# Dynamic stability and basins of attraction in the Sir Philip Sidney game

Simon M. Huttegger
Department of Logic and Philosophy of Science
University of California at Irvine
3151 Social Science Plaza A
Irvine, CA-92697

shuttegg@uci.edu

Kevin J. S. Zollman
Department of Philosophy
Carnegie Mellon University
Baker Hall 135
Pittsburgh, PA 15213-3890

kzollman@andrew.cmu.edu

## Abstract

We study the handicap principle in terms of the Sir Philip Sidney game. The handicap principle asserts that cost is required to allow for honest signalling in the face of conflicts of interest. We show that the significance of the handicap principle can be challenged from two new directions. Firstly, both the costly signalling equilibrium and certain states of no communication are stable under the replicator dynamics (i.e., standard evolutionary dynamics); however, the latter states are more likely in cases where honest signalling should apply. Secondly, we prove the existence and stability of polymorphisms where players mix between being honest and being deceptive and where signalling costs can be very low. Neither the polymorphisms nor the states of no communication are evolutionarily stable, but they turn out to be more important for standard evolutionary dynamics than the costly signalling equilibrium.

2

# 1 Introduction

The Sir Philip Sidney game (Maynard Smith, 1991) has been a locus of attention for the analysis of costly signalling between relatives. It represents perhaps the simplest game that captures the central notion of handicaps (due to Zahavi, 1975); i.e., that cost is required to maintain honest signalling in the presence of partial conflict of interest. The canonical example for the Sir Philip Sidney game is the interaction between feeding chicks and their parents. Maynard Smith introduced the discrete Sir Philip Sidney game as a simplified version of Grafen's 1990 notable costly signalling model (see also Johnstone and Grafen, 1992, 1993; Bergstrom and Lachmann, 1997, 1998; Lachmann and Bergstrom, 1998; Godfray and Johnstone, 2000; Brilot and Johnstone, 2003; Hamblin and Hurd, 2009). Many variations of the initial game have been considered, and there is an ongoing debate about how widespread this phenomenon is (Maynard Smith and Harper, 2003; Searcy and Nowicki, 2005). However, much of this debate retains the central methodology of calculating Evolutionarily Stable States (ESS) (Maynard Smith and Price, 1973; Maynard Smith, 1982).

In this paper we conduct a dynamic analysis of the original Sir Philip Sidney game, which identifies gaps in our understanding of honest signalling caused by the strong focus on ESS. We concentrate on two issues. Firstly, we contrast the evolutionary significance of states of perfect communication (the signalling ESS) and states of no communication (pooling equilibria). In most of the cases relevant for costly signalling, the pooling equilibrium appears to be the most probable evolutionary outcome in terms of having the larger basin of attraction under standard evolutionary dynamics (Section 4). Secondly, while the existence of the signalling ESS and of pooling equilibria is a well established result in the literature, we prove that, in addition, a family of polymorphic equilibria exists which allow for partial communication (Section 2). These polymorphisms can establish partly honest signalling even in the face of low signalling costs. We prove dynamic stability of the polymorphic equilibria (Section 3), and show that this outcome is in many cases more likely than the signalling ESS (Section 4). These results have some important biological implications. We

suggest in particular that one should not place to much weight on perfect communication in situations involving conflicting interests; instead, one may look for alternative equilibria such as the polymorphisms described below. We shall discuss this issue more thoroughly in the concluding section of this paper.

## 2   The Sir Philip Sidney game

The Sir Philip Sidney game proceeds in two stages. A sender can be either healthy (with probability $1 - m$) or needy (with probability $m$). In both states, the sender can send a signal bearing a cost $c > 0$, or decline sending a signal at all. In the second stage, after having (or not having) received the signal, a donor may respond by transferring a resource $d > 0$ to the sender and thus reducing her survival probability to $1 - d$, or may abstain from doing so. The donor does not know the sender's true state. The extensive form of this game is depicted in Figure 1. Without receiving the resource, a sender's probability of surviving is $1 - a$ if she is needy and $1 - b$ if she is healthy; we assume that $a > b$ throughout this paper (i.e. a needy individual profits more from receiving the resource).

The extensive form game of Figure 1 does not allow for communication since the donor's strategy of never transferring the resource is strictly dominant. Introducing a relatedness parameter $k \in [0, 1]$ permits a higher degree of common interest between the sender and the donor. At each outcome, a player receives her own payoff plus $k$ times the payoff of the other player. This gives rise to a four-by-four strategic-form game. A player's strategies are given by a rule of behavior which tells her what to do at each of her information sets. These strategies are enumerated in Figure 2.

Investigators often reduce the four-by-four strategic form structure by ignoring the strategy 'signal only if healthy' and the corresponding donor strategy (like Bergstrom and Lachmann, 1997; Maynard Smith and Harper, 2003). In this truncated game, two kinds of equilibria are usually considered. The first one is the most important Nash equilibrium in
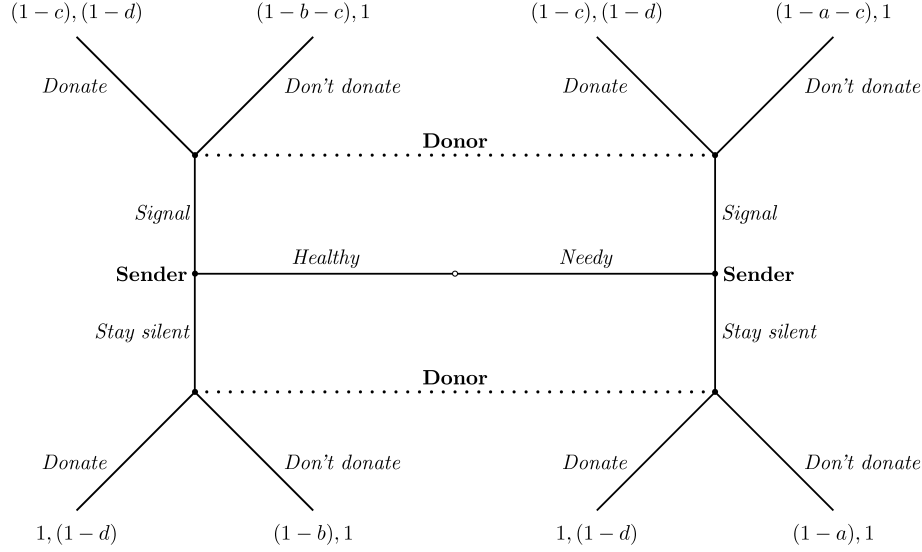
4

