On Symbols and Cooperation*

Tom Potoms[†]· Tom Truyts[‡]

Abstract

How are group symbols (e.g., a flag, a Muslim veil, a clothing style) helpful in sustaining cooperation and social norms? We study the role of symbols in an infinitely repeated public goods game with random matching, endogenous partnership termination, limited information flows and endogenous symbol choice. We characterize an efficient segregating equilibrium, in which players only cooperate with others bearing the same symbol. In this equilibrium, players bearing a scarcer symbol face a longer expected search time to find a cooperative partner upon partnership termination, and this sacrifice of outside options allows them to sustain higher levels of cooperation. We compare this equilibrium to other equilibria in terms of renegotiation proofness, and we discuss the relation this has to the evolution of intolerance.

JEL: C73, D83

*We are grateful to Bram De Rock, Heiner Schumacher, Ilona Sergant, Gilles Grandjean, Marco Mantovani, as well as seminar and conference participants in Brussels, Louvain-la-Neuve, Stony Brook, Paris, Oxford, Utrecht, Luxemburg, St.-Petersburg, Leuven, Namur and Maastricht.

†ECARES, Université Libre de Bruxelles, 42 Avenue Roosevelt, 1050 Brussels, Belgium. E-mail: tom.potoms@ulb.ac.be. Tom Potoms is a doctoral research fellow of the F.R.S.-FNRS.

‡CEREC, Saint-Louis University - Brussels, Boulevard du Jardin botanique 43, 1000 Brussels, Belgium; Center for Economic Studies, University of Leuven; CORE, Université Catholique de Louvain. E-mail: tom.truyts@usaintlouis.be, tel.: +3222117845. Tom Truyts gatefully aknowledges financial support from the Belgian French speaking community ARC project n°15/20-072, "Social and Economic Network Formation under Limited Farsightedness: Theory and Applications", Université Saint-Louis - Bruxelles, October 2015 - September 2020.

> "I didn't understand what was happening at first. People started talking to me more. (...) And I was made to feel like I was actually from this planet. Maybe I was finally fitting in? But then it became fishy. The Muslim taxi drivers who would almost always say "Assalamu Alaikum," ask me where I'm from or if I'm single, or not allow me to pay for the fare became cold and dry. (...) Have I become unfriendly? Arrogant? But other people had become even nicer. (...) Then it hit me. My knit hat and winter scarf covered my hijab (the head scarf part) entirely and all that was visible were my eyes behind my wannabe hipster glasses (...). They didn't even 'know' I was Muslim."
>
> (leenamielus, "I Took Off My Hijab", http://leenamielus.blogspot.be)

1 Introduction

Migration and globalization have given rise to a surge of nationalism, xenophobia and radicalism all over the globe. As such, cultural diversity and societal polarization are high on the policy agenda in most countries. The social tensions and debates often focus on various cultural markers or group symbols, as illustrated by the ban of the Muslim veil in several European countries. But the exact importance and functioning of these group symbols remains more obscure. These group symbols, e.g., a flag, a Muslim veil, a T-shirt of a rock band or an expensive corporate style suit, function in various ways as coordination devices. On one hand, they reveal information about underlying heterogeneity. Strangers can form a reasonably accurate idea about one's socioeconomic background and tastes from a casual observation of one's clothing and lifestyle. On the other hand, symbols strengthen group identification and loyalty. By displaying the above symbols, one is met with initial sympathy from some strangers, and with aversion from others. Tajfel and Turner's (1979) famous 'minimal group experiment' shows that symbols can give rise to a differential sympathy or hostility towards strangers, even if these symbols are understood to reflect no underlying heterogeneity.[1] While these findings gave rise to an extensive body of literature and numerous replications, the underlying mechanisms are relatively poorly understood.

Iannaccone (1992) presents an interesting interpretation of the role of symbols in the context of cults and sects. He understands these symbols

[1] Tajfel and Turner (1979) randomly allocated a minimal marker to test subjects, e.g. a colored dot on the forehead, and found that such a minimal symbol sufficed to cause participants to discriminate between complete strangers on the basis of their symbol, even in the absence of any underlying heterogeneity. Test subjects bestowed advantages on strangers bearing the same symbol, at the expense of others bearing a different symbol.

as a solution for a typical group problem: the underprovision of a club good. The standard solution in club theory is to levy membership fees, and then use the revenues to subsidize contributions to the club good (see e.g., Sandler and Tschirhart (1997)). But if such a formal scheme is unfeasible or undesirable, Iannaccone (1992) suggests, then a cult or sect can 'tax' resources spent outside the group. If members contribute time to the club good, such a 'tax' takes the form of restrictions on members' clothing, diet, haircut or language, all of which impede social interactions with non-members. By sacrificing their capacity for social interactions outside the group, these typical idiosyncrasies of religious, political or subcultural groups help members to commit their resources to the group.

The sacrifice of outside options in order to demonstrate commitment and sustain cooperation and group norms is documented by social scientists in a variety of contexts. Gambetta (2009, p.41) discusses how prisoners demonstrate their commitment to a life in crime by applying prison tattoos on visible body parts, thus ruining their chances of an honest life. Gambetta (2009, p.19) equally describes how candidate members of Colombian youth gangs are required to kill a friend or family member. Besides proving one's ability to murder, it also shatters gang members' fall back option for leaving the gang. Berman (2000) documents these sacrifice mechanisms for the case of ultra-orthodox Jews. Berndt (2007) shows how being a member of a distinct and despised ethnic or religious minority, and the implied lack of outside options, allowed e.g., 19th century Jewish peddlers to act as middlemen in high stake financial transactions. Shimizu (2011) models self-sacrifice in military and terrorist groups as a result of giving up individual autonomy. Aimone et al. (2013) find that the possibility of sacrificing private outside options enhances club good contributions in a Voluntary Contribution Mechanism experiment.

Yet, the ability of symbols to discipline group members' behavior crucially depends on their damaging effect on members' outside options, i.e., on non-members' reactions to these symbols. The literature following Iannaccone (1992) typically assumes a negative reaction to the group symbols as exogenously given. While such a reaction is inherent in some cases, such as killing family members, it is much less obvious for more arbitrary and minimal symbols, such as clothing or hair color. A first contribution of the present paper is that we derive symbol choices, the reactions to symbols and the resulting cooperation levels, all jointly from a notion of equilibrium.

We study the role of symbols in the context of an infinitely repeated public goods game, with random matching, endogenous partnership termination and limited information flows. We consider an infinite population of homogeneous players, who differ *ex ante* only in a visible but payoff-irrelevant symbol (e.g., a colored hat). Players begin each round with one partner, with whom they play a stage game consisting of two phases. First, they play a public goods game. Second, upon observing the public goods game's outcome, both players simultaneously decide whether to terminate the partnership or not, and whether to change their symbol at a certain cost. Partnerships break up if at least one partner wishes to terminate, and are otherwise terminated exogenously with a small probability. Furthermore, players whose partnership was terminated are then randomly rematched. Starting a new partnership, players have no information about their partner's past play, but only observe his symbol.

We characterize an efficient segregating equilibrium of this game, in which players exert no effort in the public goods game if their partner bears a different symbol. Consequently, such heterogeneous partnerships are a waste of time, and are immediately terminated by both players. In partnerships which are homogeneous in terms of symbols, players exert the maximal incentive compatible effort. Failure to comply with the equilibrium effort in a homogeneous partnership is punished with partnership termination, thus implying in expectation a certain search time to find a new identical symbol partner to start cooperating with. Therefore, players bearing a more scarce symbol face in expectation a longer search for a cooperative partner after a break-up, and this sacrifice of outside options allows them - in the spirit of Iannaccone (1992) - to sustain higher cooperation levels.

This paper relates to a large body of literature on cooperation in infinitely repeated public goods or prisoner's dilemma games. The central question in this literature is how to constrain the continuation payoffs of defectors in order to sustain cooperation on the equilibrium path, despite of defection being the stage game's dominant strategy. However, the present setting excludes a large number of well-known mechanisms that sustain cooperation. First, endogenous partnership termination and random rematching excludes the entire class of personal enforcement mechanisms, in which cheating triggers a punishment by the victim. Because defectors can terminate a partnership before undergoing their punishment, the usual folk theorems and trigger strategy results do no apply. Second, the absence of information about a partner's past play in previous partnerships excludes community enforcement mechanisms, in which shirkers are identified and punished by other members of the population.[2] Third, even though the contagion mechanisms of Kandori (1992) and Ellison (1994) can sustain cooperation if players are randomly rematched and only aware of their own history of play, they are excluded in this setting by the continuum population.[3]

The literature has nevertheless advanced two mechanisms to sustain cooperation in the present restrictive setting. A first solution relies on gradual trust-building or 'incubation'.[4] In these equilibria, partners only engage in full cooperation after a sufficiently long trust-building or 'incubation' phase, i.e., a number of rounds in which they defect or exert low effort. The prospect of a trust-building stage with a new partner suffices to deter players from cheating in the later stages of a partnership.[5] A second solution relies on the presence of exogenous defectors in the population, which gives the situation of having a cooperative partner sufficient scarcity value to discourage defection.[6] Ghosh and Ray (1996) show how cooperation in a public goods game is sustainable if the defectors' population share is neither too small nor too large. Adverse selection, due to the defectors always having to draw a new partner while patient cooperators lock themselves into long term partnerships, means that a small population share of defectors can suffice to sustain cooperation among patient players. Moreover, Ghosh and Ray (1996) demonstrate how their equilibrium satisfies 'bilateral rationality', i.e., it is robust against bilateral renegotiation by current partners.

The present paper also contributes to this literature in the sense that we study a setting similar to Ghosh and Ray's (1996) repeated public goods game, but in which the role of the exogenous defectors is played in

[2] Sustaining cooperation through punishments by other community members has been shown effective under various information assumptions by e.g. Greif (1993), Okuno-Fujiwara and Postlewaite (1995), Mailath and Morris (2000) or Takahaski (2010).

[3] In these equilibria, players defect in all their future partnerships if their partner cheats. If players are sufficiently patient, they are dissuaded from defecting by the foresight of eventually triggering the entire population to defect forever. In the present setting, a defection eventually infects at most countably many out of uncountably many players into defecting.

[4] See e.g. Datta (1996), Kranton (1996), Eeckhout (2006), Fujiwara-Greve and Okuno-Fujiwara (2009) or Fujiwara-Greve et al. (2012). This approach also relates to the idea of 'starting small' in Watson (1999, 2002), where the stakes of the game gradually increase with the partnership's age.

