Coordination through Status

Peter Engseld*

July 16, 2003

Abstract

The main purpose of this paper is to suggest a mechanism of coordination whereby concerns for high relative position might evolve endogenously. A simple symmetric 2×2 strategic coordination game is analyzed in an evolutionary environment where continua of bounded rational agents are randomly matched against each other. The agents are assumed to have the ability to make an imperfect observation on how their own status relates to that of their opponent. These observations are then utilized as an instrument for coordinating the strategies in the game by the conditioning of the pure strategies on whether the opponent has a higher status or a lower status. The outcome of the game is analyzed through a newly defined evolutionary criterion as the agent's observational skills come close to the limit of being perfect. It is shown that there exists an endogenous asymmetry in the model that results in just one of the conditioned strategies in the pure strategy set being evolutionary stable. Moreover, it is also shown that this strategy will prevail against strategies conditioned through other personal characteristics, e.g. size, and gender.

JEL classification: C70; C72

Keywords: Coordination; Hawk-Dove Games; Positional Concerns; Status; Evolutionary Games; Evolutionary Equilibrium

1 Introduction

The present paper takes position from observations made in ethology, where it has long been known that animals tend to yield or give precedence for other animals of higher status in terms of e.g. food and mating opportunities. This behavioral pattern indicates that status might have some coordinating role in various conflicting situations. The perhaps most intuitive explanation for the behavior in this particular example would be that higher status often is associated with higher strength, so that any resistance from an animal of lower status

---

*Department of Economics, Lund University, Box 7082, 220 07, Lund, Sweden. Phone: +46-46-2224290, email: peter.engseld@nek.lu.se. Financial support from HSFR and The Crafoord Foundation is gratefully acknowledged.

would be futile or come at a very high price. In a more formal game theoretical setting, it is obvious that the asymmetry stemming from the difference in the retaliatory power would be reflected into the payoff-matrix, so that one agent would suffer much more in case of a realized conflict. But let us look past any asymmetry caused by differences in retaliatory power between the agents, and ask if it is possible that status might serve as coordinator for a larger class of symmetric games. That is, does there exist an inherent asymmetry in strategic coordination games that benefits high status?

1.1 Background

1.1.1 Symmetric Games with Asymmetric Agents

The class of games mentioned above with heterogeneous agents was originally investigated by Maynard Smith [12] as an extension of the now famous Hawk-Dove model in his pioneering book *Evolution and the Theory of Games.* He there demonstrates that symmetric Hawk-Dove games with asymmetric agents, where the agents condition their strategies on the role of the opponent, reveals two *evolutionarily stable strategie*s, *ESS*, that in theory are equally plausible. However, the analysis of these games was made under the crucial assumption that the probability of an agent occupying a particular role is independent of its strategy. Furthermore, it is stressed that this assumption is reasonable if the agents participate only once in these games or if the outcome of the game does not influence the role the agent will have in the next game. However, it is not hard to foresee how the role of the agent could be dependent on the outcome. The most obvious would be if the agents were to condition their strategies on how well the opponent had done before, i.e. conditioning by position or status. It is obvious that this type of conditioning would violate the assumption above.

1.1.2 Positional Concerns

The question of positional concerns has attracted attention from economists since the beginning of economic literature and can be traced back to Adam Smith [11, pp. 108 - 110] who questioned why individuals seem to work more than required for their basic needs. Later, Veblen [14] claimed that the primary purpose for consumption was to signal one's status in the society, since societies at every arbitrary developmental phase establish expectations on a consumption pattern fitting to each level of status within a society.

Various attempts to explain and model positional concerns have emerged since then. An approach that is closely related to that of status can be found in the fairness literature. A basic difference between the two approaches, however, is that fairness is often associated with a specific economic situation, whereas status considers contextual factors such as the history of the individuals. But as Roth [7, pp. 329 - 330] has noted, one would expect individuals to act differently whether their opponent is a business tycoon or a manual laborer.

The status literature is difficult to survey due to the elusive character of the concept status. Nevertheless, Weiss & Fershtman [15] makes an excellent effort to capture this concept in the sociological literature.

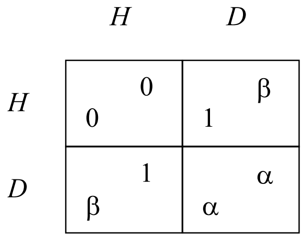
Two major approaches can be identified in the literature of modeling status; one that assumes that people have a direct preference for status, and another that argues that the concerns for status are *instrumental*. However, the former approach has been heavily criticized for being ambiguous in character since the model might explain every imaginable action just by letting the individuals have some positional concern, and thus not being able to explain anything ex ante. For a more in-depth discussion see Postlewaite [8].

In the latter approach, it is often assumed that the status can be associated with some desirable ability of the individual, since the signaling of status through different positional goods comes cheaper for individuals with this specific quality than for those without. Spence [13] seminal work of separable equilibrium was the first to highlight this approach. He found that individuals with a higher productivity were willing to "burn money" by overspending on education in order to signal their abilities. Subsequently, Frank [5] showed that people overspend relatively in positional goods compared with non-positional goods in order to "Keep up with the Joneses". In the same manner, Cole *et al.* [2] shows that agents tend to oversave in a "Rat race of the rich".

Rege [9] shows that positional goods such as Rolex watches and Armani suits can serve as signals of non-observable abilities in complementary interactions.

1.2 Outline of the Model

Consider a symmetric 2×2 game $\Gamma(\alpha, \beta)$ as depicted in the figure 1 below, where $\alpha \in (0,1)$ and $\beta \in (0,1)$.



	H	*D*
H	0 0	β 1
D	1 β	α α

Figure 1: *Symmetric* 2×2 *strategic coordination game*

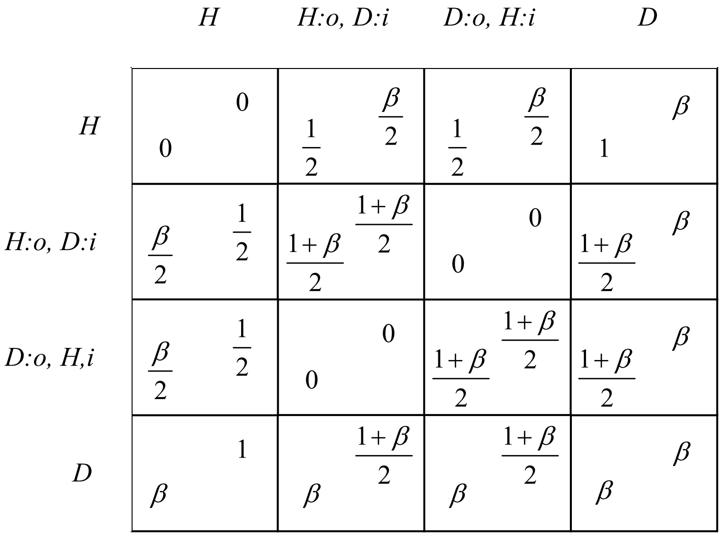
This game constitutes a simple coordination game where at least one agent has an interest in avoiding coordination failures. It can easily be verified that there exists a unique *EES* where one part of the population would be playing *hawk*, *H*, and another would be playing *dove*, *D*. But this will clearly result in coordination failure when an agent encounters another agent with the same strategy.

However, if the agents could find some way to successfully coordinate their behavior in Γ this would clearly convey a much more efficient equilibrium. It is easy to realize that the only feasible way for the agents to coordinate the actions perfectly would be if they have the ability to somehow differentiate between the various types of agents.

1.2.1 Conditioned Strategies

The pioneering work in studying this type of conditioning behavior on some observable asymmetric variable was originally done by Maynard Smith [12]. Through a *Hawk-Dove game* it was shown that the conditioning of strategies in an evolutionary setting by *static variables* such as, "*owner & intruder*", "*larger & smaller*", and "*older & younger*" results in two conditioned *ESS*. This can easily be verified in figure 2 below where a less general game ($\alpha = \beta$) is rewritten with conditioned strategies on whether the opponent is an owner meeting an intruder or an intruder meeting an owner. Note that for simplicity the game is given as the special case when the probability for being an intruder is equal to being an owner.

Consider an agent that e.g. plays H if he is an owner meeting an intruder and D if he is an intruder meeting an owner, i.e. $(H{:}o, D{:}i)$. From the game in figure 2 it can be seen that this agent will receive a higher payoff when meeting another agent with the same strategy, than an agent with any of the other strategies will receive when meeting an agent with this particular strategy. I.e. $\pi\left((H{:}o, D{:}i), (H{:}o, D{:}i)\right) > \pi\left(s, (H{:}o, D{:}i)\right)$ for $s \in \{H, D, (D{:}o, H{:}i)\}$, hence it is an *ESS.*



	H	*H:o, D:i*	*D:o, H:i*	*D*
H	0 0	$\frac{\beta}{2}$ $\frac{1}{2}$	$\frac{\beta}{2}$ $\frac{1}{2}$	β 1
H:o, D:i	$\frac{1}{2}$ $\frac{\beta}{2}$	$\frac{1+\beta}{2}$ $\frac{1+\beta}{2}$	0 0	β $\frac{1+\beta}{2}$
D:o, H,i	$\frac{1}{2}$ $\frac{\beta}{2}$	0 0	$\frac{1+\beta}{2}$ $\frac{1+\beta}{2}$	β $\frac{1+\beta}{2}$
D	1 β	$\frac{1+\beta}{2}$ β	$\frac{1+\beta}{2}$ β	β β

Figure 2: *Owner & intruder game*

The question that we will try to answer here is; does coordinating through "*high status & low status*" convey a unique equilibrium? In other words, does there exist some inherent asymmetry that benefits this strategy such that the game will result in a single evolutionary equilibrium? Let us, for clarity, informally define the status of an agent in terms of how well the agent has done in previous games; that is, a good track record is associated with a high status.

1.2.2 Status as a Coordinator

Now consider a simple example $\Gamma\left(\frac{3}{4},\frac{3}{4}\right)$, as depicted below, were two agents with high respectively low status (l & h) are met in a strategic coordination game. The agents can choose between playing H, or D. From the game it is obvious that two equilibria in conditioned pure strategies can be found.

The *Villain equilibrium* ($H{:}h, D{:}l$) where a low status agent yields for a high status agent; that is, play H against an agent with lower status and play D against an agent with higher status. Let this conditioned strategy henceforth be called *Villain strategy.*

Secondly, there is the reversed *Robin Hood equilibrium* ($D{:}h, H{:}l$) where a high status agent yields for a low status agent, i.e. play D against an agent with lower status and play H against an agent with higher status. Let this conditioned strategy henceforth be called *Robin Hood strategy.*

	H:l	*D:l*
H:h	0 0	¾ **1**
D:h	1 ¾	¾ ¾

Villain equilibrium

	H:l	*D:l*
H:h	0 0	¾ 1
D:h	**1** ¾	¾ ¾

Robin Hood equilibrium

There is obviously nothing in the game form that favors either of the two different equilibria. But could there exist another property, possibly dynamic, that would explain why high status is favored in coordination games? It should be noted that status has an interesting dynamic component, which is not included in the variables tested by Maynard Smith. This implies that if a population, in a tournament style contest, were to utilize the *Villain strategy*, this would, given the informal definition above, result in an *expansion* in the distribution of status since this strategy would favor agents that have been successful in previous games. Consequently the *Robin Hood strategy* would result in a *contraction* of the status distribution, since a less successful past is favored by this strategy. Moreover, since it would not be realistic to assume that the agents' observational skills are perfect, the *Robin Hood strategy* would over time result in more coordination failures that than *Villain strategy.*

1.3 Outline of the Study

The main objective of this paper is to suggest a mechanism for endogenous evolution of positional concerns from coordination in a class of symmetric 2×2 games. This is done by a generalization of Maynard Smiths [12, pp. 94 - 96] symmetric Hawk-Dove games with asymmetric agents, without the assumption of independence between the agents' role and strategy. It is assumed that the agents have the ability to condition their strategies by whether the opponent has a higher or lower status. Status is here defined as how well the agents have done in previous games. We assume that the agents have limited observational skill. The game is being played by a continuum of agents, where each agent is matched to play the game with an opponent randomly drawn from the population in each period. In order to avoid the difficulties associated with a complete analysis of this complex process, we only investigate whereto this process converges in the ultra long run. We define and characterize a criterion for stability in distributions of status.

Since conditioned strategies enable heterogeneous payoff, there will subsequently be a need for a slight modification of the *ESS* concept. Three alternative criteria for evolutionary stability are characterized, two within a population and one between different populations.

Given our assumptions, we find that the evolutionary criteria within populations result in multiple equilibria, whereas the criteria for evolutionary stability between populations display a unique equilibrium. This unique equilibrium will be matched against strategies conditioned by an arbitrary static variable that lacks the dynamic component of status; that is, every variable such that the role of the agent is independent of its strategy. Moreover, this unique equilibrium is also matched against conditioned hybrid strategies that comprises conditioning through compounded variables both with and without any dynamic component. Under our assumptions, we find that the unique equilibrium in status prevails for the hybrid strategies included in our analysis.

2 The Model

2.1 The Game

Consider a symmetric 2×2 game $\Gamma(\alpha, \beta)$ as described by the normal game form in figure 1. For $\alpha \in (0,1)$ and $\beta \in (0,1)$ we have a general class of symmetric 2×2 coordination games that includes both *Hawk-Dove games* $(\alpha > \beta)$ as well as *Battle of the sexes* $(\alpha \to 0)$. Let, as customary, the pure strategy set be given by $S = \{H, D\}$. It is easy to verify that three *Nash equilibria*, *NE*, can be found in this game: two in pure strategies, $(s_i, s_{-i}) = (H, D)$ and $(s_i, s_{-i}) = (D, H)$, and one in mixed strategies $\Pr(H) = \frac{1-\alpha}{1+\beta-\alpha}$, $\Pr(D) = \frac{\beta}{1+\beta-\alpha}$. If we were to loosen the assumption of fully rational agents and instead perceive the game from an evolutionary standpoint, the only remaining equilibrium (*ESS*) would be the equivalence to the mixed strategy. That is, $\frac{1-\alpha}{1+\beta-\alpha}$ of the population playing H, and the remaining $\frac{\beta}{1+\beta-\alpha}$ of the population playing D.

It should be noted that the *ESS*, as mentioned above, is associated with a probability for coordination failure, $\left(\frac{\beta}{1+\beta-\alpha}\right)^2 + \left(\frac{1-\alpha}{1+\beta-\alpha}\right)^2$, that equals or exceeds the probability for a successful coordination. Let us therefore assume that the agents in addition to the pure strategy set also have the ability to condition their strategies through status.

2.2 The Conditioned Strategy Set

Definition 1 (*Status*) *Let the* status *of agent i at t , w_i^t, be given by $w_i^t \equiv \rho \pi_i^{t-1} + (1-\rho) w_i^{t-1}$ where π_i^{t-1} is the expected payoff in the pervious period given $t \in \mathbb{Z}$ and $\rho \in (0,1)$.*

The status at the present period is identically given for every agent by some convex combination between the payoff and status of previous period. It would make for an improbable model to assume that the agents have the ability to determine from what period t a certain payoff stems. Instead, we assume that the agents observe a cumulative payoff. The status should thus primarily be interpreted as a measure on how well agents have done on average in previous games. Note that the status by this approach becomes equal to, or a very good proxy for, the average payoff and thus also to some normalization of the relative wealth.

For simplicity let henceforth, when there is not risk for confusion, w_i^t be denoted as w_i. Furthermore, make the assumption that each agent in addition to knowing their own status w_i also have the ability to make an observation of their opponent's status w_{-i}. Given that the observational skills of the agents can be considered as imperfect, it is realistic to also assume that the probability for coordination failures will increase as the distance between w_i and w_{-i} decreases.

However, there are two plausible ways to model the agents behavior as the distance between their status level decreases. Either it can be assumed that the agents adhere to their strategies in every given situation regardless of the difference between their own status and that of their opponent, $|w_i - w_{-i}|$, or that the agents at some point, when the distance in status is sufficiently close, abandons the attempts to condition their strategies and instead use an unconditioned strategy where $\Pr(H) = \frac{1-\alpha}{1+\beta-\alpha}$, $\Pr(D) = \frac{\beta}{1+\beta-\alpha}$. For clarity, let the standard unconditioned *ESS* in this particular game be called the *Maynard Smith solution* and abbreviated *MSS*.

Let us simplify the calculations by approximating this assumption with a step-function in the following manner; if $|w_i - w_{-i}| \geq \delta$ the agents will coordinate the game perfectly whereas when $|w_i - w_{-i}| < \delta$ the agents will use a strategy $\gamma \equiv \lambda H + (1-\lambda) D$ were $\lambda \in (0,1)$, alternatively λ play H and $(1-\lambda)$ play D. This approach should thereby comprise both when the agents assign equal probability to the strategies as a result of coordination failures, and when they due to lack of information adapt their strategies according to *MSS*. Note that this is consistent with *MSS* as well as with the mixed strategy which both reveal the same payoff. However, there is a qualitative difference between *MSS*, where the agents are playing a pure strategy, and the case where the agents randomize their strategies by assigning equal probability to both strategies. The former is obviously a pure strategy whereas the later is a mixed strategy.

Moreover, since the agents are assumed to observe the cumulative payoff, this implies that even though the distance in status is constant, the absolute difference in wealth will grow larger by time.

Assumption 1 $\delta \to 0$.

Note that this assumption implies that the approximation above will not change any of the results qualitatively at the limit, as $\delta \approx 0$.

The agents should, given the assumptions above, consequently be free to condition their strategies on the relation between their own and opponent's status. Hence, the pure strategy set **S** has to be enriched such that it also comprises the conditioned strategies. There are of course an innumerable amount of different ways to condition the pure strategies on the relation between the agents' own status and that of their opponent. But in this study we will restrict our attention to the case when an arbitrary agent i conditions the strategies on the relation between his own and opponent's status through either of two contingent strategy plans.[1]

$$\textit{Robin Hood strategy:}\quad \mathcal{R}_i = \begin{cases} s_h = H & \textit{if} \quad w_i \leq w_{-i} - \delta \\ s_l = D & \textit{if} \quad w_i \geq w_{-i} + \delta \\ s_u = \lambda H + (1 - \lambda)\, D & \textit{if} \quad w_{-i} + \delta > w_i > w_{-i} - \delta \end{cases}$$

$$\textit{Villain strategy:}\quad \mathcal{V}_i = \begin{cases} s_h = D & \textit{if} \quad w_i \leq w_{-i} - \delta \\ s_l = H & \textit{if} \quad w_i \geq w_{-i} + \delta \\ s_u = \lambda H + (1 - \lambda)\, D & \textit{if} \quad w_{-i} + \delta > w_i > w_{-i} - \delta \end{cases}$$

Hence, the enriched pure strategy set is given by:

$$\mathbf{S} = \{H, D, \mathcal{R}, \mathcal{V}\} \tag{1}$$

2.3 Playing the Game

Imagine a continuum of agents I where each individual agent is assumed to adhere to some pure strategy s from the enriched pure strategy set $\mathbf{S}$, $s \in \mathbf{S}$. Let the strategy mix in the population, consistent with the behavior of the population, be denoted $\mathbf{s}$ and henceforth called *the incumbent strategy.* Formally $\mathbf{s}$ is a convex combination of the pure strategies s, $s \in \mathbf{S}$, that is given by the distribution of strategies in the population. In other words, $\mathbf{s}$ is a a point in the strategy simplex Δ such that $\Delta = \left\{\mathbf{s} \in \mathbb{R}_+^4 \mid \sum_{s \in \mathbf{S}} \mathbf{s}_s = 1\right\}$.

