

1 | Results

1.1 Output

1.1.1 Typical simulations

A simulation = show picture : genes, field, network => show two picture, one with also model SS2 Genes = weighted over 8 bits : values between 0 and $2^8 - 1$

this should be much longer

so as figures don't run into one another...

otherwise learn to skip pages...

Mostly will be working with average for relative value for genes over the pop / after a lot of time One experiment => results for four genes Typically (unless specified otherwise), values averaged over 30 experiments

MATH = NE.... => en fait on en a déjà parle en section 2.... => MODIF SECTION II plutot...

1.1.2 Remembered Heroes

These output values are precise to the percentage point. For **SelfSacrifice** this poses a problem, as final values are expected to be relatively small - with typical parameter values (as in **section 2**), average probability of self-sacrifice is capped at 2-3%.

For this reason, simulations also kept track of **RememberedHeroes**, a measure for the number of heroes in the live memory of a society. At each simulation step (year), the number of "voluntary" martyrs was computed by subtracting the expected number of martyrs which could be solely at-

tributable to mutations in a population that did not engage in self-sacrifice. These were added to the number of heroes which could be assumed to have been *witnessed* by (alive) individuals in the population. In a given year, **RememberedHeroes** corresponds to the number of heroes witnessed by at least *RemThreshold* percent of the population (5 % in practice). In situations where self-sacrificial behavior can be said to be absent, **RememberedHeroes** is close to 0 (e.g. when *Admiration* is null in the first model under).

1.2 Exogenous model: simulation outputs

1.2.1 Main results

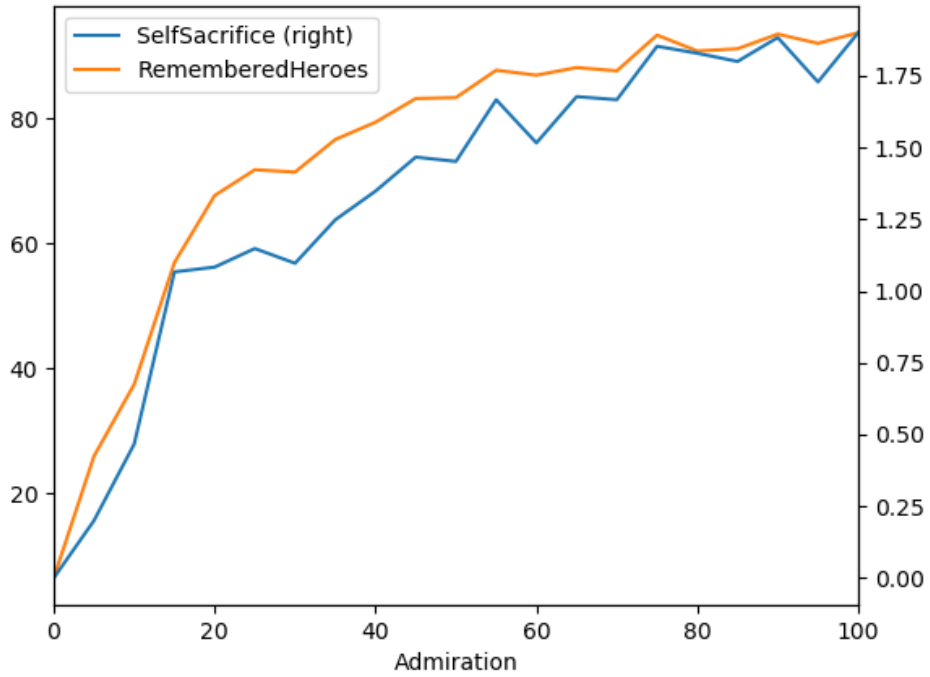


Figure 1: **SelfSacrifice** and **RememberedHeroes**, as a function of *Admiration* (typical parameter values).

Figure 1 shows results obtained by averaging over 30 simulations, according to the *Admiration* parameter (all others being kept constant at the previously described values). As could be expected, **RememberedHeroes** rises from 0

with *Admiration*, quickly reaching two thirds of its maximum value of under 100 heroes when *Admiration* exceeds 20.

SelfSacrifice follows similar dynamics, with values ranging between 0 and 2% (values to a higher precision than the percentage point being obtained artificially by averaging over experimental results). These values are far from negligible: for a typical population of 200, we expect an average of (almost) 4 martyrs each year. Such collective behavior is captured here by individuals all bearing similar genetic probability P of self-sacrifice at equilibrium (and obtaining similar **Reproductive_points** in the long-term), as seen on **Figure 2**.

