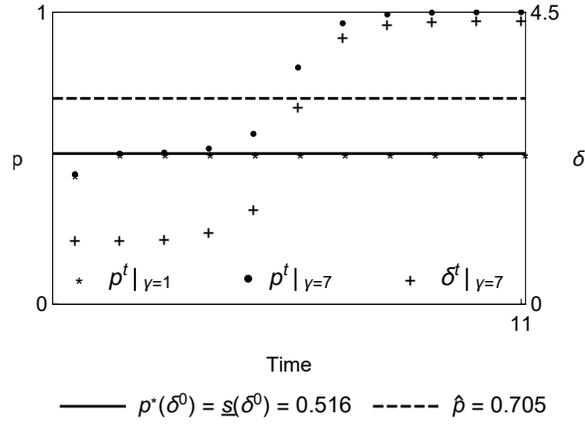
$$\Omega(\gamma) \supset \Omega(\gamma') \text{ whenever } \gamma > \gamma'.$$

First, the conditions for radicalization are qualitatively similar to the case of prestige-bias alone ($\alpha > 0, \gamma = 0$): a sufficiently large prestige bias α , high-type attachment θ_H , and club discount factor μ . However, endogenous blanket discrimination ($\gamma > 0$) produces radicalization under a larger set of conditions. Figure 6 presents two examples that compare the evolution of identity-based participation with and without niche construction. In panel (a), at $\alpha = 0.8$, there is no radicalization under $\gamma = 0$: a moderate inclusive club forms and locks in. In contrast, full radicalization occurs under niche construction, $\gamma = 7$. In panel (b), at larger $\alpha = 0.9$, there is radicalization in both cases: $\lim_{t \rightarrow \infty} p^t > \underline{s}(\delta^0)$. However, $\lim_{t \rightarrow \infty} p^t$ is larger under niche construction. Because α is larger, a smaller γ is required to produce full radicalization.

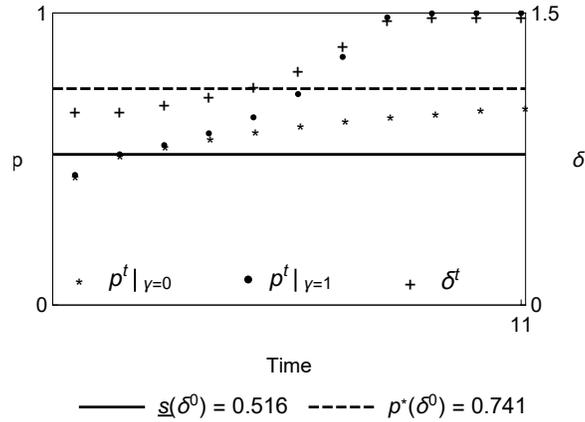
The reason endogenous δ promotes radicalization is as follows: By raising strictness today and screening out low-attachment types, the club raises blanket discrimination against the identity group (when $\alpha > 0$). Blanket discrimination lowers the opportunity cost of participation by ‘taxing’ outside activity for all group members (active or otherwise). This boosts future club participation, which further raises blanket discrimination, and so on. We view this as a form of cultural niche construction (Odling-Smee et al. 2003). Niche construction is an important factor in evolutionary history. In ecology, a classic example is the construction of dams by beavers which alter the environment in which natural selection operates. By inducing blanket discrimination against group members, the club similarly shields itself from selection pressures in the form of competition from outside alternatives.

Several additional remarks are in order. To get radicalization going from a low share of high-attachment types (p^0 low) requires prestige bias $\alpha > 0$, so that a few high-attachment types can raise identification by group members. A smaller α is required when γ is larger. However, when p^0 is high, niche construction ($\gamma > 0$) is sufficient to stabilize an extractive club at high strictness without prestige bias ($\alpha = 0$).¹⁸ Importantly, if societal concerns about growing extremism in the identity group raise γ , then they can become self-perpetuating.

¹⁸Notice that we have not claimed that an extractive club persists in the limit (as in the $\gamma = 0$ case). That is because we cannot rule out switches between exclusivity and inclusivity.



(a) $\alpha = 0.8$



(b) $\alpha = 0.9$

Figure 6: Numerical examples of niche construction: $p^0 = 0.445$, $\delta^0 = 1$, $w = 10$, $\Delta = 0.915$, $\theta_L = 3$ and $\theta_H = 4.5$, $\mu = 0$.

Of course, the niche construction path to radicalization might not work if society responded to rising participation by stigmatizing participating members only, rather than with blanket discrimination. Nevertheless, as we shall discuss in Section 5, raising the stigma to identity-based participation can backfire and also boost radicalization under certain conditions.

4 Competition

Let us now introduce competition among clubs. Entrepreneurs enter freely and compete for members of the identity group I . Does competition promote or inhibit radicalization?

The literature on religious competition uses Hotelling-style models in which individuals have fixed preferences over religious strictness and join the organization closest to their ideal point (Barros & Garoupa 2002, Montgomery 2003, McBride 2008, Iyer et al. 2014).¹⁹ When there are increasing returns to membership as in our model, however, the problem is more complex. Rather than an individual's ideal strictness being primitive, it depends on the size of the club.

Consider a countable set of clubs $K = \{1, 2, \dots\}$ indexed by k . We assume $|K| \geq 4$. The number of clubs that are 'active', i.e. attract a positive mass members, is determined endogenously.

Now in each period t , clubs simultaneously announce strictness levels at date 1. The strictness profile is denoted by $\mathbf{s}^t = (s_1^t, s_2^t, \dots, s_k^t \dots)$.

At date 2, each individual $i \in I$ chooses participation, where x_{ik}^t denotes i 's participation in club k . We make the standard assumption of participation in one club: if $x_{ik}^t > 0$ then $x_{ik'}^t = 0$ for all $k' \neq k$. As before, i is a member of club k if $x_{ik}^t \geq s_k^t$. The set of such individuals is denoted by M_k^t .

For each club k and time t , participation is summarized by $x_k^t = (x_{Lk}^t, x_{Hk}^t)$, where $x_\omega^t k$ is the (common) level of participation in club k by ω types. Due to increasing returns and coalitional deviations, either all type ω agents join club k or none do, $\omega = L, H$. Hence total participation in club k can be written $X_k^t = p^t x_{Hk}^t + (1 - p^t) x_{Lk}^t$.

Suppressing time notation for the moment, the payoff to individual $i \in M_k$ is:

$$u_i(x_i, x_k) = \pi(1 - x_{ik}) + \theta_i X_k^{1/2}. \quad (9)$$

The payoff from not being a member of any club is again $\Delta\pi$.

As before, each club k chooses a function $s_k(p)$ that maximizes the discounted sum of participation in club k over time given by:

$$G_k^T = \sum_{t=T}^{\infty} \mu^{t-T} X_k(s^t, p^t). \quad (10)$$

We examine the general case of $\alpha \in [0, 1]$ and $\gamma \geq 0$. Redefine group-average participation as the sum over all clubs: $X(s^t, p^t) = \sum_{k \in K} X_k(s^t, p^t)$. Club-average participation $\tilde{X}(s^t, p^t)$

¹⁹For related examples in local public economics, see the Salop-style models of club formation developed by Haimanko et al. (2004) and Polborn (2008).

equals $X(s^t, p^t)$ except in an exclusive SRIE, in which case it equals $x_H(s^t, p^t)$. With these revised definitions, the cultural dynamics under competition are given as before by (7) and (8). Hence, despite competition among clubs, we can still apply our definition of an Identity Equilibrium (Definition 3).

4.1 Short-Run Equilibrium

Competition among entrepreneurs changes the equilibrium structure in several ways. Under monopoly, the identity group could only fragment through non-participation by L types. Under competition, there is a new form of fragmentation — schism — which is simply a separating (short-run) equilibrium.

Definition 6. *A schismatic equilibrium is an SRIE in which all L types join one club and all H types join another.*

The most prominent examples are the long-lasting religious schisms in Christianity, Islam (Maloney et al. 2010) and Judaism (Carvalho & Koyama 2016).

Competition also affects how strict clubs can be. A monopolist pushes members to the point of indifference between joining and not joining. Competition, combined with coalitional deviations, pushes clubs toward maximizing members' welfare. In a schismatic SRIE, each active club is homogeneous. The moderate club maximizes the welfare of its L -type members by setting strictness²⁰

$$\tilde{s}_L \equiv (1 - p) \left(\frac{\theta_L}{2\pi} \right)^2. \quad (11)$$

The more extreme club maximizes the welfare of its H -type members by setting strictness

$$\tilde{s}_H \equiv p \left(\frac{\theta_H}{2\pi} \right)^2 > \tilde{s}_L. \quad (12)$$

In a cohesive SRIE, the active club's membership is heterogeneous, so there is no s that maximizes the welfare of all members. The club chooses s^* between the ideal strictness of L and H types, i.e. between

$$s_L \equiv \left(\frac{\theta_L}{2\pi} \right)^2 \text{ and } s_H \equiv \left(\frac{\theta_H}{2\pi} \right)^2. \quad (13)$$

²⁰When δ is endogenous, we need to restrict the following definitions so that strictness is no greater than one, which we do in the proofs.

The short-run equilibrium structure for a given state p is depicted by Figure 7, where $\tilde{\Delta} \equiv 1 + (1 - p) (\theta_L/2\pi)^2$. (See Proposition A2 in the Appendix for a formal characterization and proof.) When low and high commitment types are sufficiently similar (panel (a): $\theta_L/\theta_H \geq 1/2$), three types of equilibria arise as before. When they are sufficiently distinct (panel (b): $\theta_L/\theta_H < 1/2$), a fourth type of equilibrium arises. When Δ is small and the proportion of H types p is large, the unique SRIE is schismatic. Rather than forming an inclusive club or assimilating, L and H types sort into two exclusive clubs. This schism fragments the identity group where it would have been cohesive under monopoly. Because individual group members, who are short-lived, dictate strictness under competition, the short-run equilibrium structure does not depend on dynamic considerations such as prestige bias (α) and niche construction (γ).

Despite differences in the equilibrium structure, the role of high-attachment types is the same as in the monopoly case. Whether through schism or formation of a single exclusive club, a large share p of H types fragments the identity group. In addition, once a critical mass is reached and fragmentation occurs, total club participation scales superlinearly with p .²¹

4.2 Dynamics

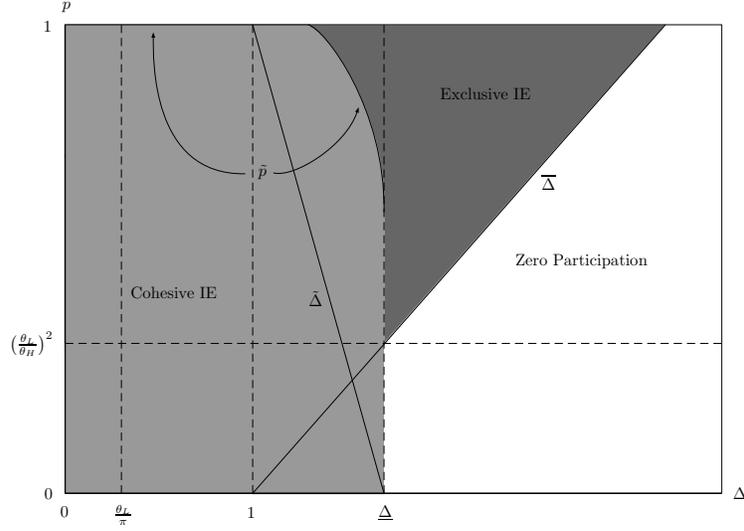
Competition rules out the dynamic radicalization strategies described in Section 3 for the following reason: Individuals live for one period and do not have dynastic preferences. Therefore if a club tries to adjust strictness over time in a way that does not (approximately) maximize its members' stage-game payoffs, then another entrepreneur could step in and steal its members by setting the welfare-maximizing level of strictness for that period. Hence competitive forces mean that each club acts as if it were myopic ($\mu = 0$). The effect on the dynamics is as follows.

