Supplementary information for

Convolution of individual and group identity: self-reliance increases polarisation in basic opinion model

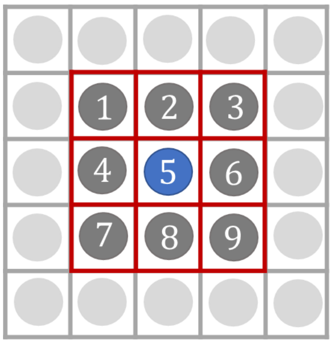
L. Quante*, A. Stechemesser*, D. Hödtke, A. Levermann

*These two authors contributed equally.

Supplementary figures

Supplementary calculations

1. Initialize toy example with 9 agents



Agent	A_i^* - conviction	γ_i - self reliance
1	0.8	0.7
2	0.84	0.8
3	0.7	0.9
4	0.75	0.6
5	0.2	0.2
6	0.75	0.3
7	0.2	0.1
8	0.1	0.3
9	0.15	0.2

2. Compute weighted difference to neighbour's attitudes for agent number 5

We assume all the weights $w_{j,i}$ are equal and set to 1/8.

$$W_5 = \sum_{j \in N_R} \frac{1}{8}(A_j - 0.2) = \underbrace{\frac{1}{8}\big((0.8-0.2) + \ldots + (0.15-0.2)\big)}_{\text{Mean difference to neighbour's attitudes}} = 0.19$$

3. Compute updated attitude for the agent

We compute the updated attitude based on the equation

$$\frac{\Delta A_i}{\Delta t} = \underbrace{\frac{|N_i|}{\tau}\left((1-\gamma_i) \cdot sign(N_i) + \gamma_i(A_i^* - A_i)\right)}_{RHS}$$

which can be modelled as

$$A_i(t+1) = A_i(t) + \Delta t(RHS(t))$$

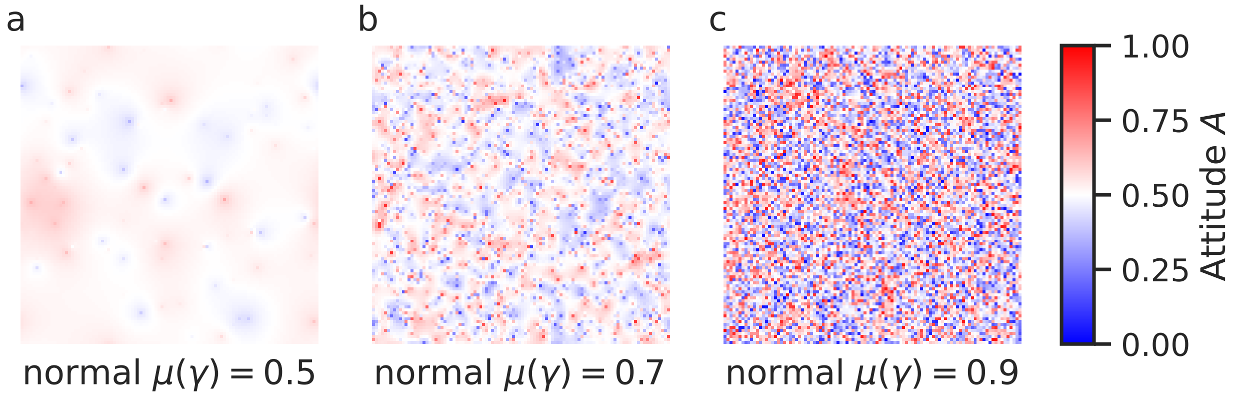
Setting $\Delta t = 1$ and $\tau = 1$ and plugging in W_5 and γ_5 yields:

