ECONOMICS

Working Paper Series

No. 17-2020

Unhappy is the land without symbols – Group symbols in infinitely repeated public good games

Tom Potoms
Department of Economics, University of Sussex Business School
T.Potoms@sussex.ac.uk

Unhappy is the land without symbols - Group symbols in infinitely repeated public good games

Tom Potoms[y] and Tom Truyts[z]

September 21, 2020

Abstract

How are group symbols (e.g. a flag, Muslim veil, 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 efficiently segregating equilbrium, in which players only cooperate with others bearing the same symbol. Players bearing a scarcer symbol face a longer expected search time to find a cooperative partner upon partnership termination, and can therefore sustain higher levels of cooperation. We compare this equilibrium to other equilibria in terms of Pareto dominance and robustness to (some form of) bilateral renegotiation.

Keywords: Endogenous segregation; repeated games; random matching ; public goods games

JEL codes: C73, D83.

1 Introduction

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, or the recent polarization surrounding the wearing of face coverings during the COVID pandemic[1]. 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. [2]

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 (1979) their 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. [3] 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 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

[1] There are many reports in the news that indicate support for face coverings in public spaces breaks down very much along partisan lines. For example, according to a NBC News/Survey Monkey Weekly Tracking Poll (July 20-26, 2020) almost 97 percent of Democrats or those who lean Democratic indicate they wear masks at least most of the time when they leave the house, whereas this drops to approximately 70 percent of Republicans (with only 48 percent of Republican-registered respondents indicating they wear masks `all the time' when they leave the home, in comparison to 86 percent for Democrats and 71 percent of Independents.)

[2] Even seemingly innocuous symbols can operate as a signal of `trustworthiness' or `friendliness' and can

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 (2011) 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 (2011) 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.

What we do. In this paper, 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 focus on stationary public perfect equilibria (PPE) of the game. 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 (with continuous effort choices). 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.[4]

We characterize a class of efficiently segregating equilibrium of this game, in which

> players exert no effort in the public goods game if their partner bears a different symbol.

[4]Notice that random matching, though quite standard, is also very natural in the present setting. Indeed, random matching confronts players with a stream of opportunities to form a beneficial partnership with

- 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.

Given that in heterogeneous (in terms of symbols) partnerships there are no positive efforts, they constitute a waste of time, and are immediately terminated by both players. Another feature of these sort of equilibria is that 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 (see our Proposition 1) . This damaging effect on members' outside options (i.e., non-members' reactions to these symbols) is a crucial mechanism for symbols to discipline group members' behavior. A nice feature of our framework is that, in contrast to the literature inspired by Iannaccone (1992), we do not take the negative reaction to group symbols as exogenously given. In particular, Proposition 2 provides the conditions on switching costs and continuation values under which no player would want to switch symbols. This extension towards endogenous symbol choice and reactions as equilibrium behavior is an important extension to real-world applications, since even though a negative 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. To the best of our knowledge, we are the first paper to study cooperative behavior with continuous action spaces, within an environment of limited information flows, repeated random matching and where symbol choices, the reactions to symbols and the resulting cooperation levels, are all jointly derived from (a notion of) equilibrium behavior.

We provide several results comparing the class of efficiently segregating PPE with another class of (`symbol-blind') cooperative PPE which are discussed in the literature. In particular, Proposition 5 shows that, for any symbol-blind PPE we can find a Pareto dominating efficiently segregating PPE. Proposition 6 strengthens this result in the sense that it shows even for the `best' symbol-blind PPE there exist Pareto dominant efficiently segregating PPE. Furthermore, Propositions 8 and 9 show that the class of efficiently segregating PPE distribution of players across symbols. In particular, we show (cfr. Proposition 7) that the uniform distribution does not, in general, yield the largest average payoffs, especially not in societies with a large number of (available) symbols.

Our contribution to the literature. 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.[5] Whereas knowledge of the full histories of players might be a plausible assumption in small communities, this is not the case for (relatively) large communities. Since we are interested in the scope for cooperation in large anonymous societies where individuals don't have full access to each other's past interactions, we don't make the assumption that players can observe private histories. Third, even though the contagion mechanisms of Kandori (1992) and Ellison (1994) can sustain cheating in the later stages of a partnership.[7]

2. Committed players. The presence of exogenous defectors in the population gives the situation of having a cooperative partner su cient scarcity value to discourage defection.[8] 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 su ce to sustain cooperation among patient players. More recently, in a richer information framework where players can observe part of their partner's past play, Heller and Mohlin (2018) show that the presence of `commitment types' can induce cooperative behavior, depending on the strategic environment (type of Prisoner's dilemma).

