

# Supplementary Materials

*The ‘Icarus effect’ of preventative health behaviors*  
S. Carrignon, R. A. Bentley, M. J. Silk, N. H. Fefferman

## Algorithms

---

**Algorithm 1: Main Algorithm**

---

**Data:**  $N$ : number of agents,  $t_{step}$ : number of time steps,  $\vec{P}(b)$ : vector with the probability of infection for different behaviors,  $b$ : the distance between two agents under which the disease can be transmitted,  $i_0$ : the number of initial infections,  $p$ : the probability of observational learning,  $r$ : radius within which individual can learn socially  $\kappa$ : transparency (steepness of the sigmoid),  $\nu$ : inflexion point of the sigmoid,  $\kappa_r$ : steepness of the reversion sigmoid,  $\nu_r$ : inflexion point of the reversion sigmoid,

**Result:** A table with the SIR distribution per timestep

```
1 pop ← generatePopulation(N, x, y)
2 pop[random(N)] = infected
3 foreach t in t_step do
4   foreach a ∈ pop do
5     a.move()
6     X ~ U(0, 1)
7     if X < p then
8       // Observational Learning
9       g_i ← getPropInfected(pop, a.age)
10      if i.behavior = "NA" then
11        p_switch ← sig(g_i, κ, ν)
12        X ~ U(0, 1)
13        if X < p_switch then a.behavior ← "A"
14      else
15        p_switch ← sig(1 - g_i, κ_r, ν_r)
16        X ~ U(0, 1)
17        if X < p_switch then a.behavior ← "NA"
18      end
19    else
20      // Social Learning
21      m = selectModel(pop, a.age, strategy, r) // select another individual
22      to copy its behavior individual within radius r
23      a.behavior = m.behavior
24    end
25    foreach i ∈ pop - a | dist(i, a) < d do
26      X ~ U(0, 1)
27      if X < p[a.behavior] then a.state ← "I"
28    end
29    if a.state == "I" then
30      a.recovery ← a.recovery - 1
31      if a.recovery < 1 then a.state = "R"
32    end
33  end
34 end
```

---

---

**Algorithm 2:** Generate a population of agents

---

**Data:**  $N$ : number of agents,  $R$ : a uniform distribution of recovery times,  $B$ : a distribution of behaviors,  $S$ : a normal distribution of speeds,  $x, y$ : spatial limits,

**Result:** a table with  $n$  agents

```
1  $pop =$  foreach  $a \in n$  do
2    $a.age \sim \{.24, .09, .12, .26, .13, .16\} \times n$  // source:
   https://www.kff.org/other/state-indicator/distribution-by-age/
3    $a.behavior \leftarrow "NA"$  // by default start with everyone is
   Non-adherent
4    $a.position \leftarrow (rand(x), rand(y))$  // random position in the
   grid
5    $a.state \leftarrow S$ 
6    $a.recovery \sim U(8, 14)$  // agents have different recovery
   times, sampled between 8 and 14 days
7    $pop.add(a)$ 
8 end
9  $return(pop)$ 
```

---

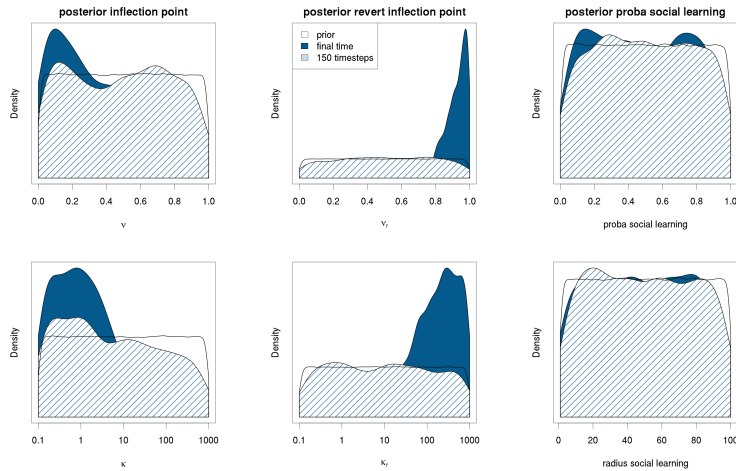


Figure S1: All posterior distributions for the 6 parameters used in our simulations. The white area represents the distribution of the parameters of all simulations (the priors), the blue area represents the distribution for the top 1000 simulations rank using  $\delta$  while the shaded area represent the parameters of the 1000 simulations with the lowest number of infected people after 150 timesteps.

## Posterior Distribution for all parameters

This section details the posteriors distribution for all parameters of the model in its variant used in the main text of the paper.

Figure S1 is a fuller rendition of Figure 4 in the main text. In Figure S1 shows the posterior distribution of parameters in switch from Non-Adherent to Adherent (left column) from Adherent to Non-Adherent (middle column), and probability and radius of social learning (right column).