$$A_5(1) = A_5(0) + \Big(0.19 \cdot \big((1-0.2) * 1 + 0.2(0.2-0.2)\big)\Big)$$

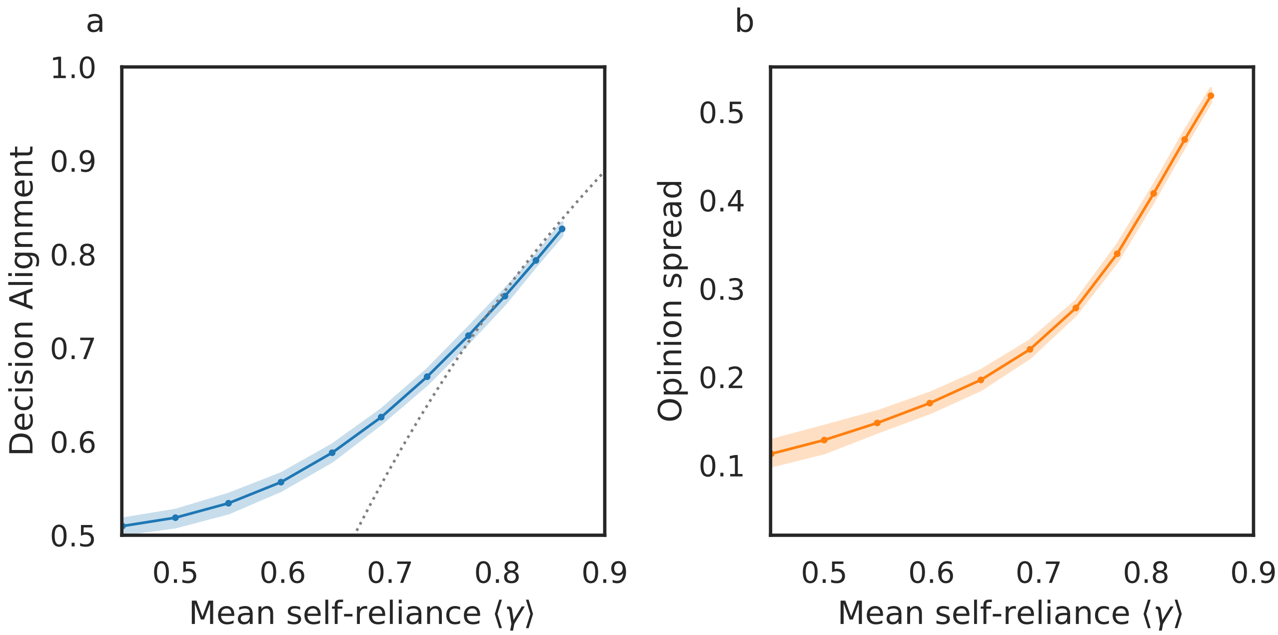
$$= 0.2 + 0.152 = 0.352$$

After one timestep, agent 5 has adjusted their opinion in the direction of their neighborhood majority (mean of neighborhood opinions = 0.53). As the agent has a relatively low self-reliance, the opinion will continue to be adjusted in the following steps and it will likely become part of a cluster around the self-reliant agents in the neighborhood.

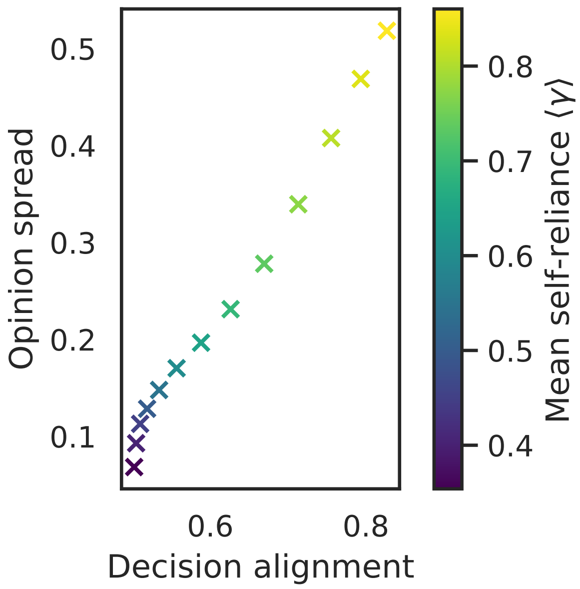
Supplementary Figure 1 Toy example: Agent evolution for an example agent. The figure illustrates the agent evolution for one agent in their neighbourhood and one model step. First, the weighted difference to the neighbour's attitudes is computed. Based on that and the agent-specific self-reliance and initial opinion, the attitude is updated.