However, in this study we restrict our investigation to when the incumbent strategy $\mathbf{s}$ initially is in either of vertices $\mathcal{R}$ and $\mathcal{V}$, i.e. a homogenous population where every single agent acts according to one pure conditioned strategy, or at

---

[1] This way of conditioning the pure strategies on the status is identical with how Maynard Smith originally defined this class of asymmetric games (see Maynard Smith [12]).

either of two points along the convex combination between the two remaining vertices H and D. The points on this line include both *MSS* and the case when the agents assign equal probability to the unconditioned strategies H and D. That is, just a part of the hull of the strategy simplex Δ is considered.

Let each agent be matched to play Γ with an opponent randomly drawn from the population I at every time t. At the end of each time period when all the agents have been matched, the status is updated according to definition 1. Note that this treatment reveals an identical expected payoff to the more traditional tournament-style contest, where each agent in a single tournament is matched with each remaining agent exactly once, which is a property of an infinite number of agents. Let the expected payoff for agent i at period t be denoted π_i^t.

The outcome for a single agent, thus also the entire game, is at any given time t not only dependent on the behavior of that agent, but also on how the population is distributed and where this agent can be found in the distribution. Since we already have assumed that every agent $i \in I$ is associated with a status w_i, let us now instead focus our attention on *how* the status is distributed in the population and thereby perhaps clarify some of the dynamics of the game. Therefore, let $w_i \in W$ for $\forall i \in I$. Hence at any given time t we have that $W \sim F_{\mathbf{s}}$, i.e. the status in the population is distributed according to $F(W;\mathbf{s})$. The payoff of an arbitrary agent i could consequently be characterized by $\pi_i^t(s, w_i^t; \mathbf{s}, F)$. Since we have assumed the incumbent strategy to be homogenous, let us for now simplify the expression by denoting the payoff as $\pi_i(w_i^t; \mathbf{s}, F)$.

The outcome of the game can be perceived from three time perspectives: *short*, *medium* and *long term*.

First consider what happens with status of an agent in the *short term* at time $t+1$. If the payoff from the previous period t is e.g. lower than the status this will result in a lower status in period $t+1$, $\pi_i(w_i^t; \mathbf{s}, F) < w_i^t \Rightarrow w_i^{t+1} < w_i^t$. Either way, if $\pi_i(w_i^t; \mathbf{s}, F) \neq w_i^t$ for any $i \in I$ the distribution F

will be different in the period $t+1$ and thus almost always reveal a different payoff for the agents at this period. If not in this period, the payoff will eventually change in the *medium term*.

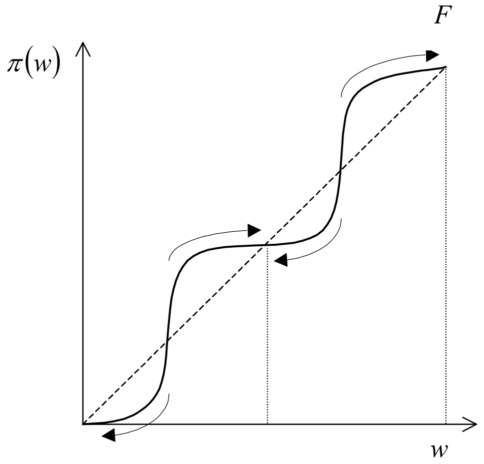
This process will continue as the distribution eventually converges to what hopefully is a stabile distribution in the *long term*. Let us circumvent the difficulties of calculating the outcome in the short and medium term and solely focus on whereto the distributions of strategies converge in the *long term* and investigate under what conditions these can be considered as stable in some sense.

2.4 Characterizing Stable Distributions in Status

A necessary condition for stability in the long term would be that the payoff of every agent π_i for status $w_i \in \operatorname{supp} F$ is constant over time a given incumbent strategy $\mathbf{s}$ and distribution of status F in the population I (*Condition 1*). This can formally be expressed as $\pi_i\left(w_i; \mathbf{s}, F\right) = w_i$. Since this is a property that has to be valid for $\forall i \in I$, the notation can be simplified by rewriting the condition as $\pi\left(w; \mathbf{s}, F\right) = w$.

However, this particular subset of every feasible distribution also includes distributions that would diverge from its support by an infinitesimal tremble in the status level. Since it would make for an unrealistic model if all mistakes were to be ruled out, distributions that are vulnerable for infinitesimal trembles should be considered as unstable. Moreover, additional perturbations in the payoff should be expected from the actual realizations of each matching, since the status here is defined through the average payoff.

Let us therefore exclude these distributions by adding a second condition; for every discrete point in status levels that satisfies *Condition 1* there must exist a surrounding by this point, such that every status level in a close neighborhood of this point converges back to the point (*Condition 2*). Let henceforth the subset points in the support that fulfills these conditions be denoted as *local attraction points*.



An example of a distribution with three *local attraction points.*

Let us formally define a stable distribution;

Definition 2 (*Stability*) *Every distribution* $F(W; \mathbf{s})$ *such that:*

Condition 1: $\pi(w; \mathbf{s}, F) = w$

Condition 2a: *If* $\exists x' < w$ *where* $(x', w) \cap \operatorname{supp} F = \emptyset \Rightarrow$
$\exists x'' < w$ *such that* $\pi(w^*; \mathbf{s}, F) > w^*, \forall w^* \in (x'', w)$

Condition 2b: *If* $\exists z' > w$ *where* $(w, z') \cap \operatorname{supp} F = \emptyset \Rightarrow$
$\exists z'' > w$ *such that* $\pi(w^*; \mathbf{s}, F) < w^*, \forall w^* \in (w, z'')$

$\Leftrightarrow$

$w \in \operatorname{supp} F$

is stable

Every discrete distribution belonging to the set of stable distributions of the incumbent strategy **s** could consequently be characterized as a distribution with support at every *local attraction point*, such that any given support in the distribution prevails over minute perturbations. Henceforth, let any distribution, discrete or continuous, that is stable according to definition 2 be denoted as $F \in \mathbb{F}_{\mathbf{s}}$.

However, it should be fair to assume that the agents are engaged in other economic activities than exclusively the coordination game described above. If this is the case, it might prove difficult for the agents to determine which part of the cumulative payoff are stemming from a coordination game and which part originates in other economic activities. It should moreover be considered as highly unlikely that the payoff distributions from other economical activities exactly mirror those of the coordination game. Hence, each distribution should repeatedly be subjected to exogenous perturbations of different magnitudes; i.e. the agents could due to these perturbations drift away from the support even though the support is a *local attraction point.*

Consider two *adjacent points* in the support such that

$$w', w'' \in \operatorname{supp} F \text{ and } (w', w'') \cap \operatorname{supp} F = \emptyset$$

where $F \in \mathbb{F}_{\mathbf{s}}$. Moreover, define the size of the *upper attraction basin of* w' as $A^{+}(w') \equiv |w' - x|$ such that

$$x = \sup \{ w^* \in (w', w'') \mid \pi(w^*; \mathbf{s}, F) < w^*, \forall w^* \in (w', w^*) \}.$$

In the same manner; define the size of the *lower attraction basin of* w'' as $A^{-}(w'') \equiv |x - w''|$ such that

$$x = \inf \{ w^* \in (w', w'') \mid \pi(w^*; \mathbf{s}, F) > w^*, \forall w^* \in (w^*, w'') \},$$

see figure 3.

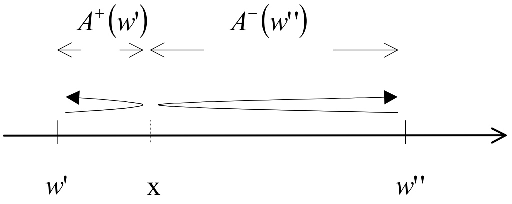


Figure 3: The attraction basins between two points in the support

There should exist a positive probability that agents in w' due to exogenous perturbations will divert away and out of $A^{+}(w')$ into $A^{-}(w'')$. This probability will be proportional to the probability mass of agents that can divert and inversely proportional to how far they have to divert in order not to converge back to the support. The same reasoning also applies for w''. The process of agents diverting away from the support will clearly never stop, but the distribution itself will asymptotically converge to a "*stable*" state, given some δ, where the expected mass of agents loss and gain, in each time period, at every support is balanced.

Even though this is the state whereat every discrete distribution eventually will converge as $t \rightarrow \infty$, it imposes a very harsh restriction on the stability of distributions since it could take an infinite time before it converges. On the other hand, it should be noted that the conditions for stability stated in definition 2 result in a set of stable distributions that includes distributions with an extremely large net flow between two adjacent points in the support, (see figure 4). Since any discrete distribution F such that $F \in \mathbb{F}_{\mathbf{s}}$ will become less stable the higher the net flow gets, it would be desirable to exclude the distributions with the most extreme net flow.

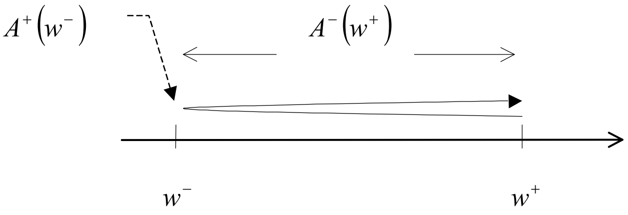


Figure 4: Example on a $F \in \mathbb{F}$ where $A^{-}(w^{+}) >> A^{+}(w^{-})$.

Another rationale can be found if the distribution is viewed from an evolutionary perspective. In an evolutionary environment it is a standard assumption that a higher payoff is associated with a higher growth and vice versa. From the case in figure 4, we should consequently expect the population to grow in a higher rate at w^{+} than in w^{-}. Depending on the density of each support, we should expect the evolutionary forces to affect the distribution in possible two ways. If the density of the support at w^{+} is higher than at w^{-}, the density at w^{-} will eventually diminish and become close to zero as $t \rightarrow \infty$. In the opposite case when the density at w^{-} is higher than at w^{+}, we should expect the support to gradually move towards w^{-}. At some point this distribution will become unstable according to definition 2 by the emergence of a stable point above w^{+}.

Let us therefore introduce a fix K such that $K \in (1, \infty)$ and define a slightly more restrictive form of stability.

Definition 3 (*K-Stability*) *Every distribution* $F(W; \mathbf{s})$ *where* $F \in \mathbb{F}_{\mathbf{s}}$ *such that:*

Condition 3a: *If* $\exists w' \in \operatorname{supp} F$ *such that* $(w', w) \cap \operatorname{supp} F = \emptyset \Rightarrow$ $\frac{\max\left(A^{+}\left(w'\right), A^{-}(w)\right)}{\min\left(A^{+}(w'), A^{-}(w)\right)} < K$

Condition 3b: *If* $\exists w'' \in \operatorname{supp} F$ *such that* $(w, w'') \cap \operatorname{supp} F = \emptyset \Rightarrow$ $\frac{\max\left(A^{+}(w), A^{-}\left(w''\right)\right)}{\min\left(A^{+}(w), A^{-}(w'')\right)} < K$

$\Leftrightarrow$

$w \in \operatorname{supp} F$

is K-stable *for some given* K *such that* $K \in (1, \infty)$.

In other words, the set of *K-stable* distributions of the incumbent strategy $\mathbf{s}$ (henceforth denoted as $\mathbb{F}_{\mathbf{s}K}$) is a subset of the set of stable distribution $\mathbb{F}_{\mathbf{s}K} \subseteq \mathbb{F}_{\mathbf{s}}$, where every *K-stable* distributions $F(W; \mathbf{s}) \in \mathbb{F}_{\mathbf{s}K}$ is characterized by that the ratio between the attraction basins in between two adjacent points in the support is upward bounded.

2.5 Stable Distributions in Status

In order to simplify the analysis of the model, let us begin to investigate if there exists some general quality associated with the stable distributions in status. Moreover let $w^{-} \equiv \inf \operatorname{supp} F$ and $w^{+} \equiv \sup \operatorname{supp} F$; also let $\gamma' \equiv \frac{1+\alpha+\beta}{4}$ and $\gamma^{*} \equiv \frac{\beta}{1+\beta-\alpha}$.

Two qualitatively different cases can be found whereto the distribution of status in the population converges, when $\gamma = \gamma^{*}$and $\gamma = \gamma'$.

2.5.1 Stable Distributions of status given $\beta > \gamma + \delta$

For the case of $\beta > \gamma + \delta$, it can be shown that two adjacent points belonging to the support of a stable distribution will have a distance of at least δ. Hence, no continuous distribution can belong to the set of stable distributions.

Claim 1 *The distance between two adjacent points in the support is characterized by* $\left|w^{j}-w^{j+1}\right| \geq \delta \ \forall F$ *such that* $F \in \mathbb{F}_{\mathbf{s}}$ *given* $\beta>\gamma+\delta$ *.*

Proof. See Appendix ■

Corollary 1 *There does not exist any continuous distribution* F *such that* $F \in \mathbb{F}_{\mathbf{s}}$ *for* $\beta>\gamma+\delta$ *.*

From definition 3 it follows that for any given pair (δ, K) there exists a maximum distance between two adjacent points in the support. Let this distance be denoted as d such that $\left|w^{j}-w^{j+1}\right| \equiv d$. It follows directly from the definition that the maximum distance d is increasing by δ. Since the upper support w^{+} in any stable distribution will be $w^{+}>1-\delta$ for *Villain strategy* and $w^{+}>\beta-\delta$ for *Robin Hood strategy*, the probability mass p in the support will decrease as δ decreases.

Claim 2 $\frac{\partial d}{\partial \delta}>0$.

Proof. See Appendix ■

Corollary 2 $\frac{\partial p}{\partial \delta}>0, \forall p\left(w^{j}\right) \in \operatorname{supp} F$ *such that* $F \in \mathbb{F}_{\mathbf{sK}}$.

Now consider the stable distributions by the conditioned strategies.

Definition 4 *Let* $\widehat{\delta}$ *be the distance between adjacent points in the support such that* $\delta \leq \widehat{\delta} \leq d$ *given* $\beta>\gamma+\delta$ *.*

Stable Distributions of *Robin Hoods* In a population where the agents condition their strategies according to the *Robin Hood strategy*, the following relation has to be fulfilled in order for the distribution to be *K-stable*:

$$\pi\left(w^{j+1} ; \mathcal{R}, F\right)-\pi\left(w^{j} ; \mathcal{R}, F\right)=\widehat{\delta} \tag{2}$$

This can be rewritten as:

$\beta p\left(w^{j}\right)+\gamma p\left(w^{j+1}\right)-\gamma p\left(w^{j}\right)-p\left(w^{j+1}\right)=\widehat{\delta}$

$\Leftrightarrow$

$p\left(w^{j}\right)(\beta-\gamma)-p\left(w^{j+1}\right)(1-\gamma)=\widehat{\delta}$

$\Leftrightarrow$

$p\left(w^{j}\right)=p\left(w^{j+1}\right) \frac{1-\gamma}{\beta-\gamma}+\frac{\widehat{\delta}}{\beta-\gamma}$

Thus, the probability mass p in each support is decreasing in the argument w^j of the distribution.

Claim 3 *$p\left(w^j\right) > p\left(w^{j+1}\right), \forall F \in \mathbb{F}_{\mathbf{s}\mathrm{K}}$ given $(\mathbf{s}, F)$ such that $\mathbf{s} = \mathcal{R}$.*

Proof. See Appendix ■

Since the probability mass is decreasing exponentially in the support, it follows that the probability mass at w^- must exceed $\frac{1-\beta}{1-\gamma}$ for every $\widehat{\delta} > 0$.

Claim 4 *$p\left(w^-\right) > \frac{1-\beta}{1-\gamma}$ for $F\left(w; \mathcal{R}\right) \in \mathbb{F}_{\mathbf{s}\mathrm{K}}$.*

Proof. See Appendix ■

Also note that the probability mass at upper support $p\left(w^+\right) \to 0$ as $\delta \to 0$.

Claim 5 *$\forall F \in \mathbb{F}_{\mathbf{s}\mathrm{K}}$ for $(\mathbf{s}, F)$ such that $\mathbf{s} = \mathcal{R} \rightsquigarrow 0 < p\left(w^+\right) \leq \frac{\delta}{1-\gamma}$.*

Proof. See Appendix ■

Stable Distributions of *Villains* In a population where the agents condition strategies according to the *Villain strategy*, the following relation has to be fulfilled:

$$\pi\left(w^{j+1}; \mathcal{V}, F\right) - \pi\left(w^j; \mathcal{V}, F\right) = \widehat{\delta} \tag{3}$$

This expression implies $p\left(w^j\right) + \gamma p\left(w^{j+1}\right) - \gamma p\left(w^j\right) - \beta p\left(w^{j+1}\right) = \widehat{\delta}$ which can be rewritten as $p\left(w^j\right) = p\left(w^{j+1}\right) \frac{\beta-\gamma}{1-\gamma} + \frac{\widehat{\delta}}{1-\gamma}$, which in turn implies that the probability mass p is increasing in the argument of the distribution as $\delta \to 0$. However from definition 2 it follows that the probability mass at the upper support $p\left(w^+\right)$ has to fulfill $0 < p\left(w^+\right) \leq \frac{\delta}{1-\gamma}$ for any stable distribution, hence also every *K-stable* distribution. Thus it follows directly that the probability mass $p\left(w\right) \to 0$ at every support that belongs to the distribution as $\delta \to 0$.

Claim 6 *$p\left(w^j\right) < p\left(w^{j+1}\right), \forall F \in \mathbb{F}_{\mathbf{s}\mathrm{K}}$ given $(\mathbf{s}, F)$ such that $\mathbf{s} = \mathcal{V}$ when $\delta \to 0$ for $\beta > \gamma + \delta$.*

Proof. See Appendix ■

Claim 7 *$\forall F \in \mathbb{F}_{\mathbf{s}\mathrm{K}}$ for $(\mathbf{s}, F)$ such that $\mathbf{s} = \mathcal{V} \rightsquigarrow 0 < p\left(w^+\right) \leq \frac{\delta}{1-\gamma}$.*

Proof. See Appendix ■

Corollary 3 *$\lim_{\delta \to 0} p(w) = 0, \forall w \in \operatorname{supp} F$ such that $F(w; \mathcal{V}) \in \mathbb{F}_{\mathbf{s}\mathrm{K}}$ for $\beta > \gamma + \delta$.*

From the claims above, the following property follows directly.

Property 1 *$\overline{\pi}(\mathcal{R}, F) < \overline{\pi}(\mathcal{V}, F)$ given $\beta > \gamma + \delta$ and $F \in \mathbb{F}_{\mathbf{s}\mathrm{K}}$ as $\delta \to 0$.*

2.5.2 Stable Distributions of status given $\beta \leq \gamma + \delta$

When $\beta \leq \gamma + \delta$, every distribution in which all agents condition their strategies according to the *Robin Hood strategy* will end up as a one-point distribution.

Property 2 *The only $F \in \mathbb{F}_{\mathbf{s}}$ for the pair $(\mathbf{s}, F)$ such that $\mathbf{s} = \mathcal{R}$ and $\beta \leq \gamma + \delta$ is a one-point distribution at $w = \gamma$.*

If the agents instead were to condition their strategies according to the *Villain strategy* it follows directly that a one-point distribution is not stable according to definition 2. It can also be shown that there are no stable continuous distributions in existence with this strategy. However the population can, as shown below, converge to discrete distributions where the distance between two arbitrary adjacent points belonging to the support is arbitrarily small. Consider two adjacent points w^j and w^{j+1} such that $w^j, w^{j+1} \in \operatorname{supp} F$ and $\left(w^j, w^{j+1}\right) \cap \operatorname{supp} F = \emptyset$ and $w^j, w^{j+1} \neq w^-, w^+$. Then consider the following distances between these two adjacent points:

- $\left|w^{j+1} - w^j\right| > 2\delta$

This can obviously not be a stable distribution, since a there exist a point w such that $w^j + \delta < w < w^{j+1} - \delta$ where $\pi(w) = w$.