[5] Deb and González-Díaz (2014) show how these results depend on the structure of the prisoner's dilemma stage game. Slight modifications in the payoff structure break down the standard results and imply a need for more sophisticated trust-building techniques.

[6] Related mechnisms are also studied by e.g. Fujiwara-Greve and Okuno-Fujiwara (2009) and Schumacher (2013).

equilibrium by endogenous group symbols. Hence, we assume no preference heterogeneity, but rather derive that players act in equilibrium much like defectors towards others bearing a different symbol. In this equilibrium, players bearing different symbols face generically different incentives. In the spirit of Iannaccone (1992), players can thus sacrifice their outside options by bearing a more scarce symbol, and this sacrifice allows them to sustain higher cooperation levels. Finally, we also characterize the segregating equilibrium in terms of a parametrized criterion of renegotiation proofness, named ε-renegotiation proofness, which eliminates the gradual trust-building equilibria and allows us to differentiate between the game's different equilibria.

The importance of payoff irrelevant group symbols for cooperation is also central in Eeckhout (2006) and Choy (2014).[7] Eeckhout (2006) studies a public correlation device such as skin color in an infinitely repeated prisoner's dilemma with endogenous partnership termination and limited information. Eeckhout compares a standard ('color-blind') incubation equilibrium to a 'segregation equilibrium', in which new partners of the same color start cooperating immediately, while other new partners play an incubation strategy. Eeckhout shows that color distributions exist for which the segregation equilibrium Pareto dominates the color-blind equilibrium.

Choy (2014) is the closest to our paper. He studies how segregation on the basis of visible group affiliations helps to sustain cooperation in an infinitely repeated public goods game. Choy assumes that players also know the group affiliation of their partners' previous partners and that groups are hierarchically ranked. He characterizes a renegotiation proof segregating equilibrium, in which players refuse to interact with members of lower groups to protect their reputation. Preserving this reputation implies higher search costs upon partnership termination, which in turn helps to sustain more cooperation. In contrast with Choy (2014), we assume no information about a partner's past play, and unlike Eeckhout (2006) and Choy (2014), we consider symbols a choice variable.

The remainder of this paper is organized as follows. The formal setting and equilibrium concept are introduced in Section 2. Section 3 discusses how symbols are helpful in sustaining cooperation by characterizing the efficient segregating equilibrium. In Section 4, we characterize the efficient segregating and other equilibria in terms of a parametrized version of renegotiation proofness, and we discuss the relationship with

[7] See also Peski and Szentes (2013) on how payoff irrelevant symbols can lead to discriminatory behavior.

various forms of chauvinistic preferences. The final Section concludes. All proofs and derivations are detailed in the Appendix.

2 Formal Setting

Assume a continuum of players. Time is indexed by $t \in \mathbb{N}$, and all players have the same discount factor $0 \leq \delta < 1$. Each player wears one publicly visible symbol out of a given set of symbols $S = \{s^i\}_{i=1,\ldots,n}$. Players start each round of the game with one partner. We call a partnership between two players homogeneous if both bear the same symbol, and we otherwise call it heterogeneous.

In each round of the game, players first play a public goods game with their current partner. After observing the public goods game's outcome and inferring their partner's contribution, players decide on symbol change and partnership termination. The symbol switching cost of a player who begins a round t with symbol s^i and ends it with symbol s^j is denoted $c_t(i,j)$, with $c_t(.) \geq 0$ and $c_t(i,j) = 0$ if $i = j$. For simplicity, we assume that these switching costs are independent of t. Partnerships are exogenously terminated with probability $\lambda \in (0,1]$. Otherwise, both players choose whether or not to continue the partnership, $l \in \{0,1\}$, where $l = 1$ means continuing the partnership. A partnership ends if at least one of the partners wishes to terminate it. If their partnership is terminated, players randomly draw a new partner from the set of players whose partnership was terminated with uniform probability. Of course, the assumption of exogenous partnership termination ensures that drawing a new partner is uninformative about past behavior on the equilibrium path.[8] When meeting a new partner, players do not observe his past behavior but only see his symbol.

We first formally characterize the public goods game. In this game, both partners choose a level of efforts $e \in \mathbb{R}_+$. A player who contributes e while his partner contributes e' obtains a stage payoff $\pi(e, e')$. If π_k and π_{kh} denote the partial derivative of π w.r.t. argument k and w.r.t. arguments k and h, respectively, then the following restrictions on the public goods game technology are imposed.

Condition 1 *Let π be a twice continuously differentiable function such that:*

1. *(Public goods game) $\pi_1(e,.) < 0$ for all $e > 0$, $\pi_2(.) > 0$ and $\pi_1(e,e) + \pi_2(e,e) > 0$ for all $e \geq 0$,*

[8] On the equilibrium path, players never terminate a homogeneous partnership such that, in equilibrium, a new partner's previous homogeneous partnership was exogenously terminated with probability 1.

2. *(Boundedness)* $\pi(e,e)$ *is bounded for all* $e \in \mathbb{R}_+$*, and* $\exists \gamma > 0$ *such that* $\pi_2 > \gamma$*,*

3. *(Initial condition) let* $\pi(0,0) = 0$*, while* $\pi_1(0,.) = 0$*, and* $\pi_{11}(0,0) < 0$.

The first part of condition 1 ensures that π represents a public goods game: the payoffs are decreasing with own effort, such that zero effort is a dominant strategy in the stage game, and increasing with the partner's effort. Moreover, coordinating symmetrically on a higher effort level is always mutually beneficial. The next two parts of condition 1 impose some regularity conditions to ensure that the players' problem and behavior is always well defined. Part 2 of condition 1 bounds the benefits of symmetric efforts, implying that $\lim_{e\to\infty} \pi_1(e,e) + \pi_2(e,e) = 0$, and bounds marginal benefits of the partner's efforts away from zero. Part 3 of condition 1 normalizes π to be zero in the absence of any contribution and ensures that our problem is well defined near zero. The following simple example shows a public goods game technology which satisfies the above condition and will serve as a closed form example in the remainder of this text.

Example 1 *The payoff function*

$$\pi(e,e') = 1 + e' - e - \frac{1}{1+e}$$

satisfies condition 1, as $\pi_1(e,.) = -1 + \frac{1}{(1+e)^2} < 0$ *for* $e > 0$*,* $\pi_2(.) = 1 > 0$ *and* $\pi_1(e,e) + \pi_2(e,e) = \frac{1}{(1+e)^2} > 0$ *for all* $e \in \mathbb{R}_+$*. Moreover,* $\pi(e,e) = \frac{e}{1+e}$ *is bounded from above by 1,* $\pi(0,0) = 0$ *and* $\pi_1(0,.) = 0$*.*

The information of a player at the beginning of round t, denoted $h_t \in H_t$, consists of the fundamentals of the game, the symbol of his current partner, and their history of play in the current partnership.[9] Let $H \equiv \cup_t H_t$ be the set of all possible information sets.

[9]Hence, these information assumptions keep players from conditioning their behavior on what happened in their previous partnerships. On one hand, note that players run into one of their past partners again with zero probability, because even for $t \to \infty$, players have met at most countably many out of uncountably many players. On the other hand, our equilibrium concept excludes players from conditioning their behavior on private histories of play (cfr. infra).

A pure strategy $\sigma : H \to \mathbb{R}_+ \times \{0,1\}^{\mathbb{R}_+} \times S^{\mathbb{R}_+}$ specifies a triplet (e_t, l_t, s_t) for all t and possible information sets h_t.[10] The first element specifies how much effort to exert given the symbol of the current partner and the history of play. The second element specifies, for each h_t, a termination decision for all possible effort levels of the partner e'_t:

$$l_t : \mathbb{R}_+ \to \{0,1\}.$$

Similarly, the last element specifies a symbol switching decision as a function of the partner's effort e'_t. We say that a strategy is public if it is conditioned only on the partners' public information.[11]

Players evaluate a strategy by considering the expected future payoff streams to which a strategy is expected to give rise; i.e., they wish to maximize

$$u(\sigma) \equiv E \sum_t \delta^t \left(\pi\left(e_t, e'_t\right) - c_t(i,j)\right),$$

in which the expectation operator E indicates the expectations over all possible future histories of play and symbols of partners to which a strategy σ may lead, given the strategies of other players as well as the stochastic processes of partnership termination and formation. We study the stationary perfect public equilibria (PPE) of this game, i.e., profiles of public strategies which yield for all t and all h_t a Nash equilibrium for round t and all consecutive rounds.

3 Symbols and cooperation

This Section characterizes one particular stationary PPE in pure strategies, designated the efficient segregating PPE, which we will focus on in this paper. This PPE is in the following sense efficient and segregating in terms of symbols.

Definition 1 (Efficient segregating PPE) *The efficient segregating PPE is a stationary PPE in pure strategies in which players*

[10] As usual, Y^X represents the set of all mappings from X to Y. Note that this formulation is equivalent with players making termination and symbol switching decisions at the second phase of round t as a function of an intermediate history of play, which comprises h_t and the effort strategies in round t's public goods game. A partner's effort choice constitutes the only new information at this intermediate stage of round t.

[11] Note that allowing players to condition their choices on their private histories of play, i.e., their play in previous partnerships, would introduce asymmetric information into the game. This would complicate the analysis considerably, as it would require us to introduce beliefs about a partner's private information as well as to have strategies equally depending on these beliefs.

1. *exert effort 0 in heterogeneous partnerships and, for all i, the maximal effort level which is incentive compatible, denoted $\bar{e}^i$, in homogeneous s^i partnerships,*

2. *never switch symbol,*

3. *never terminate a homogeneous partnership on the equilibrium path but always terminate a heterogeneous partnership,*

4. *terminate a partnership after any deviation from equilibrium play, including deviations from equilibrium reactions to deviations, etc.*

Hence, segregating indicates that players exert only strictly positive effort in a homogeneous partnership.[12] As a result, a heterogeneous partnership is a waste of time and is thus terminated immediately. Although players never switch symbols in equilibrium, the ability to switch symbols is relevant, because it allows us to characterize how symbol switching costs determine the equilibrium symbol frequencies, and ultimately which patterns of cooperation can be sustained in equilibrium. The equilibrium is efficient in the sense that players exert the maximal level of effort which can be sustained in a homogeneous partnership. Note that this form of efficiency is stronger than Pareto efficiency, because a strategy profile in which two partners respectively exert efforts on and strictly below the incentive compatibility constraint is Pareto efficient, but not effecient in the above sense. Since two players bearing the same symbol face the same incentives, the efficient segregating PPE is symmetric in the sense that two players in a homogeneous partnership exert the same effort. In the remainder of this Section, we proceed with steps to characterize the efficient segregating PPE.