The present paper also contributes to this literature in the sense that we study a setting similar to Ghosh and Ray (1996)'s repeated public goods game, but in which the role of the exogenous defectors is played in equilibrium by endogenous group symbols. Hence, we assume no preference heterogeneity (in contrast to `committed players'), but rather derive that players act in equilibrium much like defectors towards others bearing a di erent symbol. In this equilibrium, players bearing di erent symbols face generically di erent incentives. In the spirit of Iannaccone (1992), players can thus sacri ce their outside options by bearing a more scarce symbol, and this sacri ce allows them to sustain higher cooperation levels.

The importance of payo irrelevant group symbols for cooperation is also central in Eeckhout (2006) and Choy (2018).[9] Eeckhout (2006) studies a public correlation device such as skin color in an in nitely 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 (2018) is the closest to our paper. He studies how segregation on the basis of visible group a liations helps to sustain cooperation in an in nitely repeated public goods game.

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

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

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

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 (2018), we assume no information about a partner's past play, and unlike Eeckhout (2006) and Choy (2018), we consider symbols a choice variable.

Structure of the paper. 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 a class of efficiently segregating equilibria. Section 4 introduces the class of gradual trust-building equilibria in our framework and compares this class with the efficiently segregating PPE. In section 5, we discuss the optimal distribution of symbols in society in terms of average payoffs. Section 6 discusses the robustness of the class of efficiently segregating and the gradual trust-building PPE to (a notion of) bilateral rationality, i.e., joint deviations from equilibrium behavior within partnerships. Finally, section 7 concludes the paper. The main proofs and derivations are detailed in the Appendix.

2 Formal Setting

We assume a continuum of players. Time is discrete and 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 $S = \{s^i\}_{i=1,\dots,n}$. In the initial period, each player is endowed with a particular symbol, after which they can switch their symbol. The fraction of players g 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.[10]

normalizes 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*

$$(e;e') = 1 + e' \quad e \quad \frac{1}{1+e}$$

satisfies condition 1, *as* $_1(e;:) = \quad 1 + \frac{1}{(1+e)^2} < 0$ *for* $e > 0$; $_2(:) = 1 > 0$ *and* $_1(e;e) + {}_2(e;e) = \frac{1}{(1+e)^2} > 0$ *for all* e 2 $\mathbb{R}_+$. *Moreover,* $(e;e) = \frac{e}{1+e}$

Similarly, the last element specifies a symbol switching decision as a function of the partner's effort e'_t.[14]

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

$$\mathbb{E}\left(\sum_t \delta^t \left(\pi(e_t;e'_t) - c_t(i;j)\right)\right);$$

in which the expectation operator $\mathbb{E}$ 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 The role of symbols for cooperation

This Section discusses how the payoff irrelevant symbols can help to sustain cooperative Play

4.

By De nition 1, no player switches symbol in equilibrium (given the switching costs $s(:;:)$), hence, ${}^{i}_{t} = {}^{i}$ for all t. The dynamics of the shares $(x^1_t;:::;x^n_t)$ are described by the following system of equations:

$$x^i_{t+1} = \quad 1 \quad p^i_t\,(1 \qquad)\;x^i_t + \qquad {}^i \qquad x^i_t \;; \text{ for all } i = 1;:::;n: \tag{2}$$

$e^i;$ rather than defecting on his partner and starting anew with a new partner in the next period. E ciency then imposes the inequality in (6) to be satis ed with equality. Solving (4) and (5) for v^i and w^i and substituting into (6)$;$ we de ne

$$
\begin{aligned}
d \quad e; p^i \qquad & v^i(e) \qquad (0; e) \qquad w^i(e) \\
= & \frac{(e; e)}{1 \qquad (1 \qquad)(1 \quad p^i)} \qquad (0; e);
\end{aligned}
\tag{7}
$$

such that (6) can be written as $d(e^i; p^i) \quad 0:$ Note that

$$
d(e; 0) = \frac{(e; e)}{1 \qquad (1 \qquad)} \qquad (0; e) \tag{8}
$$

is the di erence between the expected actual value of the current partnership, when cooperating at e ort level $e;$ and the one shot pa1 Tf 5l48 we;

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 very asymmetric distributions over symbols, . Such extremely skewed frequencies on the population level translate in extreme values of p^i. Notice that, for $p^i \to 1$; 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 $w^i(e^i)$ to zero, despite s^i players being able to sustain the highest possible effort level in a homogenous partnership, which we denote

$$\bar{e} \quad \max\{e \mid d(e;0) = 0\}.$$

Hence, $\bar{e}$ 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, 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 $p^i \leq p^*$, $w^i(e^i)$ increases with p^i. Let e^* denote the corresponding effort, such that $d(e^*; p^*) = 0$.

Summarizing, $E(c)$ consists of all distributions such that efforts for players are given by (9), continuation values and switching costs are such that no player has incentive to change symbol, i.e. (11) is satisfied. Before turning to a further discussion of the model, it first remains to show that at least one efficient segregating PPE exists. This is done in the following Proposition.