- $\delta \leq \left|w^{j+1} - w^j\right| \leq 2\delta$

In this case, there might for a given (α, β) exist a point w such that $w^{j+1} - \delta < w < w^j + \delta$ where $\pi(w) = w$.

- $\left|w^{j+1} - w^j\right| < \delta$

The agents in w^{j+1} and w^{j} are unable to identify each others status, why the difference in expected payoff stems from having w^{j+1} unable to identify some $w > w^{j+1}$, whereas w^{j} can, and/or that w^{j+1} is able identify some $w < w^{j}$, whereas w^{j} is not. It follows directly that every adjacent support can be arbitrarily close; hence the following property.

Property 3 *The distance between two adjacent points in the support is characterized by* $\left|w^{j+1} - w^{j}\right| > 0 \; \forall F$ *such that* $F \in \mathbb{F}_{\mathbf{s}}$ *given* $\beta \leq \gamma + \delta$ *.*

Since the probability mass is determined in the same way as when $\beta > \gamma + \delta$, i.e. $\pi\left(w^{j+1}; \mathcal{V}, F\right) - \pi\left(w^{j}; \mathcal{V}, F\right) = p\left(w^{j-n}\right) + \gamma p\left(w^{j+n}\right) - \gamma p\left(w^{j-n}\right) - \beta p\left(w^{j+n}\right)$, we know that $p\left(w^{j-n}\right) < p\left(w^{j+n}\right) \rightsquigarrow p\left(w^{j}\right) < p\left(w^{\;j+1}\right)$. Consequently we could make the same conclusion as in claim 3, leading to the following property:

Property 4 $\lim_{\delta \to 0} p\left(w\right) = 0, \forall w \in \operatorname{supp} F$ *such that* $F\left(w; \mathcal{V}\right) \in \mathbb{F}_{\mathbf{s}\mathrm{K}}$*.*

Any stable distribution in status by this population must have a larger support distributed over $(\beta, 1)$. The expected payoff by a stable distribution in a population with agents exclusively conditioning their strategies according to the *Robin Hood strategy* is less than in any population with agents exclusively conditioning their strategies through the *Villain strategy.*

Property 5 $\overline{\pi}\left(\mathcal{R}, F\right) < \overline{\pi}\left(\mathcal{V}, F\right)$ *given* $\beta \leq \gamma + \delta$ *and* $F \in \mathbb{F}_{\mathbf{s}}$*.*

3 Evolutionary Criteria

The question whether a strategy could be invaded becomes harder to determine in this type of modeling, than in models with exclusively non-conditioned strategies. Note that the rationale behind the equilibrium concept of the standard *ESS* model, as defined by Maynard Smith, is built on the following assumption: agents associated with a strategy that reveals a higher expected payoff are more frequent in the population next period than agents with strategies that reveals a lower payoff. The implementation of this idea is not as straightforward in the model described above as in the standard *ESS* model. The added difficulty stems from that the strategy in this model is usually associated with a heterogeneous payoff, which leaves the question for how the population grows indeterminate.

Consequently, this unfortunate circumstance makes the standard *ESS* model indeterminate as to whether invasions are possible in cases where a mutant strategy reveals a higher payoff against some fraction of the incumbent strategy in the population, and lower payoff against the remaining fraction. There will be a need for a modified *ESS* concept, both in relation to the growth of the population and in relation to when a mutant strategy invades the incumbent strategy, in order for it to be applicable to this type of modeling.

3.1 Restrictions

In contrast to the standard evolutionary model, there are three simultaneous dynamic processes in this model: *growth of the population*, *mutations*, and the *distribution of the population*.

3.1.1 Growth of the Population

Let us begin by making the following crucial assumption.

Assumption 2 *The growth in every support of the distribution is proportional to payoff.*

3.1.2 Mutations

Let us assume that the incumbent distribution of strategies at each period with probability μ is subjected to a minor invasion of a mutant strategy m, and that with a probability $(1-\mu)$ the agent's behavior stays fixed. In case of invasion, let ε of the population change their behavior and instead play the game according to some other strategy m from the enriched pure strategy set $\mathbf{S}$, and $1-\varepsilon$ of the population remain with the incumbent strategy.

Moreover, if two different strategies are resulting in the same realized behavior and thus same payoff, it should be more likely that the less complex strategy is used since a more complex strategy should be considered as more costly.

Assumption 3 *If m is less complex than m' and $\pi(m, w; \mathbf{s}, F) = \pi(m', w; \mathbf{s}, F)$ will the agents use the m strategy.*

A conditioned strategy is regarded as more complex than an unconditioned strategy; a mixed strategy is regarded as a more complex than a pure strategy.

3.1.3 Distribution of the Population

Let μ be so small that growth in the population, due to differences in expected payoff, affects the distribution to a much higher degree than mutations do. For simplicity, assume that the populations will converge into a *K-stable* distribution before any mutations take place.

Assumption 4 *The adjustment process of the distribution is much faster than the growth process, which in turn is much faster than the process of mutations.*

3.2 Single Support Mutations

A criterion for evolutionary stability, which comprises the concept of *ESS*, would be that there does not exist any alternative strategy $m \in \mathbf{S} \backslash s$ that will do just as well or better than any support $w \in \operatorname{supp} F(\mathbf{s})$ of the incumbent strategy $\mathbf{s}$. That is:

Definition 5 (Criterion $\mathfrak{A}$) *The pair $(\mathbf{s}, F)$ where $F \in \mathbb{F}_{\mathbf{s}}$ will meet criterion $\mathfrak{A}$ for surviving a mutant strategy, $m \in \mathbf{S} \backslash s \Leftrightarrow$*

$$\pi(m, w; \mathbf{s}, F) < w \text{ for } \forall w \geq w^{-} \tag{4}$$

Let such pairs henceforth be denoted as $(\mathbf{s}, F) \in \mathbb{F}_{\mathbf{s}}^{\mathfrak{A}}$.

If there exists any alternative strategy $m \in \mathbf{S} \backslash s$ such $\pi(m, w; \mathbf{s}, F) \geq \pi(s, w; \mathbf{s}, F)$, it can be assumed that the agents through the evolutionary process in assumption 4, by mutation adjust their behavior until criterion $\mathfrak{A}$ is fulfilled.

3.3 Cluster Mutations

An alternative approach on how to compare a mutated strategy with the incumbent strategy would be to examine how an arbitrarily small randomly drawn proportion ε would perform on average against the initial population; i.e. to see how well the agents would do on average in a given distribution if they were to mutate.

Assumption 5 *The mutations are identically distributed as the incumbent strategy*

Given that other economic engagements and coordination failures combined bring about enough perturbations to the support of F, F can be seen to have a dual interpretation. Besides from being a description of how the status is "stably" distributed in a continuum of identical agents, it can also be interpreted as the probability over time whereto an arbitrary agent will be converging when playing this coordination game.

It follows that the probability for an arbitrary agent, playing the game according to the incumbent strategy mix $\mathbf{s}$, to converge to $w \in \operatorname{supp} F$ at any given time is equal to $p\left(s, w; \mathbf{s}, F\right)$.

Definition 6 *Let $p\left(s, w; \mathbf{s}, F\right)$ be the probability mass at w given a stable distributed population exclusively using the incumbent strategy $\mathbf{s}$.*

Consequently, the average payoff over time should be a highly relevant factor given that mutations are events that occur very seldom in comparison with exogenous economic events. The average payoff over time for the incumbent strategy can be denoted as:

$$\overline{\pi}\left(s; \mathbf{s}, F\right) = \textstyle\int_{w=w^-}^{w^+} p\left(s, w; \mathbf{s}, F\right) \pi\left(s, w; \mathbf{s}, F\right) \tag{5}$$

It is also necessary to take into account that an agent with a mutant strategy m might divert to a different stable status level than an agent with the incumbent strategy s, also when they initially enter the game at the same status level.

Definition 7 *Let $p\left(m, w; \mathbf{s}, F\right)$ be the probability mass at w for an infinitesimal part ε of the population using a mutant strategy m in a stable distributed population that is exclusively using strategy s.*

The average payoff over time for a mutant strategy would be:

$$\overline{\pi}\left(m; \mathbf{s}, F\right) = \textstyle\int_{w=w^-}^{w^+} p\left(m, w; \mathbf{s}, F\right) \pi\left(m, w; \mathbf{s}, F\right) \tag{6}$$

Definition 8 (Criterion $\mathfrak{B}$) *The pair* $(\mathbf{s}, F)$ *where* $F \in \mathbb{F}_{\mathbf{s}}$ *will meet criterion* $\mathfrak{B}$ *for surviving a mutant strategy,* $m \Leftrightarrow$

$$\overline{\pi}\left(m; \mathbf{s}, F\right) < \overline{\pi}\left(s; \mathbf{s}, F\right) \tag{7}$$

Let such pairs henceforth be denoted as $(\mathbf{s}, F) \in \mathbb{F}_{\mathbf{s}}^{\mathfrak{B}}$*.*

3.4 Group Evolution

Yet another possible evolutionary selection process can be found when stable distributions of strategies that fulfill criterion $\mathfrak{A}$ evolve separately in different isolated populations. For now, assume that the feasible distributions that are stable according to criterion $\mathfrak{A}$, are uniformly distributed over the different isolated populations. In addition to the conventional assumption that a higher payoff implies a larger growth, it can also be assumed that a higher average payoff is associated with a larger growth of the population. One can imagine that these different isolated populations, as they grow larger, sooner or later will come in contact with each other. When they do, they are immediately united into one pooled population. It is proposed that the probability for a population of encountering another population will be proportional to its' growth rate.

The fusion of these populations will converge to a stable distribution $F_{s'}\left(\mathbf{s}'\right)$, $F_{s''}\left(\mathbf{s}''\right) \rightsquigarrow \mathbf{F}\left(\mathbf{s}'\mathbf{s}''\right)$ where $\mathbf{F} \in \mathbb{F}_{\mathbf{s}K}$ and $\left(\mathbf{s}', F_{s'}\right), \left(\mathbf{s}'', F_{\mathbf{s}''}\right) \in \mathbb{F}_{\mathbf{s}}^{\mathfrak{A}}$. For simplicity, let s belong to the incumbent strategy $\mathbf{s}$ $\forall t$. According to assumption 4, the pooled distribution converges to a stable distribution before any growth of the population takes place, whereafter the agents adjust their behavior if $\pi\left(s, w; \mathbf{s}, \mathbf{F}\right) \leq \pi\left(m, w; \mathbf{s}, \mathbf{F}\right)$, for $s \in \mathbf{S}$,$m \in \mathbf{S} \backslash s$. Let the resulting distribution where $\nexists\left(s, w\right) \in \operatorname{supp} \mathbf{F}$ such that $\pi\left(s, w; \mathbf{s}, \mathbf{F}\right) \leq \pi\left(m, w; \mathbf{s}, \mathbf{F}\right)$ be called *strategically stable* and denoted $\widetilde{\mathbf{F}}$.

Definition 9 (Criterion $\mathfrak{C}$) *The pair* $\left(\mathbf{s}, \widetilde{\mathbf{F}}\right)$ *where* $\widetilde{\mathbf{F}} \in \mathbb{F}_{\mathbf{s}\mathrm{K}}$ *will meet criterion* $\mathfrak{C}$ *for surviving an alternative strategy,* $m \in \mathbf{S} \backslash s \Leftrightarrow$

$$\max \pi\left(m, \cdot; \mathbf{s}, \widetilde{\mathbf{F}}\right) \leq \min \pi\left(s, \cdot; \mathbf{s}, \widetilde{\mathbf{F}}\right) < \max \pi\left(s, \cdot; \mathbf{s}, \widetilde{\mathbf{F}}\right)$$

Note that criterion $\mathfrak{C}$ in essence is criterion $\mathfrak{A}$ applied twice; first *within* a population and then *between* populations.

4 Evolutionary Stability

Let us now investigate whether some combination from the enriched pure strategy set **S** fulfills any of the criteria defined above, as $\delta \rightarrow 0$. It should be noted that when $\gamma = \gamma'$, i.e. when the agents use a mixed strategy where they assign equal probability to H and D, we exclude the possibility that the agents can use any equilibrium strategy when $|w_i - w_{-i}| < \delta$.

4.1 Stable Strategies in Single Mutations

In relation to different distributions and combination of strategies for which stability according to criterion $\mathfrak{A}$ in definition 5 can be found, a cumbersome process of going through any feasible combination of strategies and distributions can be foreseen. Instead, consider the following: in this study we try to investigate whereto a population playing the game in 1above converges in the long term. It would be fair to assume that basic behaviors are developed over extremely long periods; i.e. from the creation of a population until the population eventually becomes very large or almost infinite.

Initially, at the creation of a population, we should expect (given the restrictions above) that the agents behavior converge rapidly towards some equilibrium; either to an unconditioned equilibrium, *MSS*, or to a close neighborhood of a homogenous population at a conditioned equilibrium as defined in criterion $\mathfrak{A}$. Any other distribution of behavior than the above mentioned would lead to a redistribution in the relative proportion of strategies and should consequently not be regarded as *K-stable*.

In order to capture the full dynamic of this process correctly, we would have to model the agents through a discrete model. However, since we have in this model (in order to simplify the calculations) assumed a continuum of agents, we are compelled to make a conjecture.

Conjecture 1 *The only feasible K-stable distributions of behavior that could satisfy criterion* $\mathfrak{A}$*, are populations that evolve from either the unconditioned equilibrium,* MSS*, or a homogenous population utilizing one pure conditioned strategy.*

Hence it will be sufficient to investigate the cases when the incumbent strategy consists of a homogenous conditioned strategy and when we have a convex combination of unconditioned strategies such as *MSS*.

4.1.1 ***Robin Hood Strategy*** as Incumbent Strategy

Case: $\beta > \gamma' + \delta$

It is easy to verify that the *Robin Hood strategy* is vulnerable for invasion when $\beta > \gamma' + \delta$.

Claim 8 $\mathcal{R}$ *is according to criterion* $\mathfrak{A}$ *vulnerable for invasion* $\forall \alpha, \beta \in (0,1)$, *given* $\beta > \gamma + \delta$ *for* $\gamma = \gamma'$.

Proof. See Appendix ■

Note that any invasion is restricted to either a D invasion at the upper support w^{+}, or an H invasion at the lower support w^{-}. It can be shown there exists a slightly corrupted population of the homogenous *Robin Hood strategy*, denoted $\mathcal{R}^{c}$, that consists almost exclusively of agents conditioning the strategies according to *Robin Hood strategy* besides from either the upper w^{+} or lower support w^{-} at the distribution and that this population cannot be invaded by any other strategies.

Claim 9 $\exists \mathcal{R}^{c}$ *such that this population according to criterion* $\mathfrak{A}$ *is* not *vulnerable for any invasion, given* $\beta > \gamma' + \delta$ *at the limit when* $\delta \rightarrow 0$.

Proof. See Appendix ■

Case: $\boldsymbol{\beta} > \boldsymbol{\gamma}^{*} + \boldsymbol{\delta}$

Assume that the population uses the *MSS* instead of an equal randomization between the strategies when $|w_i - w_{-i}| < \delta$. We find that even though the *Robin Hood strategy* can be invaded, there exists a $\mathcal{R}^{c}$ with an identical realized behavior of each agent as when the entire population condition through *Robin Hood strategy*, that are not vulnerable for invasion.

Claim 10 $\mathcal{R}$ *is according to criterion* $\mathfrak{A}$ *vulnerable for invasion by* MSS, *given* $\beta > \gamma + \delta$ *for* $\gamma = \gamma^{*}$.

Proof. See Appendix ■

Consider a population where the agents are conditioning the game according to *Robin Hood strategy* in every support except for w^- and w^+. Assume that $\frac{1-\alpha}{1+\beta-\alpha}$ of the agents in w^- are playing *Hawk*, and $\frac{\beta}{1+\beta-\alpha}$ of the agents in w^+ are playing *Dove*. It then follows that every agent in this population will behave identical to the *Robin Hood strategy*, but this distribution will not be vulnerable for invasion by any mutant strategy.

Claim 11 *$\exists\mathcal{R}^c$ with identical realized behavior as $\mathcal{R}$, that according to criterion $\mathfrak{A}$ is* not *vulnerable for invasion, given $\beta > \gamma + \delta$ for $\gamma = \gamma^*$.*

Proof. See Appendix ■

Case: $\boldsymbol{\beta \leq \gamma + \delta}$

From claim 2 we know that if the incumbent strategy is a homogenous population coordinating according to *Robin Hood strategy*, then the distribution will converge towards a single point.

If $\boldsymbol{\gamma = \gamma'}$, it follows directly according to definition 5 that this distribution can be partly invaded by *MSS*. Furthermore from assumption 3 it follows that it can also be completely invaded since the pure strategies have lower complexity than a randomized strategy.

In the case of $\boldsymbol{\gamma = \gamma^*}$ it follows from assumption 3 that this strategy will be completely invaded by *MSS* since a conditioned strategy has a higher complexity than a pure strategy.

Property 6 *$\mathcal{R}$ is according to criterion $\mathfrak{A}$ vulnerable for complete invasion by* MSS, *given $\beta \leq \gamma + \delta$ when $\boldsymbol{\gamma = \gamma^*}$ and $\boldsymbol{\gamma = \gamma'}$.*

4.1.2 *Villain Strategy* as Incumbent Strategy

Case: $\boldsymbol{\gamma = \gamma'}$

Similar to the case above with the *Robin Hood strategy*, it is easy to see that also the *Villain strategy* is vulnerable for invasion when $\boldsymbol{\gamma = \gamma'}$.

Claim 12 *$\mathcal{V}$ is according to criterion $\mathfrak{A}$ vulnerable for invasion $\forall\alpha, \beta \in (0, 1)$, given $\gamma = \gamma'$.*

Proof. See Appendix ■

Neither the *Villain strategy* can be completely invaded by mutant strategies, since any invasion is restricted to either an H invasion at the upper, or a D invasion at the lower support. Hence there exists a slightly corrupted population of the homogenous *Villain strategy*, denoted $\mathcal{V}^c$, that cannot be invaded by any other strategies.

Claim 13 *$\exists\mathcal{V}^c$ such that this population according to criterion $\mathfrak{A}$ is* not *vulnerable for any invasion, given $\boldsymbol{\gamma} = \boldsymbol{\gamma}'$ when $\delta \to 0$.*

Proof. See Appendix ■

Case: $\boldsymbol{\gamma} = \boldsymbol{\gamma}^*$

Now consider the case when a population uses the *MSS* instead of an equal randomization between the strategies when $|w_i - w_{-i}| < \delta$.

Claim 14 *$\mathcal{V}$ is according to criterion $\mathfrak{A}$ vulnerable for invasion by* MSS, *given $\gamma = \gamma^*$.*

Proof. See Appendix ■

Just as in the case above with the *Robin Hood strategy*, consider a population where the agents are conditioning the game exclusively according to *Villain strategy* in every support except for w^- and w^+. Assume that $\frac{\beta}{1+\beta-\alpha}$ of the agents in w^- is playing *Dove* and $\frac{1-\alpha}{1+\beta-\alpha}$ of the agents in w^+ is playing *Hawk*. It then follows that every agent in this population will behave identically to the *Villain strategy* , but this distribution will not be vulnerable for invasion by any mutant strategy.