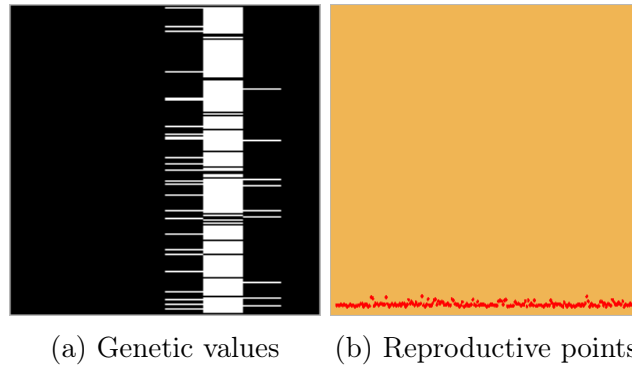


Figure 2: Snapshot of individual values, $A=50$. An individual's genome is represented by a horizontal line on the left: here most bear a **SelfSacrifice** gene of relative value $\frac{2^2}{2^8-1}$ (around 1.6%). Individuals (horizontal axis) and their reproductive points (vertical) are represented on the right.

However, one can also imagine the collectively equivalent situation where only a small fraction f of individuals engage in self-sacrificial behavior with higher probability p (with $P = f * p$), to the much more significant benefit of their families (see **Section 1.3**).

1.2.2 Influence of *ReproGainsThreshold* RGT

As a function of *Admiration*, **RememberedHeroes** resembles a function of the form: $f_{C,\tau}(t) = C * (1 - \exp(-t/\tau))$. Very approximately¹, it can best be approached by a function of this form with parameters $C = 95$ and $\tau = 20$, as visible on **Figure 3**.

¹By choosing optimal C and τ at a precision of 5 units; the idea here being simply to get a "feel" for overall variation with A . Optimal parameters are the ones for which Euclidean distance is minimal.

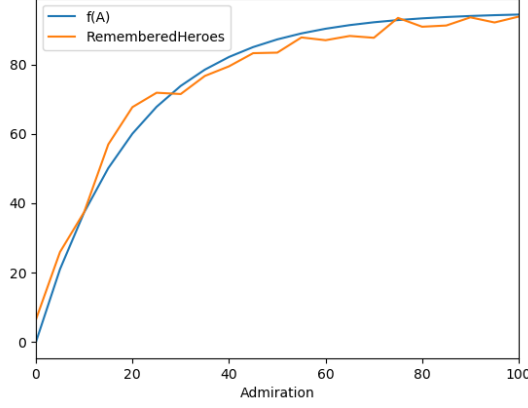
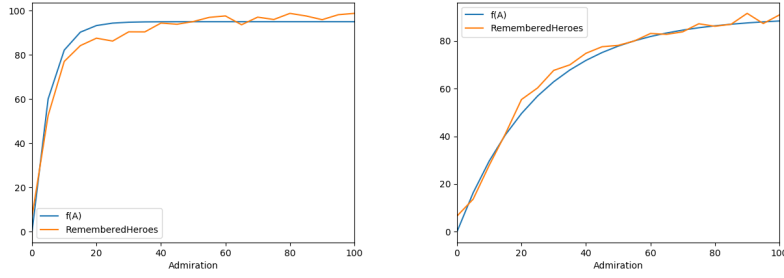


Figure 3: **RememberedHeroes** and corresponding optimal f as functions of *Admiration*, for $RGT = 10$ ($C = 95, \tau = 20$).



(a) $RGT = 5$ ($C = 95, \tau = 5$) (b) $RGT = 20$ ($C = 90, \tau = 25$)

Figure 4: **RememberedHeroes** and f for $RGT = 5, 20$

The key parameter governing relative growth of **RememberedHeroes**² thus appears to be one equal to two times *ReproGainsThreshold*. The importance of RGT is not surprising since it plays the crucial role of defining the unit in which **Reproductive_points** are counted. The relationship between RGT and final results is however non-trivial, since variation according to A for $RGT = 5$ (resp. $RGT = 20$) are best captured by $\tau = 5$ (resp. $\tau = 25$), as visible on **Figure 4**; suggesting perhaps a quadratic relationship between RGT and τ . **Section 1.3.3** returns to this issue.