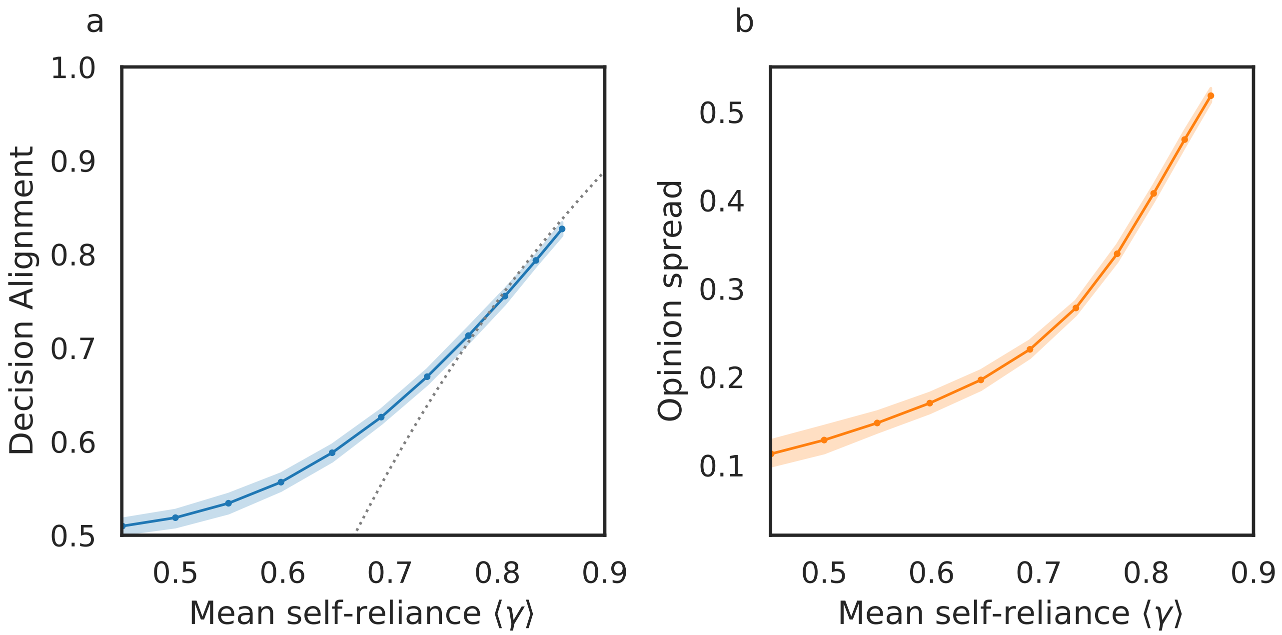
Supplementary Figure 2 Size and number of opinion clusters changes with mean self reliance. Panels show the final attitude after evolving the model for 1000 time steps for normally distributed self-reliance γ with varying means and fixed standard deviation $\sigma = 0.1$ as in the main text. Means of 0.5 (panel **a**), 0.7 (panel **b**) and 0.9 (panel **c**) are considered. The higher the mean of the normal distribution, the more fine-grained the clustering.



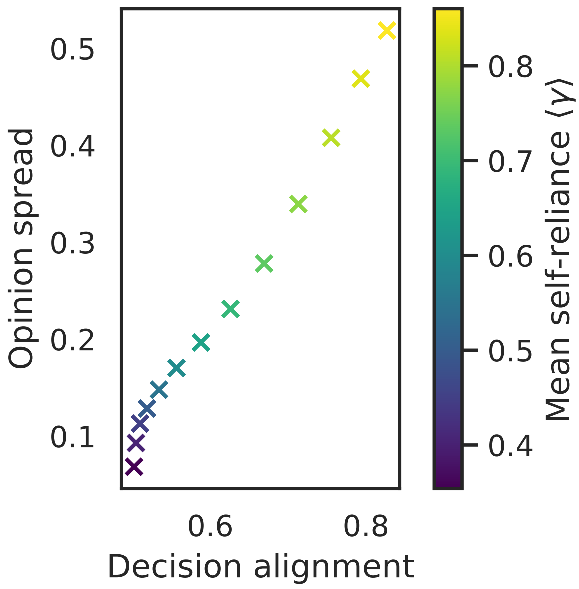
Supplementary Figure 3 Decision alignment and social cohesion for wider opinion spread. a shows the decision alignment in dependence of the average number of self-reliant agents. The decision alignment is computed as the average over the differences between the agents' initial decision and their final decision after evolving the model. The grey dots show an analytical approximation based on a mean-field approximation (Assuming 25% of agents are in equilibrium with their neighbours, i. e. $\mathbb{P}(N_i = 0) = 0.25$). **b** shows the opinion spread in dependence of the average number of self-reliant agents. In contrast to the main spread, where we consider the difference between the 90th and 10th is here measured as the difference between the 95th and 5th percentiles of the distribution of the final attitude. The average number of self-reliant agents corresponds to the mean of a normal distribution with mean γ and standard deviation σ = 0.1 from which the self-reliance was sampled. Confidence bands show the [5,95] Confidence interval based on 100 simulations with varying initial conditions. The results are robust to the wider percentile thresholds for the opinion spread.



Supplementary Figure 4 Trade-off between opinion spread and alignment with inherent decision for wider opinion spread. Decision alignment and opinion spread increase for higher mean self-reliance. Opinion spread between the 95th and 5th percentiles of final attitudes. Markers show mean values for colour-coded mean self-reliance.