These marginal posteriors may hide more complex interactions between the different parameters and may be due to the stochasticity of the model. We thus propose to look at different joint posteriors that combine interesting pairs of parameters and we varies the threshold we use to select priors in the Figure S2. In Figure 9 distributions were drawn using the parameters selected to generate the 1,000 “best” outcomes (best being the simulations that minimizes  $\delta$  as defined in equation 1). The patterns observed for the posteriors selected as the best 1,000 simulations are very robust and very unlikely to be due to chance. To illustrate this we show in Figure S2 the joint posteriors of representative pair of parameters of the model for different level of what we consider as the “best” simulations. The 2d areas represent the High Density Regions, ie the smallest regions within which falls a certain percentage of the distribution of the parameters our selected simulation. The lighter areas represent the area within which all the simulations are distributed and darker areas represent regions for smaller HDR.

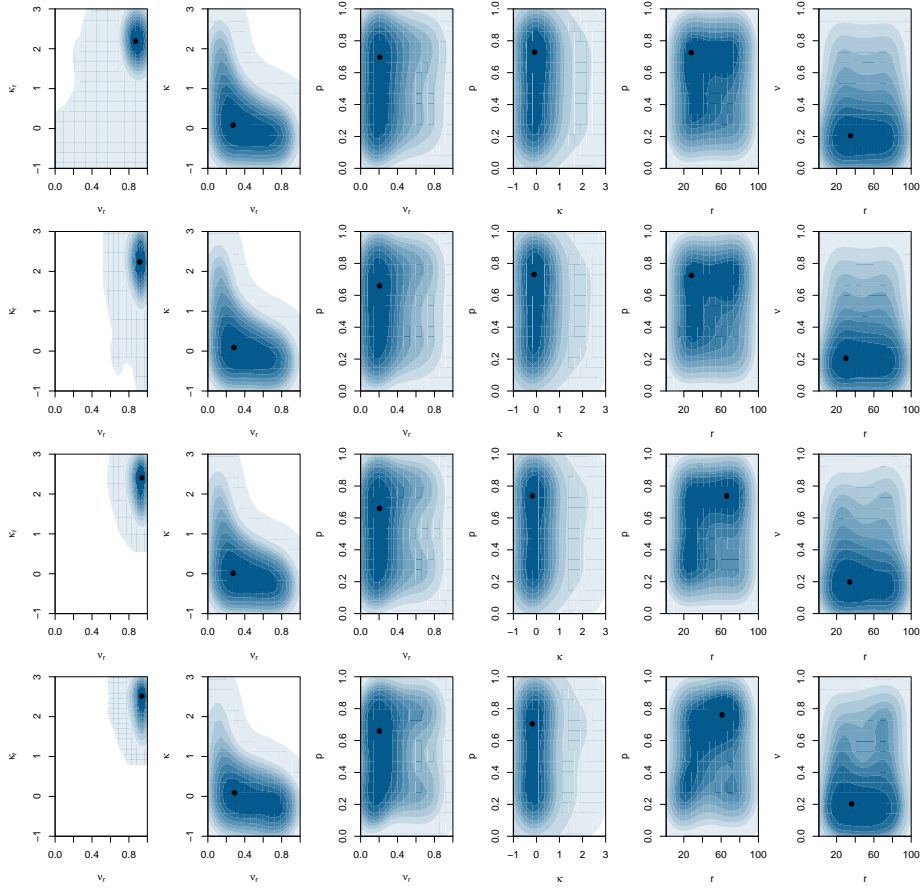


Figure S2: Joint posterior distributions for some combination our 6 parameters. Each distribution represent the parameter distribution of simulation for which the metric  $\delta$  is ranked in level respectively 10,000; 12,500; 2,500; 1,000

SLS	learning start	% Adherents at t=0			
		0	15	50	100
Healthy	0	A	B	C	
Random	0	D	E		
Healthy	100			F	

Table S1: Summarize all variants of the models

## Different Model of Social Learning and Initial distribution

Aside the simulations run with the setup described in the main paper we run sets of slight variation where:

- Instead of allowing individual to copy only healthy individual of the same age (Best SLS), the social learning strategy allows to copy the behaviour of any individuals of the same age (Random SLS),
- The initial distributions of behaviors varies (0% Adherent - the normal setup, 15% Adherent and 100% Adherent),
- Learning (individual and social learning) starts after 100 time steps.

We did not try all combinations of all variations, the table S1 summarize which one are discussed in the following sections. The variant A is the one presented in the main paper. The numbers of simulations ran for each variants are the following:

- Variant A : 826, 892 simulations
- Variant B : 158, 400 simulations
- Variant C : 251, 000 simulations
- Variant D : 127, 199 simulations
- Variant E : 153, 499 simulations
- Variant F : 309, 998 simulations

The Figure S3 show the distributions of  $\delta$ ,  $\tau$  and  $I_{max}$  for all simulations of all model's variants. As we can see all variants generate slightly similar outputs, expect for the Variants F, for wich learning (social and observational) is impossible until the time step 100. In this case the mode of  $\tau$ , the time to reach the maximum number of infected people, is slightly shifted by around 40 time steps. In this variant all individual are adherent to social distancing and stick to it for the first 100 time step, thus slowing down the infection. Nonetheless, as soon as we relax and bring back the learning abilities, the exact same properties are observed than for the other simulations: we will reach the same  $I_{max}$  and  $\delta$ , and the combination of parameters allowing to reach such results are the same than for the over variants, as the posteriors shown in the next sections illustrate.