Proposition 3 *For all matrices of symbol switching costs c, $E(c) \neq \emptyset$, i.e., we can always find a vector $(\ ^i)_{i=1,\dots,n}$ that can be sustained as an efficiently segregating PPE.*

The proof of Proposition 3 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 2 is always satisfied for uniform symbol frequencies, since in that case $w^i(e^i) = w^j(e^j)$.

4 Gradual trust building

In the previous sections, we restricted our attention to a particular class of PPE. Clearly, there are potentially many more PPE which are also `symbol-blind', 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.[20] Second, among

The strategies involved in this class of (simplified) gradual-trust building PPE are a straightforward extension of the 'incubation strategies' tailored to discrete action spaces (e.g. in prisoner's dilemma games), such as those discussed in Fujiwara-Greve and Okuno-Fujiwara (2009). We begin the analysis of these simple gradual-trust strategies by finding the continuation values. Now, we denote by v_t the (expected) continuation value for a player who has been in a partnership with another player for t periods. Then, we have that

$$v_t = \pi(e_1, e_1) + \delta[(1-\lambda)v_{t+1} + \lambda v_0] \text{ for } t \geq T;$$

that is, for a partnership that has already lasted (at least) T periods, each player obtains $\pi(e_1, e_1)$. Furthermore, with a probability of $1-\lambda$ the partnership is not broken up, yielding a continuation value of v_{t+1}. With probability λ the relationship is broken up, after which each player draws a new partner and the 'counter' (duration of partnership) is reset to zero, resulting in an expected continuation value of v_0. Notice that, as long as the partnership continues (after reaching together for at least T periods), the continuation values are constant, that is, $v_t = v$ for all $t \geq T$. We can then derive an explicit expression for this expected continuation value, in case the partnership survives the incubation phase:

$$v = \frac{\pi(e_1, e_1)}{1-\delta(1-\lambda)} + \frac{\delta\lambda}{1-\delta(1-\lambda)} v_0. \tag{15}$$

Continuing with the continuation values for the other periods, we obtain:

$$v_t = \begin{cases} \pi(e_0, e_0) + \delta[(1-\lambda)v + \lambda v_0] & \text{if } t = T-1; \\ \pi(e_0, e_0) + \delta[(1-\lambda)v_{t+1} + \lambda v_0] & \text{if } t < T-1. \end{cases}$$

Which can be solved to yield the following expression[22]:

$$v_t = \frac{1-(\delta(1-\lambda))^{T-t}}{1-\delta(1-\lambda)} \pi(e_0, e_0) + \delta\lambda\frac{1-(\delta(1-\lambda))^{T-t}}{1-\delta(1-\lambda)} v_0 + (\delta(1-\lambda))^{T-t} v. \tag{16}$$

Substituting (15) into (16) then gives:

$$v_t = \frac{(\delta(1-\lambda))^{T-t}}{1-\delta(1-\lambda)}(\pi(e_1, e_1) - \pi(e_0, e_0)) + \frac{\pi(e_0, e_0) + \delta\lambda v_0}{1-\delta(1-\lambda)}. \tag{17}$$

[22] This can be derived by solving backwards, through repeated substitution of v_{t+1} into v_t.

The effort levels $(e_0; e_1)$ then need to satisfy the following set of incentive compatibility constraints:

$$(e_0; e_0) + \quad [(1 \quad) v_t + \quad v_0] \qquad (0; e_0) + \quad v_0; \text{ for all } t < T; \tag{18}$$

$$(e_1; e_1) + \quad [(1 \quad) v + \quad v_0] \qquad (0; e_1) + \quad v_0; \text{ for } t \quad T: \tag{19}$$

At this point, we can easily show that we can reduce the set of incentive compatibility constraints in (18) by checking that no player has an incentive to deviate from the gradual-trust strategy at the start of a partnership, that is,

$$(e_0; e_0) + \quad [(1 \quad) v_t + \quad v_0] \qquad (0; e_0) + \quad v_0;$$

which, after rearranging terms, yields the following:

$$(e_0; e_0) + \quad (1 \quad) v_0 \qquad (0; e_0) + \quad (1 \quad) v_0:$$

Given our assumptions on the payoff function, cfr. Condition 1, this incentive compatibility constraint can only be satisfied by setting $e_0 = 0:$ Given we are focussing on efficient effort levels within the partnership, the effort level post-incubation, e_1 are determined by the following:

$$\max_{e_1} \sum_{T} (\ (1 \quad)) \quad (e_1; e_1);$$

subject to (19). After substituting (15) and rearranging terms, we can rewrite (19), in analogy to equation (8) for the efficiently segregating PPE, by defining the following map

$$\tilde{d}(e; T) \quad \frac{1 \quad (\ (1 \quad))^{T+1}}{1 \quad (1 \quad)} \quad (e_1; e_1) \qquad (0; e_1) \quad 0: \tag{20}$$