Proposition 5. *In every IE:*

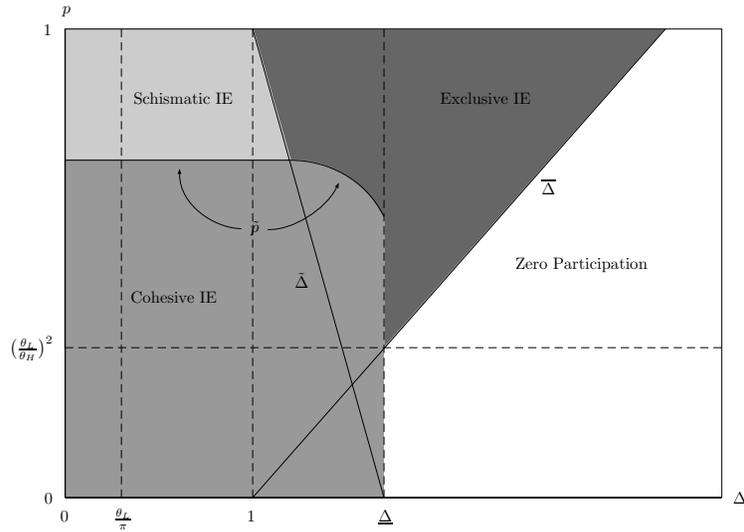
There exists a finite time T such that either an exclusive club forms or schism occurs in all $t \geq T$ if α , θ_H , p^0 , and γ are sufficiently large. In this case, $\lim_{t \rightarrow \infty} p^t = 1$ under schism, and $\lim_{t \rightarrow \infty} p^t = 1$ under an exclusive club if $\alpha > 0$ (prestige bias).

Otherwise, including whenever $\gamma = 0$, the result is the same as in Proposition 2 and radicalization does not occur.

²¹The online appendix describes how total participation varies with discrimination under competition.



(a) $\frac{\theta_L}{\theta_H} \geq \frac{1}{2}$



(b) $\frac{\theta_L}{\theta_H} < \frac{1}{2}$

Figure 7: Short-run equilibrium structure under competition

First, regardless of the degree of prestige bias α , radicalization cannot occur under competition without niche construction $\gamma > 0$. In this sense, competition limits radicalization, which can occur under monopoly for $\alpha > 0$ and $\gamma = 0$. As in the monopoly case, $\alpha > 0$ is only required when p^0 is small, to get a transition going from a moderate, inclusive group to a strict, fragmented one. A smaller α is required when γ is large. Hence prestige bias and niche construction are once again complementary. Unlike the monopoly case, p^0 cannot be too small

if radicalization is to occur.²² In addition, the clubs' discount factor μ is irrelevant, because strictness is dictated by the preferences of short-lived group members, not the preferences of forward-looking identity-based organizations. Hence radicalization emerges under competition from the decentralized actions of identity group members, with niche construction replacing the dynamic radicalization strategies of forward-looking organizations.

4.3 Competition Limits Radicalization

We shall now formalize the notion that competition limits radicalization. Recall that G^T , given by (1), is the club's discounted payoff stream under monopoly at time T . Define $H^T \equiv \sum_{k \in K} G_k^T$ as the sum across all clubs under competition, where G_k^T is given by (10).

Proposition 6. *In every IE, $G^t \geq H^t$ for all t , and strictly so for an open set of parameters.*

The basic intuition behind this result warrants repeating. In our model, identity-based organizations maximize participation by members, rather than members' welfare. As they do not internalize the cost of participation by members, clubs wish to push members beyond their welfare-maximizing level of participation. Competition limits how strict clubs can be, tying participation to individual preferences rather than the clubs' incentives. The proof, however, is more involved than this intuition suggests. One cannot simply compare the strictness of inclusive and exclusive clubs under monopoly and competition. For example, for a given set of parameters, an exclusive club might form under monopoly while schism occurs under competition (see Figure 7). The exclusive club under monopoly is stricter than the corresponding H -type club under competition. But including the schismatic L -type club, it could be that overall participation is higher under competition. We find this not to be the case.

5 Application: Muslim Communities in Europe

We have identified several broad mechanisms behind radicalization that apply to a variety of contexts, including political, racial, and religious radicalization. Here we apply our results to the case of contemporary Muslim communities in Europe.

²²That is, a high θ_H can no longer fully compensate for a low initial share of high-attachment types p^0 .

Muslims comprise the largest religious minority in most European nations. Concerns about the social isolation and radicalization of Muslim communities have been a powerful force in European politics, spurring anti-immigration movements and shaping public policy, including bans on Islamic symbols such as minarets and veils. Muslim immigrants in Europe exhibit higher levels of religiosity than natives (Norris & Inglehart 2012) and lower levels of social and economic integration than other migrant groups (Algan, Bisin, Manning & Verdier 2012). Brown (2000) and Lindley (2002) find that Muslims in the UK have worse labor market outcomes than any other religious group. Algan, Landais & Senik (2012) report that in France immigrants with a Maghrebi background exhibit the lowest level of exogamy and convergence to native cultural values and practices, in particular those relating to religion and gender (e.g. fertility, spousal age gap). Constant et al. (2012) find that Turks are the least integrated immigrant group in Germany. They have weaker German national identity, speak German less proficiently, marry earlier, and have higher fertility rates. There is, however, evidence of cultural convergence over time across a number of European nations (Norris & Inglehart 2012, Algan, Bisin, Manning & Verdier 2012).

One concern is that slow integration could aid terrorist recruitment (see also Rosendorff & Sandler 2004, 2010, Bueno de Mesquita & Dickson 2007). Around 5000 Europeans were estimated to have fought for ISIS in Syria. Benmelech & Klor (2018) present evidence that foreign fighters who joined ISIS did not do so for economic reasons. Within the European Union, they came primarily from Belgium, France, Germany and the UK, the more developed countries. Instead of economic deprivation, qualitative evidence suggests foreign fighters were motivated by difficulties they faced in integrating into these societies.

Principles of Anti-Radicalization. Our analysis has produced two normative principles for preventing radicalization, each of which involves a form of symmetry: First, active participants in identity-based activities should be accorded no greater prestige and cultural influence than inactive group members. Second, a society must not respond to rising participation by extreme group members with blanket discrimination against all members of the identity group. In this sense, it is not economic policy, but a shift in social norms that is required to limit the risk of radicalization.

To illustrate, we restate our results on prestige bias and niche construction as follows. There exist multiple equilibria, so the same identity group can evolve along different paths, either socially integrating over time or radicalizing. If p^0 is low, either a moderate inclusive club forms or there is zero identity-based participation. If p^0 is high, we get radicalization. Thus,

an exogenous increase in the share of high-attachment types can lead to radicalization of a moderate group. In the case of Muslim communities, we interpret x^t as the degree of religious participation, $1 - x^t$ as the degree of social integration, and p^t as the proportion of high-religiosity types. An (exogenous) rise in p^t is precisely what occurred during the global Islamic revival beginning in the 1970s (e.g. Berger 1999, Binzel & Carvalho 2017). Bayat (2007) describes a qualitative shift in Islamic practice in Egypt, the epicenter of Islamic movements in the Arab world: “[...] the actively pious began to judge others for what and how they believed. By privileging their own forms of devotion, they generated new lines of division and demarcation” [p. 150]. This is the process of strict club formation we analyze, which spread to Muslim communities in Europe.

To lead to radicalization, our model requires prestige-biased cultural transmission and/or niche construction. We expect that both mechanisms play a role in Muslim communities. In many Muslim societies, there exists an asymmetry in which the pious seem to be accorded greater prestige and influence by the non-observant than the non-observant are accorded by the pious. This is the prestige bias we model in this paper and is what Bayat (2007) describes in Egypt. In addition, a process akin to niche construction is uncovered by Adida, Laitin & Valfort (2016). They suggest that a discriminatory equilibrium exists in France, in which Muslims face blanket discrimination for not integrating and do not integrate because they face blanket discrimination. Evidence of (blanket) labor market discrimination against French Muslims is provided by Adida et al. (2010) and Duguet et al. (2010).

The events of 9/11 and numerous subsequent terrorist attacks raised blanket discrimination against Muslims in Europe and the United States and could have exacerbated this vicious cycle. Following the Charlie Hebdo attack in Paris, Packer (2015) describes statements on online forums such as “I fear for the Muslims of France. The narrow-minded or frightened are going to dig in their heels and make an *amalgame*” – conflate terrorists with all Muslims’. Hanes & Machin (2014) find that hate crimes in parts of England spiked following 9/11 and the subsequent 7/7 attacks on London, and did not return to pre-attack levels. Our analysis demonstrates how an otherwise temporary spike in blanket discrimination can lead to long-lasting radicalization. Forward-looking identity-based organizations play an important role in this process. Former CIA Director General David Petraeus describes the dynamic niche construction strategies of terrorist group as follows:

*The terrorists’ explicit hope has been to try to provoke a clash of civilizations
— telling Muslims that the United States is at war with them and their reli-*

gion. When Western politicians propose blanket discrimination against Islam, they bolster the terrorists' propaganda. (Petraeus 2016)

Stigma vs Blanket Discrimination. The distinction we make between stigma and blanket discrimination is important in the context of Muslim communities. For example, it provides an economic explanation for why veiling among Muslim women is supposed to have risen following 9/11 (Haddad 2007), despite the greater stigma associated with it. The cost of veiling depends on the difference between the stigma faced by a veiled Muslim woman and the (blanket) discrimination faced by an unveiled Muslim woman. Thus, veiling could increase after 9/11 if stigma was swamped by a rise in blanket discrimination.

At first glance our analysis suggests that stigmatizing participation in identity-based activities reduces radicalization, while blanket discrimination against the identity group promotes it. This would suggest that the French secular approach (*laïcité*), which restricts public expression of religiosity and distinctions based on religious affiliation, would produce less radicalization than the multicultural model of inclusiveness that protects and de-stigmatizes religious participation. France banned conspicuous religious symbols, including the *hijab*, from schools in 2004 and the face veil (*niqab*) from all public spaces in 2011. In the language of our model, these policies sought to reduce blanket discrimination, while permitting a form of conditional discrimination (stigma) against participants in strict religious groups. However, a deeper look at the dynamics of our model reveals a more complicated picture. Consider the case of prestige bias ($\alpha > 0$), but no niche construction ($\gamma = 0$). If $\Delta \leq \underline{\Delta}$, it takes a forward-looking club ($\mu > 0$) to radicalize a moderate, inclusive group. Raising stigma to $\Delta \in (\underline{\Delta}, \overline{\Delta}(p^0)]$ would eliminate the inclusive club, as low-attachment types (but not H types) voluntarily assimilate (zero participation). Hence an exclusive club forms, but not through the design of a forward-looking leader (i.e. even when $\mu = 0$), and can be stabilized for α and θ_H are sufficiently large. In addition, it could be that raising stigma increases α , if active club members gain prestige from being stigmatized. Hence raising stigma can backfire and boost radicalization under certain conditions. Of course, there are many other factors to consider and one might reasonably reject *laïcité* on pre-consequential grounds as a violation of religious freedom.