²This output's absolute value is arbitrary, chosen according to *RemThreshold* in order to provide for visible variations (see **Section 1.1.2**).

1.2.3 Variation with *SacrificeHeredity* h

Contrary to what could be expected, $h \rightarrow 1$ is a problem: probs because pop = too small and all end up with same h... - as shown by simu // because same family ? (or just expected in LT?) $h = 0$ discontinuity = expected

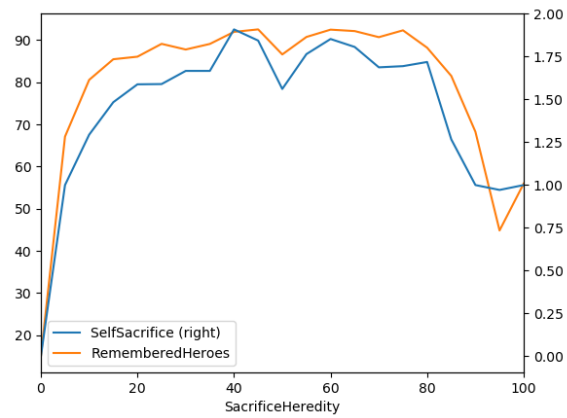


Figure 5

1.2.4 Variation with *ReproductionRate* r

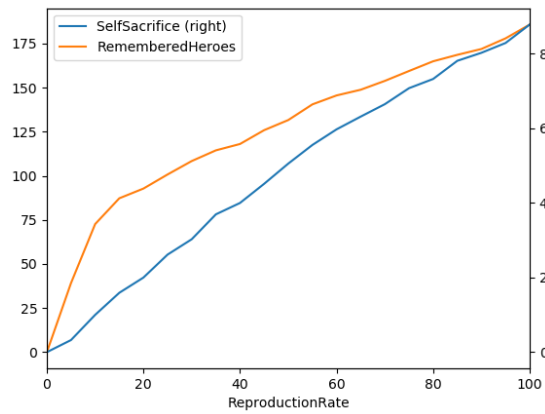


Figure 6

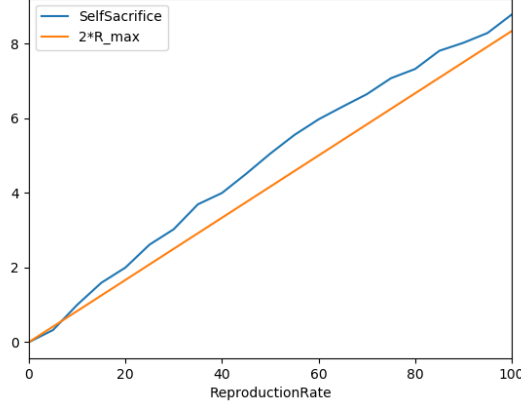


Figure 7

1.3 Exogenous model: mathematical proof of concept

1.3.1 Simplifying assumptions

Individuals live to a maximum of *AgeMax* M years, fixed at 40 here. Expected lifespan is lower however, as individuals face random accidents (in a wide sense) and is equal to $\alpha * M$, where α captures the effects of natural selection. In this model, all individuals face the same α , but may obtain differing reproductive opportunities (**Selectivity** mode).

Individuals who engage in self-sacrificial behavior with probability p shorten their expected lifespan by a multiplicative factor of (see **Appedix B.1.1**):

$$\beta(p) = \frac{(1 - p) - (1 - p)^{(M+1)}}{p * M} \quad (1)$$

As such, their reproductive window is shorter on average, which should lead to the disappearance of self-sacrificial behavior - unless this is compensated by increased reproductive opportunities.

In order to attempt to attain a mathematical proof of concept, a number of *simplifying assumptions* were made, which stand in contrast to how the previous simulation is actually played:

Assumption 1.3.1 (A1.3.1): *Probability of self-sacrifice p is fixed; what varies is the proportion of agents who engage in such a strategy f .*

Assumption 1.3.2 (A1.3.2): *Self-sacrifice is related to a dominant schematic "allele": the children of a would-be martyr and another individual are would-be martyrs for yearly probability p (and not $\frac{p}{2}$).*

Assumption 1.3.3 (A1.3.3): *p and M are large enough for such agents to be all but assured that they will die laying down their lives for the sake of the group (and not in another way) which can be written as: $(1 - p)^M \ll 1$ and $p \gg \alpha$.*