Figure 1: An extensive form representation of the Sir Philip Sidney game. The dotted lines represent the donor's information sets, i.e. the donor is unable to distinguish between the decision nodes connected by the dotted lines. The terminal nodes show the sender's and the donor's payoffs, respectively.

| | Sender strategies | | Donor strategies |
|---|---|---|---|
| $S_1$: | Signal only if healthy | $R_1$: | Donate only if no signal |
| $S_2$: | Signal only if needy | $R_2$: | Donate only if signal |
| $S_3$: | Never Signal | $R_3$: | Never donate |
| $S_4$: | Always Signal | $R_4$: | Always donate |

Figure 2: Sender and donor strategies in the Sir Philip Sidney Game.

the literature on handicap signals. It is given by the strategy pair 'signal only if needy' and 'donate only if signal,' and it is a Nash equilibrium if

$$a \geq c + kd \geq b \tag{1a}$$

$$a \geq \frac{d}{k} \geq b \tag{1b}$$

(see Bergstrom and Lachmann, 1997). At this equilibrium, cost can be said to maintain signal reliability since a needy sender is willing to pay the cost $c$ in order to get the resource, while a healthy sender does not pretend to be needy by sending the signal. If the inequalities

5

76 are strict, then the equilibrium is a strict Nash equilibrium and, therefore, also an ESS in

77 the symmetrised game (cf. Cressman, 2003). This equilibrium is usually called a signalling

78 ESS or a separating equilibrium.

79     There always exists a pooling equilibrium where no information is transferred. Bergstrom

80 and Lachmann (1997) show that the profile 'never donate' and 'never signal' is a Nash

81 equilibrium if

$$d > k(ma + (1 - m)b); \tag{2}$$

82 if this inequality is reversed, then the profile 'always donate' and 'never signal' is a Nash

83 equilibrium. Since the signal is not sent at both equilibria, these profiles can be called pooling

84 equilibria.

85     In Appendix B, we prove that 'signal only if healthy' and 'donate only if no signal' is a

86 Nash equilibrium if (1b) holds together with

$$a \geq kd - c \geq b. \tag{3}$$

87 This equilibrium is often ignored by other investigators on grounds of requiring too much

88 common interest between sender and donor to be relevant for costly signalling theory (May-

89 nard Smith, 1991; Bergstrom and Lachmann, 1997). We think that this exclusion is a

90 mistake. Firstly, the existence conditions of this equilibrium play a role in interpreting dy-

91 namical properties of the Sir Philip Sidney game as parameters are varied (see Sections 3

92 and 4). Moreover, $R_1$ is important for the equilibrium structure of one of the pooling equi-

93 libria. Once we allow all sender and donor strategies, a pooling equilibrium will never be

94 a strict Nash equilibrium. Given that the sender chooses 'never signal', $R_2$ and $R_3$ as well

95 as $R_1$ and $R_4$ are behaviorally equivalent. In Appendix A, we show that pooling equilibria

96 are elements of larger sets of Nash equilibria. In the case of $(S_3, R_3)$, the donor can play a

mixture $(1 - \lambda)R_2 + \lambda R_3$. As long as

$$\lambda \geq 1 - \frac{c}{a - kd} \tag{4}$$

the strategy profile $(S_3, (1 - \lambda)R_2 + \lambda R_3)$ is a Nash equilibrium if (2) holds; otherwise, $(S_3, (1 - \mu)R_1 + \mu R_4)$ is a Nash equilibrium as long as

$$\mu \geq 1 - \frac{c}{kd - b}. \tag{5}$$

Since pooling equilibria are elements of a line of equilibria, they cannot be evolutionarily stable. One can show, however, that they correspond to neutrally stable strategies (Maynard Smith, 1982). We will come back to this point shortly in the broader context of dynamic stability (Section 3).

To the best of our knowledge, another kind of equilibrium has been completely overlooked in the biological literature. In the corresponding economics literature (Spence, 1973), similar equilibria—called hybrid equilibria—are known to exist but were considered unimportant. Recently, Wagner (2009) has shown that, contrary to received wisdom, a hybrid equilibrium can be very significant in Spence's game from the point of view of game dynamics. In the Sir Philip Sidney game there exists a family of polymorphisms which corresponds to hybrid equilibria. Each polymorphism is a mixed Nash equilibrium where the sender mixes between 'signal only when healthy' and 'always signal', while the donor mixes between 'donate if signal' and 'never donate.' The family of polymorphic equilibria is given by $\lambda S_2 + (1 - \lambda)S_4$ and $\mu R_2 + (1 - \mu)R_3$, where

$$\lambda = \frac{k(ma + (1 - m)b) - d}{(1 - m)(kb - d)} \quad \text{and} \quad \mu = \frac{c}{b - kd}. \tag{6}$$

7

$114$ Since $\lambda, \mu$ must be well defined, the polymorphism (6) exists if

$$a > \frac{d}{k} > b \quad \text{and} \quad b - kd > c \tag{7}$$

$115$ and if (2) is met. (The proof can be found in Appendix B.) The conditions (7) show that,
$116$ under generic conditions, the mixed Nash equilibrium (6) exists if and only if the signalling
$117$ ESS (1) does not exist. The value $c^* = b - kd$ is the minimum cost of a believable signal
$118$ (Bergstrom and Lachmann, 1997). Condition (7) implies that at the polymorphism the cost
$119$ of the signal will be lower than $c^*$. However, the polymorphic equilibrium sustains some
$120$ level of meaningful communication. Thus, information transfer is possible in a wider range
$121$ of cases than was previously believed. (Notice that partly honest communication here does
$122$ not arise because of the introduction of different sender types as in Johnstone and Grafen
$123$ (1993).)