The value $\tilde{d}(e; T)$ gives the difference between the expected value of the current partnership (with a length of incubation phase given by T and efforts in the cooperation phase given by e) and the one shot payoff of cheating. Efficiency then implies that efforts will be chosen in such a way that the inequality (20) is exhausted. The following summarizes the properties of efforts in a T-GTB equilibrium:

Proposition 4 *In a T-GTB equilibrium, efforts e_1^{GTB} uniquely solve*

$$e_1^{GTB} = \max \left\{ e | \tilde{d}(e; T) = 0: \right\}: \tag{21}$$

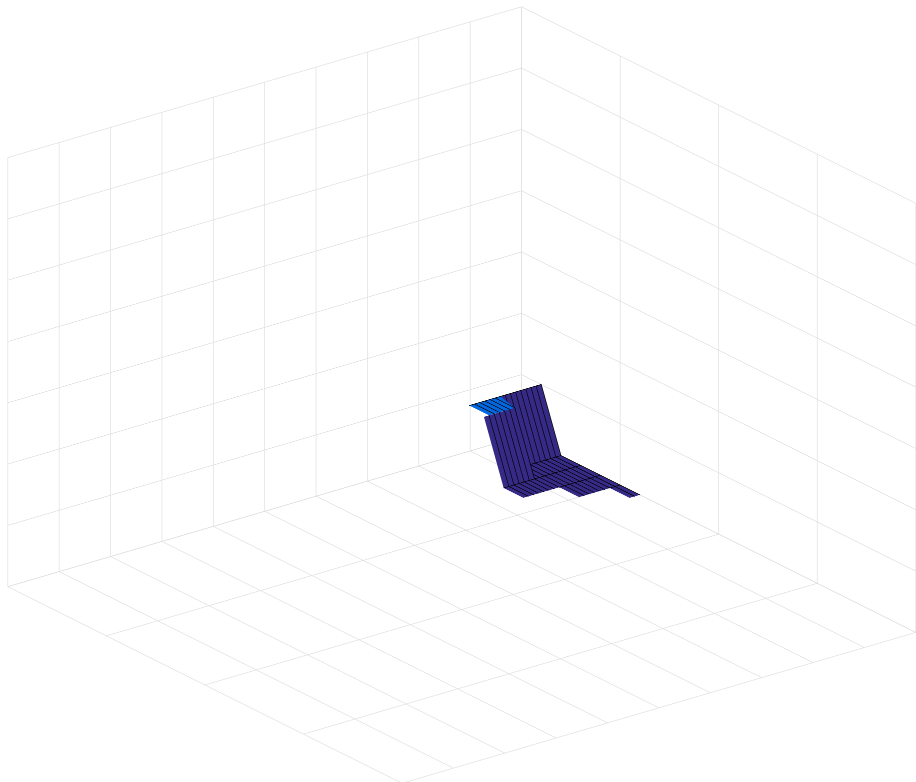
Moreover, e_1^{GTB} is strictly increasing function of T, right-continuous and strictly increasing with   and a left-continuous and decreasing function of .

The fact that e_1^{GTB} is increasing with T motivates the name `gradual-trust building', indeed,

4.1 Optimal GTB

Substituting the equilibrium e ort levels e_1^{GTB} (as de ned in (21)) into the expression for

Figure 3: T as function of ;



Larger values of and smaller values of (hence low e ective discount factors (1)*) are associated with shorter optimal incubation phases.*

Figure 3 also illustrates that T will generally take on the form of a step-function, in particular, given the integer restriction $T \in \mathbb{N}_0$ on the range of (23), if we x a value of , then for any value of there will be a range of discount factors that give the same value of T. Similarly, if we x a value of , then for all there will be th
probability of break up that gives rise to the same optimal length of the incubation phase.[24]

[24] More formally, if we x a value for , say $\in (0; 1)$, then for all *8to383;*

4.2 Comparing gradual trust-building with efficiently segregating PPE

We now want to compare the optimal gradual trust-building equilibria with the efficiently segregating PPE. To be more precise, we will compare the (expected) payoffs for the gradual trust-building equilibria with the average expected payoffs under efficient segregation. To make progress, we can compute the average expected payoffs under an efficiently segregating PPE. In particular, for a player bearing symbol s^i and a distribution of symbols this is given by[25]:

$$W^i(\;;n) = \;{}^i\; x^i\; v^i\; e^i\; + x^i$$

remainder of the proof then boils down to directly compare the resulting (average) expected payoffs from the corresponding efficiently segregating PPE and the GTB. Notice that the Proposition excludes those non-generic cases in which $= 0$ or $= 1$, since these cover the non-interesting setting in which no player exerts any effort. Furthermore, Proposition 5 states that the relative benefits of the (dominating) efficiently segregating PPE` are reduced for larger values of the effective discount factors for players, i.e. larger values of $(1 \quad):$ Intuitively, more patient players can afford to incur low payoffs in the incubation phase and thereby receive larger expected payoffs in the cooperative phase of the T-GTB equilibrium. This mechanism thereby reduces the relative advantage of the efficiently segregating PPE over the T-GTB PPE.