Religious Competition. The main prediction of the religious markets literature is that religious competition increases overall religious *affiliation* (e.g. Iannaccone 1991, McBride 2008). We focus on religious *participation* (and other identity-based activities) and show that religious competition reduces participation. Our results bear on a famous debate between

David Hume and Adam Smith. Hume proposed that a state-funded religious monopoly would limit religious extremism by bribing the clergy into indolence (Hume 1983, vol. 3, sec. 29). Smith argued to the contrary that religious competition would produce fragmentation and avert power struggles among large religious rivals (Smith 2003). We have provided here an alternative argument for religious competition: Competitive religious organizations cannot enforce strict participation requirements that push members to the point of indifference between joining and assimilating, as a monopolist does. Instead religious strictness is dictated by the preferences of an organization’s members, which reduces strictness and rules out various dynamic radicalization strategies. This is a simple point, but one missing in the literature on religious markets.

Our analysis also bears upon anti-radicalization strategies used by governments today. A common approach is to co-opt religious leaders, anointing a leader or organization as the main representative of a religious community. A severe example is Austria’s expulsion of 60 imams and closure of seven mosques in June 2018. In exchange for market power, these organizations are supposed to moderate the community. In the language of our model, the government co-opts a religious organization to reduce both strictness s and the cultural influence of extreme types in the community α . This is reminiscent of co-option strategies used for centuries by rulers in the Muslim world (Auriol & Platteau 2017, Rubin 2017). In Europe, however, this anti-radicalization strategy may neglect the moderating effects of religious competition identified here. To the extent that moderate community leaders (a) crowd out religious competition and (b) lack legitimacy in marginalized and highly religious parts of a religious community, the community leader model may grant more extreme groups a virtual local monopoly in the areas most susceptible to radicalization. Moreover, co-option may break down at critical times when the government loses legitimacy within the religious community, and moderation by community leaders would lead to their loss of legitimacy and control (see for e.g. Moustafa 2000).

6 Conclusion

This paper combines approaches from cultural evolutionary theory and the economics of religion to study conditions under which an organization can radicalize an identity group, mobilizing moderate individuals into a strict club with intensive participation in identity-based activities. Club participation scales superlinearly with the share of high-attachment types, once a critical mass is reached. A rising share of high-attachment types fragments

the identity group. By itself, this does not lead to radicalization. Based on concepts in cultural evolutionary theory, we identify two complementary mechanisms behind radicalization: prestige-biased cultural transmission and niche construction. Stigmatizing participants in an identity-based club may backfire and boost radicalization. Competition rules out dynamic radicalization strategies by clubs and moderates identity club participation more generally. Finally, we apply our model to Muslim communities in Europe and set out principles for preventing radicalization.

References

- Adida, C. L., Laitin, D. D. & Valfort, M.-A. (2010), ‘Identifying barriers to Muslim integration in France’, *Proceedings of the National Academy of Sciences* **107**(52), 22384–22390.
- Adida, C. L., Laitin, D. D. & Valfort, M.-A. (2016), *Why Muslim integration fails in Christian-heritage societies*, Harvard University Press.
- Aimone, J. A., Iannaccone, L. R., Makowsky, M. D. & Rubin, J. (2013), ‘Endogenous group formation via unproductive costs’, *Review of Economic Studies* **80**(4), 1215–1236.
- Akerlof, G. A. & Kranton, R. E. (2000), ‘Economics and identity’, *The Quarterly Journal of Economics* **115**(3), 715–753.
- Akerlof, G. A. & Kranton, R. E. (2010), *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-being*, Harvard University Press.
- Akerlof, R. (2017), ‘Value formation: The role of esteem’, *Games and Economic Behavior* **102**, 1–19.
- Algan, Y., Bisin, A., Manning, A. & Verdier, T. (2012), *Cultural integration of immigrants in Europe*, Oxford University Press, New York, NY.
- Algan, Y., Landais, C. & Senik, C. (2012), Cultural integration in France, in ‘Cultural integration of immigrants in Europe’, Oxford University Press, New York, NY, pp. 49–68.
- Auriol, E. & Platteau, J.-P. (2017), ‘Religious co-option in autocracy: A theory inspired by history’, *Journal of Development Economics* **127**, 395–412.
- Barro, R. J. & McCleary, R. M. (2003), ‘Religion and economic growth across countries’, *American Sociological Review* pp. 760–781.
- Barros, P. P. & Garoupa, N. (2002), ‘An economic theory of church strictness’, *The Economic Journal* **112**(481), 559–576.
- Bayat, A. (2007), *Making Islam Democratic*, Stanford University Press, Stanford, CA.
- Benmelech, E. & Klor, E. F. (2018), ‘What explains the flow of foreign fighters to ISIS?’, *Terrorism and Political Violence* pp. 1–24.

- Berger, P. L. (1999), The desecularization of the world: A global overview, *in* P. L. Berger, ed., ‘The desecularization of the world: resurgent religion and world politics’, Ethics and Public Policy Center, Washington D.C., pp. 1–18.
- Berman, E. (2000), ‘Sect, subsidy, and sacrifice: An economist’s view of Ultra-Orthodox Jews’, *Quarterly Journal of Economics* **115**(3), 905–953.
- Berman, E. (2009), *Radical, Religious, and Violent: The New Economics of Terrorism*, MIT Press, Cambridge, MA.
- Berman, E. & Laitin, D. D. (2008), ‘Religion, terrorism and public goods: Testing the club model’, *Journal of Public Economics* **92**(10-11), 1942–1967.
- Bernheim, B., Peleg, B. & Whinston, M. D. (1987), ‘Coalition-proof nash equilibria i. concepts’, *Journal of Economic Theory* **42**(1), 1 – 12.
- Bernstein, L. (1992), ‘Opting out of the legal system: Extralegal contractual relations in the diamond industry’, *The Journal of Legal Studies* **21**(1), 115–57.
- Binzel, C. & Carvalho, J.-P. (2017), ‘Education, social mobility, and religious movements: The Islamic revival in Egypt’, *The Economic Journal* **127**(607), 2553–2580.
- Bisin, A., Patacchini, E., Verdier, T. & Zenou, Y. (2011), ‘Formation and persistence of oppositional identities’, *European Economic Review* **55**(8), 1046–1071.
- Bisin, A. & Verdier, T. (2000), ‘A model of cultural transmission, voting and political ideology’, *European Journal of Political Economy* **16**(1), 5–29.
- Bisin, A. & Verdier, T. (2017), ‘On the joint dynamics of culture and institutions’, NBER Working paper No. 23375.
- Boyd, R. & Richerson, P. J. (1988), *Culture and the evolutionary process*, University of Chicago press.
- Brown, M. S. (2000), ‘Religion and economic activity in the South Asian population’, *Ethnic and Racial Studies* **23**(6), 1035–1061.
- Bueno de Mesquita, E. & Dickson, E. S. (2007), ‘The propaganda of the deed: Terrorism, counterterrorism, and mobilization’, *American Journal of Political Science* **51**(2), 364–381.
- Carvalho, J.-P. (2013), ‘Veiling’, *The Quarterly Journal of Economics* **128**(1), 337–370.
- Carvalho, J.-P. (2016), ‘Identity-based organizations’, *American Economic Review: Papers and Proceedings* **106**, 410–414.
- Carvalho, J.-P. (2019), Religious clubs: The strategic role of religious identity, *in* J.-P. Carvalho, S. Iyer & J. Rubin, eds, ‘Advances in the Economics of Religion’, Palgrave.
- Carvalho, J.-P. (forthcoming), ‘Sacrifice and sorting in clubs’, Forum for Social Economics.
- Carvalho, J.-P. & Koyama, M. (2014), ‘Resisting education’, *Working Paper, University of California, Irvine* .
- Carvalho, J.-P. & Koyama, M. (2016), ‘Jewish emancipation and schism: Economic development and religious change’, *Journal of Comparative Economics* **44**(3), 562–584.

- Cavalli-Sforza, L. L. & Feldman, M. W. (1981), *Cultural transmission and evolution: A quantitative approach*, number 16, Princeton University Press.
- Chen, T., McBride, M. & Short, M. B. (2018), ‘Dynamics of religious group growth and survival’, Working Paper UC Irvine.
- Chudek, M., Heller, S., Birch, S. & Henrich, J. (2012), ‘Prestige-biased cultural learning: Bystander’s differential attention to potential models influences children’s learning’, *Evolution and Human Behavior* **33**(1), 46–56.
- Constant, A., Nottmeyer, O., & Zimmermann, K. (2012), Cultural integration in Germany, in ‘Cultural integration of immigrants in Europe’, Oxford University Press, New York, NY, pp. 69–124.
- Duguet, E., Leandri, N., L’horty, Y. & Petit, P. (2010), ‘Are young French jobseekers of ethnic immigrant origin discriminated against? a controlled experiment in the Paris area’, *Annales d’Économie et de Statistique* **99/100**, 187–215.
- Fouka, V. (2017), ‘How do immigrants respond to discrimination? the case of Germans in the US during World War I’, *Working Paper, Stanford University*.
- Glazer, N. & Moynihan, D. P. (1963), ‘Beyond the melting pot’.
- Haddad, Y. Y. (2007), ‘The post-9/11 hijab as icon’, *Sociology of religion* **68**(3), 253–267.
- Haimanko, O., Breton, M. L. & Weber, S. (2004), ‘Voluntary formation of communities for the provision of public projects’, *Journal of Economic Theory* **115**(1), 1 – 34.
- Han, B., Hirshleifer, D. & Walden, J. (2018), Social transmission bias and investor behavior, Technical report, National Bureau of Economic Research.
- Hanes, E. & Machin, S. (2014), ‘Hate crime in the wake of terror attacks: Evidence from 7/7 and 9/11’, *Journal of Contemporary Criminal Justice* **30**(3), 247–267.
- Hauk, E. & Mueller, H. (2015), ‘Cultural leaders and the clash of civilizations’, *Journal of Conflict Resolution* **59**(3), 367–400.
- Henrich, J. & Gil-White, F. J. (2001), ‘The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission’, *Evolution and Human Behavior* **22**(3), 165–196.
- Hume, D. (1983), *The History of England: From the Invasion of Julius Caesar to the Revolution in 1688*, Liberty Classics.
- Hwang, S.-H., Naidu, S. & Bowles, S. (2016), ‘Social conflict and the evolution of unequal conventions’.
- Iannaccone, L. R. (1990), ‘Religious practice: A human capital approach’, *Journal for the scientific study of religion* pp. 297–314.
- Iannaccone, L. R. (1991), ‘The consequences of religious market structure: Adam Smith and the economics of religion’, *Rationality and society* **3**(2), 156–177.
- Iannaccone, L. R. (1992), ‘Sacrifice and stigma: Reducing free-riding in cults, communes, and other collectives’, *Journal of Political Economy* **100**(2), 271–91.