Let p^i denote the stationary share symbol s^i players in the set of all players who draw a new partner (i.e., the stationary probability of a randomly drawn partner having symbol s^i). In the efficient segregating PPE, the expected continuation value of an s^i player in a homogeneous partnership is recursively defined as:

$$v^i\left(\bar{e}^i\right) = \pi\left(\bar{e}^i, \bar{e}^i\right) + \delta\left(1-\lambda\right)v^i\left(\bar{e}^i\right) + \delta\lambda w^i\left(\bar{e}^i\right), \tag{1}$$

in which $w^i\left(\bar{e}^i\right)$ is the expected continuation value of an s^i player starting the round with a randomly drawn partner:

$$w^i\left(\bar{e}^i\right) = p^i v^i\left(\bar{e}^i\right) + \left(1-p^i\right)\delta w^i\left(\bar{e}^i\right). \tag{2}$$

[12] Of course, equivalent equilibria can be conceived in which players treat a subset of symbols as if it were the same symbol. In this case, one can easily understand the elements of S as partitions of a larger set of symbols.

Hence, in equilibrium both players get a stage payoff $\pi(\bar{e}^i, \bar{e}^i)$ in a homogeneous s^i partnership, after which their partnership survives to the next round with probability $1-\lambda$, and is terminated otherwise. In case of termination, the s^i players immediately draw a new s^i partner with probability p^i, in which case they start cooperating immediately and thus return to continuation value $v^i(\bar{e}^i)$. Otherwise they get stage payoff zero, terminate the partnership, and start the next round again with a randomly drawn partner, having continuation value $w^i(\bar{e}^i)$.

Incentive compatibility requires that an s^i player is not worse off when providing the equilibrium effort level $\bar{e}^i$, rather than defecting on his partner and starting anew with a new partner in the next period, i.e.,

$$v^i\left(\bar{e}^i\right) \geq \pi\left(0, \bar{e}^i\right) + \delta w^i\left(\bar{e}^i\right), \tag{3}$$

while efficiency implies that (3) must be satisfied with equality. Solving (1) and (2) for v^i and w^i and substituting into (3), we define

$$\begin{aligned} d\left(e, p^i\right) &\equiv v^i(e) - \pi(0, e) - \delta w^i(e) \\ &= \frac{\pi(e, e)}{1 - \delta(1-\lambda)(1-p^i)} - \pi(0, e), \end{aligned}$$

such that (3) can be written as $d(\bar{e}^i, p^i) \geq 0$. Note that

$$d(e, 0) = \frac{\pi(e, e)}{1 - \delta(1-\lambda)} - \pi(0, e)$$

is the difference between the expected actual value of the current partnership, when cooperating at effort level e, and the one shot payoff of cheating. If players have no hope of finding a new cooperative partner after break-up, e.g., because $p^i = 0$, then incentive compatibility requires $d(e, 0) \geq 0$. However, note that d decreases with p^i. Indeed, the possibility of a new partnership with another cooperative s^i player decreases the punishment that breaking up the present partnership constitutes, and it thus reduces the effort levels players can commit to. As a result, we obtain the following effort levels in the efficient segregating PPE.

Proposition 1 *In the efficient segregating PPE, the equilibrium effort in homogeneous s^i partnerships, $\bar{e}^i$, uniquely solves*

$$\bar{e}^i = \max\left\{e \middle| d\left(e, p^i\right) = 0\right\}. \tag{4}$$

Moreover, $\bar{e}^i$ is a left-continuous and strictly decreasing function of p^i and λ, and a right-continuous and strictly increasing function of δ.

Hence, players with more scarce symbols face a worse outside option, and this allows them to sustain higher effort levels. Moreover, effort levels are increasing with the 'effective' discount factor $\delta(1-\lambda)$, i.e., if the value of a current homogeneous partnership increases because players become more patient or because the expected longevity of their present partnership increases.

Example 2 *In the closed form example, $\frac{\pi(e,e)}{\pi(0,e)} = \frac{1}{1+e}$ strictly decreases with e, such that $\bar{e}^i$ is a continuous function of p^i. The equilibrium effort levels are*

$$\bar{e}^i = \frac{(1-p^i)\,\delta\,(1-\lambda)}{1-(1-p^i)\,\delta\,(1-\lambda)}.$$

We can now characterize the relationship between p^i and the share of the population bearing symbol s^i. Let x_t^i denote the proportion of s^i players that start round t with a new partner, and note that x_t^i follows a simple Markov dynamic:

$$x_{t+1}^i = \left(1 - p^i\,(1-\lambda)\right) x_t^i + \lambda\left(1 - x_t^i\right). \tag{5}$$

Symbol s^i players currently in a homogeneous partnership only have to draw a new partner if their partnership was exogenously terminated (with probability λ), and a fraction $1-(1-\lambda)p^i$ of s^i players who play the present round with a new partner will have to do so as well in the next round. In the stationary equilibrium, it must be that $x_t^i = x_{t+1}^i = x^i$, such that we obtain the following result.

Proposition 2 *For all i, a share $x^i = \frac{\lambda}{p^i+(1-p^i)\lambda}$ of all s^i players draw a new partner in the efficient segregating PPE. Moreover, p^i strictly exceeds the population share of s^i players iff $p^i < \sum_j (p^j)^2$.*

Hence, in equilibrium, s^i players spend on average a fraction x^i of their time meeting noncooperative partners, and this share of their time going to waste increases with the probability of break-up λ and decreases with the probability of randomly drawing another s^i player. This longer expected waiting time is precisely the sacrifice that allows players bearing a scarce symbol to sustain higher levels of cooperation. By Bayes' rule, the share of the population bearing symbol s^i, finally, is then $\frac{p^i}{x^i}\left(\sum_j \frac{p^j}{x^j}\right)^{-1}$, and this share exceeds p^i if $p^i < \sum_j (p^j)^2$. In this case, s^i players are more likely to draw another s^i partner from the pool of players searching for a new partner than they would be when drawing from the population at large. As such, they are overrepresented in the set of players drawing a new partner (i.e., p^i overstates their overall population share). Note that this selection effect differs from the adverse selection effect in Ghosh and Ray (1996), where the myopic types are overrepresented in the set of players drawing a new partner, because patient players lock themselves into long-term cooperative partnerships, unlike myopic players. In the efficient segregating PPE, the selection effect softens any asymmetries in symbol frequencies. Because they spend more time looking for a same symbol partner, players bearing a scarce symbol have more chance of finding a same symbol partner in the pool of players looking for a new partner than in the population at large. In the next Section, we show that other symbol-dependent PPE can exhibit the opposite selection effect. Finally, note also that most of the results in the above Propositions remain valid if players are disproportionately more likely to draw a same symbol partner, e.g., because real world interactions display a certain degree of homophily. In this case, p^i would be higher than if players were to draw a partner with uniform probability from the set of players searching for new partners. The share of time wasted looking for a same symbol partner after a break-up, x^i, would then be lower, and consequently the effort levels that can be sustained, $\bar{e}^i$, would decrease as well. As such, it is only the relationship between p^i and the population share of s^i players that is different if we allow for this kind of homophily in the random matching process.

Players bearing scarce symbols can commit to higher efforts due to the longer expected search time, but only if they can be trusted not to switch to another symbol after a break-up. As such, the matrix of symbol switching costs $(c_t(i,j))_{i,j=1,\ldots,n}$ determines which vectors of symbol frequencies p are compatible with an efficient segregating PPE, or, equivalently, what kind of symbol switching costs are required to sustain a certain level of cooperation in equilibrium. Players can unilaterally switch symbols in the following cases: when in a homogeneous partnership and exerting equilibrium efforts, when in a homogeneous partnership after deviating from the equilibrium effort, and before drawing a new partner (i.e., when currently in a heterogeneous partnership). The two latter options give exactly the same continuation values and are therefore payoff equivalent. Furthermore the first option is dominated by these two latter options: a player in a homogeneous s^i partnership who exerts $\bar{e}^i$ and then switches symbol without breaking up will face a partner exerting zero effort and breaking up the partnership in the next round.[13] Therefore

[13] Given that equilibrium efforts are required to be subgame perfect, this implies that a deviating reaction to symbol switching should correspond to a partner exerting zero efforts and terminating the partnership, etc.

the player can strictly improve payoffs by exerting no effort, switching symbols and terminating the partnership himself. Considering therefore the incentives in the two latter cases, we observe that an s^i player will not choose to unilaterally switch symbol (to s^j, say) if:

$$v^i\left(\bar{e}^i\right) \geq \pi\left(0, \bar{e}^i\right) + \delta w^j\left(\bar{e}^j\right) - c_t\left(i, j\right). \tag{6}$$

Substituting (3) satisfied with equality into (6), we obtain the following characterization of unilateral symbol switching.

Proposition 3 *In the segregating PPE, no s^i players switch to symbol s^j if*

$$w^j\left(\bar{e}^j\right) - w^i\left(\bar{e}^i\right) \leq \frac{c\left(i, j\right)}{\delta}.$$

The symbol frequencies compatible with the efficient segregating PPE thus depend on the switching costs and on the shape of w, the expected continuation values when beginning a partnership with a randomly drawn new partner. Using (1), (2) and (3), we write

$$w^i\left(\bar{e}^i\right) = \frac{p^i \pi\left(0, \bar{e}^i\right)}{1 - \delta}.$$

The following Lemma characterizes the shape of $w^i\left(\bar{e}^i\right)$ as a function of p^i.

Lemma 1 *In the efficient segregating PPE, $w^i\left(\bar{e}^i\right)$ is a left-continuous, positive function of p^i, which decreases at any discontinuity and is such that*

$$\lim_{p^i \to 0} w^i\left(\bar{e}^i\right) = \lim_{p^i \to 1} w^i\left(\bar{e}^i\right) = 0.$$

Moreover, $w^i\left(\bar{e}^i\right)$ increases with p^i, where $w^i\left(\bar{e}^i\right)$ is differentiable w.r.t. p^i, iff $\frac{\partial d\left(\bar{e}^i, 0\right)}{\partial \bar{e}^i} \leq 0$.