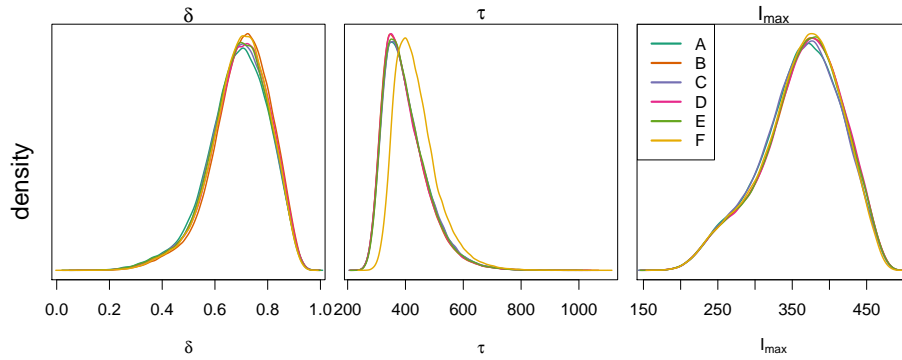


Figure S3: Distribution of the metrics described in the section Results for all simulations for all model (approx 100k simulations per model).

### Results for all variants

The Figures S4-S9 show the marginal posteriors parameters for all parameters of all model's variants. The Figures S10-S15 show the joint posteriors of the parameters used for observational learning for model's variants. All Figures from S4 to S15 confirm that the overall picture isn't affected by the Social Learning Strategies nor the initial distribution of behaviors, the selected posterior are roughly the same for all the variants tested.