Claim 15 *$\exists\mathcal{V}^c$ with identical realized behavior as $\mathcal{V}$, that according to criterion $\mathfrak{A}$ is* not *vulnerable for invasion, given $\gamma = \gamma^*$.*

Proof. See Appendix ■

That is, even though the *Villain strategy* can be invaded, there exists a slightly corrupted population of the homogenous *Villain strategy*, $\mathcal{V}^c$, with an identical realized behavior as the homogenous *Villain strategy*, that is not vulnerable for invasion.

4.1.3 *MSS* as incumbent strategy

Since *MSS* is one-point distributed, and there does not exist a mix of strategies from the enriched pure strategy set **S** that reveals a higher payoff, it follows directly from assumption 3 that it is not vulnerable for invasion.

Property 7 MSS *is according to criterion* $\mathfrak{A}$ not *vulnerable for invasion.*

4.2 Stable Strategies in Cluster Mutations

4.2.1 *Robin Hood Strategy* as Incumbent Strategy

Case: $\boldsymbol{\beta > \gamma' + \delta}$

Since any invasion of the *Robin Hood strategy* according to criterion $\mathfrak{A}$ is restricted to either the upper support or the lower support, it can be showed that this strategy is fulfill criterion $\mathfrak{B}$ in definition 8 for evolutionary stability.

Claim 16 $\mathcal{R}$ *is according to criterion* $\mathfrak{B}$ not *vulnerable for invasion by any convex combination of pure strategies or mixed strategies for any* F *as* $\delta \to 0$*, given* $\beta > \gamma' + \delta$*.*

Proof. See Appendix ■

Corollary 4 $\mathcal{R}^c$ *is according to criterion* $\mathfrak{B}$ not *vulnerable for invasion by any convex combination of pure strategies or mixed strategies for any* F *as* $\delta \to 0$*, given* $\beta > \gamma' + \delta$*.*

Case: $\boldsymbol{\beta > \gamma^* + \delta}$

In this case too, we have that invasions are restricted to either the upper support or the lower support why the *Robin Hood strategy* fulfills criterion $\mathfrak{B}$ for evolutionary stability.

Claim 17 $\mathcal{R}$ *is according to criterion* $\mathfrak{B}$ not *vulnerable for invasion by any convex combination of pure strategies or mixed strategies for any* F*, given* $\beta > \gamma^* + \delta$*.*

Proof. See Appendix ■

Corollary 5 *$\mathcal{R}^c$ is according to criterion* $\mathfrak{B}$ not *vulnerable for invasion by any convex combination of pure strategies or mixed strategies for any F as $\delta \to 0$, given $\beta > \gamma^* + \delta$.*

Case: $\boldsymbol{\beta \leq \gamma + \delta}$

Since *Robin Hood strategy* in this case is one point distributed we have that it could be invaded by e.g. *Villain strategy.* Moreover, since there exists a strategy mix with less complex pure strategies that reveals equal or higher payoff, it follows directly from assumption 3 that the *Robin Hood strategy* is vulnerable for invasion.

Property 8 *$\mathcal{R}$ is according to criterion $\mathfrak{B}$ vulnerable for invasion by some strategy, given $\beta \leq \gamma + \delta$ for both γ' and γ^*.*

4.2.2 *Villain Strategy* as Incumbent Strategy

In analogy with 4.2.1, we find that any invasion of the *Villain strategy* according to criterion $\mathfrak{A}$ is restricted to the upper support and/or the lower support. Thus it can be shown that the *Villain strategy* fulfills criterion $\mathfrak{B}$ in definition 5 for evolutionary stability.

Case: $\boldsymbol{\gamma = \gamma'}$

Claim 18 *$\mathcal{V}$ is according to criterion* $\mathfrak{B}$ not *vulnerable for invasion by any convex combination of pure strategies or mixed strategies for any F as $\delta \to 0$, given $\gamma = \gamma'$.*

Proof. See Appendix ■

Corollary 6 *$\mathcal{V}^c$ is according to criterion* $\mathfrak{B}$ not *vulnerable for invasion by any convex combination of pure strategies or mixed strategies for any F as $\delta \to 0$, given $\gamma = \gamma'$.*

Case: $\boldsymbol{\gamma = \gamma^*}$

Claim 19 *$\mathcal{V}$ is according to criterion* $\mathfrak{B}$ not *vulnerable for invasion by any convex combination of pure strategies or mixed strategies for any F, given $\gamma = \gamma^*$.*

Proof. See Appendix ■

Corollary 7 $\mathcal{V}^c$ *is according to criterion* $\mathfrak{B}$ not *vulnerable for invasion by any convex combination of pure strategies or mixed strategies for any* F *as* $\delta \to 0$*, given* $\gamma = \gamma^*$*.*

4.2.3 *MSS* as incumbent strategy

Since *MSS* according to property 7 fulfills criterion $\mathfrak{A}$ for evolutionary stability, then property 9 follows directly.

Property 9 MSS *is according to criterion* $\mathfrak{B}$ not *vulnerable for invasion* $\forall \alpha, \beta \in (0,1)$*, given* $\gamma = \gamma^*$*.*

4.3 Stable Strategies through Group Evolution

Now let us examine how two isolated populations, both stable according to criterion $\mathfrak{A}$ but at different equilibria, would perform when they eventually come in contact with each other. Note that in this section, and henceforth, we just examine for when $\boldsymbol{\gamma} = \boldsymbol{\gamma}^*$. However the results for $\boldsymbol{\gamma} = \boldsymbol{\gamma}'$ are analogue.

4.3.1 *Robin Hood Strategy* versus *MSS*

Case: $\boldsymbol{\beta} > \boldsymbol{\gamma}^* + \boldsymbol{\delta}$

Proposition 1 $\mathcal{R}$ *is according to criterion* $\mathfrak{C}$ not *vulnerable for invasion by* MSS*, given* $\boldsymbol{\beta} > \boldsymbol{\gamma}^* + \boldsymbol{\delta}$ *as* $\delta \to 0$*.*

Proof. See Appendix ■

4.3.2 *Villain Strategy* versus *MSS*

Case: $\boldsymbol{\gamma} = \boldsymbol{\gamma}^*$

Proposition 2 $\mathcal{V}$ *is according to criterion* $\mathfrak{C}$ not *vulnerable for invasion by* MSS*, given* $\boldsymbol{\gamma} = \boldsymbol{\gamma}^*$ *as* $\delta \to 0$*.*

Proof. See Appendix ■

4.3.3 *Villain Strategy* versus *Robin Hood Strategy*

Case: $\boldsymbol{\beta} > \boldsymbol{\gamma}^{*}+\boldsymbol{\delta}$

Proposition 3 *$\mathcal{V}$ is according to criterion $\mathfrak{C}$* not *vulnerable for invasion by $\mathcal{R}$, given $\boldsymbol{\beta} > \boldsymbol{\gamma}^{*}+\boldsymbol{\delta}$ as $\delta \to 0$.*

Proof. See Appendix ■

Corollary 8 *$\mathcal{V}$ is the* only *strategy* not *vulnerable for invasion under criterion $\mathfrak{C}$ for $m \in \mathbf{S} \backslash \mathcal{V}$, given $\boldsymbol{\gamma} = \boldsymbol{\gamma}^{*}$ as $\delta \to 0$.*

4.4 Conclusions

As shown above, we can find three qualitatively different equilibria as defined by criterion $\mathfrak{A}$. This set of equilibria is immeasurable with the standard *ESS* concept since the contents are of different dimensions. Moreover, the set of equilibria as defined in criterion $\mathfrak{A}$ is a subset of the larger set of equilibria as defined by criterion $\mathfrak{B}$.

Property 10 $\mathbb{F}_{\mathbf{s}}^{\mathfrak{A}} \subseteq \mathbb{F}_{\mathbf{s}}^{\mathfrak{B}}$

However, since criterion $\mathfrak{B}$ implies weaker restrictions than criterion $\mathfrak{A}$, and since we have already have found that no unique equilibrium exists by criterion $\mathfrak{A}$ alone, criterion $\mathfrak{B}$ should only be regarded to have a marginal interest.

Criterion $\mathfrak{C}$, on the other hand, results in a unique equilibrium strategy, given that the players are given choice to condition their strategies through status as an alternative to the unconditioned strategies. Since criterion $\mathfrak{C}$ is only criterion $\mathfrak{A}$ applied twice (first within a population and then between populations), property 11 follows directly.

Property 11 $\mathbb{F}_{\mathbf{s}}^{\mathfrak{A}} \supseteq \mathbb{F}_{\mathbf{s}}^{\mathfrak{C}}$

5 Coordination through Characteristics

5.1 The Basic Framework

In the previous chapter, a mechanism was suggested that will eventually result in that every agent will utilize the *villain strategy* in the general coordination game $\Gamma(\alpha, \beta)$, depicted in figure 1, given that they have the option to condition their strategies through status. However, so far we have not presented anything that will exclude the possibility that the agents might coordinate the game by e.g. the size or any other static variable.

Let us therefore construct a framework such that the strategies can be conditioned a general class of variables, including the static variables originally examined by Maynard Smith. Assume that every player i possesses a profile of characteristics $\mathbf{C}_i$ such that status is just one element in this set.

$$\mathbf{C}_i = \left\{c_i^1, c_i^2, c_i^3, ...\right\} \tag{8}$$

For convenience sake, let c_i^n and c_j^n represent the same characteristic n for $\forall i, j \in I$. The different characteristics c in the profile $\mathbf{C}$ could of course be of a great variety. In addition to status we could imagine e.g., gender, weight, wealth, and length. Clearly, any subset of the characteristics-profile $y \subseteq \mathbf{C}$ could, in principle, be utilized for conditioning of strategies.

Definition 10 *Let a static variable y^s be a subset of the characteristics-profile that does not include any of the status-bearing characteristics y^d, $y^s \subseteq \mathbf{C} \backslash y^d$, such that the "value" of the characteristics is independent of the payoff in the game.*

From this definition we can conclude that the dynamic variable y^d should consist of characteristics that are correlated to the payoff, e.g. type of living, car, and vacation habits.

Definition 11 *Let $F_{\mathbf{C}}(y)$ denote the proportion of the population that has a characteristics-profile that reveals a lower expected payoff.*

As previous, let $F_y(w)$ denote the proportion of the population that reveals a lower expected payoff. Also assume that every population conditioning their strategies by some variable y is stable according to definition 2, i.e. $F_y \in \mathbb{F}_{\mathbf{s}}$.

Conclusion 1 $F_y \in \mathbb{F}_{\mathbf{s}} \Rightarrow F_y(w) = F_{\mathbf{C}}(y), \forall i \in I.$

That is, a characteristics-profile associated with an expected high payoff is also associated with a high status.

5.1.1 Conditioning through Static Variables

The conditioning of the strategies can be performed by a static variable through a single characteristic such as "*larger & smaller*" or "*longer & shorter*" and also through a multiple characteristic such as first condition by "*male & female*". When two of the same gender meet, then condition by "*larger & smaller*". That is, the population first conditions their strategies by one characteristic. If both agents are unable to coordinate they instead condition their strategies by another characteristic. If that fails, they try yet another characteristic, and so on ad infinitum.

The output of the model is of course dependent on what order the different characteristics are used to condition the strategies of the by. Let y henceforth denote any arbitrary static variable whether it is a singleton or an ordered vector of characteristics. Denote the two conditioned strategies by static variable through a single characteristic c^n for $\mathcal{B}$ and $\mathcal{P}$.

Example 1 *If the variable consists of a single characteristic c^n that is gender; let e.g. $\mathcal{B}$ stand for playing* D *against females and* H *against males, whereas $\mathcal{P}$ stands for playing* H *against females and* D *against males.*

Example 2 *If the variable consists of a single characteristic c^n that is size; let e.g. $\mathcal{B}$ stand for playing* D *against opponents with a higher value $y_i^n \leq y_{-i}^n - \delta^n$ and* H *against agents with lower value $y_i^n \geq y_{-i}^n + \delta^n$, where δ^n is the limitation of the agents observational skills by this particular characteristic c^n. Consequently let the $\mathcal{P}$ stand for the mirrored strategy.*

5.1.2 Growth of the Population

Consider the growth in a population which is conditioning its' strategies by some static variable. In the previous chapter, we have assumed that $\delta \rightarrow 0$ since the variable the agents' are observing is the cumulative payoff. This implies that even though the distance in status is constant, the absolute difference in status between the agents increses by time. This is not the case when it comes to static variables. Even though we could assume that the observational skills become better by time, we would still find that a population conditioning their strategies by status is better at identifying their opponents.

Another possibility is that a population conditions the strategies by multiple characteristics so that the static variable y consists of a large ordered vector of characteristics. In this case, the proportion of coordination failures will gradually become smaller in the population as the ordered vector becomes larger. However, it should be fair to assume that such complex conditioning-rules come at a price. Hence we have:

Assumption 6 $\overline{\pi}\left(y; y, F_y\right) < \overline{\pi}\left(\mathcal{V}; \mathcal{V}, F_{\mathcal{V}}\right)$

5.2 Evolutionary Stability of Static Variables

Now let us consider strategies conditioned through an arbitrary static variable y^s and investigate whether some $(\mathbf{s}, F_y) \in \mathbb{F}_{\mathbf{s}}^{\mathfrak{A}}$ is also stable according to criterion $\mathfrak{C}$.

Assumption 7 *$\forall c \in y$ is i.i.d. in every population.*

Claim 20 $\overline{\pi}\left(\mathcal{B}; \mathcal{B}, F_y\right) = \overline{\pi}\left(\mathcal{P}; \mathcal{P}, F_y\right)$.

Proof. See Appendix ■

Corollary 9 *Any permutation of $\mathcal{B}$ and $\mathcal{P}$ by any arbitrary static variable y, such that y is an ordered vector of characteristics, reveals the same average payoff.*

Since any completely or partially mirrored conditioning-rule by the static variable y^s has the same average payoff, it should suffice to investigate whether an arbitrary conditioning-rule of some arbitrary static variable y^s will be stable according to criterion $\mathfrak{C}$ when encountering a population coordinating according to the *Villain strategy*. Let this arbitrary conditioning-rule be denoted $\mathcal{Z}$.

Proposition 4 *$\mathcal{V}^c$ is according to criterion $\mathfrak{C}$* not *vulnerable for invasion by any strategy conditioned on any static variable y^s, given $\boldsymbol{\gamma} = \boldsymbol{\gamma}^*$ as $\delta \rightarrow 0$.*

Proof. See Appendix ■

5.3 Evolutionary Stability of Hybrid Variables

Consider a population that conditions its' strategies through both static and dynamic variables as described above. That is, $\forall y \subseteq \mathbf{C}$. Let us denote this strategy as $\mathcal{Z}$.

Proposition 5 $\mathcal{V}^c$ *is according to criterion* $\mathfrak{C}$ not *vulnerable for invasion by any strategy conditioned on any variable* $y \subseteq \mathbf{C}$, *given* $\boldsymbol{\gamma} = \boldsymbol{\gamma}^*$ *as* $\delta \to 0$.

Proof. Analogous to Proposition 4. ■

However, it should be noted that the results above are driven by lemma 9, and that this lemma is powerless in some cases of time inconsistent strategies. It is possible to envision strategies conditioned on a set of variables $y(t)$ such that a new condition rule is implemented at each time period. In such cases, the conditions for lemma 9 would not be satisfied. The question whether strategies like this might invade, is considered to reflect an unlikely scenario and is left unanswered.

6 Summary

> "Beggars cannot be choosers about the source of their signal, or about its attractiveness compared with others that they can only wish were as conspicuous." Schelling [10]

6.1 Concluding Remarks

As it has been demonstrated above, there exist mechanisms that favor the *Villain strategy* over any other, conditioned or unconditioned, time consistent strategy. One can ask; are the results presented above not intuitively expected? In fact, the study does not aim to just demonstrate that the Villain strategy is the only evolutionarily stable strategy by some newly defined criteria, but rather emphasize the consequences hereof. First notice that in a setting where the status is partially private information, it should be in the agents interest to signal high status through e.g. consumption to ensure that their status is visible for every feasible opponent. It should be noted that this is consistent with the ideas of Veblen [14] who argued that the primary purpose for consumption was to signal one's rank in the society. Since no profound behavioral reason was mentioned for why agents should show rank, this model could, along with different variations of Spence's model, serve as a rationale for such behavior.

Secondly, and perhaps of even greater significance, the study shows that if the agents were to utilize the *Villain strategy* in simple coordination games, it is evolutionary favorable for an arbitrary agent to also have concerns for high status in other economic activities besides simple coordination games, given that this class of games has a positive probability to occur. In other words, in any strategic game the agents would have to maximize their own outcome at the same time as they minimized that of their opponent since the opponents outcome might affect the agents own outcome in a future coordination game. The concern for high status would clearly also enter other economic situations without a strategic component, such as in the choice of lottery. In a lottery, it would in the long run be favorable to consider the status effect from the different outcomes of the lottery, since it would have an effect on the agents outcome in a possible coordination game later. The qualitative difference between this model and models based on Spence's idea should be noted; the agents concerns for high status evolves endogenously in this model, whereas variations of Spence's model assume that high status comes cheaper for some individuals that are associated with some desirable quality.

In almost every economic analysis of the agents decision-making, there is an implicit assumption often made about independence between the situation analyzed and other activities. However, this study suggests that the concern for high status could, under realistic circumstances, evolve endogenously from the model and thus bring about dependence between different economic events. In order to avoid an inconsistent model in an evolutionary setting, we would thus either have to assume that the game in figure 1 never is played, or that the agents do not have the ability to condition their strategies in the game by personal characteristics. A third possible way to circumvent the making of an inconsistent model would be to assume that the effect of concern for high status is negligible, i.e. that the occurrence of coordination games as in figure 1, are very rare.

6.2 Suggested Further Studies

6.2.1 Evolutionarily Stable Mappings

In the present study, the agents are restricted to coordination of the game by one particular mapping as first suggested by Maynard Smith [12]. Even though this is a very natural way to condition strategies upon, the formal proof, showing that there does not exist any other evolutionary feasible mappings, is still lacking. Such proof would of course strengthen these findings considerably.

6.2.2 Discrete Model

A discrete model would make a more realistic model and would thus be preferable. It would provide a more adequate way of modeling the given situation, and would also provide an even greater spillover effect between the coordination game and other economic events than the continuous model, especially in small populations. With two agents meeting in a strategic game, the conflict between the agents is sharpened due to the fact that they at a later stage might be matched against each other in a coordination game as described above. Hence it would be in both agents' interest to not only maximize their own payoff but also make sure that the opposing agent receives a low payoff in a higher degree than in a large population.

6.2.3 Choosing Opponents in Games

In the modeling of economics it is often assumed that the agents are randomly matched with each other. But what would be the consequences if they could choose their opponent? The main framework of the present model could also be used in the analysis of how agents choose with whom the want to engage in a game. It would be evolutionarily stable for every agent to make sure that they are matched up with the agent that reveals the highest payoff for them. It is likely that this situation demands two contextual variables, status and reputation. A high status agent would, according to the model described above, demand more than a low status agent would. However, it should also be vital for the agent to have someone to interact with. If they fail to do so, they will eventually get a lower status.

7 Proofs

Claim 1 *The distance between two adjacent points in the support is characterized by* $\left|w^j - w^{j+1}\right| \geq \delta \; \forall F$ *such that* $F \in \mathbb{F}_{\mathbf{s}}$ *given* $\beta > \gamma + \delta$ *.*

Proof. Let $F \in \mathbb{F}_{\mathbf{s}}$. By definition we have that $w^- \in \operatorname{supp} F$.