Assumption 1.3.4 (A1.3.4): *Generations do not overlap; individuals are granted their lifetime reproductive potential R at birth.*

Even though these assumptions set our mathematical model apart from the simulation, we still expect some connection between final results in both cases. In particular, population probability of self-sacrifice p^{sim} in the first case should resemble the one here which is equal to $p * f$. Since mutation rate (0.5%) and previous results were small (1-2%) (and p is assumed not to be here), let us also assume:

Assumption 1.3.5 (A1.3.5): *Mutation rate is negligible.*

Assumption 1.3.6 (A1.3.6): *Proportion of would-be martyrs remains negligible: $f \ll 1$.*

1.3.2 Mathematical characterization of equilibrium

Let us start from a situation where there are no martyrs in recent memory, and a proportion f_0 of individuals born with a propensity to martyrdom, captured by p . These individuals obtain on average $\alpha * \beta(p) * M * r$ offspring each, when others obtain $\alpha * M * r$, where r is the population-level *ReproductionRate*.

Following (A1.3.3), a would-be martyr's child i obtains reproductive potential $R^i > r$. Martyrs children are in proportion f_1 in their generation (A1.3.4). Since $f_1 \ll 1$ (A1.3.6), R^i can be assumed to be the same for all such children, and will be written $R_+(A, f_1)$, where A is *Admiration* (we neglect the possibility of individuals being born from two would-be martyrs). Children of non-would be martyrs obtain on average R_- which, for the same reason, can be approximated to r (A1.3.6).

Since self-sacrifice is assumed to be heritable, these second-generation martyrs obtain on average $\alpha * \beta(p) * M * R_+(A, f_1)$ children, who obtain $\alpha * \beta(p) * M * R_+(A, f_2)$ children, etc. Non-martyrs continue to obtain $\alpha * M * r * children$. Thus, (potential) equilibrium between $f \ll 1$ would-be martyrs and $1 - f$ others is characterized by:

$$R_+(A, f) * \beta(p) = r \quad (2)$$

1.3.3 Necessary condition for an ESS

Let N be equal to *PopulationSize*, implicitly assumed to be large here (since f is negligible).

In the spirit of (A1.3.4), let us assume that total social admiration is borne by the children of non-heroes, and is thus equal to $A * (1 - f) * N$ at potential equilibrium. For children of martyrs, two situations are possible:

- either $\frac{A * (1 - f) * N}{fN} < RGT$, and all receive reproductive potential r ;
- or $\frac{A * (1 - f) * N}{fN} \geq RGT$, and all receive $R_+(A, f) > r$.

Thus, self-sacrifice as characterized by $p > 0$ and $f > 0$ can only constitute an ESS if:

$$A \geq A_{min} = \frac{RGT * f}{1 - f} \quad (3)$$

A condition which, since we assumed $f \ll 1$, is highly likely to be met.

1.3.4 Equilibrium value

As shown in **Appendix B.1.2**, a first-order approximation of R_+ in the latter case is (S is equal to *Selectivity*):

$$R_+(A, f) = \frac{S * r}{\log(1 + S)} * (1 - \frac{S * f}{2}) \quad (4)$$

Using (2), we deduce that both envisioned strategies are equivalent if and only if:

$$\frac{Sr}{\log(1 + S)} (1 - \frac{S * f}{2}) \beta(p) = r$$

$$\Longleftrightarrow f = \frac{2}{S} * \left(1 - \frac{\log(1+S)}{\beta(p) * S}\right) \quad (5)$$

Since $(1-p)^M \ll 1$ (A1.3.3),

$$\beta(p) \approx \frac{1-p}{p * M}$$

Which yields:

$$f \approx f_{eq}(p) = \frac{2}{S} * \left(1 - \frac{p * \log(1+S)}{(1-p) * S}\right) \quad (6)$$

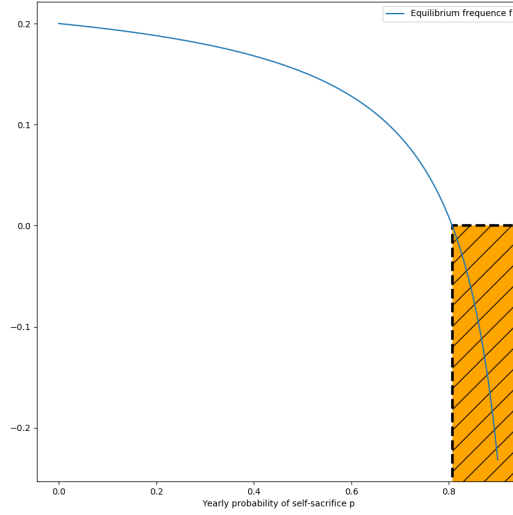


Figure 8: $f_{eq}(p)$. Forbidden values ($p \geq p_{max}$) are represented by the orange rectangle.