- Ibrahim, S. E. (1980), 'Anatomy of Egypt's militant Islamic groups: Methodological note and preliminary findings', *International Journal of Middle East Studies* **12**(4), 423–453.
- Iyer, S. (2016), 'The new economics of religion', *Journal of Economic Literature* **54**(2), 395–441.
- Iyer, S., Velu, C. & Weeks, M. (2014), 'Divine competition: Religious organisations and service provision in India', working paper, University of Cambridge.
- Kukathas, C. (2003), *The liberal archipelago: A theory of diversity and freedom*, Oxford University Press.
- Kuran, T. (2006), *Islam and Mammon*, Princeton University Press, New Jersey.
- Levy, G. & Razin, R. (2012), 'Religious beliefs, religious participation, and cooperation', *American Economic Journal: Microeconomics* **4**(3), 121–151.
- Lindley, J. (2002), 'Race or religion? The impact of religion on the employment and earnings of Britain's ethnic communities', *Journal of Ethnic and Migration Studies* **28**(3), 427–442.
- Mackie, G. (1996), 'Ending footbinding and infibulation: A convention account', *American Sociological Review* pp. 999–1017.
- McBride, M. (2008), 'Religious pluralism and religious participation: A game theoretic analysis', *American Journal of Sociology* **114**(1), 77–106.
- McBride, M. (2010), 'Religious market competition in a richer world', *Economica* **77**(305), 148–171.
- McBride, M. (2015), 'Why churches need free-riders: religious capital formation and religious group survival', *Journal of Behavioral and Experimental Economics* **58**, 77–87.
- McCleary, R. M. & Barro, R. J. (2006), 'Religion and economy', *Journal of Economic Perspectives* **20**(2), 49–72.
- Montgomery, J. D. (2003), 'A formalization and test of the religious economies model', *American Sociological Review* **68**(5), 782–809.
- Moustafa, T. (2000), 'Conflict and cooperation between the state and religious institutions in contemporary Egypt', *International Journal of Middle East Studies* **32**(1), 3–22.
- Norris, P. & Inglehart, R. F. (2012), 'Muslim integration into Western cultures: Between origins and destinations', *Political Studies* **60**(2), 228–251.
- Odling-Smee, F. J., Laland, K. N. & Feldman, M. W. (2003), *Niche construction: The neglected process in evolution*, Princeton University Press.
- Packer, G. (2015), 'The other France: Are the suburbs of Paris incubators of terrorism?', *The New Yorker*.
- Petraeus, D. (2016), 'Anti-Muslim bigotry aids Islamist terrorists', Washington Post.
- Polborn, M. K. (2008), 'Competing for recognition through public good provision', *The B.E. Journal of Theoretical Economics* **8**(1), Article 22.
- Prummer, A. & Siedlarek, J.-P. (2017), 'Community leaders and the preservation of cultural traits', *Journal of Economic Theory* **168**, 143–176.

- Rosendorff, B. P. & Sandler, T. (2004), 'Too much of a good thing? the proactive response dilemma', *Journal of Conflict Resolution* **48**(5), 657–671.
- Rosendorff, B. P. & Sandler, T. (2010), 'Suicide terrorism and the backlash effect', *Defence and Peace Economics* **21**(5-6), 443–457.
- Roy, O. (1994), *The failure of political Islam*, Harvard University Press.
- Rubin, J. (2017), *Rulers, Religion, and Riches: Why the West got rich and the Middle East did not*, Cambridge University Press.
- Sen, A. (2007), *Identity and violence: The illusion of destiny*, Penguin Books India.
- Shayo, M. (2009), 'A model of social identity with an application to political economy: Nation, class, and redistribution', *American Political Science Review* **103**(2), 147–174.
- Smith, A. (2003), *The Wealth of Nations*, Bantam Classics.
- Verdier, T. & Zenou, Y. (2015), 'The role of cultural leaders in the transmission of preferences', *Economics Letters* **136**, 158–161.
- Verdier, T. & Zenou, Y. (2018), 'Cultural leader and the dynamics of assimilation', *Journal of Economic Theory* **175**, 374–414.
- Young, H. P. (2015), 'The evolution of social norms', *Annual Review of Economics* **7**(1), 359–387.

Appendix

Propositions are proved in order of mention in the main body, including the two additional propositions A1 and A2 underlying Figures 1 and 7 respectively. Before proving Proposition A1, we state and prove:

Lemma 1. *Club membership in state p is determined by the value of the outside option as follows:*

(i) *There exists an $s \in [0, 1]$ such that $M^*(s) = I$ only if*

$$\Delta \leq \underline{\Delta} \equiv 1 + \left(\frac{\theta_L}{2\pi}\right)^2. \quad (IR_L)$$

Otherwise, L types assimilate: $M^(s) \cap I_L = \emptyset$ for all $s \in [0, 1]$ in every identity equilibrium (IE).*

(ii) *There exists an $s \in [0, 1]$ such that $I_H \subseteq M^*(s)$ only if*

$$\Delta \leq \max\{\underline{\Delta}, \bar{\Delta}\}, \text{ where } \bar{\Delta} \equiv 1 + p \left(\frac{\theta_H}{2\pi}\right)^2. \quad (IR_H)$$

Otherwise, all types assimilate: $M^(s) = \emptyset$ for all $s \in [0, 1]$ in every IE.*

Proof. (i) Consider a cohesive equilibrium. Because agents are nonatomic, $x_L(s) = x_H(s) = s$.

The payoff to i from joining this inclusive club is

$$\pi(1 - s) + \theta_i (s)^{1/2}, \quad (14)$$

which is maximized at

$$s = \left(\frac{\theta_i}{2\pi}\right)^2, \quad (15)$$

yielding a maximum of

$$\pi + \frac{1}{\pi} \left(\frac{\theta_i}{2}\right)^2. \quad (16)$$

Hence there exists an s such that $M^*(s) = I$ only if the non-participation payoff $\Delta\pi$ exceeds (16) for L types, or

$$\Delta \leq \underline{\Delta} \equiv 1 + \left(\frac{\theta_L}{2\pi}\right)^2. \quad (17)$$

(i) Consider an exclusive equilibrium. Because agents are nonatomic, $x_H(s) = s$.

The payoff to i from joining this exclusive club is

$$\pi(1 - s) + \theta_H (ps)^{1/2}, \quad (18)$$

which is maximized at

$$s = p \left(\frac{\theta_H}{2\pi} \right)^2, \quad (19)$$

yielding a maximum of

$$\pi + \frac{p}{\pi} \left(\frac{\theta_H}{2} \right)^2. \quad (20)$$

Hence there exists an s such that $M^*(s) = I_H$ only if the non-participation payoff $\Delta\pi$ exceeds (20), or

$$\Delta \geq \bar{\Delta} \equiv 1 + p \left(\frac{\theta_H}{2\pi} \right)^2. \quad (21)$$

□

We know that if $\Delta > \max \{\underline{\Delta}, \bar{\Delta}\}$, then there is zero participation (Lemma 1). For $\Delta \leq \max \{\underline{\Delta}, \bar{\Delta}\}$, the following proposition characterizes the set of short-run identity equilibria (SRIE).

Proposition A1. *The set of SRIE is as follows.*

(i) $\Delta \leq \theta_L/\pi$: *There exists a cohesive SRIE with strictness $s^* = 1$ and contributions $x_L = x_H = 1$.*

(ii) $\Delta \in (\theta_L/\pi, \underline{\Delta}]$: *There exists a unique threshold proportion of high-commitment types $\hat{p} \in ([\theta_L/\theta_H]^2, 1)$, which is strictly decreasing in Δ .*

If $p \leq \hat{p}$, there exists a cohesive SRIE.

If $p \geq \hat{p}$, there exists an exclusive SRIE.

(iii) $\Delta \in (\underline{\Delta}, \bar{\Delta}]$: *There exists an exclusive SRIE.*

There are no other SRIE in these cases.

Proof. The club maximizes (1) subject to $p^{t+1} = X^t$. H types have a greater payoff from participation than L types. Hence total participation X^{t+1} is non-decreasing in p^{t+1} . Thus in every period t , the club sets strictness s^t to maximize X^t .

$s^t \in \{\underline{s}, \bar{s}(p^t)\}$ in equilibrium. Otherwise, the club could increase strictness (and participation) without a decline in membership in period t .

By the same argument in Lemma 1, $x_i = s^* \in \{\underline{s}, \bar{s}\}$ for all $i \in M^*(s^*)$.

Now consider membership choices. Due to increasing returns, either all L types join or none do. Likewise for H types. As $\theta_H > \theta_L$, if all L types join, so do all H types. Hence $M^*(s) \in \{\emptyset, I_H, I\}$.

Suppose there exists an s' such that $M^*(s') \neq \emptyset$. The club will never set s such that $M^*(s) = \emptyset$, as $G(s) = 0$ in this case, a minimum of its objective function. Combining this fact with Lemma 1, $M^*(s^*) = \emptyset$ if and only if $\Delta > \max \{\underline{\Delta}, \bar{\Delta}\}$.

Now suppose that $\Delta \leq \max\{\underline{\Delta}, \overline{\Delta}\}$, so that $M^*(s^*)$ equals I_H or I . We have established that $s^* = \bar{s}$ in the first case and $s^* = \underline{s}$ in the second case.

Given $x_i = s$ for all $i \in M(s)$, the club prefers to be inclusive if and only if

$$\underline{s} \geq p\bar{s}. \quad (22)$$

Case 1: $\Delta \leq \theta_L/\pi$ or $p \leq (\theta_L/\theta_H)^2$. By (5) and (6), $\underline{s} \geq \bar{s}$. Hence all types are willing to join the club at $s = \underline{s}$ and (22) is satisfied, so the club prefers to be inclusive.

Case 2: $\Delta > \theta_L/\pi$ and $p > (\theta_L/\theta_H)^2$. First, $\underline{s} = \bar{s}$ for $p = (\theta_L/\theta_H)^2$ by (5) and (6). Hence (22) holds. Second, $\underline{s} < \bar{s} = p\bar{s}$ for $p = 1$. Third, \bar{s} is increasing in p and \underline{s} is independent of p . Therefore, there exists a unique $\hat{p} \in ([\theta_L/\theta_H]^2, 1)$, at which (22) binds. The club prefers to be inclusive if and only if $p \leq \hat{p}$. By Lemma 2 below, \hat{p} is strictly decreasing in Δ .

Now we check incentive compatibility. By construction of (5), an inclusive club can be implemented at \underline{s} if and only if $\Delta \leq \underline{\Delta}$ (IR_L).

For an exclusive club to be incentive compatible at \bar{s} , $\Delta \leq \overline{\Delta}$ (IR_H). In addition, there must be no profitable coalitional deviation by low types. The most profitable involves I_L joining to form an inclusive club, and contributing $x_i = \bar{s}$. But \underline{s} , the maximum strictness L types would tolerate in an inclusive club, is less than the club's strictness \bar{s} for $p > (\theta_L/\theta_H)^2$. Therefore, an exclusive club is incentive compatible when $\Delta \leq \overline{\Delta}$ and $p > (\theta_L/\theta_H)^2$.