1. Compare the associated payoff for w^- and $w^* \in (w^-, w^- + \delta)$. The associated payoff for the different status levels are equal except against opponents in the interval $[w^- + \delta, w^* + \delta)$ where an agent with status w^- in contrast to w^* will be able to coordinate the game and consequently receive a higher payoff. Hence we have that $\pi(w^-) \geq \pi(w^*) \rightsquigarrow w^* \notin \operatorname{supp} F$, i.e.

$$(w^-, w^- + \delta) \notin \operatorname{supp} F \tag{9}$$

2. Now compare the associated payoff for $w' \in \operatorname{supp} F$ and $(w^-, w') \cap \operatorname{supp} F = \emptyset$ and $w^* \in (w', w' + \delta)$. The only difference in the associated payoff for the different status levels is that the agent with status w' in contrast to w^* will be able to coordinate the game in the interval $[w' + \delta, w^* + \delta)$ and consequently receive a higher payoff. That is we have that $\pi(w') \geq \pi(w^*) \Rightarrow w^* \notin \operatorname{supp} F$, i.e.

$$(w', w' + \delta) \notin \operatorname{supp} F \tag{10}$$

3. Finally compare the associated payoffs for $w^{j+1} \in \operatorname{supp} F$ and $\left(w^j, w^{j+1}\right) \cap \operatorname{supp} F = \emptyset$ and $w^* \in \left(w^{j+1}, w^{j+1} + \delta\right)$. In the same manner as above we find that the only difference in the associated payoff between the agents is in the interval $\left[w^{j+1} + \delta, w^* + \delta\right)$ where agent with status w^{j+1} in contrast to w^* will be able to coordinate the game in the interval $\left[w^{j+1} + \delta, w^* + \delta\right)$ and consequently receive a higher payoff. That is we have that $\pi\left(w^{j+1}\right) \geq \pi(w^*) \Rightarrow w^* \notin \operatorname{supp} F$, i.e.

$$\left(w^{j+1}, w^{j+1} + \delta\right) \notin \operatorname{supp} F \tag{11}$$

Claim 1 follows by induction. ■

Claim 2 $\frac{\partial d}{\partial \delta} > 0$.

Proof. Follows from $\min\left(A^{+}(w), A^{-}(w'')\right) = \delta$ ■

Claim 3 *$p\left(w^{j}\right) > p\left(w^{j+1}\right), \forall F \in \mathbb{F}_{\mathbf{s}\mathrm{K}}$ given $(\mathbf{s}, F)$ such that $\mathbf{s} = \mathcal{R}$.*

Proof. We have that $p\left(w^{j}\right) = p\left(w^{j+1}\right)\frac{1-\gamma}{\beta-\gamma} + \frac{\widehat{\delta}}{\beta-\gamma} \rightsquigarrow \max\left(p\left(w^{j+1}\right) \mid p\left(w^{j}\right)\right)$ is given by the pair $(\sup\beta, \inf\alpha)$. But $(\sup\beta - \inf\gamma) < (1 - \inf\gamma) \Rightarrow p\left(w^{j}\right) > p\left(w^{j+1}\right), \forall\widehat{\delta} > 0$ ■

Claim 4 *$p(w^{-}) > \frac{1-\beta}{1-\gamma}$ for $F(w;\mathcal{R}) \in \mathbb{F}_{\mathbf{s}\mathrm{K}}$.*

Proof. We know that $\pi(w^{-};\mathcal{R},F) = \gamma p(w^{-}) + (1 - p(w^{-}))$,
and $\pi(w^{+};\mathcal{R},F) > \pi(w^{-};\mathcal{R},F)$. $\min p(w^{-})$ is given by $\sup\pi(w^{-};\mathcal{R},F)$.

$\min p(w^{-})$ is given by $\sup\pi(w^{+};\mathcal{R},F) > \sup\pi(w^{-};\mathcal{R},F) \Leftrightarrow$
$\beta > \sup\pi(w^{-};\mathcal{R},F) \rightsquigarrow \beta - 1 > \gamma p(w^{-}) - p(w^{-}) \Leftrightarrow \frac{1-\beta}{1-\gamma} < p(w^{-})$. ■

Claim 5 *$\forall F \in \mathbb{F}_{\mathbf{s}\mathrm{K}}$ for $(\mathbf{s}, F)$ such that $\mathbf{s} = \mathcal{R} \rightsquigarrow 0 < p(w^{+}) \leq \frac{\delta}{1-\gamma}$.*

Proof. From definition 3 it follows that the probability mass of the richest cluster must fulfill $p(w^{+})(\beta-\gamma) \leq \delta$ in order to be the richest cluster. Rewriting this expression gives us $0 < p(w^{+}) \leq \frac{\delta}{\beta-\gamma}$ ■

Claim 6 *$p\left(w^{j}\right) < p\left(w^{j+1}\right), \forall F \in \mathbb{F}_{\mathbf{s}\mathrm{K}}$ given $(\mathbf{s}, F)$ such that $\mathbf{s} = \mathcal{V}$ when $\delta \to 0$ for $\beta > \gamma + \delta$.*

Proof. $\pi\left(w^{j+1};\mathcal{V},F\right) - \pi\left(w^{j};\mathcal{V},F\right) = \widehat{\delta}$

$\Rightarrow$

$p\left(w^{j}\right) + \gamma p\left(w^{j+1}\right) - \gamma p\left(w^{j}\right) - \beta p\left(w^{j+1}\right) = \widehat{\delta}$

$\Leftrightarrow$

$p\left(w^{j}\right)(1-\gamma) - p\left(w^{j+1}\right)(\beta-\gamma) = \widehat{\delta}$

$\Leftrightarrow$

$p\left(w^{j}\right) = p\left(w^{j+1}\right)\frac{\beta-\gamma}{1-\gamma} + \frac{\widehat{\delta}}{1-\gamma}$

From claim 2 we have that $\delta \to 0 \Rightarrow \widehat{\delta} \to 0$. Since $\frac{\beta-\gamma}{1-\gamma} < 1 \rightsquigarrow p\left(w^{j}\right) < p\left(w^{j+1}\right)$. ■

Claim 7 *$\forall F \in \mathbb{F}_{\mathbf{s}\mathrm{K}}$ for $(\mathbf{s}, F)$ such that $\mathbf{s} = \mathcal{V} \rightsquigarrow 0 < p(w^{+}) \leq \frac{\delta}{1-\gamma}$.*

Proof. From definition 3 it follows that the probability mass of the richest cluster must fulfill $p(w^{+})(1-\gamma) \leq \delta$ in order to be the richest cluster. Rewriting this expression gives us $0 < p(w^{+}) \leq \frac{\delta}{1-\gamma}$ ■

Lemma 1 *$\mathcal{R}$ is according to criterion $\mathfrak{A}$ vulnerable for invasion by the* H*, given $\gamma = \gamma'$ and $\alpha + \beta \leq 1$.*

Proof. $\frac{1}{2} \geq \gamma' \Rightarrow \pi\left(H, w^-; \mathcal{R}, F\right) \geq \pi\left(\mathcal{R}, w^-; \mathcal{R}, F\right)$
$\frac{1}{2} \geq \gamma' \Leftrightarrow \frac{1}{2} \geq \frac{1+\alpha+\beta}{4} \Leftrightarrow 1 \geq \alpha + \beta$ ■

Lemma 2 *$\mathcal{R}$ is according to criterion $\mathfrak{A}$ vulnerable for invasion by the* D*, given $\gamma = \gamma'$ and $1 \leq \alpha + \beta$.*

Proof. $\frac{\alpha+\beta}{2} \geq \gamma' \Rightarrow \pi\left(D, w^+; \mathcal{R}, F\right) \geq \pi\left(\mathcal{R}, w^+; \mathcal{R}, F\right)$
$\frac{\alpha+\beta}{2} \geq \gamma' \Leftrightarrow 1 \leq \alpha + \beta$ ■

Claim 8 *$\mathcal{R}$ is according to criterion $\mathfrak{A}$ vulnerable for invasion $\forall \alpha, \beta \in (0,1)$, given $\beta > \gamma + \delta$ for $\gamma = \gamma'$.*

Proof. Follows directly from lemmae 1 and 2. ■

Claim 9 *$\exists \mathcal{R}^c$ such that this population according to criterion $\mathfrak{A}$ is* not *vulnerable for any invasion, given $\beta > \gamma' + \delta$ at the limit when $\delta \to 0$.*

Proof. From lemmae 1 and 2 it follows that the *Robin Hood strategy* is vulnerable for invasion at the extreme supports.

- $1 \geq \alpha + \beta$

  An invasion by the *Hawk strategy* is limited to
  $\max\left(0, \frac{1-\alpha}{1+\beta-\alpha} - \frac{1}{2}\right)$ in w^- for $1 \geq \alpha + \beta$.

  To see this, consider the second lowest support w^2 and compare the payoff of the *Robin Hood strategy* with an *Hawk* invader

  $$\pi\left(\mathcal{R}, w^2; \mathcal{R}^c, F\right) - \pi\left(H, w^2; \mathcal{R}^c, F\right) = \beta p\left(w^-\right) + \left(\gamma - \tfrac{1}{2}\right) p\left(w^2\right)$$

  From claim 4 we have that $p\left(w^-\right) > \frac{1-\beta}{1-\gamma} \rightsquigarrow$

  $$\begin{aligned} &\beta p\left(w^-\right) + \left(\gamma' - \tfrac{1}{2}\right) p\left(w^2\right) > \\ &\beta \frac{1-\beta}{1-\gamma} + \left(\gamma' - \tfrac{1}{2}\right) \frac{1-\beta}{1-\gamma} \left(\frac{\beta-\gamma'}{1-\gamma'} - \frac{\delta}{1-\gamma'}\right) > \\ &\beta \frac{1-\beta}{1-\gamma} + \left(\frac{\alpha+\beta-1}{4}\right) \frac{1-\beta}{1-\gamma'} = \\ &\left(\frac{\alpha+5\beta-1}{4}\right) \frac{1-\beta}{1-\gamma'} > 0 \end{aligned}$$

  since $\beta > \gamma' + \delta \Rightarrow \beta > \gamma' \Leftrightarrow \beta > \frac{1+\alpha+\beta}{4} \Leftrightarrow 3\beta - \alpha > 1 \Rightarrow \alpha + 5\beta - 1 > 0$

Hence we have that $\pi\left(\mathcal{R}, w^2; \mathcal{R}^c, F\right) - \pi\left(H, w^2; \mathcal{R}^c, F\right) > 0$. It follows directly that we have $\pi\left(\mathcal{R}, w^j; \mathcal{R}^c, F\right) - \pi\left(H, w^j; \mathcal{R}^c, F\right) > 0, \forall j \geq 2$

Note that no other strategy can invade this population.

  - *Dove strategy* will do worse against opponents within the support where it invades and equal (iff invasion at w^+) or worse against opponents outside the support. Hence, $\mathcal{R}^c$ is not vulnerable for invasion by the *Dove strategy.*
  - *Villain strategy* will do just as well against opponents within the support where it invades but will fail to coordinate against any opponents outside the support and thus earn much less. Hence, $\mathcal{R}^c$ is not vulnerable for invasion by the *Villain strategy.*

- $1 \leq \alpha + \beta$

  In the upper supports, we have that the invasion by the *Dove strategy* is limited to $\max\left(0, \frac{\beta}{1+\beta-\alpha} - \frac{1}{2}\right)$ in each support $w \in \operatorname{supp}\mathbf{F}$.

  If $\mathcal{R}$ is to be invaded at w^j by D then we have that:

$$\begin{gathered}
\pi\left(D, w^j; \mathcal{R}^c, F\right) - \pi\left(\mathcal{R}, w^j; \mathcal{R}^c, F\right) \geq 0 \\
(\alpha - 1) F\left(w^{j-1}\right) + \left(\tfrac{\alpha+\beta}{2} - \gamma'\right) p\left(w^j\right) + \left(\gamma^* - \gamma^*\right)\left(1 - F\left(w^j\right)\right) \geq 0 \\
(\alpha - 1) F\left(w^{j-1}\right) + \left(\tfrac{\alpha-1+\beta}{4}\right) p\left(w^j\right) \geq 0 \\
(1 - \alpha) F\left(w^{j-1}\right) + \left(\tfrac{1-\alpha-\beta}{4}\right) p\left(w^j\right) < 0 \\
\tfrac{1}{4}\left(\tfrac{\beta}{(1-\alpha)} - 1\right) \geq \tfrac{F\left(w^{j-1}\right)}{p\left(w^j\right)}
\end{gathered}$$

  Note that if $\frac{1}{4}\left(\frac{\beta}{(1-\alpha)} - 1\right) < 0$ then we have that just w^+ can be invaded.

  Also note that:

$$\sup \tfrac{1}{4}\left(\tfrac{\beta}{(1-\alpha)} - 1\right) \text{ subject to } \begin{cases} \beta > \gamma' + \delta \Rightarrow \beta > \frac{1+\alpha}{3} \\ 1 \leq \alpha + \beta \end{cases}$$

$$\rightsquigarrow \sup \tfrac{1}{4}\left(\tfrac{\beta}{(1-\alpha)} - 1\right) = \tfrac{1}{2} \text{ for } \lim \beta = 1, \lim \alpha = \tfrac{2}{3}$$

  i.e. $\pi\left(\mathcal{R}, w^j; \mathcal{R}^c, F\right) - \pi\left(D, w^j; \mathcal{R}^c, F\right) > 0 \Rightarrow F\left(w^{j-1}\right) < \frac{1}{2} p\left(w^j\right)$

  That is, half the probability mass of the support has to exceed the cumulative density below the support.

  But since we know that $p\left(w^j\right) = p\left(w^{j+1}\right)\frac{1-\gamma}{\beta-\gamma} + \frac{\widehat{\delta}}{\beta-\gamma}$, i.e. that the probability mass is decreasing exponentially by $\frac{\beta-\gamma}{1-\gamma}$ in the support as $\delta \to 0$, there *cannot* be any invasion at the adjacent support to w^{+}. Hence, it follows that the invasion is limited to w^{+}.

Finally note that no other strategy can invade this population.

- *Hawk strategy* will do worse against opponents within the support where it invades and equal (iff invasion at w^{-}) or worse against opponents outside the support. Hence, $\mathcal{R}^{c}$ is not vulnerable for invasion by the *Hawk strategy.*
- *Villain strategy* will do just as good against opponents within the support where it invades but will fail to coordinate against any opponents outside the support and thus earn much less. Hence, $\mathcal{R}^{c}$ is not vulnerable for invasion by the *Villain strategy.*

■

Claim 10 $\mathcal{R}$ *is according to criterion* $\mathfrak{A}$ *vulnerable for invasion by* MSS*, given* $\beta > \gamma + \delta$ *for* $\gamma = \gamma^{*}$*.*

Proof. $\mathcal{R}$ can be invaded by both H at w^{-} and D at w^{+} according to assumption 3 as they mimic the behavior of the incumbent strategy at these supports and they are of less complexity. Note that this invasion is limited to $\frac{1-\alpha}{1+\beta-\alpha}p\left(w^{-}\right)$ in w^{-} and $\frac{\beta}{1+\beta-\alpha}p\left(w^{+}\right)$ in w^{+}. ■

Claim 11 $\exists\mathcal{R}^{c}$ *with identical realized behavior as* $\mathcal{R}$*, that according to criterion* $\mathfrak{A}$ *is* not *vulnerable for invasion, given* $\beta > \gamma + \delta$ *for* $\gamma = \gamma^{*}$*.*

Proof. Consider a population where the agents are conditioning the game according to the *Robin Hood strategy* in every support but in w^{-} and w^{+} where the $\frac{1-\alpha}{1+\beta-\alpha}$ of the agents in w^{-} are playing the *Hawk strategy* and $\frac{\beta}{1+\beta-\alpha}$ of the agents in w^{+} are playing the *Dove strategy.*

It follows directly that no invasion is possible since every agent behaves optimally within every support and is perfectly coordinated between the supports.

- *Villain strategy* will do just as well against opponents within the support where it invades but will fail to coordinate against any opponents outside the support and thus earn much less. Hence, $\mathcal{R}^{c}$ is not vulnerable for invasion by the *Villain strategy.*

- *Hawk strategy* will do worse against some opponents within the support. This strategy will coordinate against opponents where $w_i + \delta \leq w_{-i}$ but will fail to coordinate against opponents where $w_i - \delta \geq w_{-i}$. Hence, $\mathcal{R}^c$ is not vulnerable for invasion by the *Hawk strategy*, $\forall w \in \operatorname{supp} F \setminus w^-$.

- *Dove strategy* will do worse against some opponents within the support. This strategy will coordinate against opponents where $w_i - \delta \geq w_{-i}$ but will fail to coordinate against opponents where $w_i + \delta \leq w_{-i}$. Hence, $\mathcal{R}^c$ is not vulnerable for invasion by the *Dove strategy*, $\forall w \in \operatorname{supp} F \setminus w^+$.

■

Lemma 3 *$\mathcal{V}$ is according to criterion $\mathfrak{A}$ vulnerable for invasion by the* D*, given $\gamma = \gamma'$ and $\alpha + \beta \geq 1$.*

Proof. $\frac{\alpha+\beta}{2} \geq \gamma' \Rightarrow \pi\left(D, w^-; \mathcal{V}, F\right) \geq \pi\left(\mathcal{V}, w^-; \mathcal{V}, F\right)$
$\frac{\alpha+\beta}{2} \geq \gamma' \Leftrightarrow 1 \leq \alpha + \beta$ ■

Lemma 4 *$\mathcal{V}$ is according to criterion $\mathfrak{A}$ vulnerable for invasion by the* H*, given $\gamma = \gamma'$ and $\alpha + \beta \leq 1$.*

Proof. $\frac{1}{2} \geq \gamma' \Rightarrow \pi\left(H, w^+; \mathcal{V}, F\right) \geq \pi\left(\mathcal{V}, w^+; \mathcal{V}, F\right)$
$\frac{1}{2} \geq \gamma' \Leftrightarrow \frac{1}{2} \geq \frac{1+\alpha+\beta}{4} \Leftrightarrow 1 \geq \alpha + \beta$ ■

Claim 12 *$\mathcal{V}$ is according to criterion $\mathfrak{A}$ vulnerable for invasion $\forall \alpha, \beta \in (0,1)$, given $\gamma = \gamma'$.*

Proof. Follows directly from lemmae 3 and 4. ■

Lemma 5 *$\exists \mathcal{V}^c$ such that this population according to criterion $\mathfrak{A}$ is* not *vulnerable for any invasion, given $\beta > \gamma' + \delta$ when $\delta \to 0$.*

Proof. From lemmae 3 and 4 it follows that the *Villain strategy* is vulnerable for invasion at the extreme supports.

- $\alpha + \beta \geq 1$

  An invasion by the *Dove strategy* is limited to
  $\max\left(0, \frac{\beta}{1+\beta-\alpha} - \frac{1}{2}\right)$in w^- for $\alpha + \beta \geq 1$. The question is how far up in the supports the invasion is possible.

If $\mathcal{V}$ is to be invaded at w^j by D then we have that:

$$\begin{gathered}
\pi\left(D, w^j; \mathcal{V}^c, F\right) - \pi\left(\mathcal{V}, w^j; \mathcal{V}^c, F\right) \geq 0 \\
\left(\gamma^* - \gamma^*\right) F\left(w^{j-1}\right) + \left(\tfrac{\alpha+\beta}{2} - \gamma'\right) p\left(w^j\right) + (\alpha - 1)\left(1 - F\left(w^j\right)\right) \geq 0 \\
\left(\tfrac{\alpha-1+\beta}{4}\right) p\left(w^j\right) + (\alpha - 1)\left(1 - F\left(w^j\right)\right) \geq 0 \\
\left(\tfrac{1-\alpha-\beta}{4}\right) p\left(w^j\right) + (1 - \alpha)\left(1 - F\left(w^j\right)\right) < 0 \\
\tfrac{1}{4}\left(\tfrac{\beta}{(1-\alpha)} - 1\right) \geq \tfrac{\left(1-F\left(w^j\right)\right)}{p\left(w^j\right)}
\end{gathered}$$

First note if $\frac{1}{4}\left(\frac{\beta}{(1-\alpha)} - 1\right) < 0$ then we have that just w^- can be invaded.