Supplementary Figure 5 Decision alignment and social cohesion for narrower opinion spread. a shows the decision alignment in dependence of the average number of self-reliant agents. The decision alignment is computed as the average over the differences between the agents' initial decision and their final decision after evolving the model. The grey dots show an analytical approximation based on a mean-field approximation (Assuming 25% of agents are in equilibrium with their neighbours, i. e. $\mathbb{P}(N_i = 0) = 0.25$). **b** shows the opinion spread in dependence of the average number of self-reliant agents. In contrast to the main spread, where we consider the difference between the 90th and 10th is here measured as the difference between the 75th and 25th percentiles of the distribution of the final attitude. The average number of self-reliant agents corresponds to the mean of a normal distribution with mean γ and standard deviation σ = 0.1 from which the self-reliance was sampled. Confidence bands show the [5,95] Confidence interval based on 100 simulations with varying initial conditions. The results are robust to the narrower percentile thresholds for the opinion spread.

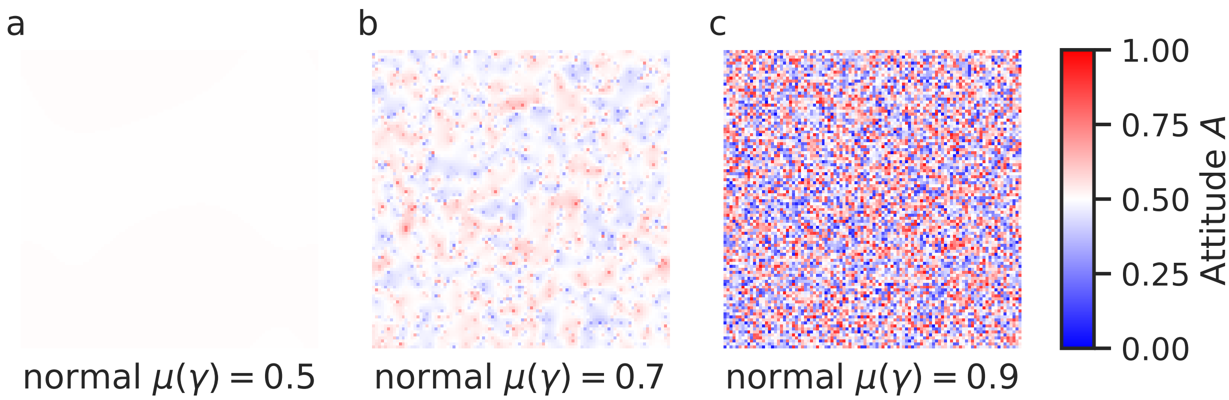


Supplementary Figure 6 Trade-off between opinion spread and alignment with inherent decision for narrower opinion spread. Decision alignment and opinion spread increase for higher mean self-reliance. Opinion spread between the 75th and 25th percentiles of final attitudes. Markers show mean values for colour-coded mean self-reliance.

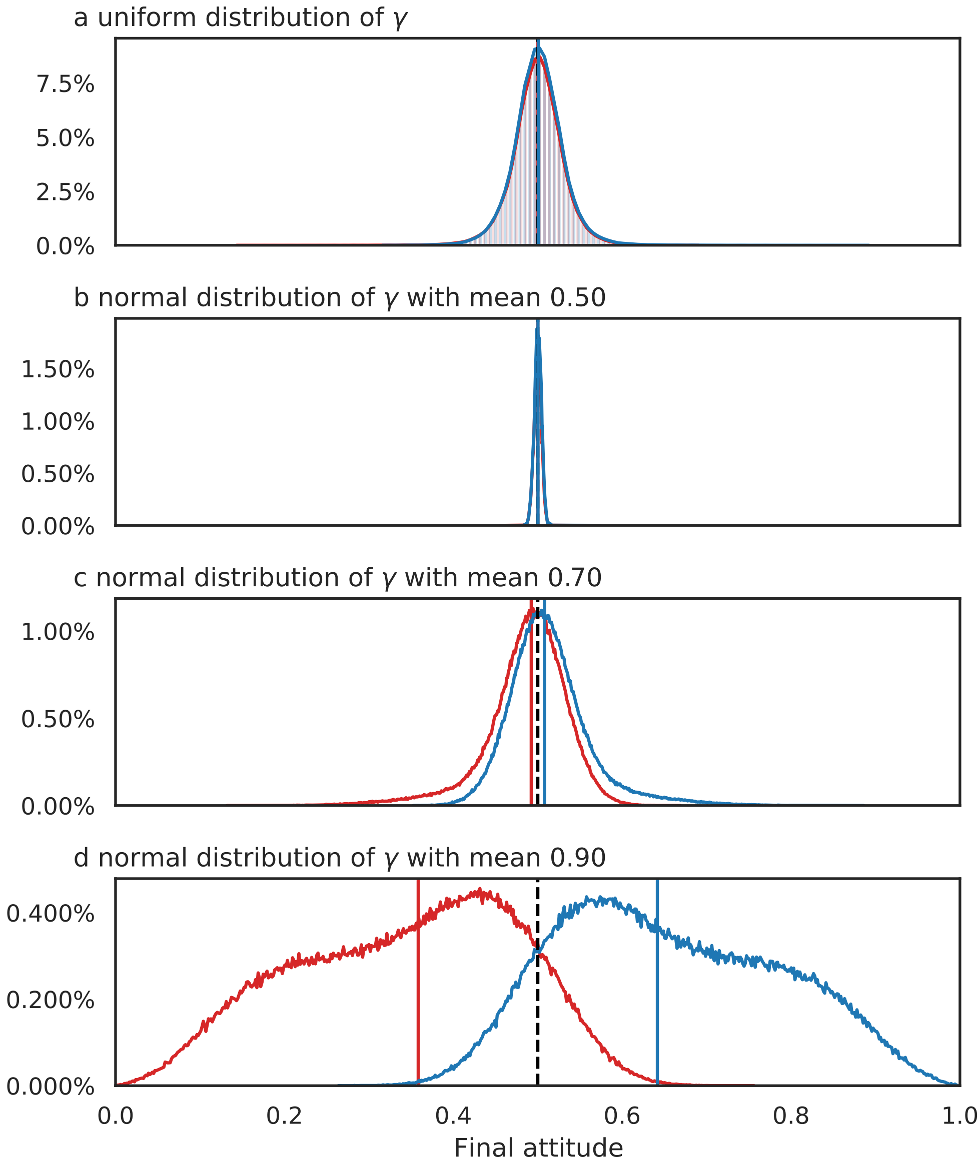
Sensitivity checks for varying standard deviations for the normal distributions of self-reliance