We know the club prefers being exclusive to zero participation. Therefore an exclusive club forms if $\Delta \in (\underline{\Delta}, \overline{\Delta}]$. This establishes the proposition. \square

Lemma 2. \hat{p} is strictly decreasing in Δ on the domain $(\frac{\theta_L}{\pi}, \underline{\Delta}]$.

Proof. By definition, $\hat{p}\bar{s}(\hat{p}) = \underline{s}$.

Suppose $\bar{s}(\hat{p}) = 1$. Then

$$\frac{d\hat{p}}{d\Delta} = \frac{d\underline{s}}{d\Delta},$$

which is negative for $\Delta > \theta_L/\pi$ by inspection of (5).

Now suppose $\bar{s}(\hat{p}) < 1$. Then

$$\begin{aligned} \frac{d\hat{p}}{d\Delta}\bar{s} + \hat{p} \left[\frac{\partial\bar{s}}{\partial\Delta} + \frac{\partial\bar{s}}{\partial\hat{p}} \frac{d\hat{p}}{d\Delta} \right] &= \frac{d\underline{s}}{d\Delta} \\ \frac{d\hat{p}}{d\Delta} \left[\bar{s} + \hat{p} \frac{\partial\bar{s}}{\partial\hat{p}} \right] &= \frac{d\underline{s}}{d\Delta} - \hat{p} \frac{\partial\bar{s}}{\partial\Delta} \\ &= \frac{\hat{p}\sqrt{\bar{s}}}{\sqrt{\hat{p}(\frac{\theta_H}{2\pi})^2 - (\Delta - 1)}} - \frac{\sqrt{\bar{s}}}{\sqrt{(\frac{\theta_L}{2\pi})^2 - (\Delta - 1)}}, \end{aligned} \quad (23)$$

so that $d\hat{p}/d\Delta$ is negative if and only if the RHS of (23) is negative. As $\hat{p} = \underline{s}/\bar{s}$, (23) is equivalent to

$$\begin{aligned}\hat{p} &< \sqrt{\hat{p}} \sqrt{\frac{\hat{p} \left(\frac{\theta_H}{2\pi}\right)^2 - (\Delta - 1)}{\left(\frac{\theta_L}{2\pi}\right)^2 - (\Delta - 1)}} \\ \hat{p} &< \frac{\hat{p} \left(\frac{\theta_H}{2\pi}\right)^2 - (\Delta - 1)}{\left(\frac{\theta_L}{2\pi}\right)^2 - (\Delta - 1)}.\end{aligned}\tag{24}$$

Since $\hat{p} \in ([\theta_L/\theta_H]^2, 1)$, (24) is satisfied. \square

Proof of Proposition 1.

Proof. Fix the state p . For $\Delta \in (\theta_L/\pi, \underline{\Delta})$, $X^* = \max\{\underline{s}, p\bar{s}\}$ by Proposition A1(ii). For $\Delta \in [\underline{\Delta}, \bar{\Delta})$, $X^* = p\bar{s}$ by Proposition A1(iii).

By (5) and (6), whenever less than one, \underline{s} is a continuous and strictly increasing function of δ and a continuous and strictly decreasing function of Δ . The same goes for \bar{s} . The result follows. \square

Proof of Proposition 2.

Proof. (i) $\Delta \leq \theta_L/\pi$. By Proposition A1, $\underline{s} = 1$. Hence an inclusive club can form at $s^t(p^t) = 1$ regardless of p^t . This maximizes X^t (and p^{t+1}) for all t . Therefore, $p^{t+1} = X^t = 1$ for all $t \geq 0$. This establishes part (i).

(ii) By Proposition A1, there are the following cases:

Case 1: $p^0 \leq \hat{p} \equiv \underline{s}/\bar{s}(p^0)$. An inclusive club forms and $p^1 = \underline{s}$. Observe that $\underline{s} \leq \hat{p} \equiv \underline{s}/\bar{s}(p^t)$, because $\bar{s}(p^t) \leq 1$. Thus, $p^1 \leq \hat{p}$. By induction, an inclusive club forms in every period and $p^t = \underline{s}$ for all $t > 0$.

Case 2: $p^0 > \hat{p}$ and $p^0 \geq \Delta \left(\frac{\pi}{\theta_H}\right)^2$. In this case, $s^0 = \bar{s}(p^0) = 1$ so that $p^1 = p^0 \bar{s}(p^0) = p^0$. By induction, an exclusive club forms at $s^t = 1$ and $p^t = p^0$ for all $t > 0$.

Case 3: $p^0 > \hat{p}$ and $p^0 < \Delta \left(\frac{\pi}{\theta_H}\right)^2$. In this case, $s^0 = \bar{s}(p^0) < 1$ so that $p^1 = p^0 \bar{s}(p^0) < p^0$. By induction, $p^{t+1} < p^t$ as long as an exclusive club is formed. Suppose this continued forever. Then by recursion,

$$p^t = p^0 \prod_{\tau=0}^{t-1} \bar{s}(p^\tau).\tag{25}$$

Because $\bar{s}(p)$ is strictly increasing in p and p^t is strictly decreasing in t ,

$$p^t = p^0 \prod_{\tau=0}^{t-1} \bar{s}(p^\tau) < p^0 [\bar{s}(p^0)]^t. \quad (26)$$

By hypothesis, $\bar{s}(p^0) < 1$. Hence if an exclusive club formed in every period, $\lim_{t \rightarrow \infty} p^t = 0$.

Therefore, there exists a finite time T such that $p^{T-1} > \hat{p}$ and $p^T \leq \hat{p}$. Therefore, an inclusive club forms at time T . From there we enter case 1: $s^t(p^t) = \underline{s}$ and $p^t = \underline{s}$ for all $t \geq T$.

(iii) $\Delta \in (\underline{\Delta}, \bar{\Delta}(p^0)]$. As $\underline{\Delta}$ is time invariant, $\Delta > \underline{\Delta}$ for all time. By Lemma 1(i) then, $M^*(s^t) \neq I$ for all t . In addition, $M^*(s^t)$ is nonempty for some s^t , only if $\Delta \leq \bar{\Delta}(p^t)$. In this case, we know equilibrium strictness s^t is $\bar{s}(p^t)$.

By (6), $\bar{s}(p^0) = 1$ for $p^0 \geq \Delta \left(\frac{\pi}{\theta_H}\right)^2$. In this case, $p^1 = p^0 \bar{s}(p^0) = p^0$. By induction, $p^t = p^0$ for all t .

By (6), $\bar{s}(p^0) < 1$ otherwise. In this case, $p^1 = p^0 \bar{s}(p^0) < p^0$. If $p^0 < \Delta \left(\frac{\pi}{\theta_H}\right)^2$ then, $p^1 < \Delta \left(\frac{\pi}{\theta_H}\right)^2$, so $\bar{s}(p^1) < 1$. By induction, $p^{t+1} < p^t$ for all $t > 0$.

From there, the argument in part (ii) case 3 establishes $\lim_{t \rightarrow \infty} p^t = 0$.

Recall $\bar{\Delta}(0) = 1$. By hypothesis, $\Delta > \underline{\Delta} > 1$. Therefore, there exists a finite time T such that $\Delta \leq \bar{\Delta}(p^{T-1})$ and $\Delta > \bar{\Delta}(p^T)$, so that zero participation occurs at time T . This establishes part (iii) of the proposition.

(iv) $\Delta > \max\{\underline{\Delta}, \bar{\Delta}(p^0)\}$. By Lemma 1(ii), $M^*(s^t) = \emptyset$ for all s^t . Hence for any p^0 , $p^1 = 0$. Recall that $\bar{\Delta}(p)$ is strictly increasing in p . Hence $\Delta > \max\{\underline{\Delta}, \bar{\Delta}(p^1)\}$. By induction, there is zero participation $p^{t+1} = 0$ for all $t \geq 1$. \square

Proof of Proposition 3

Proof. We know the club can do no better than by choosing $s^t(p^t) \in \{\underline{s}, \bar{s}(p^t)\}$.

By hypothesis, $\Delta \in (\theta_L/\pi, \max\{\underline{\Delta}, \bar{\Delta}(p^0)\})$, so a non-empty club can be formed.

Case 1: An inclusive club is formed at $t = 0$, at strictness \underline{s} . Then $p^1 = \underline{s}$.

If an inclusive club were formed at $t = 1$, again $p^2 = \underline{s}$. By induction, the same occurs in every period and $p^t = \underline{s}$ for all $t \geq 1$.

Alternatively, if an exclusive club were formed at $t = 1$, at strictness $\bar{s}(p^1)$, then $p^2 = [\alpha + (1 - \alpha)p^1] \bar{s}(p^1)$. Notice p^2 can be made arbitrarily close to $\bar{s}(p^1)$ by taking α close to one. In addition, $\bar{s}(p^1) = \bar{s}(\underline{s}) > \underline{s}$ because $p^1 = \underline{s} > \left(\frac{\theta_L}{\theta_H}\right)^2$ for θ_H sufficiently large (but not greater than 2π). Therefore, $p^2 > p^1$ for α sufficiently large. Both X^t and p^{t+1} are increasing in p^t under an exclusive club, and constant under an inclusive club. Hence if the club is exclusive at $t = 1$, it is also exclusive for all $t > 1$.

p^t converges to a fixed point p^* given by:

$$p^* = [\alpha + (1 - \alpha)p^*] \bar{s}(p^*). \quad (27)$$

If and only if $p^* \bar{s}(p^*) > \underline{s}$ is there a finite time T such that X^t is larger under an exclusive club than under an inclusive club for all $t \geq T$. This holds if and only if α and θ_H are sufficiently large (but still $\theta_H < 2\pi$).

Hence the difference between the discounted sum of participation under an exclusive club and inclusive club is:

$$\sum_{t=1}^{T-1} \mu^{t-1} \underbrace{[p^t \bar{s}(p^t) - \underline{s}]}_{<0} + \sum_{t=T}^{\infty} \mu^{t-1} \underbrace{[p^t \bar{s}(p^t) - \underline{s}]}_{>0} \quad (28)$$

For μ sufficiently close (but less than) one, the second (positive) term dominates.

Case 2: An exclusive club is formed at $t = 0$.

The same argument in case 1 applies except from $t = 0$, not $t = 1$. □

Proof of Proposition 4

Proof. Recall that $\delta^{t+1} = \delta^0 + \gamma \max\{p^{t+1} - \underline{s}(\delta^0), 0\}$.

We proceed as in Proposition 3.

Case 1: An inclusive club is formed at $t = 0$, at strictness $\underline{s}(\delta^0)$. Then $p^1 = \underline{s}(\delta^0)$ and $\delta^1 = \delta^0$.

If an inclusive club were formed at $t = 1$, again $p^2 = \underline{s}(\delta^0)$. By induction, the same occurs in every period and $p^t = \underline{s}(\delta^0)$ for all $t \geq 1$.

Hence for $\lim_{t \rightarrow \infty} p^t > \underline{s}(\delta^0)$, an exclusive club must be formed at $t = 1$. We can compute a lower bound on the club's continuation payoff G^1 by supposing the club is exclusive for all $t \geq 1$, which is the same outcome as under $\gamma = 0$ in Proposition 3.