Then note that:

$$\sup \tfrac{1}{4}\left(\tfrac{\beta}{(1-\alpha)} - 1\right) \text{ subject to } \begin{cases} \beta > \gamma' + \delta \Rightarrow \beta > \frac{1+\alpha}{3} \\ 1 \leq \alpha + \beta \end{cases}$$

$$\rightsquigarrow \sup \tfrac{1}{4}\left(\tfrac{\beta}{(1-\alpha)} - 1\right) = \tfrac{1}{2} \text{ for } \lim \beta = 1, \lim \alpha = \tfrac{2}{3}$$

i.e. $\frac{1}{2} p\left(w^j\right) \geq \left(1 - F\left(w^j\right)\right)$

From property 4 we know that $p\left(w^j\right) \rightarrow 0$ as $\delta \rightarrow 0 \rightsquigarrow \left(1 - F\left(w^j\right)\right) > \frac{1}{2} p\left(w^j\right)$. Hence, $\mathcal{V}^c$ is not vulnerable for invasion by the *Dove strategy* outside w^-.

Note that no other strategy can invade this population.

- *Hawk strategy* will do worse against opponents within the support where it invades and equal (iff invasion at w^+) or worse against opponents outside the support. Hence, $\mathcal{V}^c$ is not vulnerable for invasion by the *Hawk strategy.*
- *Robin Hood strategy* will do just as well against opponents within the support where it invades but will fail to coordinate against any opponents outside the support and thus earn much less. Hence, $\mathcal{V}^c$ is not vulnerable for invasion by the *Robin Hood strategy.*

- $\alpha + \beta \leq 1$

  In the upper supports, we have that the invasion by the *Hawk strategy* is limited to $\max\left(0, \frac{1-\alpha}{1+\beta-\alpha} - \frac{1}{2}\right)$ in each support $w \in \operatorname{supp} \mathbf{F}$.

  If $\mathcal{V}$ is to be invaded at w^j by H then we have that:

$$\pi\left(H, w^j; \mathcal{V}^c, F\right) - \pi\left(\mathcal{V}, w^j; \mathcal{V}^c, F\right) \geq 0$$
$$(1-1)\, F\left(w^{j-1}\right) + \left(\tfrac{1}{2} - \gamma'\right) p\left(w^j\right) + (0-\beta)\left(1 - F\left(w^j\right)\right) \geq 0$$
$$\left(\tfrac{1}{2} - \gamma\right) p\left(w^j\right) + (0-\beta)\left(1 - F\left(w^j\right)\right) \geq 0$$
$$\left(\frac{1-\alpha-\beta}{4}\right) p\left(w^j\right) \geq \beta\left(1 - F\left(w^j\right)\right)$$
$$\frac{1}{4}\left(\frac{1-\alpha}{\beta} - 1\right) \geq \frac{\left(1 - F\left(w^j\right)\right)}{p\left(w^j\right)}$$

As before note that:

$$\sup \frac{1}{4}\left(\frac{1-\alpha}{\beta} - 1\right) \text{ subject to } \begin{cases} \beta > \gamma' + \delta \Rightarrow \beta > \frac{1+\alpha}{3} \\ 1 \geq \alpha + \beta \end{cases}$$

$$\rightsquigarrow \sup \frac{1}{4}\left(\frac{1-\alpha}{\beta} - 1\right) = \frac{1}{2} \text{ for } \lim \beta = \frac{1}{3}, \lim \alpha = 0$$

$$\tfrac{1}{2} p\left(w^j\right) \geq \left(1 - F\left(w^j\right)\right)$$

From property 4 we know that $p\left(w^j\right) \to 0$ as $\delta \to 0 \rightsquigarrow \left(1 - F\left(w^j\right)\right) > \frac{1}{2} p\left(w^j\right)$. Hence, $\mathcal{V}^c$ is not vulnerable for invasion by the *Hawk strategy* outside w^+.

Note that no other strategy can invade this population.

- *Dove strategy* will do worse against opponents within the support where it invades and equal (iff invasion at w^-) or worse against opponents outside the support. Hence, $\mathcal{V}^c$ is not vulnerable for invasion by the *Dove strategy.*
- *Robin Hood strategy* will do just as well against opponents within the support where it invades but will fail to coordinate against any opponents outside the support and thus earn much less. Hence, $\mathcal{V}^c$ is not vulnerable for invasion by the *Robin Hood strategy.*

■

Lemma 6 *$\exists \mathcal{V}^c$ such that this population according to criterion $\mathfrak{A}$ is* not *vulnerable for any invasion, given $\boldsymbol{\beta} \leq \boldsymbol{\gamma'} + \boldsymbol{\delta}$ when $\delta \to 0$.*

Proof. From lemmae 3 and 4 it follows that the *Villain strategy* is vulnerable for invasion at the extreme supports.

- $\alpha + \beta \geq 1$

  An invasion by the *Dove strategy* is limited to $\max\left(0, \frac{\beta}{1+\beta-\alpha} - \frac{1}{2}\right)$in w^- for $\alpha + \beta \geq 1$. The question is how far up in the supports the invasion is possible.

If $\mathcal{V}$ is to be invaded at w^j by D then we have that

$$\pi\left(D, w^j; \mathcal{V}^c, F\right) - \pi\left(\mathcal{V}, w^j; \mathcal{V}^c, F\right) \geq 0$$
$$\Rightarrow$$
$$\tfrac{1}{4}\left(\tfrac{\beta}{(1-\alpha)} - 1\right) \geq \tfrac{\left(1-F\left(w^j\right)\right)}{p(w^j)}$$

First note if $\frac{1}{4}\left(\frac{\beta}{(1-\alpha)} - 1\right) < 0$ then we have that just w^- can be invaded.

Then note that:

$$\sup \tfrac{1}{4}\left(\tfrac{\beta}{(1-\alpha)} - 1\right) \text{ subject to } \begin{cases} \beta \leq \gamma' + \delta \Rightarrow \beta \leq \frac{1+\alpha}{3} \\ 1 \leq \alpha + \beta \end{cases}$$
$$\rightsquigarrow \sup \tfrac{1}{4}\left(\tfrac{\beta}{(1-\alpha)} - 1\right) = \infty \text{ for } \lim \alpha = 1$$

However from property 4 we know that $p\left(w^j\right) \to 0$ as $\delta \to 0 \rightsquigarrow \exists \delta, \forall (\alpha, \beta)$ such that $\frac{1}{2}p\left(w^j\right) \ngeq \left(1 - F\left(w^j\right)\right)$. Hence, $\mathcal{V}^c$ is not vulnerable for invasion by the *Dove strategy* outside w^-.

Note that no other strategy can invade this population.

- *Hawk strategy* will do worse against opponents within the support where it invades and equal (iff invasion at w^+) or worse against opponents outside the support. Hence, $\mathcal{V}^c$ is not vulnerable for invasion by the *Hawk strategy.*
- *Robin Hood strategy* will do just as well against opponents within the support where it invades but will fail to coordinate against any opponents outside the support and thus earn much less. Hence, $\mathcal{V}^c$ is not vulnerable for invasion by the *Robin Hood strategy.*

- $\alpha + \beta \leq 1$

  In the upper supports, we have that he invasion by the *Hawk strategy* is limited to $\max\left(0, \frac{1-\alpha}{1+\beta-\alpha} - \frac{1}{2}\right)$ in each support $w \in \operatorname{supp} \mathbf{F}$.

  If $\mathcal{V}$ is to be invaded at w^j by H then we have that:

$$\pi\left(H, w^j; \mathcal{V}^c, F\right) - \pi\left(\mathcal{V}, w^j; \mathcal{V}^c, F\right) \geq 0$$
$$\Rightarrow$$
$$\tfrac{1}{4}\left(\tfrac{1-\alpha}{\beta} - 1\right) \geq \tfrac{\left(1-F\left(w^j\right)\right)}{p(w^j)}$$

  As before note that:

$$\sup \frac{1}{4}\left(\frac{1-\alpha}{\beta}-1\right) \text{ subject to } \begin{cases} \beta > \gamma' + \delta \Rightarrow \beta > \frac{1+\alpha}{3} \\ 1 \geq \alpha + \beta \end{cases}$$
$$\rightsquigarrow \sup \frac{1}{4}\left(\frac{1-\alpha}{\beta}-1\right) = \infty \text{ for} \lim \beta = 0$$

Just as above, we know that $p\left(w^j\right) \to 0$ as $\delta \to 0 \rightsquigarrow \exists\delta, \forall(\alpha,\beta)$ such that $\frac{1}{2}p\left(w^j\right) \ngeq \left(1-F\left(w^j\right)\right)$. Hence, $\mathcal{V}^c$ is not vulnerable for invasion by the *Hawk strategy* outside w^+.

Note that no other strategy can invade this population.

- *Dove strategy* will do worse against opponents within the support where it invades and equal (iff invasion at w^-) or worse against opponents outside the support. Hence, $\mathcal{V}^c$ is not vulnerable for invasion by the *Dove strategy.*
- *Robin Hood strategy* will do just as well against opponents within the support where it invades but will fail to coordinate against any opponents outside the support and thus earn much less. Hence, $\mathcal{V}^c$ is not vulnerable for invasion by the *Robin Hood strategy.*

■

Claim 13 *$\exists\mathcal{V}^c$ such that this population according to criterion $\mathfrak{A}$ is* not *vulnerable for any invasion, given $\gamma = \gamma'$ when $\delta \to 0$.*

Proof. Follows directly from lemmae 5 and 6. ■

Claim 14 *$\mathcal{V}$ is according to criterion $\mathfrak{A}$ vulnerable for invasion by* MSS, *given $\gamma = \gamma^*$.*

Proof. Similar to the case above we find that $\mathcal{V}$ can be invaded by both D at w^- and H at w^+ according to assumption 3 as they mimic the incumbent strategy at these supports and they are of less complexity. Here too is invasion is limited to $\frac{\beta}{1+\beta-\alpha}p\left(w^-\right)$ in w^- and $\frac{1-\alpha}{1+\beta-\alpha}p\left(w^+\right)$ in w^+. ■

Claim 15 *$\exists\mathcal{V}^c$ with identical realized behavior as $\mathcal{V}$, that according to criterion $\mathfrak{A}$ is* not *vulnerable for invasion, given $\gamma = \gamma^*$.*

Proof. Consider a population where the agents are conditioning the game according to the *Villain strategy* in every support but in w^- and w^+ where the

$\frac{\beta}{1+\beta-\alpha}$ of the agents in w^- are playing the *Dove strategy* and $\frac{1-\alpha}{1+\beta-\alpha}$ of the agents in w^+ are playing the *Hawk strategy.*

It follows directly that no invasion is possible since every agent behaves optimally within every support and are perfectly coordinated between the supports.

- *Robin Hood Strategy* will do just as well against opponents within the support where it invade but will fail to coordinate against any opponents outside the support and thus earn much less. Hence, $\mathcal{V}^c$ is not vulnerable for invasion by the *Robin Hood Strategy.*

- *Hawk strategy* will do worse against some opponents within the support. This strategy will coordinate against opponents where $w_i - \delta \geq w_{-i}$ but will fail to coordinate against opponents where $w_i + \delta \leq w_{-i}$. Hence, $\mathcal{V}^c$ is not vulnerable for invasion by the *Hawk strategy*, $\forall w \in \operatorname{supp} F \setminus w^+$.

- *Dove strategy* will do worse against some opponents within the support. This strategy will coordinate against opponents where $w_i + \delta \leq w_{-i}$ but will fail to coordinate against opponents where $w_i - \delta \geq w_{-i}$. Hence, $\mathcal{V}^c$ is not vulnerable for invasion by the *Dove strategy*, $\forall w \in \operatorname{supp} F \setminus w^-$.

■

Claim 16 $\mathcal{R}$ *is according to criterion* $\mathfrak{B}$ not *vulnerable for invasion by any convex combination of pure strategies or mixed strategies for any* F *as* $\delta \to 0$, *given* $\beta > \gamma' + \delta$.

Proof. Consider the following invaders:

Case D: If $\alpha + \beta \geq 1 \Rightarrow \pi\left(D, w^+; \mathcal{R}, F\right) > \pi\left(\mathcal{R}, w^+; \mathcal{R}, F\right)$ and $\pi\left(D, w; \mathcal{R}, F\right) < \pi\left(\mathcal{R}, w; \mathcal{R}, F\right)$

$, \forall w \in \operatorname{supp} F \setminus w^+ \Rightarrow \overline{\pi}\left(D; \mathcal{R}, F\right) < \overline{\pi}\left(\mathcal{R}; \mathcal{R}, F\right)$

Case H: If $\alpha + \beta \leq 1 \Rightarrow \pi\left(H, w^-; \mathcal{R}, F\right) > \pi\left(\mathcal{R}, w^-; \mathcal{R}, F\right)$ and $\pi\left(H, w; \mathcal{R}, F\right) < \pi\left(\mathcal{R}, w; \mathcal{R}, F\right)$

$, \forall w \in \operatorname{supp} F \setminus w^- \Rightarrow \overline{\pi}\left(H; \mathcal{R}, F\right) < \overline{\pi}\left(\mathcal{R}; \mathcal{R}, F\right)$

Case $\mathcal{V}$: $\pi\left(\mathcal{V}, w; \mathcal{R}, F\right) < \pi\left(\mathcal{R}, w; \mathcal{R}, F\right), \forall w \in \operatorname{supp} F \Rightarrow \overline{\pi}\left(\mathcal{V}; \mathcal{R}, F\right) < \overline{\pi}\left(\mathcal{R}; \mathcal{R}, F\right)$

$\blacksquare$

Claim 17 *$\mathcal{R}$ is according to criterion $\mathfrak{B}$* not *vulnerable for invasion by any convex combination of pure strategies or mixed strategies for any F, given $\beta > \gamma^* + \delta$.*

Proof. Consider the following invaders:

Case H: $\pi\left(H, w^{+}; \mathcal{R}, F\right) = \pi\left(\mathcal{R}, w^{+}; \mathcal{R}, F\right)$ and $\pi\left(H, w; \mathcal{R}, F\right) < \pi\left(\mathcal{R}, w; \mathcal{R}, F\right)$ $, \forall w \in \operatorname{supp} F \setminus w^{+} \Rightarrow \overline{\pi}\left(H; \mathcal{R}, F\right) < \overline{\pi}\left(\mathcal{R}; \mathcal{R}, F\right)$

Case D: $\pi\left(D, w^{-}; \mathcal{R}, F\right) = \pi\left(\mathcal{R}, w^{-}; \mathcal{R}, F\right)$ and $\pi\left(D, w; \mathcal{R}, F\right) < \pi\left(\mathcal{R}, w; \mathcal{R}, F\right)$ $, \forall w \in \operatorname{supp} F \setminus w^{-} \Rightarrow \overline{\pi}\left(D; \mathcal{R}, F\right) < \overline{\pi}\left(\mathcal{R}; \mathcal{R}, F\right)$

Case $\mathcal{V}$: $\pi\left(\mathcal{V}, w; \mathcal{R}, F\right) < \pi\left(\mathcal{R}, w; \mathcal{R}, F\right), \forall w \in \operatorname{supp} F \Rightarrow \overline{\pi}\left(\mathcal{V}; \mathcal{R}, F\right) < \overline{\pi}\left(\mathcal{R}; \mathcal{R}, F\right)$

$\blacksquare$

Claim 18 *$\mathcal{V}$ is according to criterion $\mathfrak{B}$* not *vulnerable for invasion by any convex combination of pure strategies or mixed strategies for any F as $\delta \to 0$, given $\gamma = \gamma'$.*

Proof. Consider the following invaders:

Case D: If $\alpha + \beta \geq 1 \Rightarrow \pi\left(D, w^{-}; \mathcal{V}, F\right) > \pi\left(\mathcal{V}, w^{-}; \mathcal{V}, F\right)$ and $\pi\left(D, w; \mathcal{V}, F\right) < \pi\left(\mathcal{V}, w; \mathcal{V}, F\right), \forall w \in \operatorname{supp} F \setminus w^{-} \Rightarrow \overline{\pi}\left(D; \mathcal{V}, F\right) < \overline{\pi}\left(\mathcal{V}; \mathcal{V}, F\right)$

Case H: If $\alpha + \beta \leq 1 \Rightarrow \pi\left(H, w^{+}; \mathcal{V}, F\right) > \pi\left(\mathcal{V}, w^{+}; \mathcal{V}, F\right)$ and $\pi\left(H, w; \mathcal{V}, F\right) < \pi\left(\mathcal{V}, w; \mathcal{V}, F\right), \forall w \in \operatorname{supp} F \setminus w^{+} \Rightarrow \overline{\pi}\left(H; \mathcal{V}, F\right) < \overline{\pi}\left(\mathcal{V}; \mathcal{V}, F\right)$

Case $\mathcal{R}$: $\pi\left(\mathcal{R}, w; \mathcal{V}, F\right) < \pi\left(\mathcal{V}, w; \mathcal{V}, F\right), \forall w \in \operatorname{supp} F \Rightarrow \overline{\pi}\left(\mathcal{R}; \mathcal{V}, F\right) < \overline{\pi}\left(\mathcal{V}; \mathcal{V}, F\right)$

$\blacksquare$

Claim 19 *$\mathcal{V}$ is according to criterion $\mathfrak{B}$* not *vulnerable for invasion by any convex combination of pure strategies or mixed strategies for any F, given $\gamma = \gamma^*$.*

Proof. Consider the following invaders:

Case D: $\pi(D, w^-; \mathcal{V}, F) = \pi(\mathcal{V}, w^-; \mathcal{V}, F)$ and $\pi(D, w; \mathcal{V}, F) < \pi(\mathcal{V}, w; \mathcal{V}, F), \forall w \in \operatorname{supp} F \setminus w^+ \Rightarrow \overline{\pi}(D; \mathcal{V}, F) < \overline{\pi}(\mathcal{V}; \mathcal{V}, F)$

Case H: $\pi(H, w^+; \mathcal{V}, F) = \pi(\mathcal{V}, w^+; \mathcal{V}, F)$ and $\pi(H, w; \mathcal{V}, F) < \pi(\mathcal{V}, w; \mathcal{V}, F), \forall w \in \operatorname{supp} F \setminus w^- \Rightarrow \overline{\pi}(H; \mathcal{V}, F) < \overline{\pi}(\mathcal{V}; \mathcal{V}, F)$

Case $\mathcal{R}$: $\pi(\mathcal{R}, w; \mathcal{V}, F) < \pi(\mathcal{V}, w; \mathcal{V}, F), \forall w \in \operatorname{supp} F \Rightarrow \overline{\pi}(\mathcal{R}; \mathcal{V}, F) < \overline{\pi}(\mathcal{V}; \mathcal{V}, F)$

■

Lemma 7 $\boldsymbol{\beta} > \boldsymbol{\gamma}^* \Leftrightarrow \boldsymbol{\alpha} < \boldsymbol{\gamma}^*$.