We are now interested in whether we can extend the result in Proposition 5 to the optimal GTB, i.e. those gradual trust-building PPE in which the length of the incubation phase is given by T , as defined by (23). The main complication in this case is that the length of the incubation phase automatically (and optimally) adjusts in response to different environments, that is, to different values for and , which makes direct comparison between the best GTB and the class of efficiently segregating PPE more difficult. However, we are able to show that the result in Proposition 5 still holds:

Proposition 6 *Consider a given T GTB. We can always find an efficiently segregating PPE with initial symbols' distribution $2\ E(c)$ such that:*

$$W^i(\;;n) \quad W_{GTB}(T(\;;\;);\;;\;);\ \text{for all } s^i; \tag{26}$$

as long as $;\ 2\ (0;1):$ The difference $W^i(\;;n) \quad W_{GTB}(T\;;\;;\;)$ is decreasing with and increasing with $:$

The result in Proposition 6 has essentially the same content as in Proposition 5. The results in Propositions 5 and 6 motivates our focus on the class of efficiently segregating PPE in the present paper, in the sense that one can always find an efficiently segregating PPE which Pareto dominates any given gradual trust-building PPE. Also notice that, even though the proof of this result uses a uniform symbols' distribution, the result itself doesn't preclude that non-uniform symbol distributions can give Pareto dominant (average) expected payoffs to all players. Indeed, moving probability mass in from one symbol, say s^j to another symbol, s^i, would decrease (by the enveloppe theorem) the average expected payoffs W^j for the s^j-bearing players. However, if (25) holds with strict inequalities, then the new symbols' distribution would still yield average expected payoffs which are Pareto dominating those for the T-GTB PPE. We can repeat this exercise with an arbitrary combination of symbols, until

(25) is satisfied with equality for at least one symbol. Note that the symbol switching costs c play an important role in these iterations, since every newly obtained distribution over symbols should yield continuation values that satisfy (11), i.e., no player would be willing to switch symbols and pay the cost of doing so.

5 Optimality

break-up probability for a partnership increase this optimal number of symbols. The logic behind this result comes from the fact that lower values for and thus higher e ective discount factor, (1) imply higher sustainable e orts in the e ciently segregating PPE. This increases the continuation values as a rst order e ect. A natural question which arises is whether (27) is maximized by the uniform distribution, $= \frac{1}{n}; \ldots; \frac{1}{n}$. The answer is in general negative.To see this, rst for given 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. Such a refinement of `bilateral rationality' is overly strong, indeed, it eliminates all PPE (Kranton (1996), Ghosh and Ray (1996)). This then also carries through to our class of efficiently segregating PPE. However, not all PPE might be equally sensitive to `some' form of bilateral deviation. More formally, we will quantify the `size' of bilateral deviations for a particular (class of) PPE to survive. This allows us to compare different (classes of) PPE in terms of these bounds to bilateral renegotiations.

To make progress, we need to have a measure for the size of deviations. We opt for a definition of distance on the strategy space, in particular for two strategies σ and σ' let $m(.)$ represent, 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.[28] The distance between σ and σ' is then given by

$$M(\sigma;\sigma') = \sup_{h_t} m(\sigma(h_t);\sigma'(h_t)) \tag{31}$$

Our first result with respect to bilateral deviations shows that the gradual trust-building PPE are vulnerable to *any* sort of bilateral deviation, no matter how small:

Proposition 8 *Let σ be an equilibrium strategy for a player consistent with a GTB PPE with associated length for the incubation phase given by T. Then, for all $\varepsilon > 0$ there exists a jointly profitable (bilaterally agreed) deviating strategy σ' with $M(\sigma;\sigma') \leq \varepsilon$.*

The result in Proposition 8 states that, for a T-GTB type equilibrium with a positive length of the incubation phase ($T > 0$), one can always find a jointly profitable deviating strategy. To provide some intuition: note that (regarding what players can expect from a new partnership) the equilibrium effort sequence $\left(\underbrace{0;\ldots;0}_{\text{length}\,=\,T};e_1;\ldots;e_1\right)$ 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

[28] The particular choices for the component-distance functions can be generalized and our results don't fundamentally rely on these choices.

e orts. And such a joint deviation is pro table for both partners even if we only allow for a very small deviation. Hence, robustness against even the smallest joint deviations joint deviation to $\hat{e}(\varepsilon)$ to not be viable reads: [30]

$$(\hat{e}(\varepsilon);\hat{e}(\varepsilon))$$

development and integration as well as advances in transportation and communication technology decreased impediments on mobility of contacts across individuals of different groups. Voltaire (2002) famously applauded in his Letters Concerning the English Nation the ability of trade and economic interactions to break down walls between communities, and thus fostering greater freedom:

> “Take a view of the Royal Exchange in London, a place more venerable than many courts of justice, where the representatives of all nations meet for the benefit of mankind. There the Jew, the Mahometan, and the Christian transact together, as though they all professed the same religion, and give the name of infidel to none but bankrupts.”