$124$ # 3 Dynamic stability of equilibria

$125$ The multiplicity of equilibria in the Sir Philip Sidney game makes a priori conclusions about
$126$ its evolutionary outcomes difficult. In order to assess the evolutionary significance of the
$127$ different equilibria, we shall investigate their dynamic stability properties in terms of the
$128$ two-population replicator dynamics (Hofbauer and Sigmund, 1998). If $x_i$ is the relative
$129$ frequency of sender type $i$ and $y_j$ is the relative frequency of donor type $j$, $i, j = 1, \ldots, 4$,
$130$ then this dynamics is given by:

$$\dot{x}_i = x_i(\pi_i(\mathbf{y}) - \pi(\mathbf{x}, \mathbf{y})) \tag{8a}$$

$$\dot{y}_j = y_j(\pi_j(\mathbf{x}) - \pi(\mathbf{y}, \mathbf{x})) \tag{8b}$$

Here, $\mathbf{x} = (x_1, \ldots, x_4), \mathbf{y} = (y_1, \ldots, y_4)$, $\pi_i(\mathbf{y})$ is the payoff of $i$ against $\mathbf{y}$ and $\pi(\mathbf{x}, \mathbf{y})$ is the average payoff in the sender population; $\pi_j(\mathbf{x})$ is the payoff of $j$ against $\mathbf{x}$ and $\pi(\mathbf{y}, \mathbf{x})$ is the average payoff in the donor population. Many of our results also apply to the one-population replicator dynamics if we consider the symmetrised Sir Philip Sidney game (for details, see Cressman, 2003).

Under the relevant conditions ((1), and (1b) and (3) with strict inequalities, respectively) $(S_2, R_2)$ and $(S_1, R_1)$ are strict Nash equilibria and therefore asymptotically stable population states for (8). In Appendix A, we show that if the inequalities in (4) and (5) are strict, then the equilibria in the components (4) or (5) are quasi-strict. (An equilibrium is quasi-strict if there is no best repsonse to any of its components outside of its support.) In the symmetrised version of the game, these profiles are neutrally stable. Quasi-strictness, in turn, implies the following result (for details on why quasi-strictness implies dynamics stability in this case cf. Cressman, 2003).

**Theorem 1.** *The pooling equilibria given by (4) and (5) attract an open set of nearby population states under the appropriate conditions.*

Theorem 1 tells us that pooling equilibria are meaningful for the replicator dynamics since a non-negligible portion of initial populations will end up in the set of pooling equilibria.

The dynamic stability properties of the polymorphic equilibrium (6) are particularly interesting. In Appendix B, we prove two results: (i) All transversal eigenvalues of the Jacobian matrix $J$ of (8) evaluated at (6) are negative; and (ii), the two remaining eigenvalues of $J$ are purely imaginary. Both results hold if we assume that conditions (2) and (7) are met. (i) and (ii) imply that the polymorphism (6) is a spiraling centre; i.e., initial population states close to the polymorphism in the interior of the state space converge towards $K = \mathrm{span}(S_2, S_4) \times \mathrm{span}(R_2, R_3)$ and, once they get close enough, spiral around the polymorphism forever. The following theorem summarises the stability properties of the polymorphic equilibrium.
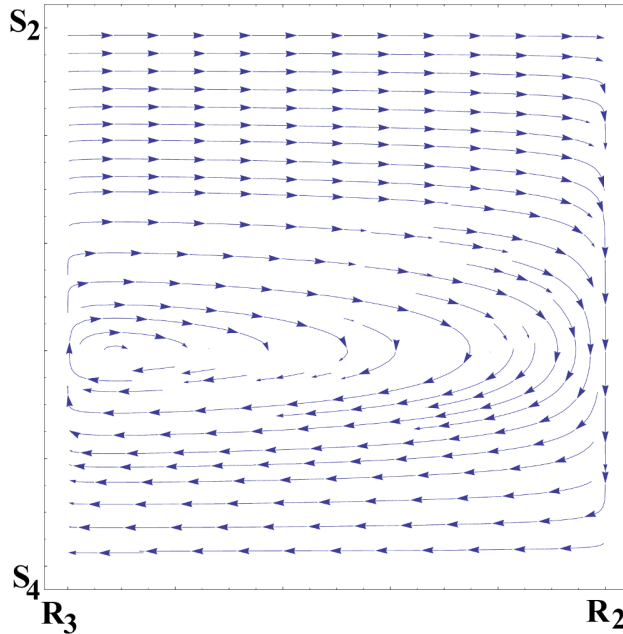
Figure 3: A phase portrait of one boundary face of the space of population states (face $K$) of the Sir Philip Sidney game illustrating the motion around the polymorphic equilibrium. The strategies are given in Figure 2. Close to the unique rest point, interior trajectories (of the full state space) converge to $K$. Notice the resemblance between this phase portrait and the phase portrait of the well known game of Matching Pennies under the replicator dynamics.

**Theorem 2.** *If* (2) *and* (7) *hold, then the polymorphic equilibrium* (6) *is Liapunov stable under the replicator dynamics* (8).

Since all transversal eigenvalues are negative, the interesting dynamical behavior is confined to $K$. The phase portrait of this face for one set of parameter values is depicted in Figure 3. We should note that this dynamical behavior is structurally unstable (Guckenheimer and Holmes, 1983); perturbations of the dynamics (8) will either result in an asymptotically stable or an unstable polymorphism. The issue of structural instability is a subtle one and needs more space; see Hofbauer and Huttegger (2008) for a case study. Suffice it to say here that many perturbations will result in an asymptotically stable polymorphism.

# 4    Basins of Attraction

The results of the previous section suggest that one cannot immediately use the ESS analysis of the Sir Philip Sidney game to conclude that signalling is likely to evolve. Although for many parameter configurations the separating equilibrium $(S_2, R_2)$ is an ESS, there are several other outcomes which have a non-negligible basin of attraction. Both pooling equilibria and also the other separating equilibrium can be outcomes of evolution. If we are interested in determining the likelihood that signalling will evolve, we have to estimate the relative sizes of the basins of attraction for the various outcomes. In this section, we will use numerical simulations to resolve this issue.