Hence, the matrix of symbol switching costs imposes a bound on the maximal difference in continuation values with a randomly drawn partner. The continuation value of s^i players with a new randomly drawn partner approaches zero for extreme values of p^i. If $p^i \to 1$, the almost certainty of finding a new s^i partner in the next round prevents them from committing to significant effort levels. If $p^i \to 0$, then the inability of finding a new s^i partner after a partnership termination drives $w^i\left(\bar{e}^i\right)$ to zero, despite s^i players being able to sustain the highest possible effort level in a homogenous partnership, which we denote

$$\tilde{e} \equiv \max\left\{e | d\left(e, 0\right) = 0\right\}.$$

Hence, $\tilde{e}$ is the effort level that can only be sustained by partners who know they will never again find a cooperative partner after the termination of their present partnership. Starting from $p^i = 0$, it is plausible to see the continuation value of s^i players, w^i, initially increase with p^i, because the decrease in sustainable efforts, $\bar{e}^i$, is initially more than compensated for by an increased likelihood of finding a new s^i partner. In particular, and for future purposes, we define p^* as the highest share such that, for all $p^i \leq p^*$, $w^i(\bar{e}^i)$ increases with p^i. Let e^* denote the corresponding effort, such that $d(e^*, p^*) = 0$.

Of particular interest is the question of how very scarce symbols can be sustained in equilibrium. Such symbols permit small radical groups to commit to very high levels of cooperation as well as very high payoffs during a homogeneous partnership, but they also imply a very high reward for cheating, which can only be countered by the prospect of a long search time after break-up. For the case of unilateral symbol switching, such scarce symbols crucially require high 'outgoing' switching costs. Such high switching costs can, e.g., reflect ethnic markers, or to a lesser extent language or religion. Caselli and Coleman (2013) argue that conflicts often develop along ethnic divides, since ethnic markers cannot easily be switched by members of the losing side. Besides symbols that are physically hard to switch, such as skin color, this equally applies to religion, where religious people may be extremely unwilling to give up their beliefs in face of a divine judgement or an eternal afterlife.[14]

Of course, although the matrix of switching costs c is taken as exogenously given in the present analysis, (groups of) players can choose a symbol in function of the desired costs in the real world. Berman (2000) argues that the Ultra-Orthodox Jews only developed their very distinctive and stringent traditional practices in the 19th century, when the economic emancipation of European Jews gave them access to many new economic opportunities in society at large.

Likewise, Berndt (2007) argues that 19th century Jewish peddlers cultivated their despised and distinctive status to enable them to act as middlemen in high stake financial transactions. Commitment to high stake interactions requires high exit costs, and this is also the case in criminal environments. Prison tattoos and initiation rites requiring a murder allow criminals to trust their partners when their life and freedom are at stake. But the same mechanism also offers policy makers a

[14] Alternatively, even in the presence of very low switching costs, very small groups can exist in equilibrium if all other symbols have extreme frequencies. For instance, if $c(.) = 0$, we must have $w^j\left(\bar{e}^j\right) = w^i\left(\bar{e}^i\right)$ for all i and j in equilibrium. For finitely many symbols, this is compatible with a $p^i \to 0$ if almost the entire population bears another symbol s^j, while all other symbol frequencies equal p^i or zero.

way to fight crime: subsidizing the outgoing switching costs. Ignoring other moral, economic and practical issues, offering criminals a way out, and even allowing them to keep a part of their spoils if they turn on their accomplices, impedes criminals in committing to criminal cooperations. The legislation on *pentiti* ensures in that sense a double blow to organized crime. First, it facilitates convictions, by motivating criminals to testify against former colleagues in exchange for reduced sentences, witness protection and a new identity. Secondly, it offers an exit option, thus reducing criminals' potential to credibly commit to a criminal endeavour.

It remains to show that at least one efficient segregating PPE exists. This is done in the following Proposition.

Proposition 4 *If π satisfies condition 1, then for all matrices of symbol switching costs c, a vector $(p^i)_{i=1,\ldots,n}$ can be found for which an efficient segregating PPE exists.*

The proof of Proposition 4 relies on the fact that the effort levels in Proposition 1 are well defined if condition 1 is satisfied, demonstrates the subgame perfection of the efficient segregating PPE, and argues that Proposition 3 is always satisfied for uniform symbol frequencies.

4 Other equilibria and renegotiation

The previous Section characterized one particular PPE out of potentially many PPE. In this Section, we compare the efficient segregating PPE to some other important PPE, on the basis of tractability and renegotiation proofness. A first class of alternative equilibria consists of 'symbol-blind' PPE, in which the contributed efforts are not conditioned on symbols. A first example of such a symbol-blind PPE is one in which players never exert positive effort. Clearly, such a PPE always exists. Second, among the main candidates for a symbol-blind PPE with strictly positive efforts are those which involve a form of gradual trust-building or incubation. In these equilibria, the equilibrium efforts depend on the age of a partnership, denoted $\tau \in \mathbb{N}$. In the early rounds of a partnership, equilibrium efforts are low, and the prospect of facing these low continuation values in a new partnership enables partners to sustain high efforts in later rounds. Thus, a symmetric incubation PPE is characterized by a sequence $(e_\tau)_{\tau=0,1,2,\ldots}$ which satisfies infinitely many incentive compatibility constraints, such that for all $\tau = 0, 1, 2, ...$:[15]

$$\pi(e_\tau, e_\tau) - \pi(0, e_\tau) + \sum_{j=1}^{\infty} (\delta(1-\lambda))^j (\pi(e_{j+\tau}, e_{j+\tau}) - \pi(e_{j-1}, e_{j-1})) \geq 0. \tag{7}$$

The incentive compatibility constraints in (7) require that, for all τ, the expected future benefits of being $\tau + 1$ rounds further in a partnership exceed the benefits of shirking in the τ-th round (and having to start 'trust-building' anew). In addition, players only terminate a partnership if their partner deviates from the equilibrium play. Efficiency then implies maximizing

$$\sum_{\tau=0}^{\infty} (\delta(1-\lambda))^\tau \pi(e_\tau, e_\tau), \tag{8}$$

subject to (7). This characterization illustrates that finding gradual symbol-blind equilibria can be quite tedious in the present public goods game setting, and that finding efficient equilibria will be very difficult.

Hence, in the context of an infinitely repeated public goods game, the efficient segregating PPE is more simple and intuitive than the main contenders, i.e., the gradual trust-building PPE. Moreover, Ghosh and Ray (1996) show that gradual trust-building equilibria are vulnerable to renegotiation by current partners. In the present context of repeated bilateral interactions, this seems particularly disturbing. If we consider, e.g., an equilibrium without cooperation, then by definition no player wants to unilaterally deviate to exerting effort while his partner sticks to the equilibrium zero effort strategy. But if the two current partners can mutually improve themselves by jointly deviating to strictly positive efforts, and if such a joint deviation is incentive compatible, then we expect both players to take advantage of it. Unfortunately, none of the PPE seems to satisfy standard renegotiation proofness in the current setting. Therefore, we introduce a parametrized version of renegotiation proofness, called ε-renegotiation proofness, in order to measure the vulnerability of different PPE w.r.t. renegotiation within a partnership. ε-renegotiation proofness bounds the 'size' of a renegotiation from above, and encompasses standard notions of renegotiation proofness as a special case. We define the distance between two strategies σ and σ' as

$$\sup_{h_t} m(\sigma(h_t), \sigma'(h_t)),$$

[15]The derivation of (7) and (8) is presented in Appendix A.2.

in which $m(.)$ represents, for the sake of simplicity, a continuously differentiable function that strictly increases with respect to the differences in efforts, measured by the Euclidian metric, and differences in termination decisions and symbol switching, which are both measured by a discrete metric. We then define ε-renegotiation proofness w.r.t. bilateral deviations by current partners as follows.

Definition 2 (ε-renegotiation proofness) *A PPE in pure strategies is ε-renegotiation proof w.r.t. bilateral deviations by current partners if no pair of current partners can mutually benefit from an incentive compatible joint deviation, such that for each player's equilibrium strategies σ and deviating strategies σ', we have* $\sup_{h_t} m(\sigma(h_t), \sigma'(h_t)) \leq \varepsilon$.

Note that if $\varepsilon \to \infty$, ε-renegotiation proofness reduces to Ghosh and Ray's (1996) criterion of bilateral rationality, albeit without its restriction to players who have not deviated in their past arrangements with each other. Thus, we use the notion of ε-renegotiation proofness as a yardstick to compare PPEs in terms of their robustness w.r.t. joint deviations.

First, we consider the 'symbol-blind' equilibria. The vulnerability of gradual trust-building equilibria to bilateral deviations by current partners was shown by Ghosh and Ray (1996) and naturally extends to the present case of joint deviations of a restricted size.

Proposition 5 *For all* $\varepsilon > 0$, *only PPE in which effort levels are constant with respect to the age of the partnership τ satisfy ε-renegotiation proofness. Symbol-blind PPE generically never satisfy ε-renegotiation proofness.*

To provide some intuition: note that, regarding what players can expect from a new partnership, the equilibrium effort sequence e_τ fixes the same outside option for all players, independent of the age of their current partnership. If this outside option makes high efforts enforceable in later rounds of a partnership, then these high efforts are equally enforceable in the first round. As long as all others play the equilibrium strategies, two current partners can mutually improve themselves by jointly deviating to higher efforts in the first round of their partnership, up to the point where they exert the highest sustainable efforts. And such a joint deviation is profitable for both partners even if we only allow for a very small deviation. Hence, robustness against even the smallest joint deviations requires that the equilibrium efforts are independent of the partnership's age. The non-existence of ε-renegotiation proof PPE for any ε is then immediate. If the constant equilibrium efforts are high, then they violate incentive compatibility. If they are low, then they constitute a bad outside option, giving rise to bilateral deviations to higher efforts. An ε-renegotiation proof symbol-blind PPE typically only exists if $\delta(1-\lambda) = 0$, i.e., if players effectively play a sequence of one shot public goods games.