Proof. $\boldsymbol{\beta} > \boldsymbol{\gamma}^* \Leftrightarrow \boldsymbol{\beta} > \frac{\beta}{1+\beta-\alpha} \Rightarrow \boldsymbol{\beta} > \boldsymbol{\alpha}$
$\boldsymbol{\alpha} < \boldsymbol{\gamma}^* \Leftrightarrow \boldsymbol{\alpha} > \frac{\beta}{1+\beta-\alpha} \Leftrightarrow \boldsymbol{\alpha}(\boldsymbol{\beta} - \boldsymbol{\alpha}) > \boldsymbol{\beta} - \boldsymbol{\alpha}$ and $\boldsymbol{\alpha} \in (0, 1)$ ■

Proposition 1 *$\mathcal{R}^c$ is according to criterion $\mathfrak{C}$* not *vulnerable for invasion by* MSS*, given $\boldsymbol{\beta} > \boldsymbol{\gamma}^* + \boldsymbol{\delta}$ as $\delta \to 0$.*

Proof. Assume that the size of the population using *MSS*, denoted γ, at the time of the fusion is λ, and $\mathcal{R}^c$ is $1 - \lambda$. Since $\overline{\pi}(\gamma; \gamma, F_\gamma) < \overline{\pi}(\mathcal{R}^c; \mathcal{R}^c, F_{\mathcal{R}^c}) \rightsquigarrow \lambda < \frac{1}{2}$. Note that γ consist of $\frac{1-\alpha}{1+\beta-\alpha}\lambda$ with H and $\frac{\beta}{1+\beta-\alpha}\lambda$ with D.

- Now consider when the populations meet; $F_{\mathcal{R}^c}(\mathcal{R}^c), F_\gamma(\gamma) \to \mathbf{F}(\mathcal{R}^c\gamma)$
  - The payoff before the fusion are:

$$\pi(H, \cdot; \gamma, F_\gamma) = \frac{\beta}{1+\beta-\alpha}\lambda$$
$$\pi(D, \cdot; \gamma, F_\gamma) = \frac{\beta}{1+\beta-\alpha}\lambda$$
$$\pi(\mathcal{R}^c, w^-; \mathcal{R}^c, F_{\mathcal{R}^c}) = \left(p(w^-)\frac{\beta}{1+\beta-\alpha} + (1 - p(w^-))\right)(1 - \lambda)$$
$$\pi(\mathcal{R}^c, w^+; \mathcal{R}^c, F_{\mathcal{R}^c}) = \left(p(w^+)\frac{\beta}{1+\beta-\alpha} + (1 - p(w^+))\right)\beta(1 - \lambda)$$
$$\rightsquigarrow \pi(H, \cdot; \gamma, F_\gamma) = \pi(D, \cdot; \gamma, F_\gamma) < \pi(\mathcal{R}^c, w^-; \mathcal{R}^c, F_{\mathcal{R}^c})$$

  - The payoff immediately after the time of the fusion, but before the distribution has become stable, are:

$$\pi(H, w_H; \mathcal{R}^c\gamma, \cdot) = (1 - \lambda) + \frac{\beta}{1+\beta-\alpha}\lambda$$
$$\pi(D, w_D; \mathcal{R}^c\gamma, \cdot) = \alpha(1 - \lambda) + \frac{\beta}{1+\beta-\alpha}\lambda$$

$$\pi\left(\mathcal{R}^c, w_{\mathcal{R}^c}^-; \mathcal{R}^c\gamma, \cdot\right) =$$
$$\left(p\left(w_{\mathcal{R}^c}^-\right)\tfrac{\beta}{1+\beta-\alpha} + \left(1 - p\left(w_{\mathcal{R}^c}^-\right)\right)\right)(1-\lambda) + \tfrac{\beta}{1+\beta-\alpha}\lambda$$

$$\pi\left(\mathcal{R}^c, w_{\mathcal{R}^c}^+; \mathcal{R}^c\gamma, \cdot\right) =$$
$$\left(p\left(w_{\mathcal{R}^c}^+\right)\tfrac{\beta}{1+\beta-\alpha} + \left(1 - p\left(w_{\mathcal{R}^c}^+\right)\right)\right)\beta(1-\lambda) + \tfrac{\beta}{1+\beta-\alpha}\lambda$$

From lemma 7 and above, it follows that:

$$\pi\left(D, w_D; \mathcal{R}^c\gamma, \mathbf{F}\right) < \pi\left(\mathcal{R}^c, w_{\mathcal{R}^c}^-; \mathcal{R}^c\gamma, \mathbf{F}\right).$$

Furthermore it is easy to see that $\pi\left(\mathcal{R}^c, w_{\mathcal{R}^c}^-; \mathcal{R}^c\gamma, \mathbf{F}\right) < \pi\left(H, w_H; \mathcal{R}^c\gamma, \mathbf{F}\right)$. Note that initially is $w_H = w_D$, and $w_H < w_{\mathcal{R}^c}^-$.

However, as the $\left(H, w_H; \mathcal{R}^c\gamma, \cdot\right)$ moves up in the distribution into $w_{\mathcal{R}^c}^-$ they will face $(1-\lambda)\,p\left(w_{\mathcal{R}^c}^-\right)\frac{1-\alpha}{1+\beta-\alpha}$ playing H and $(1-\lambda)\,p\left(w_{\mathcal{R}^c}^-\right)\frac{\beta}{1+\beta-\alpha}$ playing D. As $\left(H, w_H; \mathcal{R}^c\gamma, \mathbf{F}\right)$ mergers with $\left(\mathcal{R}^c, w_{\mathcal{R}^c}^-; \mathcal{R}^c\gamma, \mathbf{F}\right)$ will $p\left(w_{\mathcal{R}^c}^-\right)$ grow to $\widehat{p}\left(\widetilde{w}_{\mathcal{R}^c}^-\right)$.

Note, from claim 11 we have that:

$$\pi\left(\mathcal{R}^c : s_u = H, w_{\mathcal{R}^c}^-; \mathcal{R}^c\gamma, \cdot\right) = \pi\left(\mathcal{R}^c : s = H, w_{\mathcal{R}^c}^-; \mathcal{R}^c\gamma, \cdot\right)$$

That is, $(1-\lambda)\frac{1-\alpha}{1+\beta-\alpha}$ of the agents in $w_{\mathcal{R}^c}^-$ will be utilizing H.

- The payoff in the critical areas immediately after the time of the merger between $w_H, w_{\mathcal{R}^c}^- \to \widetilde{w}_{\mathcal{R}^c}^-$ are:

$$\pi\left(\mathcal{R}^c : s = H, \widetilde{w}_{\mathcal{R}^c}^-; \mathcal{R}^c\gamma, \cdot\right) =$$
$$(1-\lambda)\left(\widehat{p}\left(\widetilde{w}_{\mathcal{R}^c}^-\right)\tfrac{\beta}{1+\beta-\alpha} + \left(1 - \widehat{p}\left(\widetilde{w}_{\mathcal{R}^c}^-\right)\right)\right) + \tfrac{\alpha\beta}{1+\beta-\alpha}\lambda$$

$$\pi\left(\mathcal{R}^c : s_u = D, \widetilde{w}_{\mathcal{R}^c}^-; \mathcal{R}^c\gamma, \cdot\right) =$$
$$(1-\lambda)\left(\widehat{p}\left(\widetilde{w}_{\mathcal{R}^c}^-\right)\tfrac{\beta}{1+\beta-\alpha} + \left(1 - \widehat{p}\left(\widetilde{w}_{\mathcal{R}^c}^-\right)\right)\right) + \tfrac{\beta}{1+\beta-\alpha}\lambda$$

$$\pi\left(H, \widetilde{w}_{\mathcal{R}^c}^-; \mathcal{R}^c\gamma, \cdot\right) =$$
$$(1-\lambda)\left(\widehat{p}\left(\widetilde{w}_{\mathcal{R}^c}^-\right)\tfrac{\beta}{1+\beta-\alpha} + \left(1 - \widehat{p}\left(\widetilde{w}_{\mathcal{R}^c}^-\right)\right)\right) + \tfrac{\beta}{1+\beta-\alpha}\lambda$$

Thus, we have that the payoff relates as follows:

$$\pi\left(H, \widetilde{w}_{\mathcal{R}^c}^-; \mathcal{R}^c\gamma, \mathbf{F}\right) = \pi\left(\mathcal{R}^c : s = H, \widetilde{w}_{\mathcal{R}^c}^-; \mathcal{R}^c\gamma, \mathbf{F}\right)$$

$$\pi\left(\mathcal{R}^c : s = H, \widetilde{w}_{\mathcal{R}^c}^-; \mathcal{R}^c\gamma, \mathbf{F}\right) < \pi\left(\mathcal{R}^c : s_u = D, \widetilde{w}_{\mathcal{R}^c}^-; \mathcal{R}^c\gamma, \mathbf{F}\right)$$

This implies that $\pi\left(H, w; \mathcal{R}^c\gamma, \mathbf{F}\right) < \pi\left(\mathcal{R}^c, w; \mathcal{R}^c\gamma, \mathbf{F}\right), \forall w > \widetilde{w}_{\mathcal{R}^c}^-$ where $w \in \operatorname{supp}\mathbf{F}$.

- Mutation

  Note that $\nexists m \in \mathbf{S} \setminus \mathcal{R}^c$ such that $\pi\left(\mathcal{R}^c, w; \mathcal{R}^c\gamma, \mathbf{F}\right) < \pi\left(m, w; \mathcal{R}^c\gamma, \mathbf{F}\right)$.

  Thus, we have that the strategies open for successful mutations, $\curvearrowright$, are $(\gamma, \cdot; \mathcal{R}^c\gamma, \mathbf{F})$, and $(\mathcal{R}^c : s_u = D, \cdot; \mathcal{R}^c\gamma, \mathbf{F})$.

  - $(D, \cdot; \mathcal{R}^c\gamma, \mathbf{F}) \curvearrowright (H, \cdot; \mathcal{R}^c\gamma, \mathbf{F}), (\mathcal{R}^c : s_u = D, \cdot; \mathcal{R}^c\gamma, \mathbf{F})$
  - $(H, \cdot; \mathcal{R}^c\gamma, \mathbf{F}) \curvearrowright (\mathcal{R}^c : s_u = D, \cdot; \mathcal{R}^c\gamma, \mathbf{F})$

  and eventually

  - $(\mathcal{R}^c : s_u = D, \cdot; \mathcal{R}^c\gamma, \mathbf{F}) \curvearrowright (\mathcal{R}^c : s_u = H, \cdot; \mathcal{R}^c\gamma, \mathbf{F})$

  As the evolution proceeds, will the remaining $\mathcal{R}^c$ grow in accordance to assumption 2 and 4. Thus as $t \to \infty$ we will, through growth and mutations, have that $\max \pi\left(m, \cdot; \mathbf{s}, \widetilde{\mathbf{F}}\right) \leq \min \pi\left(\mathcal{R}^c, \cdot; \mathbf{s}, \widetilde{\mathbf{F}}\right) < \max \pi\left(\mathcal{R}^c, \cdot; \mathbf{s}, \widetilde{\mathbf{F}}\right)$for $m \in \mathbf{S} \backslash \mathcal{R}^c$.

■

Proposition 2 *$\mathcal{V}^c$ is according to criterion $\mathfrak{C}$* not *vulnerable for invasion by* MSS, *given $\boldsymbol{\gamma} = \boldsymbol{\gamma}^*$ as $\delta \to 0$.*

Proof. Assume that the size of the population using *MSS*, denoted γ, at the time of the fusion is λ and $\mathcal{V}^c$ is $1 - \lambda$. Since $\overline{\pi}\left(\gamma; \gamma, F_\gamma\right) < \overline{\pi}\left(\mathcal{V}^c; \mathcal{V}^c, F_{\mathcal{V}^c}\right) \rightsquigarrow \lambda < \frac{1}{2}$.

- Now consider when the populations meet; $F_{\mathcal{V}^c}\left(\mathcal{V}^c\right), F_\gamma\left(\gamma\right) \to \mathbf{F}\left(\mathcal{V}^c\gamma\right)$

  - The payoff before the fusion are:

  $$\pi\left(H, \cdot; \gamma, F_\gamma\right) = \tfrac{\beta}{1+\beta-\alpha}\lambda$$
  $$\pi\left(D, \cdot; \gamma, F_\gamma\right) = \tfrac{\beta}{1+\beta-\alpha}\lambda$$
  $$\pi\left(\mathcal{V}^c, w^-; \mathcal{V}^c, F_{\mathcal{V}^c}\right) = \Big(p\left(w^-\right)\tfrac{\beta}{1+\beta-\alpha} + \left(1 - p\left(w^-\right)\right)\Big)\beta\left(1 - \lambda\right)$$
  $$\pi\left(\mathcal{V}^c, w^+; \mathcal{V}^c, F_{\mathcal{V}^c}\right) = \Big(p\left(w^+\right)\tfrac{\beta}{1+\beta-\alpha} + \left(1 - p\left(w^+\right)\right)\Big)\left(1 - \lambda\right)$$

  - The payoff immediately after the time of the fusion, but before the distribution has become stable, are:

$\pi\left(H, w_H; \mathcal{V}^c\gamma, \cdot\right) = \frac{\beta}{1+\beta-\alpha}\lambda$

$\pi\left(D, w_D; \mathcal{V}^c\gamma, \cdot\right) = \beta\left(1-\lambda\right) + \frac{\beta}{1+\beta-\alpha}\lambda$

$\pi\left(\mathcal{V}^c, w^-_{\mathcal{V}^c}; \mathcal{V}^c\gamma, \cdot\right) =$
$\left(p\left(w^-_{\mathcal{V}^c}\right)\frac{\beta}{1+\beta-\alpha} + \left(1-p\left(w^-_{\mathcal{V}^c}\right)\right)\right)\beta\left(1-\lambda\right) + \frac{\beta}{1+\beta-\alpha}\lambda$

$\pi\left(\mathcal{V}^c, w^+_{\mathcal{V}^c}; \mathcal{V}^c\gamma, \cdot\right) =$
$\left(p\left(w^+_{\mathcal{V}^c}\right)\frac{\beta}{1+\beta-\alpha} + \left(1-p\left(w^+_{\mathcal{V}^c}\right)\right)\right)\left(1-\lambda\right) + \frac{\beta}{1+\beta-\alpha}\lambda$

Note that initially is: $w_H = w_D$.

The payoff in the critical area relates as follows:

- If $\boldsymbol{\beta} > \boldsymbol{\gamma}^*+\boldsymbol{\delta} \Rightarrow \boldsymbol{\beta} > \boldsymbol{\alpha} \rightsquigarrow$

  $\pi\left(\mathcal{V}^c, w^-_{\mathcal{V}^c}; \mathcal{V}^c\gamma, \mathbf{F}\right) < \pi\left(D, w_D; \mathcal{V}^c\gamma, \mathbf{F}\right) < \pi\left(\mathcal{V}^c, w^+_{\mathcal{V}^c}; \mathcal{V}^c\gamma, \mathbf{F}\right), \forall w_D \in \operatorname{supp}\mathbf{F}$

  $\pi\left(H, w_H; \mathcal{V}^c\gamma, \mathbf{F}\right) < \pi\left(\mathcal{V}^c, w^-_{\mathcal{V}^c}; \mathcal{V}^c\gamma, \mathbf{F}\right), \forall w_H \in \operatorname{supp}\mathbf{F}$

- If $\boldsymbol{\beta} \leq \boldsymbol{\gamma}^*+\boldsymbol{\delta} \Rightarrow \boldsymbol{\beta} < \boldsymbol{\alpha} \rightsquigarrow$

  $\pi\left(H, w_H; \mathcal{V}^c\gamma, \mathbf{F}\right) < \pi\left(\mathcal{V}^c, w^-_{\mathcal{V}^c}; \mathcal{V}^c\gamma, \mathbf{F}\right), \forall w_H \in \operatorname{supp}\mathbf{F}$

  $\pi\left(D, w_D; \mathcal{V}^c\gamma, \mathbf{F}\right) = \pi\left(\mathcal{V}^c, w^-_{\mathcal{V}^c}; \mathcal{V}^c\gamma, \mathbf{F}\right), \forall w_D \in \operatorname{supp}\mathbf{F}$

- Mutation

  Note that $\nexists m \in \mathbf{S} \setminus \mathcal{V}^c$ such that $\pi\left(\mathcal{V}^c, w; \mathcal{V}^c\gamma, \mathbf{F}\right) < \pi\left(m, w; \mathcal{V}^c\gamma, \mathbf{F}\right)$ for $w > w_D$; also, $\nexists m \in \mathbf{S} \setminus \mathcal{V}^c$ such that $\pi\left(\mathcal{V}^c, w_D; \mathcal{V}^c\gamma, \mathbf{F}\right) < \pi\left(m, w; \mathcal{V}^c\gamma, \mathbf{F}\right)$ for $w < w_D$. Assume that the entire population where $w < w_D$ will begin to change their strategies to D. As the population of D will grow in accordance to assumption 2 and 4, this will open for mutations where pure D will begin to play H against other D at w_D, i.e. $\left(D, \cdot; \mathcal{V}^c\gamma, \mathbf{F}\right) \curvearrowright \left(D : s_u = H, \cdot; \mathcal{V}^c\gamma, \mathbf{F}\right)$. Note that this mutation still will be playing D against $\forall w \neq w_D$. In either case, as the evolution proceeds, will the remaining $\left(\mathcal{V}^c, w; \mathcal{V}^c\gamma, \mathbf{F}\right), w > w_D$ grow in accordance to assumption 2 and 4. Thus as $t \to \infty$ we will, through growth and mutations, have that $\max\pi\left(m, \cdot; \mathbf{s}, \widetilde{\mathbf{F}}\right) \leq \min\pi\left(\mathcal{V}^c, \cdot; \mathbf{s}, \widetilde{\mathbf{F}}\right) < \max\pi\left(\mathcal{V}^c, \cdot; \mathbf{s}, \widetilde{\mathbf{F}}\right)$for $m \in \mathbf{S}\setminus\mathcal{V}^c$.

■

Proposition 3 *$\mathcal{V}^c$ is according to criterion $\mathfrak{C}$* not *vulnerable for invasion by $\mathcal{R}^c$, given $\boldsymbol{\beta} > \boldsymbol{\gamma}^*+\boldsymbol{\delta}$ as $\delta \to 0$.*

Proof. Assume that the size of the population coordinating the game according by $\mathcal{R}^c$ at the time of the fusion is λ and $\mathcal{V}^c$ is $1-\lambda$. Since $\overline{\pi}\left(\mathcal{R}^c; \mathcal{R}^c, F_{\mathcal{R}^c}\right) < \overline{\pi}\left(\mathcal{V}^c; \mathcal{V}^c, F_{\mathcal{V}^c}\right) \rightsquigarrow \lambda < \frac{1}{2}$.

- Now consider when the populations meet; $F_{\mathcal{V}^c}\left(\mathcal{V}^c\right), F_{\mathcal{R}^c}\left(\mathcal{R}^c\right) \to \mathbf{F}\left(\mathcal{V}^c\mathcal{R}^c\right)$
  - The payoff in the critical area are:
    $\pi\left(\mathcal{R}^c, w^+_{\mathcal{R}^c}; \mathcal{V}^c\mathcal{R}^c, \mathbf{F}\right) = \beta\lambda$
    $\pi\left(\mathcal{V}^c, w^-_{\mathcal{V}^c}; \mathcal{V}^c\mathcal{R}^c, \mathbf{F}\right) = \beta\left(1-\lambda\right)$
- Mutation

  Note that both $\left(\mathcal{V}^c, w_{\mathcal{V}^c}; \mathcal{V}^c\mathcal{R}^c, \mathbf{F}\right)$ and $\left(\mathcal{R}^c, w_{\mathcal{R}^c}; \mathcal{V}^c\mathcal{R}^c, \mathbf{F}\right)$ are vulnerable for invasion.