Investigators generally regard the region where $k$ is high to be uninteresting because signalling is stable for all values of $c$, including $c = 0$. We nevertheless start with the limiting case $k = 1$ as an important benchmark case. Figure 4 shows the size of the basins of attraction for *both* separating equilibria. The initial population states not converging to them are of non-negligible size. In fact, cost appears to hinder the evolution of successful signalling. Understanding this result is not difficult. Suppose that we have a population of donors who adopt either the strategy 'donate only if signal' or 'never donate'. Let the proportion of the former be represented by $\alpha$. The sender prefers the strategy 'signal only if needy' to the strategy 'never signal' only when $(15/32)\alpha > c$. So, as $c$ grows there must be a proportionally larger percentage of the donor population who will respond to a signal in order for signalling to be beneficial for the sender.

While costs generally hurt the evolution of signalling for these parameter values, it does not do so in every case. For extreme values of $m$, moderate cost appears to assist the evolution of signalling. The mechanism by which this result is produced is relatively complicated, but looking at a slightly simplified situation is insightful. Consider the initial population state illustrated in Figure 5 for $m = 0.1$ (i.e., the sender is rarely needy). If $c = 0$, there is no selection acting on the sender population at all. However, there is selection in the donor population in favor of the strategy 'never donate'. So without cost, the population converges
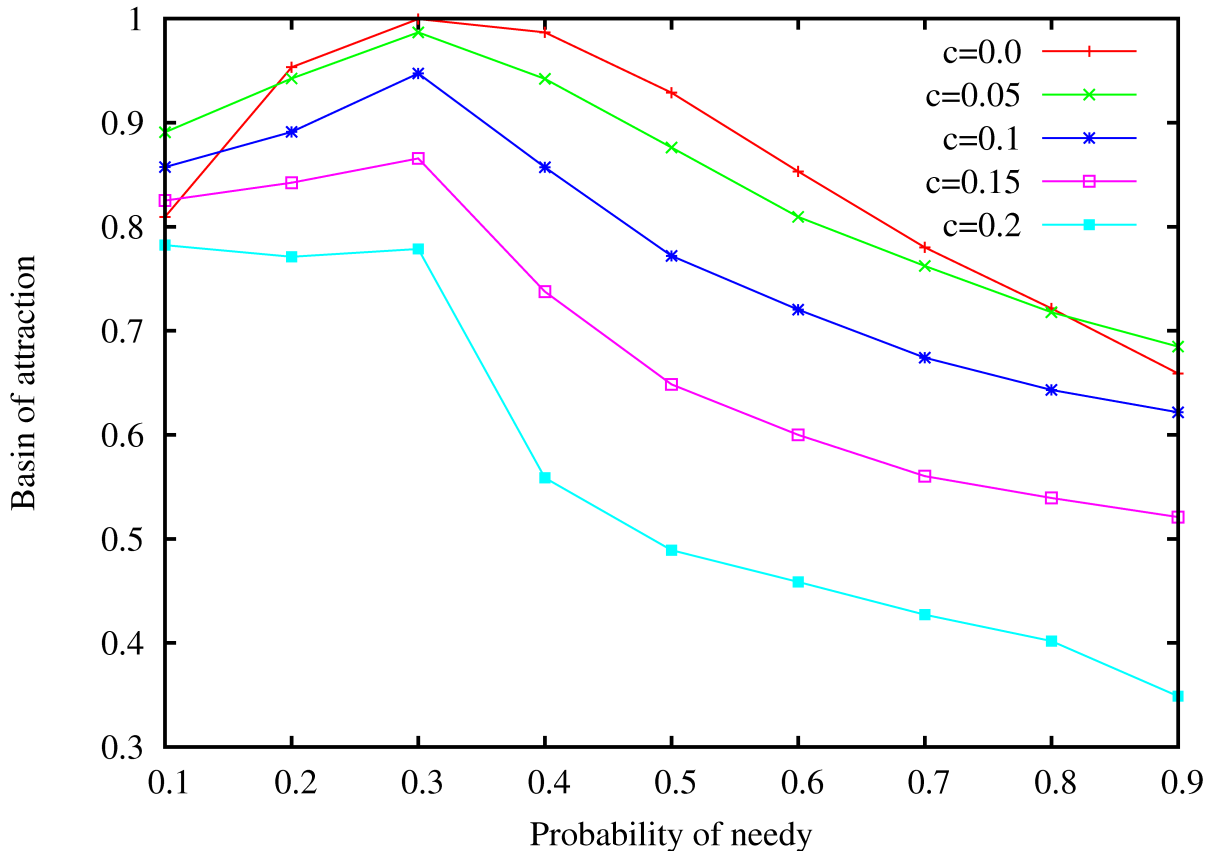
11

Figure 4: Simulation results showing the cumulative size of the basins of attraction for the two signaling equilibria when $k = 1$ for several differing values of $c$. The simulation utilizes the discrete time replicator dynamics when $a = 31/32$, $b = 9/32$, and $d = 1/2$.

| Strategy | Proportion | Payoff | Selective force |
|---|---|---|---|
| $S_1$ Signal if healthy | 0.05 | $1.626 - 0.9c$ | $-0.254c$ |
| $S_2$ Signal if needy | 0.01 | $1.626 - 0.1c$ | $0.546c$ |
| $S_3$ Never signal | 0.34 | $1.626$ | $0.646c$ |
| $S_4$ Signal always | 0.6 | $1.626 - c$ | $-0.354c$ |
| $R_1$ Donate if no signal | 0.09 | $2.22713 - 0.646c$ | |
| $R_2$ Donate if signal | 0.09 | $2.21388 - 0.646c$ | |
| $R_3$ Never donate | 0.75 | $2.3265 - 0.646c$ | |
| $R_4$ Always donate | 0.07 | $2.115 - 0.646c$ | |

Figure 5: An initial population and its associated payoffs that illustrates how a smaller $c$ might retard the evolution of signaling for extreme values of $m$.

to a pooling equilibrium. This occurs because there is little information in the sender's signal. When $c > 0$, there is selection acting on the sender population, primarily in favor of the strategy 'never signal'. Although the separating strategy 'signal only if needy' does worse than 'never signal', it does better than average and is initially selected for. (This is illustrated in the "Selective force" column, which is the payoff of that type minus the average payoff in the population.) As it increases in proportion it changes the selective force exercised on the donor population sufficiently that 'donate only if signal' becomes superior and grows. Once a sufficiently large percentage of the donor population is playing 'donate only if signal', the selective force exercised on the sender population shifts in favor of 'signal only if needy' over 'never signal', and the system evolves to a state with perfect information transfer. A similar situation obtains for the case of $m$ being sufficiently high.