Let us now turn to the efficient segregating PPE. In order to satisfy ε-renegotiation proofness, the efficient segregating PPE needs to be robust to two kinds of (sufficiently small) joint deviations. First, the equilibrium has to be robust to joint deviations in homogeneous partnerships: effort levels should be sufficiently high in homogeneous partnerships to avoid that low equilibrium efforts allow two same symbol players to credibly commit to a joint deviation to higher efforts. From (1), (2) and (3), and keeping $w^i(\bar{e}^i)$ fixed at the efficient segregating PPE level, we have that a joint deviation to higher efforts e in a homogeneous s^i partnership with equilibrium effort $\bar{e}^i$ is viable if

$$d(e,0) = \frac{\pi(e,e)}{1-\delta(1-\lambda)} - \pi(0,e) \geq \delta \frac{(1-\delta)(1-\lambda)}{1-\delta(1-\lambda)} w^i\left(\bar{e}^i\right). \qquad (9)$$

Furthermore, (9) is satisfied with equality for $e = \bar{e}^i$. By Lemma 1, the left hand side of (9) is decreasing with e for $e \geq e^*$, such that robustness to joint deviations in homogeneous partnerships requires that p^i is smaller than p^* for all i.[16]

Second, the efficient segregating PPE crucially depends on segregation to sustain cooperation. As such, this PPE can only be ε-renegotiation proof if players are not tempted to break out of this segregation by jointly deviating to cooperation in a heterogeneous partnership, in order to avoid waiting for a same symbol partner. Consider the case of a pair of s^i and s^j players, and assume without loss of generality that $p^i \leq p^j < p^*$. Clearly, joint deviations to efforts above $\bar{e}^j$ are not credible, because they fail to satisfy incentive compatibility for the s^j player. The s^i player should thus settle for less efforts than his equilibrium efforts in homogeneous partnerships but can nevertheless be tempted to avoid waiting for a homogeneous partnership. Because cooperation in heterogeneous partnerships requires a deviation in termination decisions and a deviation from zero efforts, this is where the parametrization of ε-renegotiation proofness has a bite. Let $\hat{e}(\varepsilon)$ denote the maximal effort allowed in a deviation of size ε from zero effort which also deviates in the termination decision. Note that $\hat{e}(.)$ is well defined by the assumptions

[16] Note that the generic non-existence of an ε-renegotiation proof symbol-blind PPE can be understood as a special case, where $p^i = 1$.

imposed on the functions m and π. Now ε constrains the temptation of heterogeneous cooperation, up to the point where $\hat{e}(\varepsilon) = e^*$, after which further increases in ε are not binding. It can be shown that a sufficient condition for a joint deviation to $\hat{e}(\varepsilon)$ to not be viable reads:[17]

$$\frac{\pi(\hat{e}(\varepsilon), \hat{e}(\varepsilon))}{1-\delta(1-\lambda)} - \pi(0, \hat{e}(\varepsilon)) < \frac{\delta(1-\lambda)(1-\delta)}{1-\delta(1-\lambda)} w^i(\bar{e}^i). \qquad (10)$$

Note that $w^i(\bar{e}^i)$ strictly increases with p^i for $p^i < p^*$ by Lemma 1, such that we can define $\hat{p}(\varepsilon)$ as the infimum of the set of p^i's for which (10) is satisfied. Summarizing, we have the following result:

Proposition 6 *The efficient segregating PPE is ε-renegotiation proof if $p^i \in (\hat{p}(\varepsilon), p^*]$ for all i. Moreover, $\hat{p}(\varepsilon)$ strictly increases with ε and this up to the point where $\hat{e}(\varepsilon) = e^*$. For higher ε, no efficient segregating PPE satisfies ε-renegotiation proofness.*

Hence, whereas the former symbol-blind PPE is vulnerable w.r.t. renegotiation for every $\varepsilon > 0$, the efficient segregating PPE is more robust in the sense that only for values of ε that are too large, the segregation PPE fails ε- renegotiation proofness. In our closed form example, $\hat{p}(\varepsilon)$ takes a simple form.

Example 3 *In the closed form example,*

$$d(e,0) = \frac{1}{1-\delta(1-\lambda)} \frac{e}{1+e} - e$$

is a continuous and strictly concave function of e, with a unique maximum at

$$e^* = \frac{1}{\sqrt{1-\delta(1-\lambda)}} - 1,$$

such that $p^* = \frac{\sqrt{1-\delta(1-\lambda)} - (1-\delta(1-\lambda))}{\delta(1-\lambda)}$. Moreover, $\hat{p}(\varepsilon)$ solves (10) with equality, such that

$$\hat{p}(\varepsilon) = \left(\frac{1}{(1-\lambda)\delta} - 1\right)\hat{e}(\varepsilon).$$

This means that $p^* < \hat{p}(\varepsilon)$ as long as $\hat{e}(\varepsilon) < \frac{1}{\sqrt{1-\delta(1-\lambda)}} - 1$. Figure 1 illustrates the p^i viable with an ε-renegotiation proof efficient segregating PPE for $\hat{e}(\varepsilon) = 0.1$.

[17] See the proof of Proposition 6 in Appendix A.3.

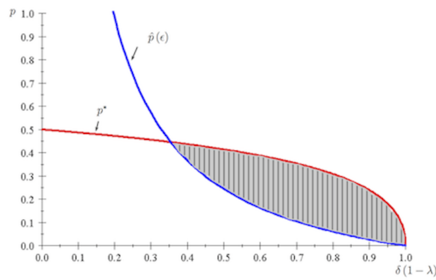

Figure 1*: Closed form example: p^*, $\hat{p}(\varepsilon)$ and the p^i consistent with an ε-renegotiation proof efficient segregating PPE (vertically hatched area) for $\hat{e}(\varepsilon) = 0.1$.*

Of course, the principal reason that the efficient segregated PPE is robust with respect to sufficiently small joint deviations by current partners, unlike symbol-blind PPE, lies in the radically different equilibrium efforts in homogeneous and heterogeneous partnerships. Restrictions on the size of joint deviations constrain in particular the profitability of deviations to cooperation in heterogeneous partnerships. As such, the restriction to small joint deviations can be understood as a reduced form of a model where other factors impede joint deviations in heterogeneous partnerships, and unbounded renegotiation proofness can thus be obtained by, e.g., permutations in the players preferences.[18] One example of such impediments are chauvinist or intolerant preferences, as modelled by Corneo and Jeanne (2009, 2010). Corneo and Jeanne study the evolution of intolerance, i.e., how the appreciation for diversity ('tolerance') evolves, by considering how parents socialize their children by dividing a unit of 'symbolic' valuation over different types people, e.g., different professions, sexual orientations or ethnicities. Their offspring's later evaluation of their own life then depends on how this distribution of valuations matches their own profession, sexual orientation or frequent interaction with other ethnicities. Corneo and Jeanne show that parents choose a narrow distribution of valuations (intolerance) in static and predictable environments. This maximizes their offspring's expected valuation of their own life, but can also induce them to avoid

[18]In a related setting, Choy (2014) shows how reputational concerns can discourage joint deviations to heterogeneous cooperation in a repeated public goods game setting with endogenous partnership termination, in which players observe the group affiliations of their partners' previous partners. Choy characterizes an equilibrium in which identifiable groups are hierarchically ordered and members of higher groups protect their reputation by refusing any contact with members of lower groups, which in turn ensures a sufficiently low outside option to sustain cooperation.

certain professional choices or interactions at the cost of foregoing economic opportunities. In dynamic environments, parents choose to raise their offspring in tolerance.

The present paper adds a second rationale for intolerance. If cooperation is sustained by segregation and if its sustainability is at risk due to the possibility of heterogeneous cooperation, then certain groups can choose to promote intolerance, i.e., stimulate members to develop a disutility from cooperating with players bearing a different symbol. Assume that a round of cooperation in a heterogeneous partnership comes at a disutility cost $\Delta \geq 0$, such that a player receives stage payoff

$$\pi \left(e, e'\right) - \Delta$$

from a round in which he exerts strictly positive effort in a heterogeneous partnership. Define then a threshold value,

$$\bar{\Delta} = \max_{e} \left\{d\left(e, 0\right)\right\} - \min_{\bar{e}^i} \left\{d\left(\bar{e}^i, 0\right)\right\}, \tag{11}$$

in order to state the following result.

Proposition 7 *If an efficient segregating PPE is ε-renegotiation proof for some ε, then it is ε-renegotiation proof for all $\varepsilon > 0$ if all players suffer a disutility $\frac{\Delta}{1-\delta(1-\lambda)} > \bar{\Delta}$ in each stage of heterogeneous cooperation.*

If e^* globally maximizes $w^i\left(\bar{e}^i\right)$ and $p^i = p^*$ for all i, then $\bar{\Delta} = 0$, and any $\Delta > 0$ suffices to make the efficient segregating PPE ε-renegotiation proof for all $\varepsilon > 0$. Moreover, one can further weaken the degree of intolerance required to avoid joint deviations to heterogeneous cooperation in several interesting ways. First, to avoid bilateral deviations to heterogeneous cooperation between s^i and s^j players, it suffices that only one of both partners is intolerant. Second, if only a share of s^i players have such intolerant preferences and if these preferences are private information, then a relatively small share of intolerant players with a sufficiently high Δ can be an effective deterrent of heterogeneous cooperation. The possibility of being defected upon by an intolerant player and suffering a negative stage payoff can be quite effective in preventing players from committing to a joint deviation in heterogeneous cooperation. This offers an interesting perspective on terrorist attacks, by which radical proponents of strong homogeneous communities use the power of mass media to suggest that a significant share of, e.g., Muslims hold intolerant preferences. A sufficiently large share of one group coming to believe in the existence of a share of intolerant players in the another group can already suffice to preclude all players from both groups from credibly committing to heterogeneous cooperation. If Islamic State or Al Qaeda change how all Muslims are perceived by enough people in the West, then their activities facilitate the enforcement of cooperation and social norms in Muslim communities in Western countries.

Note that the above considerations regarding ε-renegotiation proofness assumed strategy profiles and joint deviations without symbol switching. Yet, the possibility of joint deviations also affects symbol switching. Assume again a new partnership between an s^i and an s^j player, with $p^i < p^j$. Where unilateral symbol switching is excluded by Proposition 3 if the switching costs exceed the difference continuation values $|w^j(\bar{e}_j) - w^i(\bar{e}^i)|$, allowing for joint deviations enables current partners to jointly deviate by continuing the partnership and having one switching symbol. After this joint deviation, they continue cooperating in a now homogeneous partnership at equilibrium effort levels. Note that this joint deviation qualifies as a small renegotiation if metric m puts little weight on changes in termination and symbol switching decisions, because it requires no changes in effort decisions. In this case, ε-renegotiation proofness imposes additional restrictions on the symbol frequencies or on the matrix of symbol switching costs c. Consider again a small p^i, entailing high efforts $\bar{e}^i$ in homogeneous partnerships. If players can jointly deviate in termination and symbol switching decisions, then they can circumvent the high search costs of finding another s^i player, thus making $\bar{e}^i$ unsustainable. Therefore, a small p^i is in this case only viable in an ε-renegotiation proof efficient segregating PPE if the outgoing symbol switching costs $c(i,.)$ are high enough to avoid s^i players unilaterally leaving the group (cfr. supra), and if the incoming symbol switching costs $c(.,i)$ is sufficiently high, to avoid players moving into the group of s^i players by means of joint deviations.