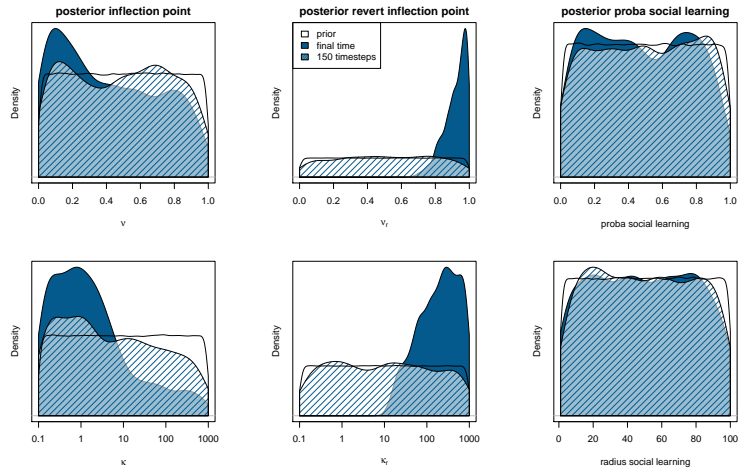


Figure S4: Same than figure 9 but for simulation run with variant A

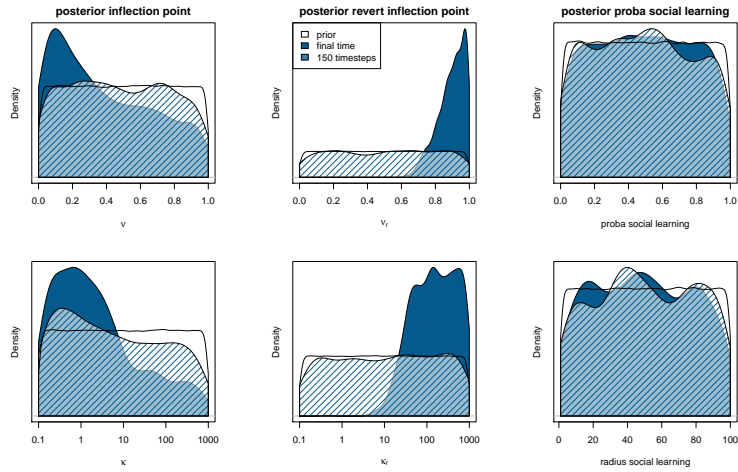


Figure S5: Same than figure 9 but for simulation run with variant B