As before, we require α to be sufficiently large so that $p^2 > p^1$, and thereby $\delta^2 > \delta^0$. There are also two conditions that depend on δ which need to hold. First, there must be a finite time T such that for all $t \geq T$

$$p^t \bar{s}(p^t, \delta^t(p^t)) > \underline{s}(\delta^0). \quad (29)$$

The RHS of (29) is independent of γ . The LHS is increasing in γ , and strictly so when $\bar{s}(p^t, \delta^t(p^t))$ is less than one. Hence the critical time T is no less under γ than under $\gamma' < \gamma$.

Second:

$$\sum_{t=T}^{\infty} \mu^{t-1} \underbrace{[p^t \bar{s}(p^t, \delta^t(p^t)) - \underline{s}(\delta^0)]}_{>0} > - \sum_{t=1}^{T-1} \mu^{t-1} \underbrace{[p^t \bar{s}(p^t, \delta^t(p^t)) - \underline{s}(\delta^0)]}_{<0}, \quad (30)$$

Recall that $\bar{s}(p^t, \delta^0) < 1$ for all p^t because of the maintained assumption $2\pi = \frac{2w}{1+\delta^0} > \theta_H$. Hence even if T and p^t are invariant to γ , $\bar{s}(p^t, \delta^t(p^t)) > \bar{s}(p^t, \delta^0)$ for all $t \geq 1$ when $\gamma > 0$. Thus, if (30) holds with weak inequality for $\gamma = 0$, it holds with strict inequality for all $\gamma > 0$.

Applying this argument to an arbitrary pair (γ, γ') , it follows that $\Omega(\gamma') \subset \Omega(\gamma)$ whenever $\gamma' < \gamma$.

Case 2: An exclusive club is formed at $t = 0$.

The same argument in case 1 applies except from $t = 0$, not $t = 1$. \square

Proposition A2. *Suppose $\theta_L/\theta_H \geq 1/2$. Then the set of SRIE is as follows.*

- (i) $\Delta \leq \underline{\Delta}$: *There exists a unique threshold of high-commitment types \tilde{p} , which is weakly increasing in Δ .*

If $p \leq \tilde{p}$, there exists a cohesive SRIE.

If $p > \tilde{p}$, there exists an exclusive SRIE.

- (ii) $\Delta \in (\underline{\Delta}, \bar{\Delta}(p)]$: *There exists an exclusive IE*

Define $\tilde{\Delta} \equiv 1 + (1-p)(\theta_L/2\pi)^2 \leq \underline{\Delta}$. *Suppose $\theta_L/\theta_H < 1/2$. Then the set of SRIE is as follows.*

- (iii) $\Delta \leq \tilde{\Delta}$: *There exists a unique threshold proportion of high-commitment types $\tilde{p} \in ([\theta_L/\theta_H]^2, 1]$ which is weakly decreasing in Δ .*

If $p \leq \tilde{p}$, there exists a cohesive SRIE.

If $p \geq \tilde{p}$, there exists a schismatic SRIE.

- (iv) $\Delta \in (\tilde{\Delta}, \underline{\Delta}]$: *Again, there exists a unique threshold proportion of high-commitment types $\tilde{p} \in ([\theta_L/\theta_H]^2, 1]$ which is weakly decreasing in Δ .*

If $p \leq \tilde{p}$, there exists a cohesive SRIE.

If $p \geq \tilde{p}$, there exists an exclusive SRIE.

- (v) $\Delta \in (\underline{\Delta}, \bar{\Delta}(p)]$: *There exists an exclusive SRIE.*

There are no other types of SRIE in these cases.

Proof of Proposition A2

Proof. We shall establish the proposition by identifying the conditions under which each class of SRIE exists for a given p and π . We make use of the fact that $x_{\omega k}^*$ equals s_k or zero for each type $\omega \in \{L, H\}$ and club $k \in K$.

Define $\tilde{\Delta} \equiv 1 + (1-p)(\theta_L/2\pi)^2 \leq \underline{\Delta}$. It is straightforward to show that if $\Delta < \tilde{\Delta}$, L types prefer schism to zero participation. Therefore, the SRIE is either cohesive or schismatic. If $\Delta > \tilde{\Delta}$, L types prefer zero participation to schism. Therefore, the SRIE is either cohesive or exclusive.

Cohesive SRIE. Let $s^* \in [s_L, s_H]$ be the strictness of the unique active group. The equilibrium payoff to i is:

$$\pi(1 - s^*) + \theta_i(s^*)^{1/2}. \quad (31)$$

Due to increasing returns, if there is a profitable deviation by a subset of type ω agents, then there is an even more profitable deviation by the full set of type ω agents I_ω , $\omega = L, H$. Thus only three types of deviations need to be ruled out: (I) another club attracts all agents to form a new inclusive club, (II) at least one other club forms an exclusive club by attracting all individuals of type ω only, and (III) at least one type ω chooses zero participation.

To be profitable, a type-I deviation requires there be an $s \in [0, 1]$ such that:

$$\pi(1 - s^*) + \theta_L(s^*)^{1/2} < \pi(1 - s) + \theta_L s^{1/2}, \quad (32)$$

and

$$\pi(1 - s^*) + \theta_H(s^*)^{1/2} < \pi(1 - s) + \theta_H s^{1/2}. \quad (33)$$

Let $s^* \in [s_L, s_H]$ (the interval is defined by (13)). Because the RHS of (32) is strictly concave and maximized at s_L , (32) is violated for $s \geq s^*$. Because the RHS of (33) is strictly concave and maximized at s_H , (33) is violated for $s \leq s^*$. If $s^* \notin [s_L, s_H]$, both (32) and (33) hold. Hence no such deviation is profitable if and only if $s^* \in [s_L, s_H]$.

Now consider a type-II deviation to an exclusive group. The most a competing entrepreneur can do to attract L types is to set $s = \tilde{s}_L$ (defined by (11)), which yields

$$\max_{s \in [0, 1]} \pi(1 - s) + \theta_L((1 - p)s)^{1/2} = \pi \left[1 + (1 - p) \left(\frac{\theta_L}{2\pi} \right)^2 \right].$$

The most a competing entrepreneur can do to attract H types is to set $s = \tilde{s}_H$ (defined by (12)), which yields

$$\max_{s \in [0, 1]} \pi(1 - s) + \theta_H(ps)^{1/2} = \pi \left[1 + p \left(\frac{\theta_H}{2\pi} \right)^2 \right].$$

Hence the following conditions rule out a profitable type-II deviation:

$$\pi(1 - s^*) + \theta_L(s^*)^{1/2} \geq \pi \left[1 + (1 - p) \left(\frac{\theta_L}{2\pi} \right)^2 \right], \quad (34)$$

$$\pi(1 - s^*) + \theta_H(s^*)^{1/2} \geq \pi \left[1 + p \left(\frac{\theta_H}{2\pi} \right)^2 \right]. \quad (35)$$

The following two participation constraints rule out a profitable type-III deviation:

$$\pi(1 - s^*) + \theta_L(s^*)^{1/2} \geq \Delta\pi \quad (36)$$

$$\pi(1 - s^*) + \theta_H(s^*)^{1/2} \geq \Delta\pi. \quad (37)$$

Case 1: $\Delta \leq \underline{\Delta}$ and $p \leq (\theta_L/\theta_H)^2$. $\underline{\Delta}\pi$ is an upper bound on the RHS of conditions (34)-(37). The LHS of (34) is a lower bound on the LHS of conditions (34)-(37). Therefore, it is sufficient to show there exists $s^* \in [s_L, s_H]$ such that

$$\begin{aligned} \pi(1 - s^*) + \theta_L(s^*)^{1/2} &\geq \underline{\Delta}\pi \\ &= \max_{s \in [0,1]} \pi(1 - s) + \theta_L s^{1/2}. \end{aligned} \quad (38)$$

Hence (38) is satisfied for $s^* = s_L$ and there exists a cohesive SRIE.

Case 2: $\Delta \leq \underline{\Delta}$ and $p > (\theta_L/\theta_H)^2$. First, note that (37) is satisfied strictly whenever (36) is satisfied. Hence it is sufficient to show there exists $s^* \in [s_L, s_H]$ that satisfies (34)-(36).

Case 2a: $\Delta \leq 1 + (1 - p) [\theta_L/(2\pi)]^2$. In this case, the RHS of (34) is no less than the RHS of (36), so the relevant constraints are (34) and (35). Denote the smallest s^* that satisfies (35) by z_H . We have:

$$z_H = \left(\frac{\theta_H}{2\pi} (1 - \sqrt{1 - p}) \right)^2.$$

Denote the largest s^* that satisfies (34) by z_L . We have:

$$z_L = \left(\frac{\theta_L}{2\pi} (1 + \sqrt{p}) \right)^2.$$

Notice from (13) that $z_L \geq s_L$ and $z_H \leq s_H$.

For case 2a then, it suffices to show that $z_L > z_H$. In this case, there exists an $s^* \in [z_H, z_L]$ which satisfies both constraints and also lies in $[s_L, s_H]$. Comparing:

$$\begin{aligned} z_L &\geq z_H \\ \left(\frac{\theta_L}{2\pi} [1 + \sqrt{p}] \right)^2 &\geq \left(\frac{\theta_H}{2\pi} [1 - \sqrt{1 - p}] \right)^2 \\ \frac{\theta_L}{\theta_H} &\geq \frac{1 - \sqrt{1 - p}}{1 + \sqrt{p}}. \end{aligned} \quad (39)$$

The RHS of (39) increases monotonically from 0 to 1/2 as p goes from 0 to 1. Therefore, (39) is satisfied for all p if $\theta_L \geq (1/2)\theta_H$.

Now suppose that $\theta_L < (1/2)\theta_H$. Evaluating (39) at $p = (\theta_L/\theta_H)^2$ yields

$$\frac{\theta_L}{\theta_H} \geq \frac{1 - \sqrt{1 - \left(\frac{\theta_L}{\theta_H}\right)^2}}{1 + \frac{\theta_L}{\theta_H}}. \quad (40)$$

Note:

$$\begin{aligned}
\left(\frac{\theta_L}{\theta_H}\right)^2 + 1 - \left(\frac{\theta_L}{\theta_H}\right)^2 &= 1 \\
\left(\frac{\theta_L}{\theta_H}\right)^2 + \sqrt{1 - \left(\frac{\theta_L}{\theta_H}\right)^2} &> 1 \\
\left(\frac{\theta_L}{\theta_H}\right)^2 &> 1 - \sqrt{1 - \left(\frac{\theta_L}{\theta_H}\right)^2} \\
\frac{\theta_L}{\theta_H} &> \frac{1 - \sqrt{1 - \left(\frac{\theta_L}{\theta_H}\right)^2}}{\frac{\theta_L}{\theta_H}},
\end{aligned}$$

which implies that (40) holds. Thus (39) is satisfied at $p = (\theta_L/\theta_H)^2$. But not at $p = 1$. Therefore, there exists a threshold \tilde{p} such that (39) is satisfied if and only if $p \leq \tilde{p}$.

Case 2b: $\Delta \in (1 + (1 - p)[\theta_L/(2\pi)]^2, \underline{\Delta}]$. Now the RHS of (34) is less than the RHS of (36), so the relevant constraints are (36) and (35). (The community now fragments through zero participation not schism.)