In the main analysis, self-reliance γ is sampled from a normal distribution with standard deviation $\sigma = 0.1$. Here, we show variations for this parameter to check sensitivity. Specifically, we replicate the main figures for values of $\sigma = 0.05, 0.15, 0.20$.

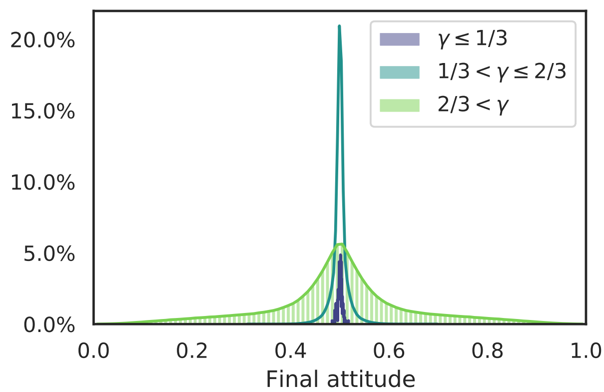
Standard deviation $\sigma = 0.05$



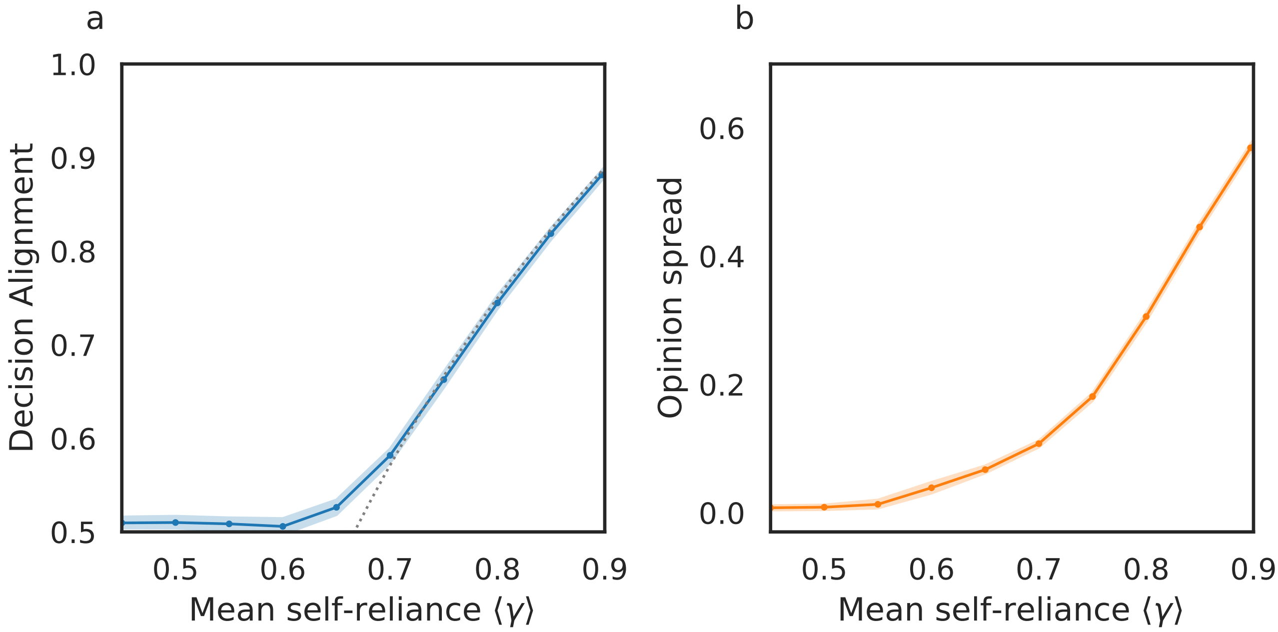
Supplementary Figure 7 Size and number of opinion clusters changes with mean self reliance; $\sigma = 0.05$. Panels show the final attitude after evolving the model for 1000 time steps for normally distributed self-reliance γ with varying means and fixed standard deviation $\sigma = 0.1$ as in the main text. Means of 0.5 (panel **a**), 0.7 (panel **b**) and 0.9 (panel **c**) are considered. The higher the mean of the normal distribution, the more fine-grained the clustering.