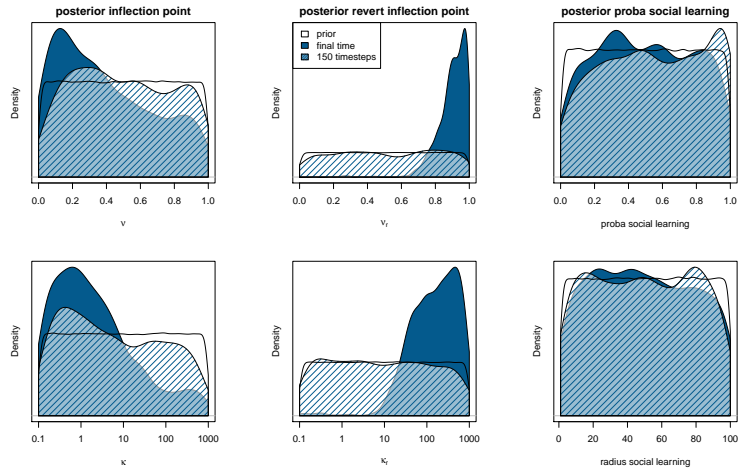


Figure S6: Same than figure 9 but for simulation run with variant C

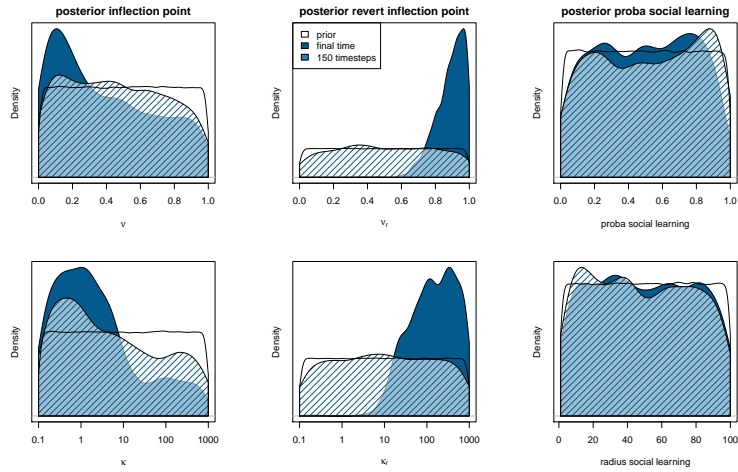


Figure S7: Same than figure 9 but for simulation run with variant D