As shown by **Figure 8**, $f_{eq}(p)$ is positive for $p < p_{max} = \frac{S}{\log(1+S)+S} \approx 0.81\%$ for $S = 10$.

Proposition 1: *Under assumptions (A1.3.1-1.3.6), for:*

- $A \geq A_{min}$;
- and $p \leq p_{max}$;

a self-sacrifice Nash equilibrium (NE) exists, where a first-order approximation of proportion of would-be martyrs at equilibrium is:

$$f(S, p) = \frac{2}{S} * (1 - \frac{p * \log(1 + S)}{(1 - p) * S})$$

For $p = \frac{1}{2}$ and $S = 10$, we find for instance: $f(S, p) \approx 15\%$, for which assumption (A1.3.6) does not seem completely absurd. For $p > \frac{3}{4}$, we further find $f \leq 5\%$.

1.3.5 Brief Discussion

p = + function of S = duh // - of M : more to sacrifice But **no r**, no A --- > *shouldbecapturedbyf...* : OK for A in $f_m ax...$ Maybe cheat : admiration goes to heroes = the parents... then to children ? How ? = need to have a generational gap (but cool here that f between 0 and 1)

+ RGT : here = linear law... probs because of approx $f \ll 1$, all receive the same. . .

+ h: not here, vu le modele

+ lim for JLD : does not depend on nb of heroes... -> visibility $\log(\text{nbheroes} \dots ?)$

1.4 Two-tier model simulation outputs

1.5 Two-tier model: mathematical analysis

1.5.1 Self-sacrifice

For would-be martyrs M , the situation can be approached as in **Section 1.3**: the question comes down to whether a non-trivial set of ESS strategies comprising a proportion f of agents engaging in self-sacrifice with probability p can emerge³, and meet the requirements of equation (2).

³And $1 - f$ agents not doing so.

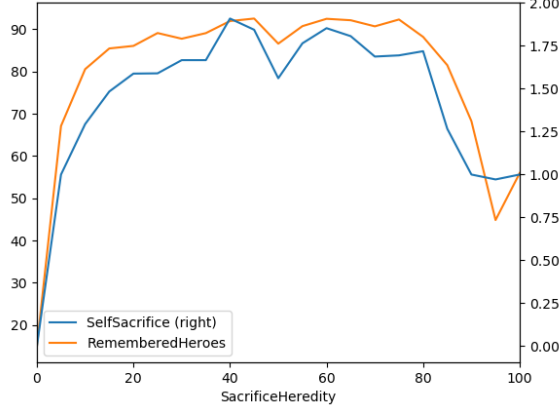


Figure 9

Total social admiration A is now however *endogenous*, as it is determined by how much non-martyrs \overline{M} engage in second-order signaling (honoring h):

$$A = \sum_{h(\overline{M}) \geq V_T} h(\overline{M}) \quad (7)$$

At (potential) equilibrium, A will thus be determined by (potential) equilibrium levels of honoring by patriots P and non-patriots \overline{P} and how high *VisibleThreshold* V_T is set. This, in turn, should determine if a self-sacrifice ESS is possible, and, if so, at what proportion of martyrs f can emerge at equilibrium, as before. Both signals are thus connected via equation (2), which can be tentatively rewritten according to equilibrium levels:

$$R_+(h_P, h_{\overline{P}}, f) * \beta(p) = r \quad (2b)$$

1.5.2 Honoring and social score

This is motivation to investigate the possibility of an honoring equilibrium. Honoring h and demand d serve as bases for establishing friendships, which are mutually beneficial (as controlled by *FriendshipValue* F). However, inter-group conflict raises the stakes: friends may be true patriots, to one's benefit (as controlled by *PatriotFriendBonus* P), or not be, leading to the possibility of betrayal with probability t (*NbTraitors*) and at potentially enormous cost DC (*DenunciationCost*).