Denote the largest s^* that satisfies (36) by \tilde{z}_L . We have:

$$\tilde{z}_L = \left(\frac{\theta_L}{2\pi} + \frac{1}{2} \sqrt{\left(\frac{\theta_L}{\pi}\right)^2 - 4(\Delta - 1)} \right)^2.$$

Notice from (13) that $\tilde{z}_L \geq s_L$ since $\Delta \leq \underline{\Delta}$.

Hence it suffices to show that $\tilde{z}_L > z_H$. In this case, there exists an $s^* \in [z_H, \tilde{z}_L]$ which satisfies both constraints and also lies in $[s_L, s_H]$. Comparing:

$$\begin{aligned}
&\tilde{z}_L \geq z_H \\
\left(\frac{\theta_L}{2\pi} + \frac{1}{2} \sqrt{\left(\frac{\theta_L}{\pi}\right)^2 - 4(\Delta - 1)} \right)^2 &\geq \left(\frac{\theta_H}{2\pi} [1 - \sqrt{1 - p}] \right)^2 \\
\theta_H \sqrt{1 - p} + \pi \sqrt{\left(\frac{\theta_L}{\pi}\right)^2 - 4(\Delta - 1)} &\geq \theta_H - \theta_L. \tag{41}
\end{aligned}$$

Note that the LHS is strictly decreasing in both Δ and p .

Evaluating (41) at $p = (\theta_L/\theta_H)^2$ and $\Delta = \underline{\Delta}$:

$$\begin{aligned}\theta_H \sqrt{1 - \left(\frac{\theta_L}{\theta_H}\right)^2} &\geq \theta_H - \theta_L \\ \sqrt{1 - \left(\frac{\theta_L}{\theta_H}\right)^2} &\geq 1 - \frac{\theta_L}{\theta_H} \\ 1 - \left(\frac{\theta_L}{\theta_H}\right)^2 &\geq 1 - 2\frac{\theta_L}{\theta_H} + \left(\frac{\theta_L}{\theta_H}\right)^2 \\ \frac{\theta_L}{\theta_H} &\geq \left(\frac{\theta_L}{\theta_H}\right)^2,\end{aligned}$$

which is true since $\theta_H > \theta_L$. Evaluating (41) at $\Delta = \underline{\Delta}$ and $p = 1$ yields $0 \geq \theta_H - \theta_L$, a contradiction. Thus when Δ is at its maximum point, there exists a value $\tilde{p} \in ([\theta_L/\theta_H]^2, 1)$ such that a cohesive IE exists if and only if $p \leq \tilde{p}$.

Where it exists, define $\tilde{p}(\Delta)$ as the value of p that equates the two sides of (41) for a given value of Δ . As the LHS of (41) is strictly decreasing in both Δ and p , $\tilde{p}(\Delta)$ is strictly decreasing in Δ .

At $\Delta = 1 + (1 - p) [\theta_L/(2\pi)]^2$, (41) is the same as (39). By continuity of the LHS of (41), if $\theta_L > (1/2)\theta_H$, there exists a value $\hat{\Delta} \in (1 + (1 - p) [\theta_L/(2\pi)]^2, \underline{\Delta})$ such that (i) for $\Delta \leq \hat{\Delta}$ a cohesive SRIE exists for all p and (ii) for $\Delta > \hat{\Delta}$ a cohesive SRIE exists if and only if $p \leq \tilde{p}(\Delta)$. If $\theta_L < (1/2)\theta_H$, for all $\Delta \in (1 + (1 - p) [\theta_L/(2\pi)]^2, \underline{\Delta}]$, a cohesive SRIE exists if and only if $p \leq \tilde{p}(\Delta)$.

Schismatic SRIE. The following conditions are necessary and sufficient. To rule out emergence of an inclusive club, there must not exist an $s^* \in [0, 1]$ such that (34) and (35) hold.

In addition, there are the participation constraints

$$\pi \left[1 + (1 - p) \left(\frac{\theta_L}{2\pi}\right)^2 \right] \geq \Delta\pi, \quad (42)$$

$$\pi \left[1 + p \left(\frac{\theta_H}{2\pi}\right)^2 \right] \geq \Delta\pi. \quad (43)$$

For $\Delta \leq \underline{\Delta}$ and $p < (\theta_L/\theta_H)^2$, we have established that a club can break up a schismatic state and form an inclusive club at $s^* = s_L$. For $\Delta > 1 + (1 - p) [\theta_L/(2\pi)]^2$, clearly (42) is violated. That leaves $\Delta \leq 1 + (1 - p) [\theta_L/(2\pi)]^2$ and $p \geq (\theta_L/\theta_H)^2$. For such Δ , (42) is satisfied. For such p , the LHS of (42) is less than the LHS of (43). Hence (43) is also satisfied. We established in case 2a above that either (34) or (35) are violated for all $s^* \in [0, 1]$ if and only if $p \geq \tilde{p}$, where \tilde{p} equates (39). Therefore, a schismatic IE exists for $\Delta \leq 1 + (1 - p) [\theta_L/(2\pi)]^2$ and $p \geq \tilde{p}$.

Exclusive SRIE. The conditions are the same as for a schismatic SRIE except that the weak inequality in (42) is reversed. Hence an exclusive SRIE exists for $\Delta \leq \underline{\Delta}$ wherever neither a cohesive or schismatic SRIE exists. \square

Proof of Proposition 5

Proof. Case 1: $\gamma = 0$.

We will show there is no radicalization in this case. First, we show that cohesive and zero participation SRIE are absorbing classes. We then show that p^t is strictly decreasing, tending toward zero, whenever the IE implements schism or exclusivity.

By Proposition A2, an inclusive club forms when $p^t \leq \tilde{p}$ and $\Delta \leq \underline{\Delta}$, setting strictness s^* . In this case, $p^{T+1} = s^*$.

We claim that $s^* \leq \tilde{p}$. In this case, $p^{T+2} = s^* \leq \tilde{p}$. By induction, $p^t = s^*$ for all $t > T$.

To establish the claim, note that z_H is an upper bound on s^* , is strictly increasing in p , and, at \tilde{p} , $z_L = z_H = s^*$ (see the proof of Proposition A2). Hence we need only show

$$\begin{aligned} z_H &\leq \tilde{p} \\ \left(\frac{\theta_H}{2\pi}[1 - \sqrt{1 - \tilde{p}}]\right)^2 &\leq \tilde{p} \\ \frac{\theta_H}{2\pi}[1 - \sqrt{1 - \tilde{p}}] &\leq \sqrt{\tilde{p}}, \end{aligned}$$

which is true because $\sqrt{p} \geq p \geq 1 - \sqrt{1 - p}$ for $p \in [0, 1]$ and $\frac{\theta_H}{2\pi} = s_H \leq 1$ by the upper bound on strictness. This establishes the claim. Hence a cohesive SRIE is absorbing.

Clearly, zero participation is also absorbing.

We now show that p^t is decreasing whenever there is an exclusive club with strictness \tilde{s}_H . Under exclusivity:

$$\begin{aligned} p^{t+1} &= \alpha \tilde{s}_H + (1 - \alpha)p^t \tilde{s}_H = [\alpha + (1 - \alpha)p^t] \tilde{s}_H \\ &= \underbrace{[\alpha + (1 - \alpha)p^t]}_{< 1} \left(\frac{\theta_H}{2\pi}\right)^2 p^t, \end{aligned} \tag{44}$$

which is less than p^t . Thus $p^{t+1} < p^t$ for all t under exclusivity, tending toward zero.

Now suppose that there is a schismatic SRIE, with strictness levels \tilde{s}_L and \tilde{s}_H , respectively. Then

$$p^{t+1} = p^t \tilde{s}_H + (1 - p^t) \tilde{s}_L. \tag{45}$$

Recall from Proposition A2 that if a schismatic SRIE exists, it must be that $\tilde{s}_H > \tilde{s}_L$, otherwise an H type could benefit by deviating and joining the all L club. Thus

$$p^{t+1} = p^t \tilde{s}_H + (1 - p^t) \tilde{s}_L < \tilde{s}_H = p^t \left(\frac{\theta_H}{2\pi}\right)^2.$$

As $\frac{\theta_H}{2\pi} < 1$,

$$p^{t+1} < p^t \left(\frac{\theta_H}{2\pi}\right)^2 < p^t.$$

Hence p^t is strictly decreasing under schism, tending toward zero.

Therefore, whenever the IE specifies schism or exclusivity, p^t is decreasing and tends toward zero. Whenever the IE specifies cohesion, p^t is constant and below \tilde{p} . Under zero participation, p^t falls to zero for all subsequent t . The result for $\gamma = 0$ follows.

Case 2: $\gamma > 0$.

Because $s^*(\delta^0) \leq \underline{s}(\delta^0)$, if a cohesive SRIE is in place at time t , $\delta^{t+1} = \delta^0 + \gamma \max\{p^{t+1} - \underline{s}(\delta^0), 0\} = \delta^0$. Hence the proof that a cohesive SRIE is absorbing from parts (i)-(ii) follows through with $\gamma > 0$. Clearly, zero participation is also absorbing.

Suppose an exclusive club forms at $t = 0$. This requires $p^0 \geq \tilde{p}$, which holds for θ_H sufficiently large. Then $p^1 = [\alpha + (1 - \alpha)p^0]\tilde{s}_H(p^0, \delta^0) = [\alpha + (1 - \alpha)p^0] \left(\frac{\theta_H(1+\delta^0)}{2w}\right)^2 p^0$. By assumption, $\frac{2w}{1+\delta^0} > \theta_H$, so $p^1 < p^0$. However, for α close to one, p^1 is approximately $\tilde{s}_H(p^0, \delta^0)$. This is greater than \tilde{p} for p^0 and θ_H sufficiently large, so an exclusive SRIE is implemented at $t = 1$. In addition, we know $\tilde{s}_H(p^0, \delta^0) > \underline{s}(\delta^0)$ for $p^0 > (\theta_L/\theta_H)^2$. Therefore, for α close to one, $\delta^1 \approx \delta^0 + \gamma[\tilde{s}_H(p^0, \delta^0) - \underline{s}(\delta^0)] > \delta^0$. Thus δ^1 can be made arbitrarily large for γ large. In particular, for γ sufficiently large, $\tilde{s}_H(p^1, \delta^1) = 1$.

In this case, $p^2 = [\alpha + (1 - \alpha)p^1] \geq p^1$ and $\delta^2 \geq \delta^1$. By induction, we have $p^t \geq \tilde{p}$ for all $t > 0$. That is, the exclusive club is stable. In addition, $\lim_{t \rightarrow \infty} p^t = 1$ for $\alpha > 0$.

Otherwise, if $\tilde{s}_H(p^1, \delta^1) < 1$, p^t is strictly decreasing until cohesion or zero participation emerges and locks in, by the argument above for $\gamma = 0$.

The same argument can be applied in case of schism at $t = 0$, which requires $p^0 \geq \tilde{p}$, to produce $\tilde{s}_H(p^1, \delta^1) = 1$. In this case, $p^2 = p^1 + (1 - p^1)\tilde{s}_L \geq p^1$ (see (45)). By induction, we have $p^t \geq \tilde{p}$ for all $t > 0$. That is, schism is stable. In addition, $\lim_{t \rightarrow \infty} p^t = 1$.