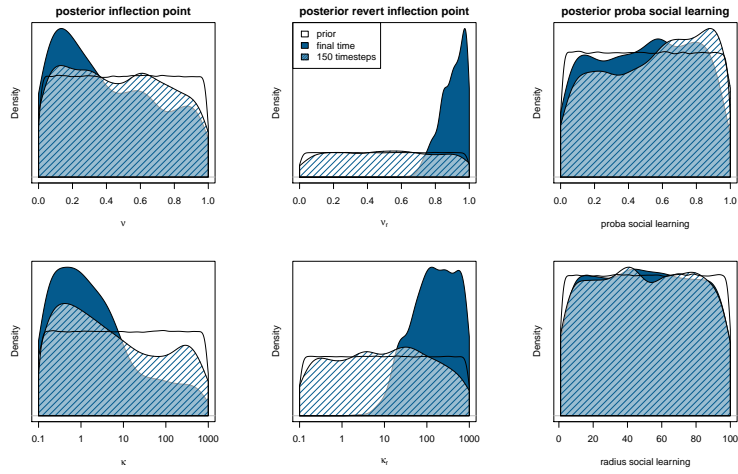


Figure S8: Same than figure 9 but for simulation run with variant E

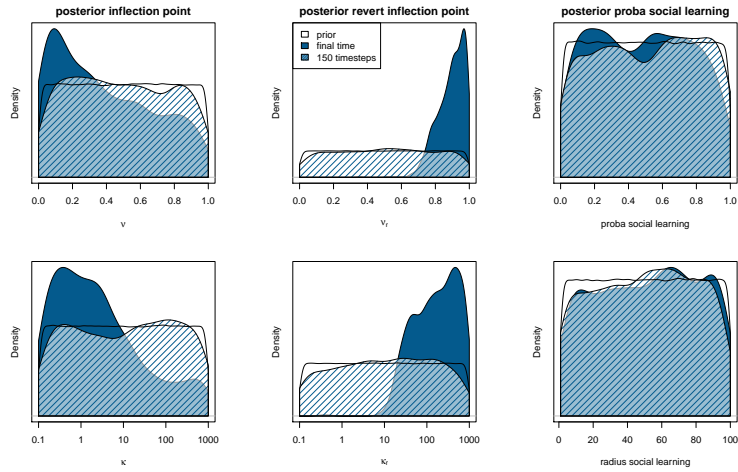


Figure S9: Same than figure 9 but