Moreover, whereas uniform zero symbol switching costs, $c(.) = 0$, are compatible with an efficient segregating PPE if we only consider unilateral symbol switching (e.g., if $p^i = p$ for all i), allowing for joint deviations in termination and symbol switching decisions always requires strictly positive switching costs. To see this, note that an s^j player is not willing to switch to s^i in a joint deviation if

$$\delta\left(\lambda w^i\left(\bar{e}^i\right) + (1-\lambda)v^i(\bar{e}^i)\right) - c(j,i) < \delta w^j\left(\bar{e}^j\right),$$

which can be arranged into an inequality

$$\left(\lambda + (1-\lambda)\left(\delta + \frac{1-\delta}{p^i}\right)\right) w^i\left(\bar{e}^i\right) - w^j\left(\bar{e}^j\right) < \frac{c(j,i)}{\delta}.$$

Clearly, $\lambda + (1-\lambda)\left(\delta + \frac{1-\delta}{p^i}\right) > 1$ for all $p^i < 1$, such that this inequality represents the seduction of immediate cooperation. Without switching costs and if players only cooperate in homogeneous partnerships, one player is always willing to switch to his partner's symbol in a heterogeneous partnership by means of a joint deviation, in order to avoid waiting for a homogenous partnership. Thus, allowing for joint deviations in termination and symbol switching decisions implies that cooperation can only be sustained if symbol switching is costly.

To conclude this Section, we briefly consider another family of efficient symbol-dependent PPE. Consider a PPE in which players only exert effort in a heterogeneous partnership with one particular other symbol. E.g., the s^i players only exert positive efforts when partnered with an s^j player, and *vice versa*. Of course, this equilibrium is largely equivalent to the efficient segregating PPE if $p^i = p^j$. Otherwise, say if $p^i < p^j$, efficiency no longer implies symmetry in cooperative partnerships. If both players exert efforts on the incentive compatibility constraint, then the s^j is worse off than the s^i player for several reasons. First, the s^j player, bearing the more frequent symbol, can credibly commit to higher efforts than the s^i player. This higher effort level makes him worse off than his partner as well as worse off than he would be if both exerted the highest incentive compatible effort of the s^i player. Second, the more frequent symbol s^j player on average faces a longer search time for an s^i partner, which is what constitutes his ability to commit to higher efforts. Third, this selection effect now makes that the vector of p values exacerbates the differences in symbol frequencies at the population level, rather than smoothening them as in the efficient segregating PPE (see Proposition 2). For all these reasons, the conditions for the existence and ε-renegotiation proofness of such a PPE are much more stringent than for the efficient segregating PPE.

5 Conclusion

By inducing segregation, payoff irrelevant symbols can help to sustain cooperation in an infinitely repeated public goods game with endogenous partnership termination and no information about a partner's past play. If players only cooperate with partners bearing the same symbol, then the consequent search for a new homogenous partnership constrains the continuation value after a defection. Due to the fact that a more scarce symbol implies on average a longer search for a homogeneous partnership, it also enables players to sustain higher efforts. This idea of symbols inducing a sacrifice of outside options is closer to Iannaccone's (1992) club theoretic approach than, e.g., Eeckhout's (2006) analysis, where players eventually cooperate with all others after an incubation period, and symbols allow for Pareto improvements upon this incubation equilibrium by serving as a public correlation device. At the same time, the reaction to symbols which pose no intrinsic impediment to out-group interactions (e.g., clothing, hairstyles, veils) is fixed by a notion of equilibrium, where the reaction to this kind of symbols is typically exogenously assumed in Iannaccone's (1992) club theoretic setting.

Ghosh and Ray (1996) show that equilibria featuring a gradual increase of efforts throughout partnerships are vulnerable to bilateral renegotiation. As long as the rest of the population plays the equilibrium strategy of low initial cooperation, and thus provides the incentives to sustain cooperation in a later stage, two partners can benefit from jointly deviating to immediate cooperation. In the present setting, the efficient segregating equilibrium is equally vulnerable to such joint deviations but satisfies a weaker parametrized version of renegotiation proofness, ε-renegotiation proofness, as long as joint deviations are not too big. Symbol-blind equilibria, on the other hand, fail to satisfy ε-renegotiation proofness even for the smallest joint deviations.

The restriction of renegotiation proofness to sufficiently small joint deviations mainly affects joint deviations to cooperation in heterogeneous partnerships, and this can thus be understood as a reduced form of a model with other obstacles to heterogeneous cooperation. The present analysis thus offers an additional perspective on the rationale for intolerant preferences, as analyzed in Corneo and Jeanne (2009, 2010). Communities can encourage intolerant preferences, by which players suffer an intrinsic disutility cost from cooperating with partners bearing a different symbol, in order to enforce cooperation and norm adhesion. In this context, it is worth noting that robustness against joint deviations to heterogeneous cooperation does not require all players to object to heterogeneous cooperation. Players cannot credibly commit to a joint deviation to heterogeneous cooperation if one of the two players is intolerant, or even if players are uncertain about whether a partner bearing a different symbol may face such objections. This suggests that terrorist attacks or xenophobe political discourses can be interpreted to exploit the power of mass media in order to cast doubt on the true preferences of members of another community. If sufficiently many members in one community doubt the true intentions of members of the other community, or if members of one community fear that too many people in another community have doubts about the true intentions of the former community's members, then this prevents that players from committing to cooperation in a heterogeneous partnership. Impeding integration between the different communities then facilitates the preservation of traditional community norms.

A group symbol of particular interest in this context is the Muslim veil, due to the controversy surrounding the legal ban on veiling in public schools in France and other Western European countries. One principal motivation for such measures is the integration and emancipation of Muslim women in Western countries.[19] However, Carvalho (2013) presents a compelling interpretation of veiling as facilitating the emancipation of Muslim women. By wearing a veil, Muslim women reduce their exposure to the temptations of Western society as well as their community's beliefs that they might give in to such temptations. Therefore, this allows them to seize the social and economic opportunities without social repercussions from their community. Carvalho argues that a ban on veiling can cause Muslim women to substitute veiling by segregation and a withdrawal from an integrated public life. The present paper adds to this line of reasoning and presents a specific formal mechanism where Carvalho's analysis on a veil reducing temptation remains implicit. If a legal ban on veiling or merely the public debate about it makes wearing a veil more controversial, and thus casts more doubt in the majority population's mind about the true preferences of Muslim women, then it will be harder for the latter to establish interactions outside their own community. In this fashion, debates about a veil ban as well as Salafist extremism and xenophobe rhetorics contribute to the ability of Muslim communities to enforce traditional social norms upon their members.

We conclude with a final caveat on possible policy implications of the present analysis. Social norms and group cooperation can be both a source of emancipation and conflict, as well as a source of both creation and destruction of welfare. In light of the present analysis and abstracting from all other concerns, the welfare assessment of segregation and its effects on cooperation depends foremost on whether and how higher effort levels contribute to societal welfare. The same mechanisms can be employed by some groups to cooperate in enhancing science and emancipation, while at the same time by others to expropriate the poor or support a totalitarian regime.

References

Aimone, J. A., L. R. Iannaccone, M. D. Makowsky, and J. Rubin (2013): "Endogenous Group Formation via Unproductive

[19] For instance, the prominent French feminist and politician Fadela Amara, an Algerian-born Muslim woman, defends the French ban of the niqab because "women wearing a niqab or burqa will never become pilot, doctor or teacher. In reality, they are bound to remain confined to their homes." (see: "*Fadela Amara dénonce «le fascisme» de «l'intégrisme musulman»*" in *Le Parisien*, 29/12/2010)

Costs," *Review of Economic Studies*, 80(4), 1215–1236.

Berman, E. (2000): "Sect, subsidy, and sacrifice: An economist's view of ultra-orthodox jews," *Quarterly Journal of Economics*, 115, 905–953.

Berndt, C. E. H. (2007): "Hostile territory - High-tension religion and the Jewish peddler," *American Journal of Economics and Sociology*, 66(5), 1005–1027.

Carvalho, J.-P. (2013): "Veiling," *The Quarterly Journal of Economics*, 128(1), 337–270.

Caselli, F., and W. J. Coleman (2013): "On the theory of ethnic conflict," *Journal of the European Economic Association*, 11(s1), 161–192.

Choy, J. P. (2014): "Constructing Social Division to support Cooperation: Theory and Evidence from Nepal," Working paper, Warwick University.

Corneo, G., and O. Jeanne (2009): "A theory of tolerance," *Journal of Public Economics*, 93(5-6), 691 – 702.

——— (2010): "Symbolic values, occupational choice, and economic development," *European Economic Review*, 54(2), 237 – 251.

Datta, S. (1996): "Building Trust," Working paper, London School of Economics.

Deb, J., and J. Gonzalez-Diaz (2014): "Enforcing Social Norms: Trust-building and community enforcement," Working paper, Stanford University.

Eeckhout, J. (2006): "Minorities and endogenous segregation," *Review of Economic Studies*, 73, 31–53.

Ellison, G. (1994): "Cooperation in the prisoner's dilemma with anonymous random matching," *The Review of Economic Studies*, 61(3), 567–588.

Fujiwara-Greve, T., and M. Okuno-Fujiwara (2009): "Voluntarily Separable Repeated Prisoner's Dilemma," *Review of Economic Studies*, 76(3), 993–1021.

Fujiwara-Greve, T., M. Okuno-Fujiwara, and N. Suzuki (2012): "Voluntarily separable repeated Prisoner's Dilemma with reference letters," *Games and Economic Behavior*, 74(2), 504–516.

Gambetta, D. (2009): *Codes of the Underworld: How Criminals Communicate*. Princeton University Press, Princeton, New Jersey.

Ghosh, P., and D. Ray (1996): "Cooperation in community interaction without information flows," *Review of Economic Studies*, 63(3), 491–519.