This considers only the case of $k = 1$. Figure 6 illustrates what happens as $k$ is reduced. One will notice from Figure 4 that the basins of attraction for separating equilibria are maximised when $m = 0.3$. This represents the situation where, if the sender provides no information (by playing either $S_3$ or $S_4$), the donor is indifferent between donating or not. This indifference point changes as $k$ is reduced. For each value of $k$ in Figure 6 we set $m$ equal to this point of indifference so as to maximise the basins of attraction for signalling. This provides the most benign scenario for the evolution of signalling. In Figure 6 we see that so long as cost free signalling is an ESS, it always evolves. However, similar to what
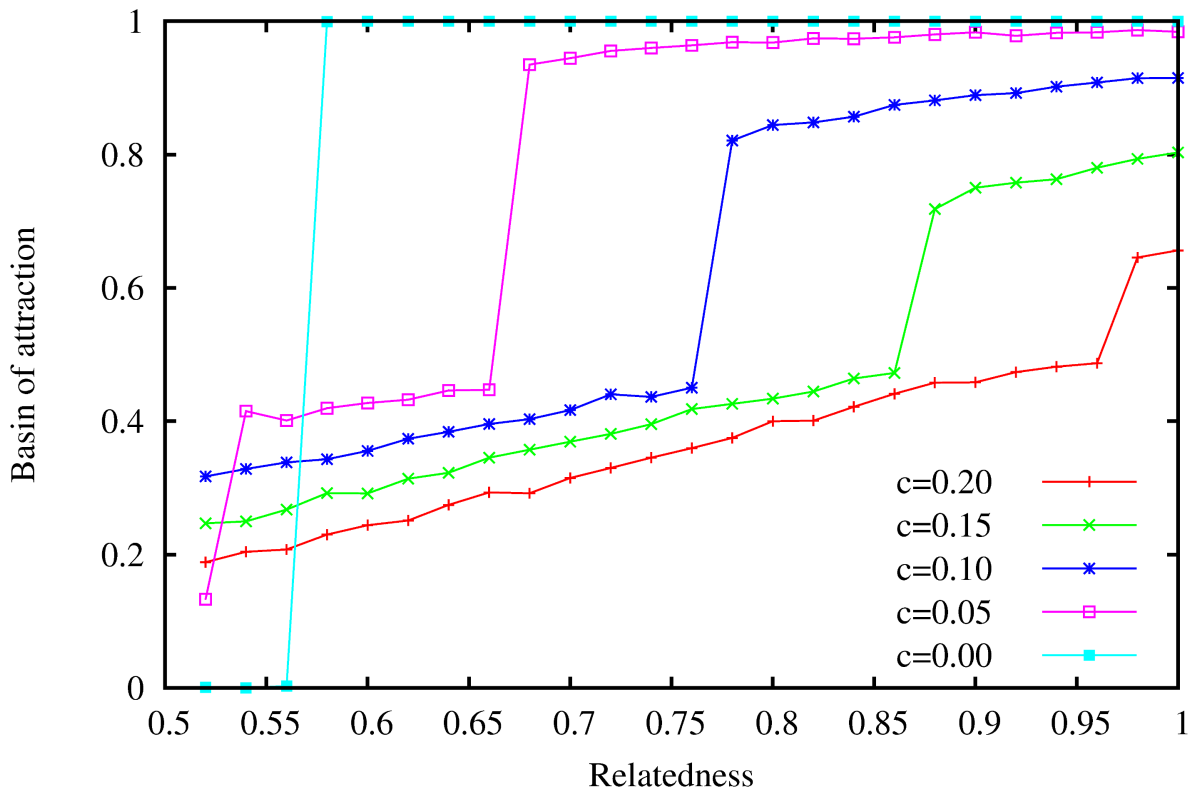
Figure 6: Simulation results showing the cumulative size of the basins of attraction for the two signaling equilibria as $k$ and $m$ vary. $m$ is set so as to maximize the basins of attraction of signaling given the specified $k$. As before, the simulation utilizes the discrete time replicator dynamics when $a = 31/32$, $b = 9/32$, and $d = 1/2$.

happens if $k = 1$, the presence of cost hinders the evolution of signalling and the higher costs result in smaller basins of attraction. One will notice that, as $k$ is reduced, there are sudden drops in the basins of attraction of costly signalling. This occurs as one crosses the boundaries in (3), when the equilibrium where the signal is used to indicate health no longer exists.

If $0.51 < k < 0.56$, then $k$ is sufficiently low that $(S_2, R_2)$ is only stable with some signal cost. Here one will notice a few interesting features. In the first place, although it is the only ESS, the separating equilibrium has a relatively small basin of attraction – it is always less than half the state space. This suggests that an analysis, based on the ESS concept, which ignores aspects of evolutionary processes can lead to misleading implications. Moreover, we see that the lower costs usually produce larger basins of attraction for signalling, similar to what happens in the common interest case. For brevity's sake, we cannot report results for other values of $k$ and $m$; but we have found that, like in the $k = 1$ case, the relationship between cost and the evolution of signalling can be very complex.