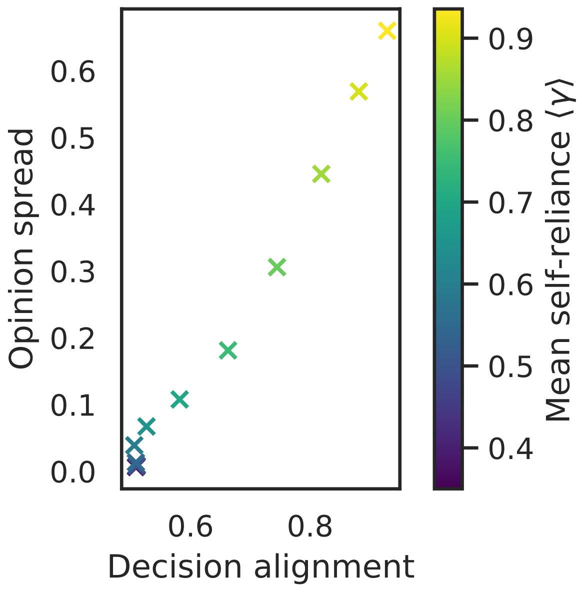
Supplementary Figure 8 Opinion spread increases in more self-reliant societies – standard deviation $\sigma = 0.05$. Panels show histograms of final attitude: **a** for uniform parameterised population, **b** for normal distributed γ with mean 0.5, **c** for normal distributed γ with mean 0.7 and **d** for normal distributed γ with mean 0.9. Red shows initial conviction < 0.5, blue shows initial conviction ≥ 0.5. Solid lines show medians, dashed black line 0.5.



Supplementary Figure 9 Highly self-reliant individuals have more polarised opinions – standard deviation $\sigma = 0.05$. Histograms of final attitude for bins of individual self-reliance γ based on ensemble of normal distributions of γ with means between 0.5 and 0.9.