Greif, A. (1993): "Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition," *American Economic Review*, 83(3), 525–548.

Iannaccone, L. R. (1992): "Sacrifice and Stigma - Reducing Free-Riding in Cults, Communes, and Other Collectives," *Journal of Political Economy*, 100, 271–291.

Kandori, M. (1992): "Social norms and community enforcement," *The Review of Economic Studies*, 59(1), 63–80.

Kranton, R. E. (1996): "The formation of cooperative relationships," *Journal of Law, Economics & Organization*, 12(1), 214–233.

Mailath, G. J., L. Samuelson, and A. Shaked (2000): "Endogenous inequality in integrated labor markets with two-sided search," *American Economic Review*, pp. 46–72.

Okuno-Fujiwara, M., and A. Postlewaite (1995): "Social Norms and Random Matching Games," *Games and Economic Behavior*, 9, 79–109.

Peski, M., and B. Szentes (2013): "Spontaneous discrimination," *The American economic review*, 103(6), 2412–2436.

Sandler, T., and J. Tschirhart (1997): "Club theory: Thirty years later," *Public Choice*, 93(3-4), 335–355.

Schumacher, H. (2013): "Imitating cooperation and the formation of long-term relationships," *Journal of Economic Theory*, 148(1), 409–417.

Shimizu, H. (2011): "Social cohesion and self-sacrificing behavior," *Public Choice*, 149(3-4), 427–440.

Tajfel, H., and J. C. Turner (1979): *The Social Psychology of Intergroup Relations*chap. An Integrative Theory of Intergroup Conflict. Brooks/Cole, Monterey.

Takahashi, S. (2010): "Community enforcement when players observe partners' past play.," *Journal of Economic Theory*, 145(1), 42–62.

Watson, J. (1999): "Starting small and renegotiation," *Journal of economic Theory*, 85(1), 52–90.

——— (2002): "Starting small and commitment," *Games and Economic Behavior*, 38(1), 176–199.

A Mathematical Appendix

A.1 Proof of Proposition 1

Solving (1) and (2) for $v^i(\bar{e}^i)$ and $w^i(\bar{e}^i)$, one obtains

$$v^i(\bar{e}_i) = \frac{1}{1-\delta} \frac{1-(1-p^i)\delta}{1-\delta(1-p^i)(1-\lambda)} \pi\left(\bar{e}^i, \bar{e}^i\right) \tag{12}$$

and

$$w^i\left(\bar{e}^i\right) = \frac{1}{1-\delta} \frac{p^i}{1-\delta(1-p^i)(1-\lambda)} \pi\left(\bar{e}^i, \bar{e}^i\right). \tag{13}$$

Substituting (12) and (13) into the incentive compatibility constraint in (3), and noting that efficiency implies that the incentive constraint in (3) is satisfied with equality, we obtain after rearranging terms:

$$d\left(\bar{e}^i, p^i\right) = \frac{\pi\left(\bar{e}^i, \bar{e}^i\right)}{1-\delta(1-\lambda)(1-p^i)} - \pi\left(0, \bar{e}^i\right) = 0. \tag{14}$$

Note that this also means that $\bar{e}^i$ solves

$$\frac{\pi\left(\bar{e}^i, \bar{e}^i\right)}{\pi\left(0, \bar{e}^i\right)} = 1 - \delta(1-\lambda)\left(1-p^i\right).$$

Under condition 1, it cannot be excluded that $\frac{\pi(e,e)}{\pi(0,e)}$ strictly increases with e on some intervals of $\mathbb{R}_+$. By efficiency, we select the highest e satisfying (14) by means of the maximum operator in (4).

This characterization of $\bar{e}^i$ is well defined if π satisfies condition 1, as the ratio $\frac{\pi(e,e)}{\pi(0,e)}$ continuously maps $\mathbb{R}_+$ on the entire unit interval. First, continuity is implied by the continuous differentiability of π, and $\frac{\pi(e,e)}{\pi(0,e)} \in [0,1]$ for all e because $\pi_1(.) \leq 0$ and $\pi(e,e) \geq 0$ for all e, and because π_2 is bounded away from zero. Second, $\lim_{e \to 0^+} \frac{\pi(e,e)}{\pi(0,e)} = 1$ by the third part of condition 1. This also means that $\lim_{p^i \to 1^-} \bar{e}^i = 0$. And third, $\lim_{e \to \infty} \frac{\pi(e,e)}{\pi(0,e)} = 0$ by the boundedness of $\pi(e,e)$ and because $\pi_2(.)$ is bounded away from zero.

Finally, note that $\bar{e}^i$ decreases continuously with p^i, except where $\bar{e}^i$ constitutes a local maximum of $\frac{\pi(e,e)}{\pi(0,e)}$. At such point, (4) decreases discontinuously to a lower effort level, such that $\bar{e}^i$ constitutes left-continuous and strictly decreasing function of p^i. In a similar fashion, $\bar{e}^i$ is right-continuous and strictly increasing function of $\delta(1-\lambda)$.

Proof of Proposition 2

First, $x^i = \frac{\lambda}{p^i+(1-p^i)\lambda}$ follows directly from (5) and $x^i_t = x^i_{t+1} = x^i$. By Bayes rule, the population share of s^i-players then equals:

$$\frac{\frac{p^i}{x^i}}{\sum_j \frac{p^j}{x^j}}.$$

Substituting $x^i = \frac{\lambda}{p^i+(1-p^i)\lambda}$, p^i exceeds this population share if:

$$\begin{aligned} p^i &> \frac{p^i\left(p^i+\left(1-p^i\right)\lambda\right)}{\sum_j p^j\left(p^j+\left(1-p^j\right)\lambda\right)} \\ &\Leftrightarrow (1-\lambda)\sum_j \left(p^j\right)^2 + \lambda > (1-\lambda)\,p^i + \lambda \\ &\Leftrightarrow p^i < \sum_j \left(p^j\right)^2. \end{aligned}$$

Proof of Proposition 3

An s^i player in a heterogeneous partnership wishes to switch to symbol s^j if

$$c(i,j) + \delta w^j\left(\bar{e}^j\right) > \delta w^i\left(\bar{e}^i\right),$$

while an s^i player in a homogeneous partnership switches to symbol s^j if

$$\pi\left(0,\bar{e}^i\right) + c(i,j) + \delta w^j\left(\bar{e}^j\right) > v^i\left(\bar{e}^i\right).$$

Because (3) is satisfied with equality in equilibrium, both these inequalities are equivalent. In a homogeneous partnership, switching symbol after a defection strictly dominates switching symbol after exerting the equilibrium effort, as argued in the text. Rearranging terms gives the inequality provided in Proposition 3.

Proof of Lemma 1

Using the expected continuation value when matched with a randomly drawn partner, $w^i\left(\bar{e}^i\right)$,

$$w^i\left(\bar{e}^i\right) = \frac{p^i\pi\left(0,\bar{e}^i\right)}{1-\delta},$$

and noting that $\pi(0,.)$ is a continuous map (by Condition 1), $w^i\left(\bar{e}^i\right)$ is a continuous composition of $\bar{e}^i$, which itself, by Proposition 1, is a left continuous function of p^i. Therefore, $w^i\left(\bar{e}^i\right)$ is a left-continuous function of p^i. As $\bar{e}^i$ decreases at each discontinuity, and $\pi(0,.)$ is increasing,

$w^i\left(\bar{e}^i\right)$ is decreasing at each discontinuity. Using the characterization of $\bar{e}^i$, we have that:

$$\lim_{p^i \to 0} \pi\left(0, \bar{e}^i\right) = \lim_{p^i \to 0} \frac{\pi\left(\bar{e}^i, \bar{e}^i\right)}{1 - \left(1 - p^i\right)\delta(1-\lambda)}.$$

And for fixed $\delta(1-\lambda)$ bounded away from 1, the latter limit exists and is bounded, because $\pi\left(e, e\right)$ is bounded for all $e \in \mathbb{R}_+$ (by Condition 1). Therefore,

$$\lim_{p^i \to 0} w^i\left(\bar{e}^i\right) = \frac{1}{1-\delta} \lim_{p^i \to 0} \frac{p^i \pi\left(\bar{e}^i, \bar{e}^i\right)}{1 - \left(1 - p^i\right)\delta(1-\lambda)} = 0.$$

In case $p^i \to 1$, we have that $\bar{e}^i \to 0$ such that $\pi\left(0, \bar{e}^i\right)$ converges to zero, and therefore, $\lim_{p^i \to 1} w^i\left(\bar{e}^i\right) = 0$. Consider a p^i at which $w^i\left(\bar{e}^i\right)$ is differentiable. Now, solving for $v^i\left(\bar{e}^i\right)$ in (1) gives us:

$$v^i\left(\bar{e}^i\right) = \frac{\pi\left(\bar{e}^i, \bar{e}^i\right)}{1-\delta(1-\lambda)} + \frac{\delta\lambda}{1-\delta(1-\lambda)} w^i\left(\bar{e}^i\right). \tag{15}$$

Now substitute (3) with equality for $v^i\left(\bar{e}^i\right)$ in (15). Rearranging terms, we obtain:

$$\frac{\delta(1-\delta)(1-\lambda)}{1-\delta(1-\lambda)} w^i\left(\bar{e}^i\right) = \frac{\pi\left(\bar{e}^i, \bar{e}^i\right)}{1-\delta(1-\lambda)} - \pi\left(0, \bar{e}^i\right).$$

Since $\bar{e}^i$ is decreasing with p^i, $w^i\left(\bar{e}^i\right)$ increases with p^i if and only if,

$$\frac{\partial d\left(\bar{e}^i, 0\right)}{\partial \bar{e}^i} \leq 0.$$

Proof of Proposition 4

We now show that the efficient segregating PPE, as defined in Definition 1, is indeed a PPE in pure strategies, and that vector $\left(p^i\right)$ exists for all matrices of symbol switching costs, $\left(c\left(i,j\right)\right)_{i,j}$, such that the efficient segregating PPE can be sustained as an equilibrium. First, by Proposition 1, the equilibrium efforts in ((4)) are always well defined under condition 1. Given that all deviations are met with a partnership termination, no unilateral deviation from $\bar{e}^i$ can increase a player's continuation value by construction. And because all deviations in termination decisions after a deviation are met with partnership termination, this equilibrium is also subgame perfect. Finally, we need that the condition in Proposition 3 that excludes symbol switching,

$$w^j\left(\bar{e}^j\right) - w^i\left(\bar{e}^i\right) \leq \frac{c\left(i,j\right)}{\delta},$$

is satisfied for some vector (p^i) for all matrices of symbol switching costs, $(c(i,j))_{i,j}$. Note that for a vector of equal components, $p^i = p^j = p$ for all i, j, we have $w^i(\bar{e}^i) = w^j(\bar{e}^j)$, which implies that the inequality in Proposition 3 is satisfied for all $(c(i,j))_{i,j}$. ■