A patriot individual characterized by d and h , who obtains a number of friends k on the bases of these (genetic) characteristics will see his social score vary according to:

$$\mathbf{E}(\Delta_P(d, h)) = k * F + k_P * P - k_{\bar{P}} * t * DC \quad (8)$$

In contrast, a non-patriot individuals also stand to gain *Judas* J if and when they betray their friends:

$$\mathbf{E}(\Delta_{\bar{P}}(d, h)) = k * F + k_P * P - k_{\bar{P}} * t * DC + t * J \quad (\bar{8})$$

1.5.3 Equilibrium when $t \leq \frac{F}{FC}$

Let us suppose that individuals of the same quality all signal at the same level s_P or $s_{\bar{P}}$, and have equal demand d_P or $d_{\bar{P}}$ (noted d when there is no ambiguity).

Since individuals encounter patriots and non-patriots with equal probability, and simply chooses whether or not to accept them based on d (potential friends are not ranked), a patriot with demand d can be in one of four situations (where k notably depends on d and time allocated to encountering potential partners):

$\mathbf{E}(\Delta_P(d, h))$	$d \leq s_{\bar{P}}$	$d > s_{\bar{P}}$
$d \leq s_P$	$\frac{k}{2} * (2F + P - t * DC)$	$k * (F + P)$
$d > s_P$	$k * (F - t * DC)$	0

Table 1: Expected social score (patriot)

When t is sufficiently small, **Table 1**'s first three cells are strictly positive: when betrayal is sufficiently improbable, friendship can always be expected to pay off for patriots - as well as for non-patriots, since $\mathbf{E}(\Delta_{\bar{P}}(d, h)) \geq \mathbf{E}(\Delta_P(d, h))^4$. For both, the optimal "strategy" is thus $d = 0$. In addition, since honoring is costly, the best response to $d = 0$ is to not invest in honoring:

$$t \leq \frac{F}{DC} \implies (d = 0, h_P = 0, h_{\bar{P}} = 0) \text{ is the only NE}^5 \quad (9)$$

⁴At least when h is small - see after

With typical parameter values, when $NbTraitors$ is smaller than 10%, no second-order signal - and therefore no first-order signal - should emerge.

1.5.4 "Honest" equilibrium when $t \geq \frac{2F+P}{DC}$

When t is sufficiently large, both cells on the left of **Table 1** are negative: patriots cannot afford to have any non-patriot friends. Let us assume, following the first-version of the proposed simulation, that $s_{\bar{P}}$ is bounded by $MaxOffer$ MO , and that non-patriots and patriots pay the same cost for signaling.

For patriots, it is then always better to have $d > MO$ than $d \leq MO$. Indeed, in the latter case where all patriots have demand $d < MO$, we are in a domain which is beneficial for non-patriots: they stand to gain more from signaling (by potentially betraying), at no extra cost. Thus, if $d < MO$, we expect $s_{\bar{P}} \geq s_P$ ⁶: the optimal situation represented by **Table 1**'s top right cell is unattainable, and, on average, patriots stand to lose from friendship.

Conversely, if $d > MO$, patriots either obtain no friends or only patriot friends and thus stand to benefit. Any strategy d^+ verifying $d^+ > MO$ thus weakly dominates any other strategy d^- , where $d^- \leq MO$.

Yet, as evoked in introducing equation (8), an individual's number of friends k depends on d . Given two demands above MO , the smaller is always best, since decreasing demand can only lead to increasing number of (patriotic) friends. For patriotic individuals, the optimal strategy is thus the smallest available d which is larger than MO , which we will note $d = MO^+$.

Let us assume in addition that individuals have ample opportunity to meet potential friends, forming Max_F friendships when possible, at cost c , and that $c * MO^+ < Max_F * (F + P)$ ⁷. Patriots stand to gain from signaling at levels above of MO , and for them, the optimal strategy is thus MO^+ , following the same reasoning as above.

Non-patriots, however, can only attract other non-patriots:

$$\mathbf{E}(\Delta_{\bar{P}}(d, h)) = k * (f - t * DC + t * J)$$

$$\mathbf{E}(\Delta_{\bar{P}}(d, h)) \geq 0 \iff t \leq \frac{F}{DC - J}$$

⁶And $d_{\bar{P}} \leq d_P$

⁷Otherwise signaling is never beneficial and the only equilibrium is trivially (0,0,0).