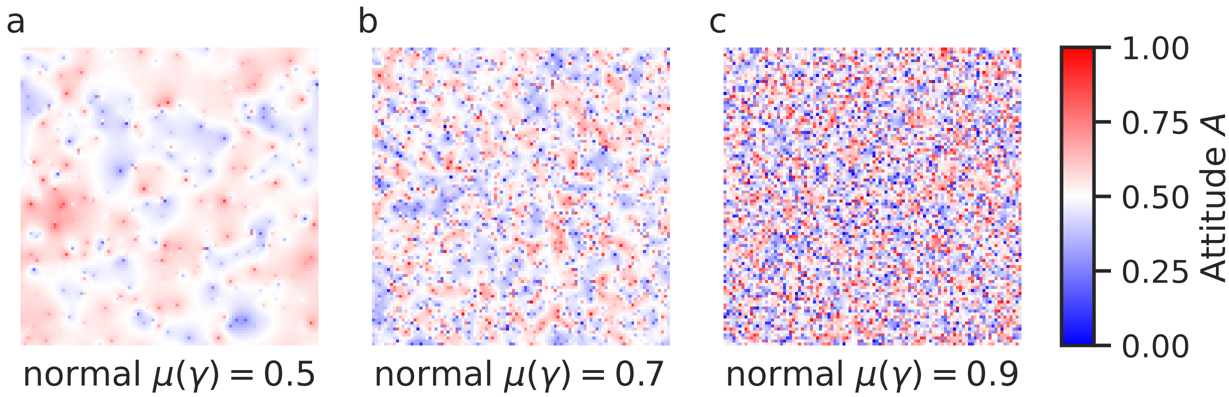


Supplementary Figure 10 Decision alignment and social cohesion for narrower opinion spread. – standard deviation $\sigma = 0.05$. a shows the decision alignment in dependence of the average number of self-reliant agents. The decision alignment is computed as the average over the differences between the agents' initial decision and their final decision after evolving the model. The grey dots show an analytical approximation based on a mean-field approximation (Assuming 25% of agents are in equilibrium with their neighbours, i. e. $\mathbb{P}(N_i = 0) = 0.25$). **b** shows the opinion spread in dependence of the average number of self-reliant agents measured as the difference between the 90th and 10th percentiles of the distribution of the final attitude. (The initial opinion spread is about 0.8 due to the uniform distribution of the initial attitude.) If the opinion spread is large societal opinions are drifting apart and social cohesion lowers. Hence, there is a trade-off between higher personal decision alignment with more self-reliant agents and more social cohesion with less self-reliant agents. In contrast to the main text, the average number of self-reliant agents here corresponds to the mean of a normal distribution with mean γ and standard deviation $\sigma = 0.05$ from which the self-reliance was sampled. Confidence bands show the [5,95] Confidence interval based on 100 simulations with varying initial conditions.