for simulation run with variant F

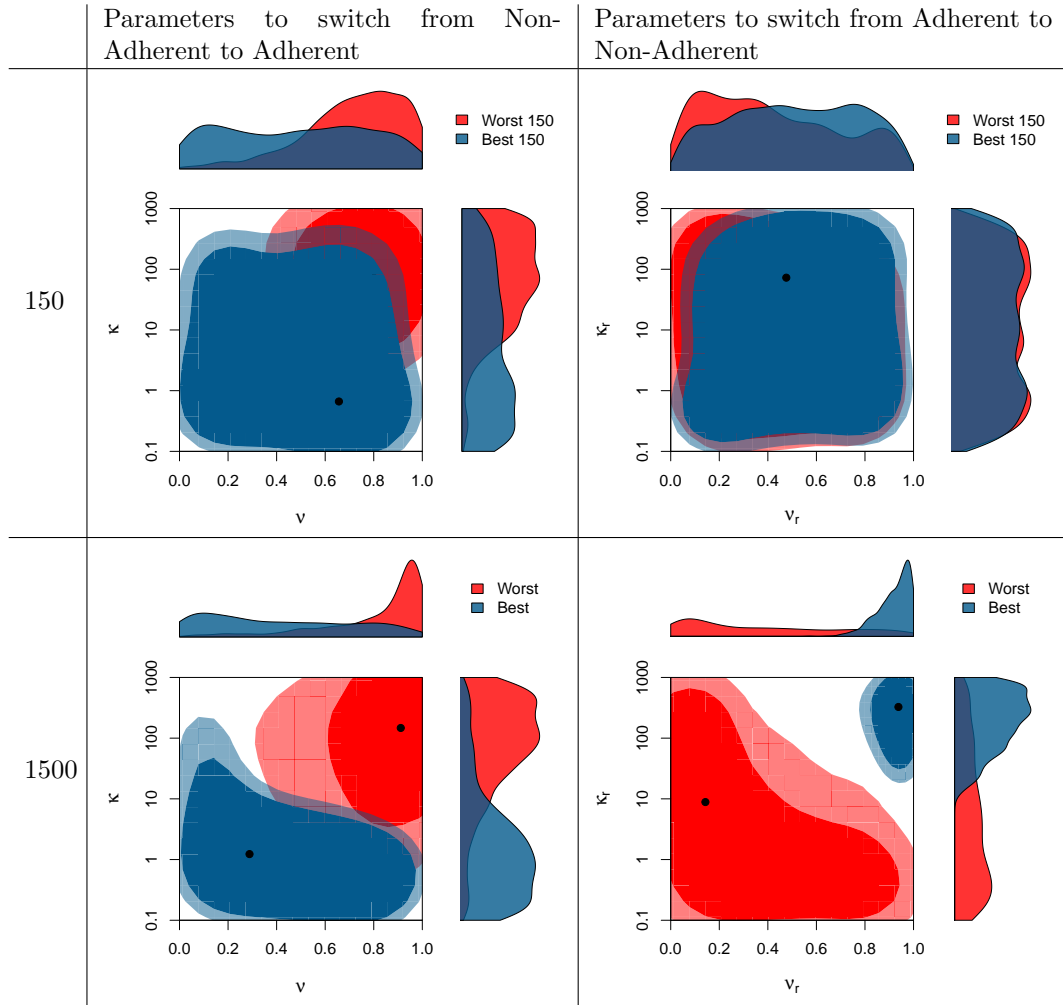


Figure S10: Variant A

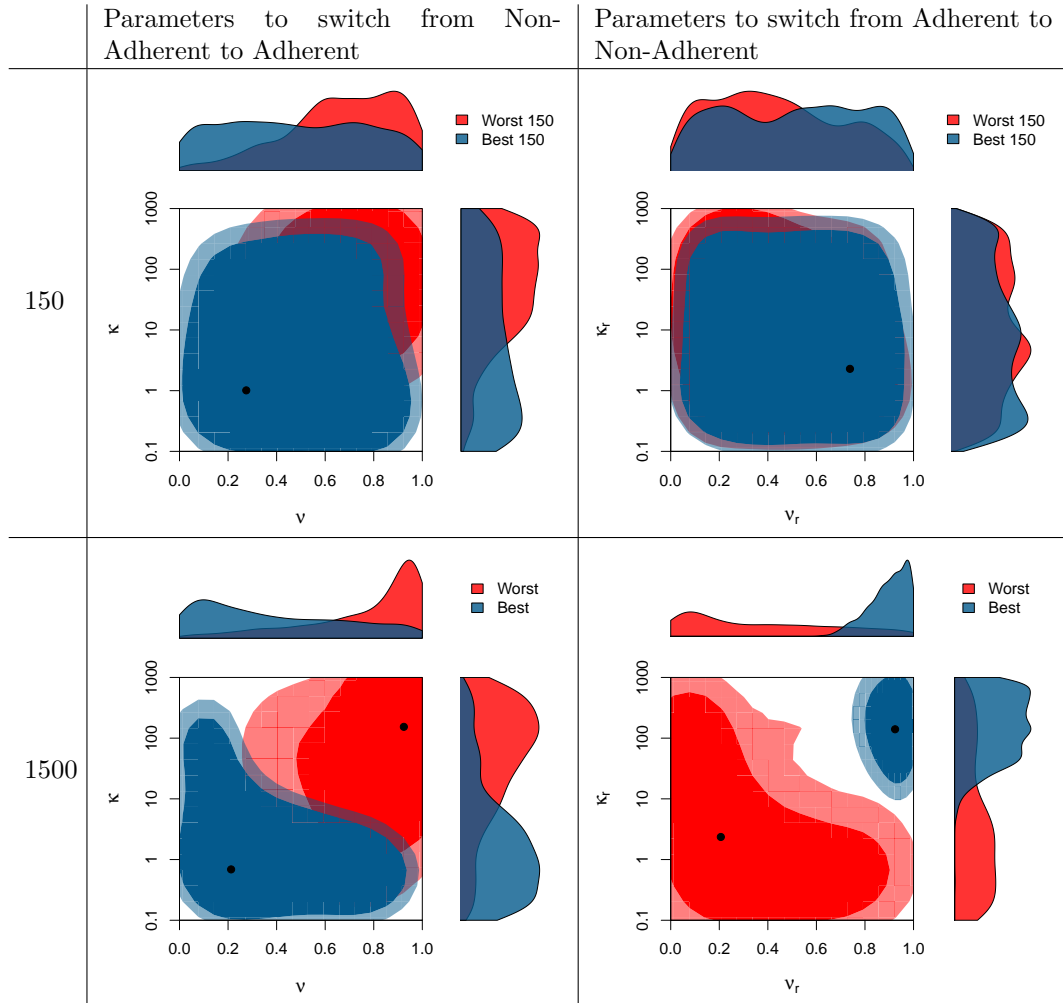


Figure S11: Variant B

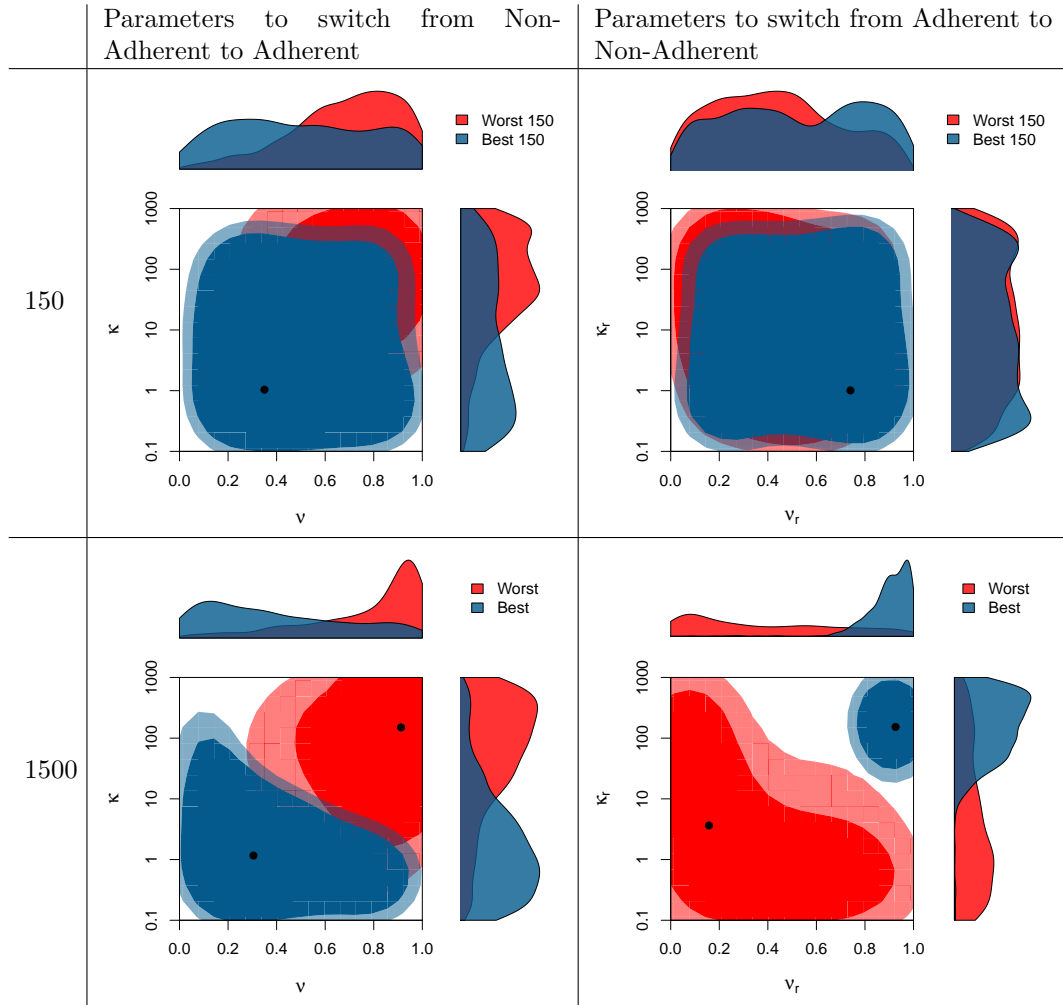