Finally, we turn to the evolutionary significance of the polymorphic equilibrium. Figure 7 illustrates the basin of attraction of the face $K$ (described in Section 3). These basins of attraction are for parameters where signal cost is required in order to sustain full communication. For these settings, the minimum believable cost $c^* \approx 0.011$. However, the corresponding polymorphic equilibrium has a significant basin of attraction for costs an order of magnitude smaller. When comparing the results here to Figure 6, one sees that, by considering the best cases, the polymorphic equilibrium has a larger basin of attraction (for fixed $k$), than does the signalling ESS with higher signal costs. Because of this result we believe that the polymorphic equilibrium may be more relevant than the signalling ESS to the study of signalling in existing populations.
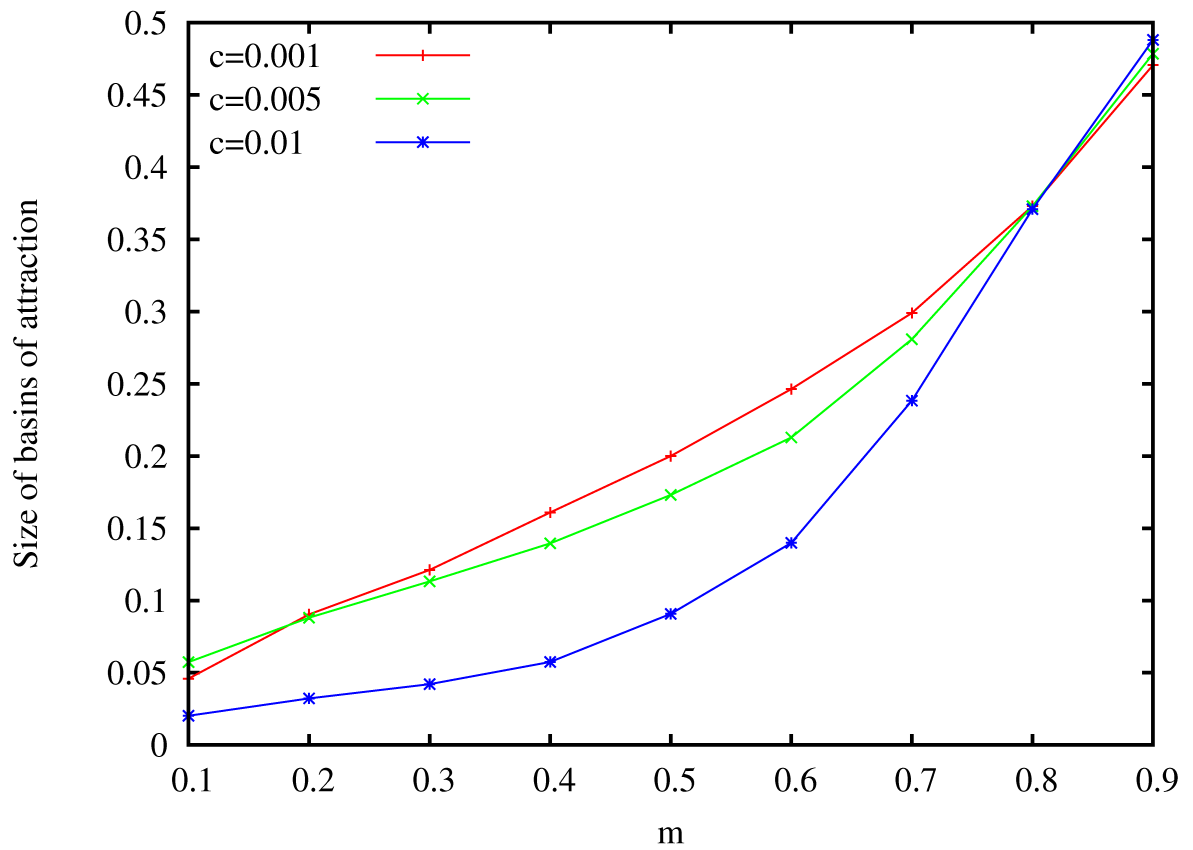
15

Figure 7: Simulation results showing the basin of attraction for the hybrid equilibrium for different values of $m$ and $c$. Here $k = 0.54$, $a = 31/32$, $b = 9/32$, and $d = 1/2$.

# 5 Discussion

Our results point out certain limitations of the ESS methodology as to the identification of long-run evolutionary outcomes with ESS. By focusing on ESS, traditional investigations of the Sir Philip Sidney game suggest that the evolution of costly signalling is far more likely than it appears to be for standard evolutionary dynamics. While the signalling equilibrium is the only ESS of the game for conflicting interests between the players, our dynamic analysis indicates that there are other states, namely pooling equilibria, which attract a significant portion of the initial populations. From the point of view of standard evolutionary dynamics, pooling equilibria also have a larger basin of attraction for many specifications of the parameters, indicating that they are the most probable evolutionary outcome in these cases. This means that the pooling equilibrium will be observed more often than the signaling equilibrium since there are more initial populations converging to the former than to the latter. We would like to emphasize that this holds for those parameters where the handicap principle should apply because of conflicts of interest between sender and donor.

The existence and stability of the polymorphic equilibria (6) illustrates another problem of the ESS methodology. By exclusively investigating the conditions under which a signalling ESS exists, one looses sight of other evolutionarily significant outcomes. This also poses problems for empirical studies of signalling, since some theoretically well founded evolutionary outcomes may not be known; consequently, the corresponding predictions of the existence of, e.g., polymorphisms such as (6) are not tested in the field.

We believe that the existence of partial information transfer even in situations where communication was believed to be impossible, as exemplified by the polymorphic equilibrium, has broad biological implications. For the relevant parameters, the polymorphic equilibrium appears to be a more likely evolutionary outcome than the traditional costly signalling ESS. As an example one may consider the paradigmatic signalling interaction between relatives, the begging of chicks for food from their parents. Parents would like to know the state of the chick while the chick would prefer food regardless of its state. If signalling need is associated

with a sufficiently high cost $c$, then the Sir Philip Sidney game (and many other signalling models) predicts that the state where only needy chicks send the signal and where parents react to the signal is a possible evolutionary outcome. Our new results predict that even if $c$ is very low, an alternative equilibrium is possible which is weakly stable and allows some information transfer between chicks and their parents. More specifically, the population dynamics may lead to a state where some, but not all chicks will signal need honestly; and some parents will respond to the signal by transferring food. The other chicks will always use the signal regardless of their state, and the other parents will never transfer food.