  - $\left(\mathcal{R}^c, w_{\mathcal{R}^c}; \mathcal{V}^c\mathcal{R}^c, \mathbf{F}\right) \curvearrowright \left(\mathcal{V}^c, \cdot; \mathcal{V}^c\mathcal{R}^c, \mathbf{F}\right)$
  - $\left(\mathcal{V}^c, w_{\mathcal{V}^c}; \mathcal{V}^c\mathcal{R}^c, \mathbf{F}\right) \curvearrowright \left(D, \cdot; \mathcal{V}^c\mathcal{R}^c D, \mathbf{F}\right)$

  Let the proportion playing $\mathcal{V}^c$ or $\mathcal{R}^c$ be denoted η and D be $\left(1-\eta\right)$.

  - The payoff in the critical area relates as follows:
    $\pi\left(D, w_D; \mathcal{V}^c\mathcal{R}^c D, \mathbf{F}\right) = \beta\eta\lambda + \alpha\eta\left(1-\lambda\right) + \alpha\left(1-\eta\right)$
    $\pi\left(\mathcal{V}^c, w^+_{\mathcal{V}^c}; \mathcal{V}^c\mathcal{R}^c D, \mathbf{F}\right) = \eta\left(1-\lambda\right) + \alpha\left(1-\eta\right)$
    $\rightsquigarrow \pi\left(D, w^D; \mathcal{V}^c\mathcal{R}^c D, \mathbf{F}\right) \geq \pi\left(\mathcal{V}^c, w^+_{\mathcal{V}^c}; \mathcal{V}^c\mathcal{R}^c D, \mathbf{F}\right)$
    $\Leftrightarrow \beta\lambda \geq \left(1-\alpha\right)\left(1-\lambda\right)$
  - If $\beta\lambda < \left(1-\alpha\right)\left(1-\lambda\right)$ then $\nexists m \in \mathbf{S} \setminus \mathcal{V}^c$ such that $\pi\left(\mathcal{V}^c, w; \mathcal{V}^c\mathcal{R}^c D, \mathbf{F}\right) < \pi\left(m, w; \mathcal{V}^c\mathcal{R}^c D, \mathbf{F}\right)$ for $w > w^D$ as $t \to \infty$, and the proposition follows as in proposition 2.
  - Now consider $\beta\lambda \geq \left(1-\alpha\right)\left(1-\lambda\right)$.
    Since $\pi\left(\mathcal{R}^c, w^+_{\mathcal{R}^c}; \mathcal{V}^c\mathcal{R}^c, \mathbf{F}\right) < \pi\left(\mathcal{V}^c, w^-_{\mathcal{V}^c}; \mathcal{V}^c\mathcal{R}^c, \mathbf{F}\right) \rightsquigarrow \lambda \to 0$ asymptotically
    Similar $\pi\left(\mathcal{V}^c, w^+_{\mathcal{V}^c}; \mathcal{V}^c\mathcal{R}^c D, \mathbf{F}\right) \leq \pi\left(D, w_D; \mathcal{V}^c\mathcal{R}^c D, \mathbf{F}\right) \rightsquigarrow \eta \to 0$ asymptotically
    As soon as $\eta < 1$, i.e. $\left(\mathcal{V}^c, w_{\mathcal{V}^c}; \mathcal{V}^c\mathcal{R}^c, \mathbf{F}\right) \curvearrowright \left(D, \cdot; \mathcal{V}^c\mathcal{R}^c D, \mathbf{F}\right)$, we have that $\pi\left(D, w_D; \mathcal{V}^c\mathcal{R}^c D, \mathbf{F}\right) < \pi\left(D : s_u = H, w_D; \mathcal{V}^c\mathcal{R}^c D, \mathbf{F}\right)$. That is $\left(D, w_D; \mathcal{V}^c\mathcal{R}^c D, \mathbf{F}\right)$ are vulnerable for invasion.

- $(D, w_D; \mathcal{V}^c\mathcal{R}^c D, \mathbf{F}) \curvearrowright (D : s_u = H, w_D; \mathcal{V}^c\mathcal{R}^c D, \mathbf{F})$

  However since we have assumed that mutation compared with growth is slow $\rightsquigarrow p\,(D : s_u = H, w_D) < \frac{1-\alpha}{1+\beta-\alpha}$. Let this mix at w_D be denoted *DH*. Furthermore it follows from assumption 2 and 4, that $\exists t$ such that $\eta > 0$, where λ is small enough such that $\beta\lambda = (1-\alpha)(1-\lambda)$ $\Rightarrow \pi\,(D, w_{DH}; \mathcal{V}^c\mathcal{R}^c DH, \mathbf{F}) = \pi\left(\mathcal{V}^c, w^+_{\mathcal{V}^c}; \mathcal{V}^c\mathcal{R}^c DH, \mathbf{F}\right)$. However note that $\pi\left(\mathcal{V}^c : s_u = H, w^+_{\mathcal{V}^c}; \mathcal{V}^c\mathcal{R}^c DH, \mathbf{F}\right) > \pi\,(D, w_{DH}; \mathcal{V}^c\mathcal{R}^c DH, \mathbf{F})$. But in difference to agents with $(D : s_u = H)$ will $(\mathcal{V}^c : s_u = H)$ not converge back to w_{DH} but remain above at $w^+_{\mathbf{F}}$.

  As soon as $(\mathcal{V}^c : s_u = H)$ has established themselves above w_{DH}, we have that:

  $\pi\left(\mathcal{V}^c : s_u = H, w^+_{\mathcal{V}^c}; \mathcal{V}^c\mathcal{R}^c DH, \mathbf{F}\right) < \pi\left(\mathcal{V}^c : s_u = D, w^+_{\mathcal{V}^c}; \mathcal{V}^c\mathcal{R}^c DH, \mathbf{F}\right)$. Hence we have that $\left(\mathcal{V}^c : s_u = H, w^+_{\mathcal{V}^c}; \mathcal{V}^c\mathcal{R}^c DH, \mathbf{F}\right)$ are vulnerable for invasion.

- $\left(\mathcal{V}^c : s_u = H, w^+_{\mathbf{F}}; \mathcal{V}^c\mathcal{R}^c DH, \mathbf{F}\right) \curvearrowright \left(\mathcal{V}^c : s_u = D, w^+_{\mathbf{F}}; \mathcal{V}^c\mathcal{R}^c DH, \mathbf{F}\right)$

  As λ becomes smaller more and more $\mathcal{V}^c$ will established themselves above w_{DH}. Moreover, since the $(\mathcal{V}^c, w; \mathcal{V}^c\gamma, \mathbf{F}), w > w_D$ will grow in accordance to assumption 2 and 4. Thus, as $t \to \infty$ we will, through growth and mutations, have that $\max\pi\left(\mathcal{R}^c, \cdot; \mathbf{s}, \widetilde{\mathbf{F}}\right) < \max\pi\left(DH, \cdot; \mathbf{s}, \widetilde{\mathbf{F}}\right) \leq \min\pi\left(\mathcal{V}^c, \cdot; \mathbf{s}, \widetilde{\mathbf{F}}\right) < \max\pi\left(\mathcal{V}^c, \cdot; \mathbf{s}, \widetilde{\mathbf{F}}\right)$ and thus also $\max\pi\left(m, \cdot; \mathbf{s}, \widetilde{\mathbf{F}}\right) \leq \min\pi\left(\mathcal{V}^c, \cdot; \mathbf{s}, \widetilde{\mathbf{F}}\right) < \max\pi\left(\mathcal{V}^c, \cdot; \mathbf{s}, \widetilde{\mathbf{F}}\right)$ for $m \in \mathbf{S}\backslash\mathcal{V}^c$.

■

Claim 20 $\overline{\pi}\,(\mathcal{B}; \mathcal{B}, F_y) = \overline{\pi}\,(\mathcal{P}; \mathcal{P}, F_y)$.

Proof. Consider two arbitrary points in the distribution.

$\left|y_i^n - y_{-i}^n\right| < \delta^n \Rightarrow \pi\,(\cdot, y_i; \cdot, F_y) = \pi\,(\cdot, y_{-i}; \cdot, F_y)$

$\left|y_i^n - y_{-i}^n\right| > \delta^n \Rightarrow \frac{\pi(\cdot, y_i; \cdot, F_y) + \pi(\cdot, y_{-i}; \cdot, F_y)}{2} = \frac{1+\beta}{2}$ ■

Lemma 8 $\exists\,(\mathcal{V}^c, w, y) \in \operatorname{supp}\mathbf{F}\,(\mathcal{V}^c\mathcal{Z})$ *such that*

$\pi\,(\mathcal{V}^c, w; \mathcal{V}^c\mathcal{Z}, \mathbf{F}) \nless \pi\,(\mathcal{Z}, w; \mathcal{V}^c\mathcal{Z}, \mathbf{F})$.

Proof.

$$\pi\left(\mathcal{V}^c, y, w; \mathcal{V}^c\mathcal{Z}, \mathbf{F}\right) =$$
$$(1-\lambda)\left(F_w + \beta\left(1-F_w\right)\right) + \lambda\left(F_wF_y + \alpha F_y\left(1-F_w\right) + \beta\left(1-F_w\right)\left(1-F_y\right)\right)$$

Now let $\mathcal{V}^c \curvearrowright \mathcal{Z} \rightsquigarrow$

$$\pi\left(\mathcal{Z}, y, w; \mathcal{V}^c\mathcal{Z}, \mathbf{F}\right) =$$
$$(1-\lambda)\left(F_wF_y + \alpha F_w\left(1-F_y\right) + \beta\left(1-F_w\right)\left(1-F_y\right)\right) + \lambda\left(F_y + \beta\left(1-F_y\right)\right)$$

$\pi\left(\mathcal{V}^c, w; \mathcal{V}^c\mathcal{Z}, \mathbf{F}\right) < \pi\left(\mathcal{Z}, w; \mathcal{V}^c\mathcal{Z}, \mathbf{F}\right) \Leftrightarrow$

$$(1-\lambda)\left(F_w + \beta\left(1-F_w\right)\right) + \lambda\left(F_wF_y + \alpha F_y\left(1-F_w\right) + \beta\left(1-F_w\right)\left(1-F_y\right)\right)$$
$$< (1-\lambda)\left(F_wF_y + \alpha F_w\left(1-F_y\right) + \beta\left(1-F_w\right)\left(1-F_y\right)\right) + \lambda\left(F_y + \beta\left(1-F_y\right)\right)$$

$$\pi\left(\mathcal{V}^c, w; \mathcal{V}^c\mathcal{Z}, \mathbf{F}\right) < \pi\left(\mathcal{Z}, w; \mathcal{V}^c\mathcal{Z}, \mathbf{F}\right)$$
$$\Leftrightarrow$$
$$F_w\left(1-F_y\right)(1-\alpha) + \beta F_y\left(1-F_w\right) < \lambda(1-\alpha+\beta)\left(F_y\left(1-F_w\right) + F_w\left(1-F_y\right)\right)$$

If v and w are *i.i.d.* $\Rightarrow \exists F_y, F_w, \forall \alpha, \beta$ such that

$F_w\left(1-F_y\right)(1-\alpha) + \beta F_y\left(1-F_w\right) \not< \lambda(1-\alpha+\beta)\left(F_y\left(1-F_w\right) + F_w\left(1-F_y\right)\right)$

■

Lemma 9 $\pi\left(\mathcal{Z}, w; \mathcal{V}^c\mathcal{Z}, \mathbf{F}\right) < \pi\left(\mathcal{V}^c, w; \mathcal{V}^c\mathcal{Z}, \mathbf{F}\right), \forall\left(w, \mathcal{Z}\right) \in \operatorname{supp}\mathbf{F}\left(\mathcal{V}^c\mathcal{Z}\right)$.

Proof. Assume the opposite; $\exists\left(w, \mathcal{Z}\right) \in \operatorname{supp}\mathbf{F}\left(\mathcal{V}^c\mathcal{Z}\right)$ such that

$\pi\left(\mathcal{Z}, w; \mathcal{V}^c\mathcal{Z}, \mathbf{F}\right) \geq \pi\left(\mathcal{V}^c, w; \mathcal{V}^c\mathcal{Z}, \mathbf{F}\right), \forall\left(w, \mathcal{Z}\right)$

$$\pi\left(\mathcal{Z}, y, w; \mathcal{V}^c\mathcal{Z}, \mathbf{F}\right) =$$
$$(1-\lambda)\left(F_wF_y + \alpha F_w\left(1-F_y\right) + \beta\left(1-F_w\right)\left(1-F_y\right)\right) + \lambda\left(F_y + \beta\left(1-F_y\right)\right)$$

Now let $\mathcal{Z} \curvearrowright \mathcal{V}^c \rightsquigarrow$

$$\pi\left(\mathcal{V}^c, y, w; \mathcal{V}^c\mathcal{Z}, \mathbf{F}\right) =$$
$$(1-\lambda)\left(F_w + \beta\left(1-F_w\right)\right) + \lambda\left(F_y + \beta\left(1-F_y\right)\right)$$

$$\pi\left(\mathcal{Z}, w; \mathcal{V}^c\mathcal{Z}, \mathbf{F}\right) \geq \pi\left(\mathcal{V}^c, w; \mathcal{V}^c\mathcal{Z}, \mathbf{F}\right), \forall\left(w, \mathcal{Z}\right)$$
$$\Leftrightarrow$$
$$0 \geq (1-\alpha)F_w\left(1-F_y\right) + \beta\left(1-F_w\right)F_y$$

This is a contradiction, why the lemma should be proven. ■

Proposition 4 *$\mathcal{V}^c$ is according to criterion $\mathfrak{C}$* not *vulnerable for invasion by any strategy conditioned on any static variable y^s, given $\boldsymbol{\gamma} = \boldsymbol{\gamma}^*$ as $\delta \to 0$.*

Proof. Assume that the size of the population conditioning the strategies according to $\mathcal{Z}$ at the time of the fusion is λ and $\mathcal{V}^c$ is $1-\lambda$. From assumption 6 we have that $\overline{\pi}\left(y; y, F_y\right) = \overline{\pi}\left(\mathcal{Z}; \mathcal{Z}, F_{\mathcal{Z}}\right) < \overline{\pi}\left(\mathcal{V}^c; \mathcal{V}^c, F_{\mathcal{V}^c}\right) \rightsquigarrow \lambda < \frac{1}{2}$.

- Now consider when the populations meet; $F_{\mathcal{V}^c}\left(\mathcal{V}^c\right), F_{\mathcal{Z}}\left(\mathcal{Z}\right) \rightarrow \mathbf{F}\left(\mathcal{V}^c\mathcal{Z}\right)$

- Mutation

  $F_{\mathcal{Z}} \in \mathbb{F}_{\mathbf{s}}^{\mathfrak{A}}$ leads to that the agents conditioning the strategies according to $\mathcal{Z}$ will play H against $w_{-i} < w_i$ and play D against $w_{-i} > w_i, \forall w_i \in \operatorname{supp} F_{\mathcal{Z}}\left(w; \mathcal{Z}\right)$. Hence from lemma 9 we have that $\pi\left(\mathcal{Z}, w; \mathcal{V}^c\mathcal{Z}, \mathbf{F}\right) < \pi\left(\mathcal{V}^c, w; \mathcal{V}^c\mathcal{Z}, \mathbf{F}\right), \forall\left(w, \mathcal{Z}\right) \in \operatorname{supp} \mathbf{F}\left(\mathcal{V}^c\mathcal{Z}\right)$. That is, every agent conditioning the game according to $\mathcal{Z}$ will receive a higher payoff if the strategy is changed to $\mathcal{V}^c$, since they then will be able to coordinate with the entire population. Also note that these mutants initially will be able to coordinate the game perfectly with each agent and thus receive a higher payoff and subsequently rise to a higher status in the distribution. If the agent would mutate back, he would receive a decrease in the payoff compared to the original payoff. No further successful mutations are thus feasible once the agent has mutated from $\mathcal{Z}$ to $\mathcal{V}^c$.

  Furthermore, from lemma 8 we know that there always $\exists\left(\mathcal{V}^c, w, y\right) \in \operatorname{supp} \mathbf{F}\left(\mathcal{V}^c\mathcal{Z}\right)$ such that $\pi\left(\mathcal{V}^c, w; \mathcal{V}^c\mathcal{Z}, \mathbf{F}\right) \nless \pi\left(\mathcal{Z}, w; \mathcal{V}^c\mathcal{Z}, \mathbf{F}\right)$. Nevertheless, any agent mutating from $\mathcal{V}^c$ to $\mathcal{Z}$ will, according to lemma 9, prefer to change back from $\mathcal{Z}$ to $\mathcal{V}^c$ as soon as the new distribution has settled.

  Finally note that $\lambda < \frac{1}{2} \Rightarrow \max \pi\left(\mathcal{Z}, w; \mathcal{V}^c\mathcal{Z}, \mathbf{F}\right) < \max \pi\left(\mathcal{V}^c, w; \mathcal{V}^c\mathcal{Z}, \mathbf{F}\right)$

  Thus as $t \rightarrow \infty$ we will, through growth and mutation, have that
  $\max \pi\left(\mathcal{Z}, \cdot; \mathcal{Z}, \widetilde{\mathbf{F}}\right) \leq \min \pi\left(\mathcal{V}^c, \cdot; \mathcal{V}^c, \widetilde{\mathbf{F}}\right) < \max \pi\left(\mathcal{V}^c, \cdot; \mathcal{V}^c, \widetilde{\mathbf{F}}\right)$

■

References

[1] Kenneth G. Binmore. *Game Theory and the Social Contract II: Just Playing.* MIT Press, Cambridge, Massachusetts, 1994.

[2] Harold L. Cole, George J. Mailath, and Andrew Postlewaite. Social norms, savings behavior, and growth. *Journal of Political Economy*, 100(6):1092–1125, 1992.

[3] Harold L. Cole, George J. Mailath, and Andrew Postlewaite. Incorporating concern for relative wealth into economic models. *Federal Reserve Bank of Minneapolis Quarterly Review*, 19(3):12–21, 1995.

[4] Robert H. Frank. *Choosing the Right Pond.* Oxford University Press, New York, New York, 1985.

[5] Robert H. Frank. The demand for unobservable and other nonpositional goods. *American Economic Review*, 75(1):101–116, 1985.

[6] Robert H. Frank. *Passions Within Reason.* W. W. Norton, New York, New York, 1988.

[7] John H. Kagel and Alvin E. Roth, editors. *Handbook of Experimantal Economics.* Princeton University Press, Princeton, 1995.

[8] Andrew Postlewaite. The social basis of interdependent preferences. *European Economic Review*, 42(3-5):779–800, 1998.

[9] Mari Rege. Why do people care about social status? Working paper, Department of Economics, Weatherhead School of Managment, Case Western Reserve University, 2002.

[10] Thomas C. Schelling. *The Strategy of Conflict.* Harvard University Press, Cambridge, Massachusetts, 1960.

[11] Adam Smith. *The Theory of Moral Sentiments.* A. Millar, London, 1759.

[12] John Maynard Smith. *Evolution and the Theory of Games.* Cambridge University Press, Cambridge, Massachusetts, 1982.

[13] Michael A. Spence. *Market Signaling: Informational Transfer in Hiring and Related Screening Processes.* Harvard University Press, Cambridge, Massachusetts, 1974.

[14] Thorstein Veblen. *The Theory of the Leisure Class.* Houghton Mifflin, Boston, 1899. Reprint 1973.

[15] Yoram Weiss and Chaim Fershtman. Social status and economic performance: A survey. *European Economic Review*, 42:802–820, 1998.