A.2 Derivation of (7) and (8)

We briefly illustrate the derivation of (7). Note that after k rounds of equilibrium play, the expected continuation value on the equilibrium path is

$$\begin{aligned}
&\pi(e_k, e_k) + \delta((1-\lambda)\pi(e_{k+1}, e_{k+1}) + \lambda\pi(e_0, e_0)) \\
&+ \delta^2\left((1-\lambda)^2\pi(e_{k+2}, e_{k+2}) + \lambda(1-\lambda)\pi(e_1, e_1) + \lambda^2\pi(e_0, e_0)\right) \\
&+ \delta^3\left((1-\lambda)^3\pi(e_{k+3}, e_{k+3}) + \lambda(1-\lambda)^2\pi(e_2, e_2) + \lambda^2(1-\lambda)\pi(e_1, e_1) + \lambda^3\pi(e_0, e_0)\right) \\
&+ \ldots
\end{aligned} \tag{16}$$

The expected continuation value of cheating in the k-th is

$$\begin{aligned}
&\pi(0, e_k) + \delta\pi(e_0, e_0) + \delta^2((1-\lambda)\pi(e_1, e_1) + \lambda\pi(e_0, e_0)) \\
&+ \delta^3\left((1-\lambda)^2\pi(e_2, e_2) + (1-\lambda)\lambda\pi(e_1, e_1) + \lambda^2\pi(e_0, e_0)\right) \\
&+ \ldots
\end{aligned} \tag{17}$$

Incentive compatibility requires that the difference between ((16)) and ((17)) is positive. After some algebraic manipulation, we obtain the constraint in (7).

For the objective function in (8), note that at the 0-th round of cooperation, a player wishes to maximize

$$\begin{aligned}
&\pi(e_0, e_0) + \delta(1-\lambda)\pi(e_1, e_1) + \delta\lambda\pi(e_0, e_0) \\
&+ \delta^2(1-\lambda)^2\pi(e_2, e_2) + \delta^2\left(\lambda^2 + (1-\lambda)\lambda\right)\pi(e_0, e_0) + \delta^2(1-\lambda)\lambda\pi(e_1, e_1) \\
&+ \ldots \\
&= \pi(e_0, e_0)\left(1 + \lambda\sum_{t=1}^{\infty}\delta^t\right) + (1-\lambda)\delta\pi(e_1, e_1)\left(1 + \lambda\sum_{t=1}^{\infty}\delta^t\right) \\
&+ ((1-\lambda)\delta)^2\pi(e_2, e_2)\left(1 + \lambda\sum_{t=1}^{\infty}\delta^t\right) + \ldots \\
&= \sum_{k=0}^{\infty}\pi(e_k, e_k)(\delta(1-\lambda))^k\left(1 + \lambda\sum_{t=1}^{\infty}\delta^t\right) \\
&= \left(1 + \frac{\lambda\delta}{1-\delta}\right)\sum_{k=0}^{\infty}\pi(e_k, e_k)(\delta(1-\lambda))^k.
\end{aligned}$$

That means that we seek to maximize $\sum_{k=0}^{\infty}(\delta(1-\lambda))^k\pi(e_k, e_k)$

Proof of Proposition 5

Suppose that players agree on a non-constant effort plan $(e_\tau)_{\tau=0,1,\dots}$. This equilibrium sequence constitutes the outside option at all moments of the current partnership (i.e., what to expect in the next partnership) and is thus independent of how far a player is in his current partnership. If at some τ a certain effort e is sustainable, then all efforts in time periods τ for which $e_\tau \le e$ can be renegotiated to level e. Repeat this argument and conclude that the only effort that is robust to $\varepsilon-$renegotiation is a constant and efficient effort level, where efficiency means exhausting the incentive compatibility constraints. Note that this argument holds for any $\varepsilon > 0$. Now, denote such constant efficient effort level by $\tilde{e}$. Let $v(\tilde{e})$ denote the expected continuation value. Then,

$$v(\tilde{e}) = \pi(\tilde{e}, \tilde{e}) + \delta v(\tilde{e}).$$

In order for $\tilde{e}$ to be incentive compatible and efficient, we should have that:

$$v(\tilde{e}) = \pi(0, \tilde{e}) + \delta v(\tilde{e}).$$

But this means that $\pi(0, \tilde{e}) = \pi(\tilde{e}, \tilde{e})$, which can only hold for $\tilde{e} = 0$. Now, consider $e(\varepsilon) > 0$ a positive effort level that is allowed under our criterion of ε-renegotiation proofness. A joint deviation from zero efforts to $\hat{e}(\varepsilon)$ gives the following expected continuation payoff:

$$v(\hat{e}(\varepsilon)) = \frac{\pi(\hat{e}(\varepsilon), \hat{e}(\varepsilon))}{1 - \delta(1-\lambda)},$$

and incentive compatibility requires that

$$v(\hat{e}(\varepsilon)) \ge \pi(0, \hat{e}(\varepsilon)),$$

which is clearly viable. Hence, we can only sustain $\tilde{e} = 0$ in an ε-renegotiation proof PPE in the non-generic case where $\delta(1-\lambda) = 0$. Generically, there does not exist a symbol-blind ε-renegotiation proof PPE. ■

A.3 Proof of Proposition 6

First, in homogeneous s^i partnerships, two players who jointly deviate to an effort level e, permitted as an ε-renegotiation, obtain a continuation value

$$v^h(e) = \pi(e, e) + \delta(1-\lambda) v^h(e) + \delta \lambda w^i(\bar{e}^i). \tag{18}$$

Solving (18) for $v^h(e)$ and substituting into the incentive compatibility constraint

$$v^h(e) \ge \pi(0, e) + \delta w^i(\bar{e}^i),$$

we obtain that a joint deviation to e is only viable if

$$d(e,0) \geq \delta \frac{(1-\delta)(1-\lambda)}{1-\delta(1-\lambda)} w^i\left(\bar{e}^i\right). \tag{19}$$

Note then that (19) is satisfied with equality for $\bar{e}^i$ by construction, and that $d(e,0)$ decreases with e for all $e \geq e^*$, such that if $p^i \leq p^*$ for all i, it cannot be that (19) is satisfied for any $e \geq \bar{e}^i$. Hence, a joint deviation to a higher effort level in a homogeneous partnership is never viable if $p^i \leq p^*$ for all i.

Second, consider a pair of players with different symbols, say an s^i and s^j player. Suppose without loss of generality that $p^i \leq p^j < p^*$. From Lemma 1, it follows that $w^i(\bar{e}^i) \leq w^j(\bar{e}^j)$. We now show that the assumption (10) is sufficient to ensure that the efficient segregating PPE is ε-renegotiation proof. The continuation value in a joint heterogeneous deviation with effort level $\hat{e}(\varepsilon)$ is

$$v^H(\hat{e}(\varepsilon)) = \pi(\hat{e}(\varepsilon), \hat{e}(\varepsilon)) + \delta(1-\lambda) v^H(\hat{e}(\varepsilon)) + \delta\lambda w^i\left(\bar{e}^i\right),$$

from which we obtain that joint deviation to heterogeneous cooperation is viable if

$$\frac{\pi(\hat{e}(\varepsilon), \hat{e}(\varepsilon))}{1-\delta(1-\lambda)} - \pi(0, \hat{e}(\varepsilon)) \geq \frac{\delta(1-\lambda)(1-\delta)}{1-\delta(1-\lambda)} w^i\left(\bar{e}^i\right). \tag{20}$$

As a result of Lemma 1, the outside continuation value $w(\bar{e}^i)$ is increasing with p^i; therefore, if we use $\hat{p}(\varepsilon)$ to denote the infimum of all $p^i \leq p^*$ satisfying (10) (i.e., not satisfying (20)), we see that our efficient segregating PPE is sustainable if $p^i > \hat{p}(\varepsilon)$. Moreover, $\hat{p}(\varepsilon)$ is increasing in ε. By increasing ε, the upper bound of joint deviations, $\hat{e}(\varepsilon)$ increases. An efficient segregating PPE is sustained as long as $\hat{e}(\varepsilon) < \bar{e}^i \leq e^*$ for all i. When $\hat{e}(\varepsilon) = e^*$, then $\hat{p}(\varepsilon) = p^*$, and therefore, the interval of feasible shares p^i becomes empty and the efficient segregating PPE is no longer ε-renegotiation proof.

■

Proof of Proposition 7

Consider an s^i and s^j player considering a joint deviation. Note that the s^j-player can at most exert effort up until $\bar{e}^j$ (by incentive compatibility), whereas the s^i-player can exert efforts at a higher level, until a maximum level of $\bar{e}^i$. The (expected) continuation value of a joint deviation to some level e for s^i reads:

$$v^i_\Delta(e) = \frac{\pi(e,e) - \Delta}{1-\delta(1-\lambda)} + \frac{\delta\lambda w^i(\bar{e}^i)}{1-\delta(1-\lambda)}. \tag{21}$$

Now, let $\Delta > \bar{\Delta}$, where $\bar{\Delta}$ is defined as in (11). The s^i-player will not comply to the deviation if

$$v_{\Delta}^{i}(e) < \pi(0, \bar{e}^{i}) + \delta w^{i}\left(\bar{e}^{i}\right), \tag{22}$$

which, after substitution of (21) and using (9) (where the inequality is replaced by an equality) becomes:

$$d(e, 0) - d\left(\bar{e}^{i}, 0\right) < \frac{\Delta}{1 - \delta(1 - \lambda)}. \tag{23}$$

Notice that the left-hand side of (23) is smaller than $\bar{\Delta}$; hence, for $\frac{\Delta}{1-\delta(1-\lambda)} > \bar{\Delta}$, (23) is satisfied. Furthermore, as only the s^i-player can exert higher efforts in an asymmetric deviation, if (23) is fulfilled, then no such asymmetric deviations are viable, as the consequent (expected) continuation value of such asymmetric deviations would be even lower than the corresponding continuation value in the symmetric deviation considered here. ■