Almost three centuries later, many cheer the internet and its anonimity (and ease of adopting new online identities) for promoting freedom, while many others deplore the degradation of courtesy and good manners that appears to proceed from it. Today's large-scale urban societies are less fragmented in small geographically defined groups, but display instead an impressive cultural and subcultural diversity. The above analysis suggests that these (sub-)cultural identities can be helpful to sustain cooperation and social norms, and that groups that are in need of strong commitment to the common cause need to find group symbols that are very costly to get rid off, e.g. prison tattoos in the context of mutual protection or crime (Gambetta, 2011).

In the second half of the paper, we compared the class of efficiently segregating PPE with symbol-blind strategies to obtain cooperative behavior In particular, the so-called gradual trust-building strategies (Datta (1996), Kranton (1996), Fujiwara-Greve and Okuno-Fujiwara (2009) and Fujiwara-Greve et al. (2012)). In these, players build ‘trust' by starting at low (or zero) effort levels and then increasing it to a higher level. Though appealing, we have shown that, for every such gradual trust PPE we can find a Pareto dominating efficiently segregating PPE. Finally, we have shown that, even though all PPE in the game environment we are studying are not robust against bilateral deviations, i.e., ‘bilateral rationality'-type refinements as presented in Kranton (1996), Ghosh and Ray (1996), the class of efficiently segregating PPE are robust to allowing players to jointly deviate (within partnerships) in a neighborhood around the equilibrium strategies. This is in contrast to the $\bar{T}$-GTB, where players will always want to jointly deviate from the corresponding equilibrium strategies, no matter how small-size deviations are allowed for. These arguments reconfirm a strong motivation to study and analyze the class of efficiently segregating PPE.

Finally, we also provided results in terms of how average payoffs of players are affected by the number of symbols in society and we have shown that in general the uniform distribution of players across symbols is not necessarily optimal for the average payo of players across all symbols.

References

Aimone, J., L. Iannaconne, M. Makowski, and J. Rubin (2013). Endogenous group formation via unproductive costs. *Review of Economic Studies 80*, 1215{1236.

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

Berndt, C. E. (2007). Hostile territory: Hightension religion and the jewish peddler. *The American Journal of Economics and Sociology 66* (5), 1005{1027.

Bhaskar, V. and C. Thomas (2019). Observations on cooperation. *Review of Economic Studies 86* (3), 1010{1032.

Choy, J. P. (2018). Social division with endogenous hierarchy. *The Economic Journal 128*, 2711{2742.

Corneo, G. and O. Jeanne (2009). A theory of tolerance. *Journal of Public Economics 93*, 691{702.

Corneo, G. and O. Jeanne (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. (2020). Cooperation and community responsibility. *Journal of Political Economy 128* (5), 1976{2009.

Deb, J. and J. Gonzalez-D az (2019). Enforcing social norms: Trust-building and community enforcement. *Theoretical Economics 14* (4), 1387{1433.

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. *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. (2011). *Codes of the Underworld: How Criminals Communicate*. Princeton, New Jersey: Princeton University Press.

Ghosh, P. and D. Ray (1996). Cooperation in community interaction without information ows. *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.

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

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

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

Voltaire (1731 (2002)). *Letters concerning the English Nation*. Electronic Classics Series. http://www.naturalthinker.net/trl/texts/Voltaire,Francois/Voltaire - Letters on England.pdf: Pennsylvania State University.

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

Watson, J. (2002). Starting small and commitment. *Games and Economic Behavior 38* (1), 176–199.

A Proofs and derivations

Proof of Proposition 1

Solving (4) and (5) for $v^i(e^i)$ and $w^i(e^i)$, one obtains

$$v^i(e_i) = \frac{1}{1} \frac{1 \quad (1 \quad p^i)}{1 \quad (1 \quad p^i)(1 \quad)} \quad e^i; e^i \tag{34}$$

and

$$w^i \quad e^i \quad = \frac{1}{1} \frac{p^i}{1 \quad (1 \quad p^i)(1 \quad)} \quad e^i; e^i \quad : \tag{35}$$

Substituting (34) and (35) into the incentive compatibility constraint in (6) *;* and noting that efficiency implies that the incentive constraint in (6) is satisfied with equality, we obtain after rearranging terms:

$$d \quad e^i; p^i \quad = \frac{(e^i; e^i)}{1 \quad (1 \quad)(1 \quad p^i)} \qquad 0; e^i \quad = 0: \tag{36}$$

Note that this also means that e^i solves

$$\frac{(e^i; e^i)}{(0; e^i)} = 1 \qquad (1 \quad) \;\; 1 \quad p^i \quad :$$

Under condition 1, it cannot be excluded that $\frac{(e;e)}{(0;e)}$ strictly increases with e on some intervals of $\mathbb{R}_+$. By e ciency, we select the highest e satisfying (36) by means of the maximum operator in (9):