In practice, the latter condition is incompatible with the one presented at the beginning of this section, since, with typical values, $\frac{F}{DC-J} = \frac{2}{15}$ and $\frac{2F+P}{DC} = 0.3$. Thus:

$$t \geq \frac{2F+P}{DC} \implies (d = MO^+, h_P = MO^+, h_{\bar{P}} = 0) \text{ is the only NE} \quad (10)$$

When *NbTraitors* is greater than 30%, a purely honest equilibrium should thus emerge at the second-order - leading to self-sacrifice at a level corresponding to $A = MO^+/2$, since only half the population honors martyrs.

1.5.5 Dishonest equilibrium when $t \geq \frac{2F+P}{DC}$

In practice not so much: - some dishonest, and not because stabilize at a level below permitted by J: cf ... - RemT = "more optimal" than expected

1.5.6 $\frac{F}{DC} < t < \frac{2F+P}{DC}$

1.6 Second sacrifice model

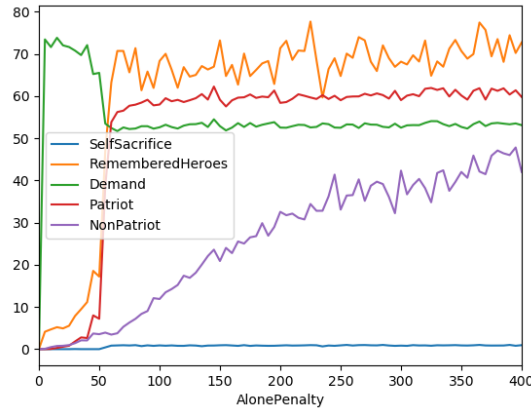


Figure 10

1.7 Limitations

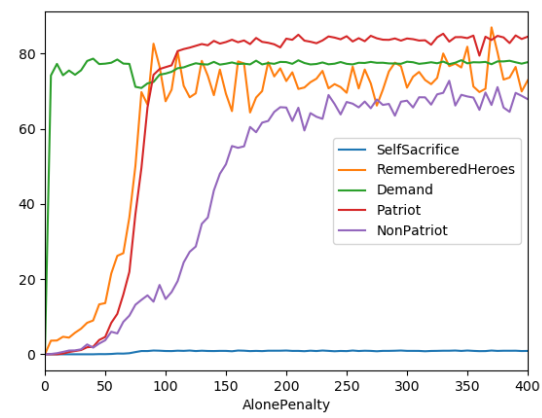


Figure 11

2 | Perspectives

References

A | Python stuff

B | Mathematical demonstrations

B.1 Exogenous model

B.1.1 $\beta(p)$

Disregarding the effects of natural selection (which are the same for all individuals here¹), an individual who bears a **SelfSacrifice** gene of relative value p , has a probability p of dying in his first year (before being able to foster any descendants), a probability $(1 - p) * p$ of dying at age 1 ... a probability $(1 - p)^n * p$ of dying at age $n < M$... and is certain to die at age M , should he or she reach it. His/her expected life span is thus:

$$ELS = p * 0 + (1 - p) * p * 1 + \dots + (1 - p)^{M-1} * (M - 1) + (1 - p)^M * M$$

Let $f: \mathbf{R} \rightarrow \mathbf{R}$ be the polynomial function defined by the expression:

$$f(x) = \sum_{n=0}^{M-1} p * (1 - p)^n * x^n + (1 - p)^M * x^M$$

By deriving f , one can note that:

$$ELS = f'(1) \tag{11}$$

$f(x)$ involves a geometric sum and can be simplified to (when $(1 - p) * x$ is different than 1):

$$f(x) = p * \frac{1 - ((1 - p) * x)^M}{1 - (1 - p) * x} + ((1 - p) * x)^M$$

¹And will therefore appear on both sides of an equation comparing the benefits of either strategy such as (2).