Our results also suggest that, for the relevant parameters, the polymorphic equilibrium is more likely to be observed than the signaling ESS. This resonates one of the issues found in many experiments which seek to determine the cost in signalling interactions such as solicitation (Searcy and Nowicki, 2005). Metabolic costs often do not seem to be high enough to accord with the existence of a signalling ESS (Bachmann and Chapell, 1998). Cost by risk of predation has also not been shown to be consistently high (Haskell, 1994). Even if costs are not sufficiently high to sustain full communication, one may find communication in a polymorphic state. This opens a new avenue for empirical research which might prove insightful.

# Acknowledgements

# A Geometry and dynamic stability of pooling equilibria

Suppose that the donor mixes between $R_2$ and $R_3$. Then $S_2$ and $S_3$ earn the same payoff against $(1-\lambda)R_2 + \lambda R_3$ if

$$
\begin{aligned}
(1-m)(1-b) + m(1-a) + k =&\lambda\left[(1-m)(1-b) + m(1-a-c) + k\right] \\
&+ (1-\lambda)\left[(1-m)(1-b+k) + m(1-c+k(1-d))\right],
\end{aligned}
$$

or if

$$
\lambda = \lambda^* = 1 - \frac{c}{a - kd}.
$$

If $\lambda > \lambda^*$, then $S_3$ gets a higher payoff than $S_2$; the reverse relation obtains if the inequality is reversed. Concerning $S_4$ we have to know when $\pi(S_4, (1-\lambda)R_2 + \lambda R_3) \geq \pi(S_3, (1-\lambda)R_2 + \lambda R_3)$. This equation is equivalent to

$$
(1-\lambda)(1+k-c-kd) + \lambda(1-ma-b+mb-c+k) \geq (1-m)(1-b) + m(1-a) + k
$$

or

$$
\lambda \leq 1 - \frac{c}{ma + (1-m)b - kd}.
$$

Since $a \geq b$, it is clear that the right-hand side of this inequality is less than or equal to $\lambda^*$. Hence, if $\lambda > \lambda^*$, then $S_4$ will also earn less payoff than $S_3$. As to $S_1$, $\pi(S_1, (1-\lambda)R_2 + \lambda R_3) \leq \pi(S_3, (1-\lambda)R_2 + \lambda R_3)$ if

19

$$\lambda\left[(1-m)(1-b-c)+m(1-a)\right]+(1-\lambda)\left[(1-m)(1-b-kd)+m(1-a)+k)\right]$$

$$\leq (1-m)(1-b)+m(1-a)+k,$$

which is clearly always the case since $c, kd \geq 0$. We already know that (2) implies that there is no donor strategy which does better against $S_3$ than either $R_2$ or $R_3$. The relation (5) can be proved similarly. These arguments also show that both kinds of lines of Nash equilibria are quasi-strict (and neutrally stable in the symmetrised game) under the appropriate existence conditions, since donor strategies not in the support of those equilibria earn less payoff than the strategies in their support.

Let us consider the line of strategy profiles given by $(S_3, (1-\lambda)R_2 + \lambda R_3)$. Rest points in the relative interior of this line attract an open set of nearby initial population states. This follows from the centre-manifold theorem (Carr, 1981), together with the fact that all transversal eigenvalues of the Jacobian matrix of (8) evaluated at those rest points are negative. The latter fact is shown by the calculations above.

# B  Alternative equilibria

Under certain conditions, $(S_1, R_1)$ is a strict Nash equilibrium. Note that the condition for donors to transfer the resource must be the same as for the equilibrium (1), i.e. $a \geq d/k \geq b$. The reason for this is that the decision problem of the donor is the same in both cases. As to the sender strategies, we have to find the conditions under which $S_i$, $i = 2, 3, 4$, earn less payoff against $R_1$ than $S_1$ does. Concerning $S_3$, $\pi(S_1, R_1) \geq \pi(S_3, R_1)$ if

$$(1-m)(1-b-c+k)+m(1+k(1-d)) \geq 1+k(1-d).$$

This is the case if and only if $kd - c \geq b$. Similarly, $\pi(S_1, R_1) \geq \pi(S_4, R_1)$ if

$$(1-m)(1-b-c+k) + m(1+k(1-d)) \geq (1-m)(1-b-c+k) + m(1-a-c+k),$$

which is equivalent to $a \geq kd - c$. The relation $b \leq kd - c$ implies $d \geq (b+c)/k$, and thus

$$1 + k(1-d) \leq 1 - b - c + k.$$

Furthermore, $a \geq kd - c$ implies

$$1 - a - c + k \leq 1 + k(1-d).$$

From these two inequalities it follows that

$$(1-m)(1-b-c+k) + m(1+k(1-d)) \geq (1-m)(1+k(1-d)) + m(1-m)(1-a-c+k),$$

which is the same as $\pi(S_1, R_1) \geq \pi(S_2, R_1)$.

We next prove the location of $(\mathbf{p}, \mathbf{q}) = ((\lambda S_2 + (1-\lambda)S_4), (\mu R_2 + (1-\mu)R_3))$, $0 < \lambda, \mu < 1$. In order to do this, we follow Hofbauer and Sigmund (1998, 10.4). The relevant payoff information for the game restricted to $K = \mathrm{span}(S_2, S_4) \times \mathrm{span}(R_2, R_3)$ is contained in the two payoff matrices

$$A = \begin{bmatrix} 0 & a_{12} \\ a_{21} & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & b_{12} \\ b_{21} & 0 \end{bmatrix},$$

where

21

$$a_{12} = \pi(S_2, R_3) - \pi(S_4, R_3) = (1 - m)c$$

$$a_{21} = \pi(S_4, R_2) - \pi(S_2, R_2) = (1 - m)(b - kd - c)$$

$$b_{12} = \pi(R_2, S_4) - \pi(R_3, S_4) = k((1 - m)b + ma) - d$$

$$b_{21} = \pi(R_3, S_2) - \pi(R_2, S_2) = m(d - ka)$$

As shown in Hofbauer and Sigmund (1998, 10.4), a unique interior equilibrium exists if $a_{12}a_{21} > 0$ and $b_{12}b_{21} > 0$. Since $c > 0$, we must have $b - kd > c$. If $b_{12} > 0$, then $(\mathbf{p}, \mathbf{q})$ is a saddle (provided that $b_{12}b_{21} > 0$). We will instead focus on the more interesting case where $b_{12}, b_{21} < 0$. This means that condition (2) is met and that $ka > d$. The rest point $(\mathbf{p}, \mathbf{q})$ is given by

$$\lambda = \frac{b_{12}}{b_{12} + b_{21}} = \frac{k((1 - m)b + ma)) - d}{(1 - m)(kb - d)} \quad \mu = \frac{a_{12}}{a_{12} + a_{21}} = \frac{c}{b - kd}.$$