Otherwise, if $\tilde{s}_H(p^1, \delta^1) < 1$, p^t is strictly decreasing until cohesion emerges and locks in, by the argument above for $\gamma = 0$. \square

Proof of Proposition 6

Proof. Recall that each club k maximizes G_k^t at each time t . A monopolist can recreate any competitive SRIE on its own except the schismatic one. Hence $G^t \geq H^t$ whenever an inclusive or exclusive club is formed under competition.

The inequality is strict for an open set of parameters. For example, the SRIE is cohesive in both the monopoly and competitive cases when $p < (\theta_L/\theta_H)^2$ and $\Delta < 1$. In such a case, $\underline{s}(p, \delta) > s^*$.

Hence we only need to check that $G^t \geq H^t$ when schism occurs under competition.

Define $X_M(p^t, \delta^t)$ as total participation under monopoly and $X_C(p^t, \delta^t)$ as total participation under competition in state (p^t, δ^t) . We claim that $X_M(p^t, \delta^t) \geq X_C(p^t, \delta^t)$ whenever the competitive SRIE is schismatic. In addition, this implies that for each such state (p^t, δ^t) , p^{t+1} is at least as high under monopoly. Taken together, this would establish the proposition.

Let us now prove the claim.

Case 1: In state (p, δ) , the SRIE under monopoly is exclusive at strictness $\bar{s}(p)$, whereas schism occurs under competition with strictness levels \tilde{s}_L and \tilde{s}_H .

Total participation is no less under monopoly if

$$\begin{aligned} X_C(p, \delta) &\leq X_M(p, \delta) \\ (1-p)\tilde{s}_L + p\tilde{s}_H &\leq p\bar{s}(p). \end{aligned}$$

Recall from Proposition A2 that schism occurs only when $\Delta \in (\theta_L/\pi, \tilde{\Delta}]$, where $\tilde{\Delta} \equiv 1 + (1-p)(\theta_L/2\pi)^2$. Hence the RHS is minimized at $\Delta = \tilde{\Delta}$, in which case $\bar{s}(p) = (\sqrt{\tilde{s}_H} + \sqrt{\tilde{s}_H - \tilde{s}_L})^2$ (see (6)). Thus it is sufficient to verify

$$\begin{aligned} (1-p)\tilde{s}_L + p\tilde{s}_H &\leq p \left(\sqrt{\tilde{s}_H} + \sqrt{\tilde{s}_H - \tilde{s}_L} \right)^2 \\ (1-p)\tilde{s}_L + p\tilde{s}_H &\leq p \left(\tilde{s}_H + \tilde{s}_H - \tilde{s}_L + 2\sqrt{\tilde{s}_H(\tilde{s}_H - \tilde{s}_L)} \right) \\ \tilde{s}_L &\leq p \left(\tilde{s}_H + 2\sqrt{\tilde{s}_H(\tilde{s}_H - \tilde{s}_L)} \right). \end{aligned} \tag{46}$$

Multiplying both sides by $[p/(1-p)](1/\tilde{s}_H)$ yields

$$\left(\frac{\theta_L}{\theta_H} \right)^2 \leq \frac{p^2}{1-p} + \frac{2p}{1-p} \sqrt{\frac{\tilde{s}_H - \tilde{s}_L}{\tilde{s}_H}}.$$

Where schism occurs, $\tilde{s}_H > \tilde{s}_L$. Hence the second term on the right-hand side is nonnegative. Therefore it suffices that

$$\left(\frac{\theta_L}{\theta_H} \right)^2 \leq \frac{p^2}{1-p}. \tag{47}$$

We also know from Proposition A2 that $p > \tilde{p}$ where schism occurs. The right-hand side of (47) is strictly increasing in p . Therefore, if the inequality holds at \tilde{p} , it holds for all $p > \tilde{p}$. Given that schism occurs, $2\theta_L < \theta_H$, so by the proof of Proposition A2, \tilde{p} is defined implicitly by

$$\frac{\theta_L}{\theta_H} = \frac{1 - \sqrt{1 - \tilde{p}}}{1 + \sqrt{\tilde{p}}}. \tag{48}$$

Substituting (48) into (47), it remains to show that

$$\begin{aligned} \left(\frac{1 - \sqrt{1 - \tilde{p}}}{1 + \sqrt{\tilde{p}}} \right)^2 &\leq \frac{\tilde{p}^2}{1 - \tilde{p}} \\ \frac{1 - \sqrt{1 - \tilde{p}}}{1 + \sqrt{\tilde{p}}} &\leq \frac{\tilde{p}}{\sqrt{1 - \tilde{p}}} \\ \sqrt{1 - \tilde{p}} - (1 - \tilde{p}) &\leq \tilde{p} + \tilde{p}^{\frac{3}{2}} \\ \sqrt{1 - \tilde{p}} &\leq 1 + \tilde{p}^{\frac{3}{2}}, \end{aligned}$$

which is true for all $\tilde{p} \in [0, 1]$.

Case 2: In state (p, δ) , the SRIE is cohesive under monopoly and schismatic under competition. Therefore, total participation is no less under monopoly if

$$\underline{s} \geq (1 - p^t)\tilde{s}_L + p^t\tilde{s}_H.$$

Because schism occurs, we know $p > \tilde{p}$. In *case 1*, we showed for $p > \tilde{p}$ that

$$p\bar{s}(p) \geq (1 - p)\tilde{s}_L + p\tilde{s}_H.$$

As an inclusive club forms under monopoly, we know $\underline{s} \geq p\bar{s}(p)$. Hence

$$\underline{s} \geq p\bar{s}(p) \geq (1 - p)\tilde{s}_L + p\tilde{s}_H.$$

This establishes the claim and the proposition. □

Online Appendix

Discrimination & Participation under Competition

Competition alters some, but not all, of the effects of discrimination on participation. To see the effects of competition most clearly, consider the case of p large and $\theta_L/\theta_H > 1/2$, so that a schismatic SRIE exists. The following effects are derived using Proposition A2 and Figure 7.

The effect of blanket discrimination δ is illustrated by Figure A1. As δ rises and the opportunity cost of participation falls, H types are *willing to* participate more intensively. That is, the welfare-maximizing strictness for H types, \tilde{s}_H , rises. Eventually, the opportunity cost becomes low enough to make it worthwhile for L types to form their own exclusive club at lower strictness \tilde{s}_L . The formation of an exclusive L type club at δ'' appears in panel (b). Thenceforth, the strictness levels of both exclusive clubs rise with δ . As under monopoly, blanket discrimination raises participation.

The effect of stigma Δ is illustrated by Figure A2. Consider the extensive margin. For Δ small, a schismatic equilibrium arises. At $\tilde{\Delta}$, zero participation becomes sufficiently attractive for L types and the exclusive L type club dissolves. Note that $\tilde{\Delta} < \underline{\Delta}$, so this requires a lower level of stigma than under monopoly. At some $\Delta \in (\tilde{\Delta}, \overline{\Delta})$, even H types choose not to participate. Again this requires, a lower level of conditional discrimination than under monopoly. Hence conditional discrimination promotes social integration at the extensive margin even more effectively than under monopoly. The same is not true of the intensive margin. Under monopoly, identity-based organizations pushed members to the point of indifference between joining and not joining. This point was determined by Δ . In contrast, competition aligns strictness choices with members' preferences, which take into account the cost of participation, so the participation constraint almost nowhere binds. Hence stigma has no effect on participation at the intensive margin under competition.

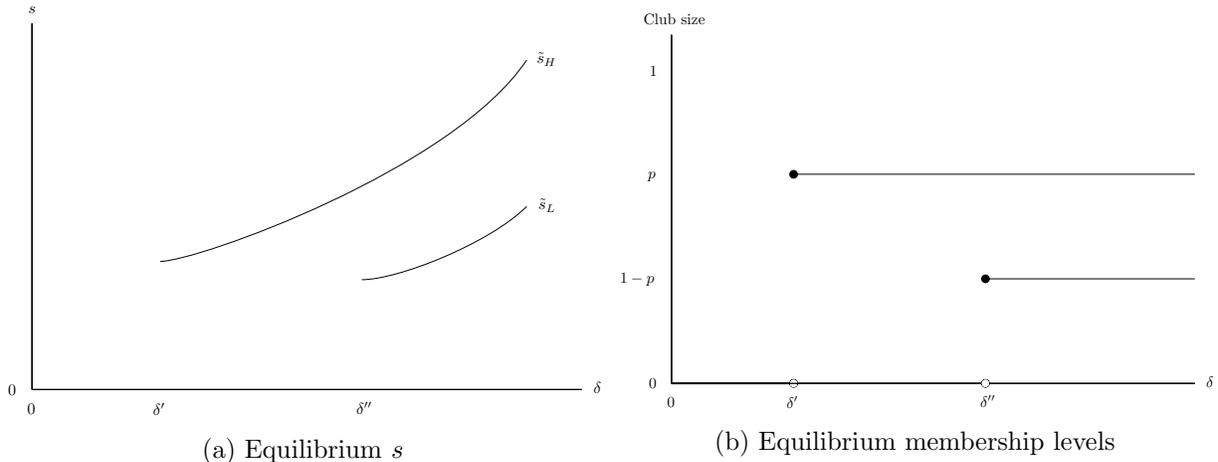


Figure A1: The effect of blanket discrimination under competition. *Case: p large, $\frac{\theta_L}{\theta_H} < \frac{1}{2}$*

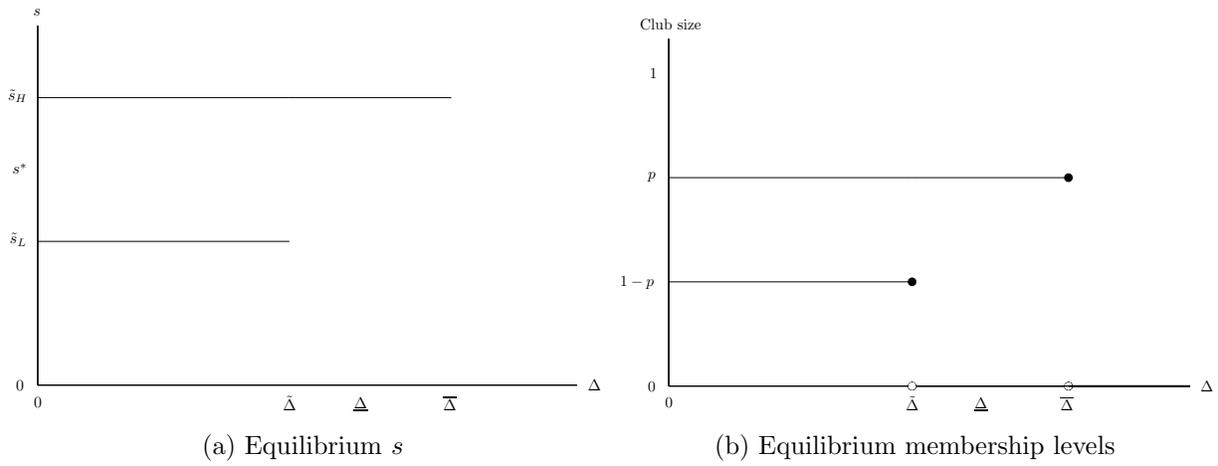


Figure A2: The effect of conditional discrimination under competition. *Case:* p large, $\frac{\theta_L}{\theta_H} < \frac{1}{2}$