This characterization of e^i is well de ned if satis es condition 1, as the ratio $\frac{(e;e)}{(0;e)}$ continuously maps $\mathbb{R}_+$ on the entire unit interval. First, continuity is implied by the continuous di erentiability of ; and $\frac{(e;e)}{(0;e)}$ 2 $[0;1]$ for all e because $_1(:)$ 0 and $(e;e)$ 0 for all e; and because $_2$ is bounded away from zero. Second, $\lim_{e\,!\,0^+} \frac{(e;e)}{(0;e)} = 1$ by the third part of condition 1. This also means that $\lim_{p^i\,!\,1}$ (e;e)
[((e);T2F65 7.97552 Tf 9.062 0 T1 [(- Td5)]TJ/F20 11.9552 Tf 124.558- T436(e)]TJ/F65 7.9701 Tf 5.425 4.339 Td [(i)

(0;e) = 3050 This aness F20 11.9552 Tf 96.822 82.072( )]TJ/F17 11.9552 Tf 9.062 0 Td [(()]TJ/F20 11.9552 Tf 4.552 0 Td

$w^i(e^i)$ is a left-continuous function of p^i. As e^i decreases at each discontinuity, and $\quad(0;\cdot)$ is increasing, $w^i(e^i)$ is decreasing at each discontinuity. Using the characterization of e^i, we have that:

$$\lim_{p^i \to 0} \quad 0; e^i \quad = \lim_{p^i \to 0} \frac{(e^i; e^i)}{1 \quad (1 \quad p^i) \quad (1 \quad\quad)}.$$

And for xed $(1 \quad\quad)$ bounded away from 1, the latter limit exists and is bounded, because $\quad(e; e)$ is bounded for all $e \in \mathbb{R}_+$ (by Condition 1). Therefore,

$$\lim_{p^i \to 0} w^i \quad e^i \quad = \frac{1}{1} \lim_{p^i \to 0} \frac{p^i \quad (e^i; e^i)}{1 \quad (1 \quad p^i) \quad (1 \quad\quad)} = 0.$$

In case $p^i \to 1$, we have that $e^i \to 0$ such that $\quad(0; e^i)$ converges to zero, and therefore, $\lim_{p^i \to 1} w^i(e^i) = 0$. Consider a p^i at which $w^i(e^i)$ is di erentiable. Now, solving for $v^i(e^i)$ in (4) gives us:

$$v^i \quad e^i \quad = \frac{(e^i; e^i)}{1 \quad (1 \quad\quad)} + \frac{\quad}{1 \quad (1 \quad\quad)} w^i \quad e^i \quad . \tag{37}$$

Now substitute (6) with equality for $v^i(e^i)$ in (37). Rearranging terms, we obtain:

$$\frac{(1 \quad\quad)(1 \quad\quad)}{1 \quad (1 \quad\quad)} w^i \quad e^i \quad = \frac{(e^i; e^i)}{1 \quad (1 \quad\quad)} \quad\quad 0; e^i \quad .$$

Since e^i is decreasing with p^i, $w^i(e^i)$ increases with p^i if and only if,

$$\frac{\partial d \,(e^i; 0)}{\partial e^i} \quad\quad 0.$$

Proof of Proposition 3

symbol switching,

$$w^j \quad e^j \qquad w^i \quad e^i \qquad \frac{c(i;j)}{};$$

is satis ed 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 $w^i(e^i) = w^j(e^j)$, which implies that the inequality in Proposition 3 is satis ed for all $(c(i;j))_{i;j}$:

Derivation of (13) and (14)

We brie y illustrate the derivation of (13). Note that after k rounds of equilibrium play, the expected continuation value on the equilibrium path is

$$\begin{aligned}
&(e_k;e_k) + \quad ((1 \quad) \quad (e_{k+1};e_{k+1}) + \quad (e_0;e_0)) \\
&+ {}^2 \quad (1 \quad)^2 \quad (e_{k+2};e_{k+2}) + \quad (1 \quad) \quad (e_1;e_1) + {}^2 \quad (e_0;e_0) \\
&+ {}^3 \quad (1 \quad)^3 \quad (e_{k+3};e_{k+3}) + \quad (1 \quad)^2 \quad (e_2;e_2) + {}^2 (1 \quad) \quad (e_1;e_1) + {}^3 \quad (e_0;e_0) \\
&+ \ldots
\end{aligned} \tag{38}$$

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

$$\begin{aligned}
&(0;e_k) + \quad (e_0;e_0) + {}^2((1 \quad) \quad (e_1;e_1) + \quad (e_0;e_0)) \\
&+ {}^3 \quad (1 \quad)^2 \quad (e_2;e_2) + (1 \quad) \quad (e_1;e_1) + {}^2 \quad (e_0;e_0) \\
&+ \ldots
\end{aligned} \tag{39}$$

Incentive compatibility requires that the di erence between (38) and (39) is positive. After some algebraic manipulation, we obtain the constraint in (13).