Figure S12: Variant C

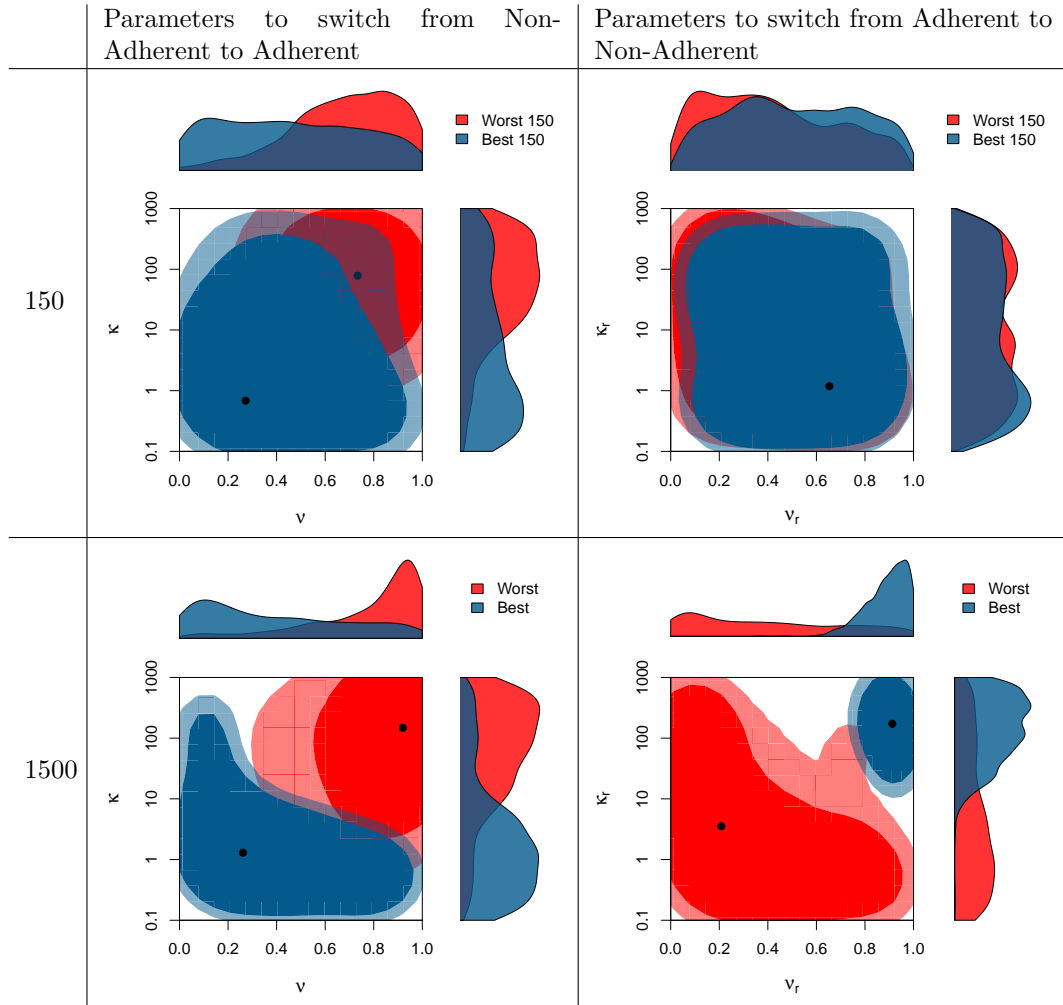


Figure S13: Variant D

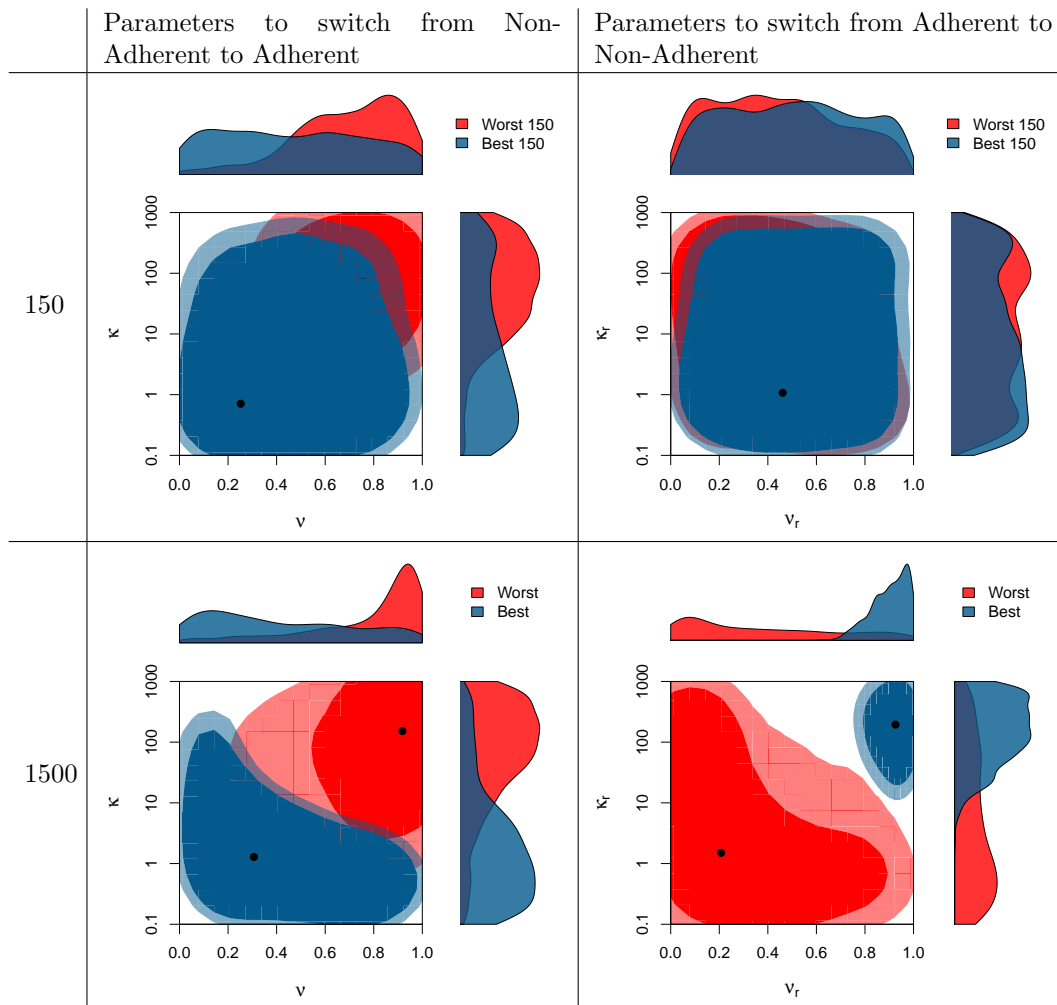


Figure S14: Variant F

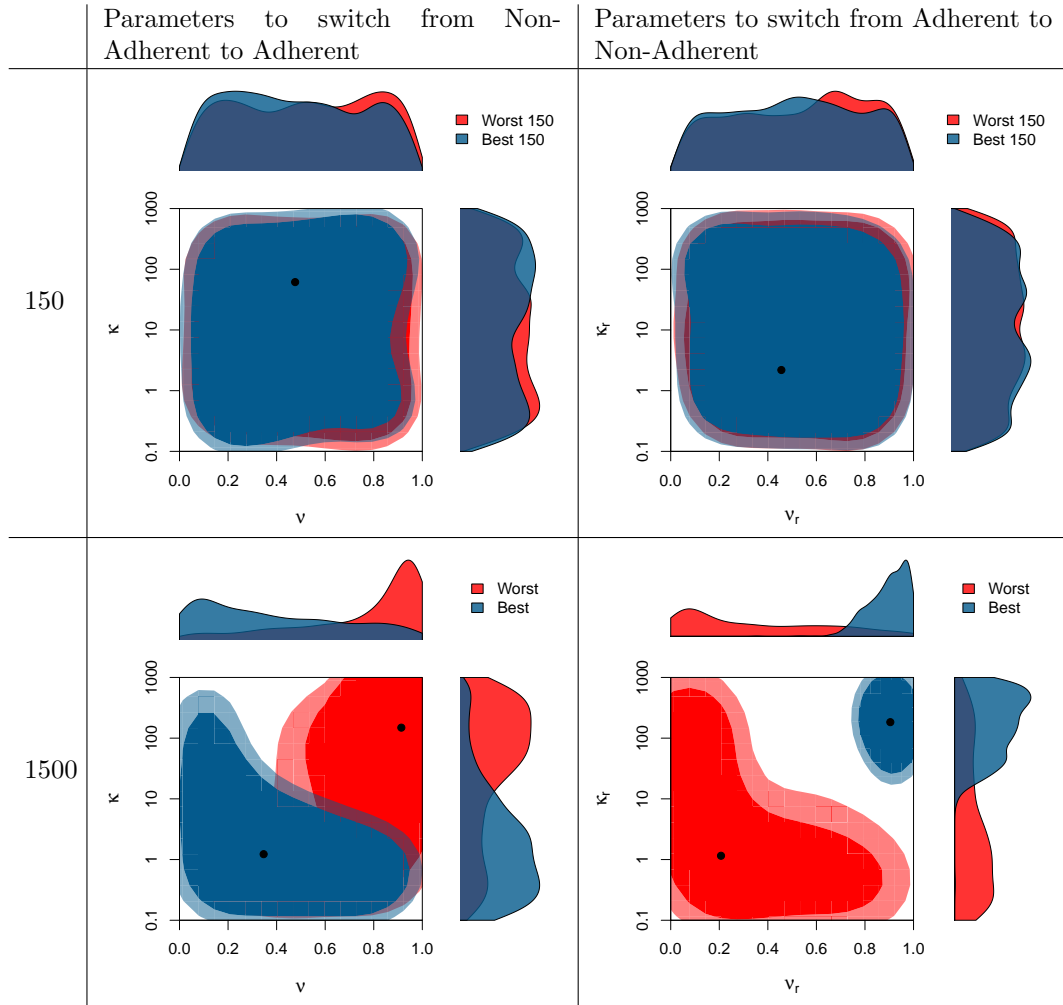


Figure S15: Variant F