Now $a_{12}b_{12} < 0$, and hence the Jacobian matrix of (8) evaluated at $(\mathbf{p}, \mathbf{q})$ has purely imaginary eigenvalues (Hofbauer and Sigmund, 1998, 10.4).

We next show that all transversal eigenvalues of the Jacobian matrix of (8) at $(\mathbf{p}, \mathbf{q})$ are negative. For the sender, the transversal eigenvalues are given by $\pi_i(\mathbf{q}) - \pi(\mathbf{p}, \mathbf{q})$ for $i = 1, 3$. It is easy to show that $\pi_1(\mathbf{q}) < \pi(\mathbf{p}, \mathbf{q})$ and $\pi_3(\mathbf{q}) < \pi(\mathbf{p}, \mathbf{q})$ if $\mu > c/(a - kd)$. The transversal eigenvalue $\pi_1(\mathbf{p}) - \pi(\mathbf{q}, \mathbf{p})$ is negative if and only if $d > bk$. Both conditions follow straightforwardly from our assumptions. Similarly, $\pi_4(\mathbf{p}) < \pi(\mathbf{q}, \mathbf{p})$ if (2) holds. That the polymorphism (6) is a spiraling centre follows from the centre-manifold theorem (Carr, 1981). In our case, this manifold coincides with an open neighborhood in $K$ around the polymorphism (6). Nearby solution trajectories approach the centre manifold exponentially. But on $K$, solution trajectories spiral around (6) forever. We also note that the polymorphic equilibrium is a Nash-Pareto pair (Hofbauer and Sigmund, 1998, 11.4). This means that the polymorphic equilibrium is almost as stable as a strict Nash equilibrium.

# References

Bachmann, G. C. and M. A. Chapell (1998). The energetic cost of begging behaviour in nestling house wrens. *Anim. Behav*, 55: 1607–1618.

Bergstrom, C. T. and M. Lachmann (1997). Signalling among relatives I. Is costly signalling too costly? *Philos. T. Roy. Soc. Lond. B*, 352: 609–617.

Bergstrom, C. T. and M. Lachmann (1998). Signalling among relatives III. Talk is cheap. *Proceedings of National Academy of Sciences*, 95: 5100–5105.

Brilot, B. O. and R. A. Johnstone (2003). The limits to cost-free signalling of need between relatives. *Proceedings of the Royal Society London, Series B*, 270: 1055–1060.

Carr, J. (1981). *Applications of Centre Manifold Theory*. New York: Springer-Verlag.

Cressman, R. (2003). *Evolutionary Dynamics and Extensive Form Games*. Cambridge, MA: MIT Press.

Godfray, H. C. J. and R. A. Johnstone (2000). Begging and bleating: The evolution of parent-offspring signalling. *Philosophical Transactions of the Royal Society London, Series B*, 355:1581–1591.

Grafen, A. (1991) Biological signals as handicaps. *Journal of Theoretical Biology*, 144:517–546.

Guckenheimer, J. and P. Holmes (1983). *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. New York: Springer.

Hamblin, S. and P. L. Hurd (2009). When will evolution lead to deceptive signalling in the Sir Philip Sidney game? *Theoretical Population Biology*, 75:176–182.

Haskell, D. (1994). Experimental evidence that nestling begging behaviour incurs a cost due to nest predation. *Proc. Roy. Soc. Lond. B*, 257: 161–164.

364 Hofbauer J., and S. M. Huttegger (2008). Feasibility of communication in binary signalling
365     games. *Journal of Theortical Biology*, 254:843–849.

366 Hofbauer, J. and K. Sigmund (1998). *Evolutionary Games and Population Dynamics*. Cam-
367     bridge: Cambridge University Press.

368 Johnstone, R. A. and A. Grafen (1992). The continuous Sir Philip Sidney game: A simple
369     model of biological signalling. *Journal of Theoretical Biology*, 156:215–234.

370 Johnstone, R. A. and A. Grafen (1993). Dishonesty and the handicap principle. *Animal*
371     *Behavior*, 46:759–764.

372 Lachmann, M. and C. T. Bergstrom (1998). Signalling among relatives II. Beyond the Tower
373     of Babel. *Theoretical Population Biology*, 54:146–160.

374 Maynard Smith, J. (1982). *Evolution and the Theory of Games*. Cambridge: Cambridge
375     University Press.

376 Maynard Smith, J. (1991) Honest signalling: the Philip Sidney game. *Animal Behavior*,
377     42:1034–1035.

378 Maynard Smith, J. and D. Harper (2003). *Animal Signals*. Oxford, New York: Oxford
379     University Press.

380 Maynard Smith, J. and G. Price (1973). The logic of animal conflict. *Nature*, 146: 15–18.

381 Searcy, W. A. and S. Nowicki (2005). *The Evolution of Animal Communication: Reliability*
382     *and Deception in Signalling Systems.* Princeton, NJ: Princeton University Pess.

383 Spence, M. (1973). Job market signalling. *The Quarterly Journal of Economics*, 87:355–374.

384 Wagner, E. (2009). Dynamics of costly signalling. *Working paper, University of California*
385     *at Irvine.*

386   Zahavi, A. (1975). Mate selection – a selection for a handicap. *Journal of Theoretical Biology,*

387     53:205–214.