For the objective function in (14), note that at the 0-th round of cooperation, a player wishes to maximize

$$(e_0; e_0) + \quad (1 \qquad) \quad (e_1; e_1) +$$

equilibrium are given by (21). Now, notice that the same e ort levels would arise in an e ciently segregating PPE with a uniform symbols' distribution, $= (1/n; ::: ; 1/n)$; if

$$n = n = \frac{1 \quad (\quad (1 \qquad))^{T+1}}{(\quad (1 \qquad))^{T} [1 \qquad (1 \qquad)]} ; \tag{40}$$

The rest of the proof is very similar to the proof of Proposition 5, in that for the given levels of $;$ we again obtain the expression (43). Notice that, given that T is now a function of $;$, the comparative statics are a bit more involved. In particular, given that for a xed value for $;$ say $=$, the map $T\ :;$ is a step function. Hence, the di erence $W^i\left((1/n;\ldots;1/n);n\right)\quad W_{GTB}\left(T;\ ;\ \right)$ is continuously decreasing, except for jumps downward at levels for which T changes.

Proof of Proposition 8

Suppose that players agree on a non-constant e ort plan $(e\ )_{=0;1;\ldots}$. 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 a certain e ort e is sustainable, then all e orts in time periods for which $e \quad e$ can be renegotiated to level e. Repeat this argument and conclude that the only e ort that is robust to $"$ renegotiation is a constant and e cient e ort level, where e ciency means exhausting the incentive compatibility constraints. Note that this argument holds for any $" > 0$. Now, denote such constant e cient e ort level by $\bar{e}$. Let $v(\bar{e})$ denote the expected continuation value. Then,

$$v(\bar{e}) = \quad (\bar{e};\bar{e}) + \quad v(\bar{e}):$$

In order for $\bar{e}$

which is clearly viable. Hence, we can only sustain $\hat{e} = 0$ for $M(\ ;\ ^{0})\quad ''$ in the non-generic case where $(1\quad) = 0$. Generically, there does not exist a symbol-blind PPE which is robust to a a joint deviation of (arbitrarily small) size $''$.

Proof of Proposition 9

Let be an equilibrium strategy for a player as part of an e ciently segregating PPE. We will put bounds on $'' > 0$ such that, the e ciently segregating PPE is robust against joint deviations prescribing for each player a strategy ^{0} and $M(\ ;\ ^{0})\quad ''$, i.e. for all $^{0} \; 2 \; N(\ ;'')$: First, notice that, in homogeneous s^i partnerships, two players who jointly deviate to ^{0}, with $M(\ ;\ ^{0})\quad ''$, with an associated e ort level e, obtain a continuation value

$$v^h(e) = \quad (e;e) + \quad (1\quad) v^h(e) + \quad w^i \quad e^i \quad : \tag{45}$$

Solving (45) for $v^h(e)$ and substituting into the incentive compatibility constraint

$$v^h(e) \qquad (0;e) + \quad w^i \quad e^i \quad ;$$

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

$$d(e;0) \qquad \frac{(1\quad)(1\quad)}{1\quad(1\quad)} w^i \quad e^i \quad : \tag{46}$$

Note then that (46) is satis ed with equality for e^i by construction, and that $d(e;0)$ decreases with e for all $e \quad e$; such that if $p^i \quad p$ for all i; it cannot be that (46) is satis ed for any $e \quad e^i$: Hence, a joint deviation to a higher e ort level in a homogeneous partnership is never viable if $p^i \quad p$ for all i:

Second, consider a pair of players with di erent symbols, say an s^i and s^j player. Suppose without loss of generality that $p^i \quad p^j < p$. From Lemma 2, it follows that $w^i(e^i) \quad w^j(e^j)$. We now show that the assumption (33) is su cient to ensure that players don't want to deviate jointly to $\hat{e}('')$. To that end, note that the continuation value in a joint heterogeneous deviation with e ort level $\hat{e}('')$ is given by:

$$v^H(\hat{e}('')) = \quad (\hat{e}('');\hat{e}('')) + \quad (1\quad) v^H(\hat{e}('')) + \quad w^i \quad e^i \quad ;$$

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

$$\frac{(\hat{e}('');\hat{e}(''))}{1\quad(1\quad)} \qquad (0;\hat{e}('')) \qquad \frac{(1\quad)(1\quad)}{1\quad(1\quad)} w^i \quad e^i \quad : \tag{47}$$

As a result of Lemma 2, the outside continuation value $w(e^i)$ is increasing with p^i; therefore, if we use $\hat{p}(")$ to denote the in mum of all p^i $p^\bullet$ satisfying (33) (i.e., not 5.875yi6p