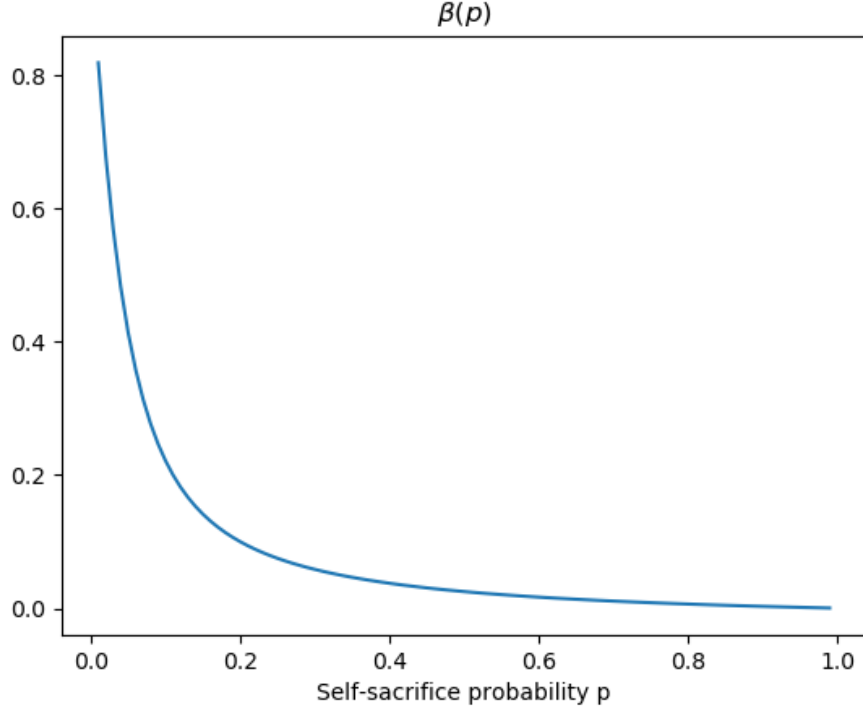


Figure 12: Loss of expected life-span due to self-sacrifice.

For $x \neq \frac{1}{(1-p)}$ ($p = 1$ trivially yields $ELS = 0$):

$$f'(x) = p * \frac{-M(1-p)^M x^{M-1} * (1 - (1-p)x) + (1-p) * (1 - ((1-p)x)^M)}{(1 - (1-p)x)^2} + M(1-p)^M x^{M-1}$$

And thus:

$$f'(1) = p * \frac{(-M(1-p)^M * p + (1-p) * (1 - (1-p)^M))}{p^2} + M(1-p)^M$$

$$f'(1) = \frac{(1-p) * (1 - (1-p)^M)}{p} \quad (12)$$

Combining these two expressions for $f'(1)$ proves equation (1). **Figure 12** shows $\beta(p)$ for p between 0 and 1. Factoring in natural selection, an individual's expected life span is therefore equal to: $\alpha * \beta(p) * M$.

B.1.2 $R_+(A)$

In **Selectivity** mode, individuals obtain reproductive potential R according to their **Reproductive_points** RP : each individual obtains a rank k according to RP , and receives reproductive potential:

$$R = \frac{r}{2} * \left(\frac{S}{(S * k + N') * \log(1 + S)} + \frac{S}{(S * (k + 1) + N') * \log(1 + S)} * N' \right)$$

where $N' < N$ is the number of eligible parents (non-martyrs) and S is equal to *Selectivity* (the average reproductive potential over eligible parents being *ReproductionRate* r). When $N \gg 1$, R verifies:

$$R \in [R_{min}; R_{max}], \text{ with } R_{max} \approx \frac{S * r}{\log(1 + S)} \text{ and } R_{min} \approx \frac{R_{max}}{1 + S} \quad (13)$$

With typical parameters ($S = 10$ and $r = 15\%$), we obtain: $R_{max} \approx 63\%$ and $R_{min} \approx 5,6\%$.

In a case where a negligible proportion f of individuals engage in self-sacrificial behavior (which is assured to end up in their martyrdom), their children each receive, on average:

- reproductive potential r , when $A < A_{min}$;
- $R_+(A, f)$ otherwise, as seen in **Section 1.3.3**.

When $N \gg 1$, expected $R_+(A, f)$ is equal to the average between the "luckiest" ($k = 0$) and "unluckiest" child, which is approximately:

$$R_+(A, f) \approx \frac{S * r}{2 * \log(1 + S)} * \left(1 + \frac{1}{S * f + 1} * \frac{N}{N} \right)$$

Which yields, for $f \ll 1$ (neglecting terms of order 2 and above in f):

$$R_+(A, f) \approx \frac{S * r}{\log(1 + S)} * \left(1 - \frac{S * f}{2} \right) = R_{max} * \left(1 - \frac{S * f}{2} \right) \quad (4)$$

B.1.3 Nash equilibrium