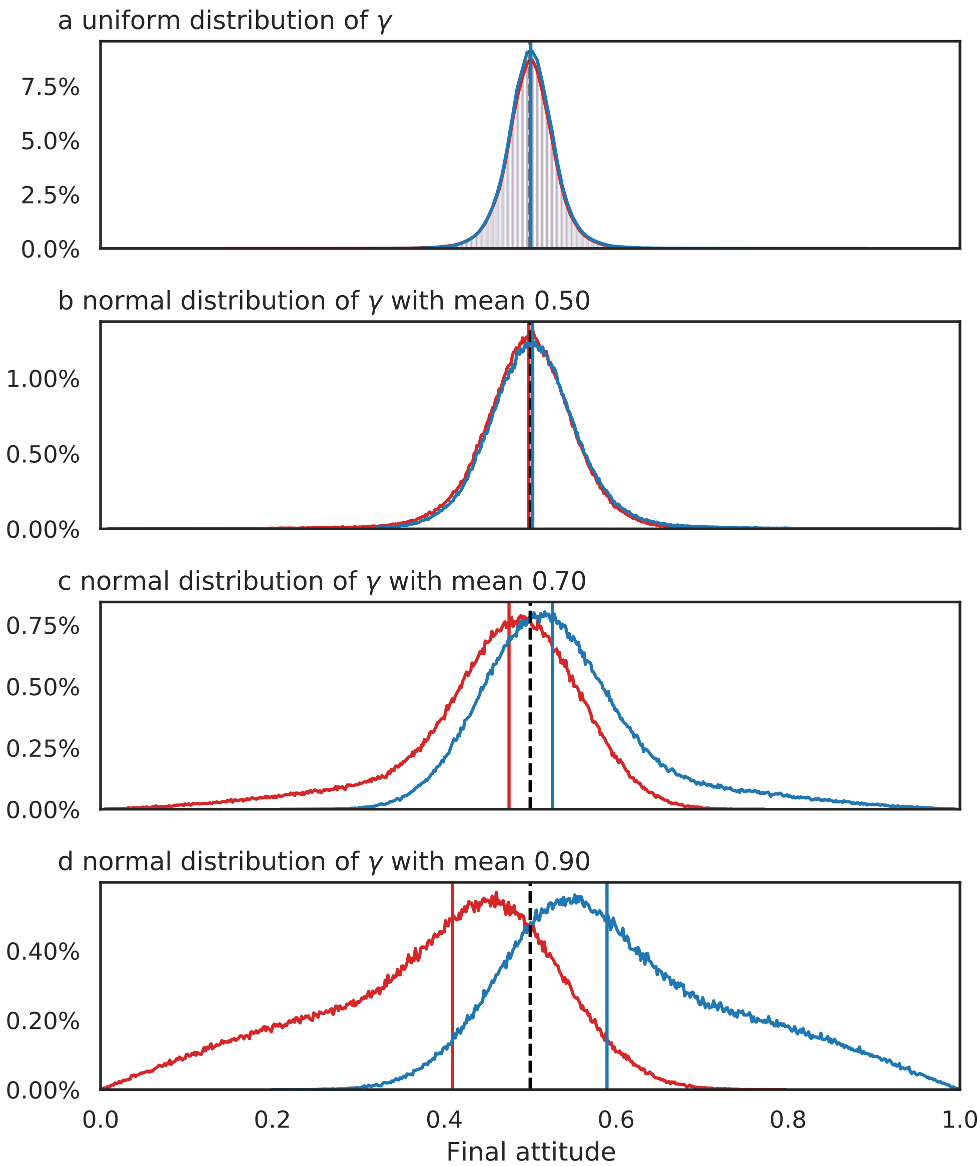


Supplementary Figure 11 Trade-off between opinion spread and alignment with inherent decision – standard deviation $\sigma = 0.05$. Decision alignment and opinion spread increase for higher mean self-reliance. Markers show mean values for colour-coded mean self-reliance.

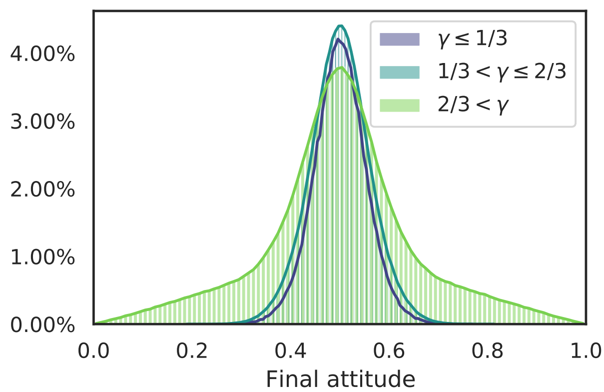
Standard deviation $\sigma = 0.15$



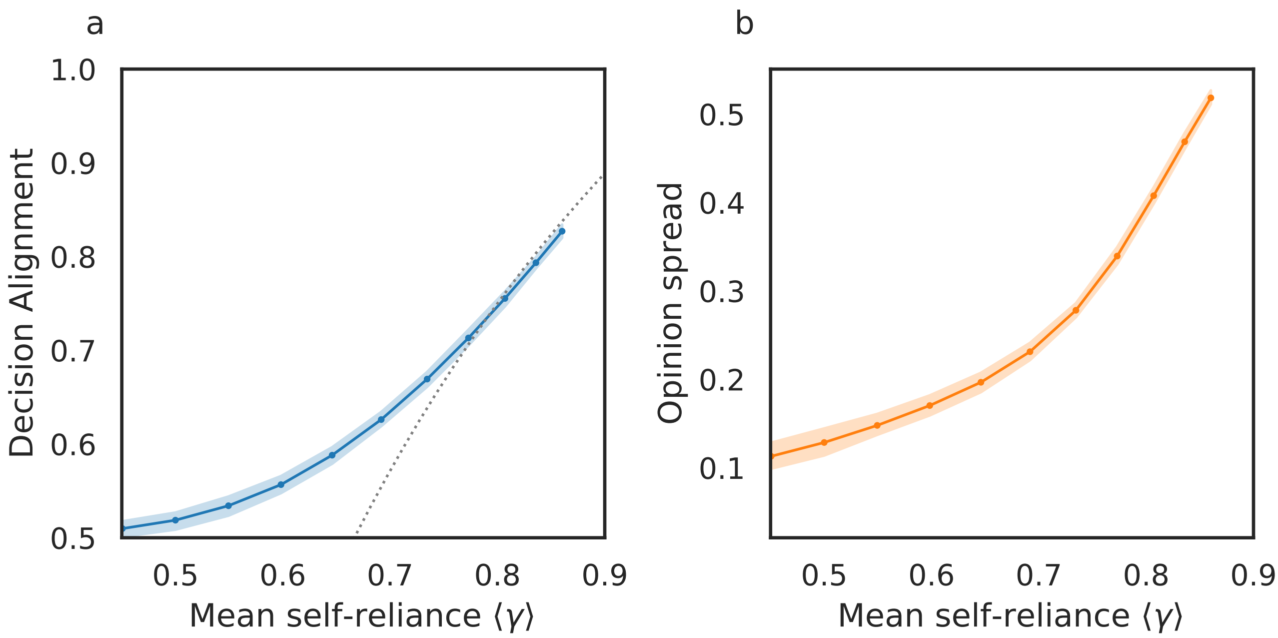
Supplementary Figure 12 Size and number of opinion clusters changes with mean self reliance; $\sigma = 0.15$. Panels show the final attitude after evolving the model for 1000 time steps for normally distributed self-reliance γ with varying means and fixed standard deviation $\sigma = 0.1$ as in the main text. Means of 0.5 (panel **a**), 0.7 (panel **b**) and 0.9 (panel **c**) are considered. The higher the mean of the normal distribution, the more fine-grained the clustering.



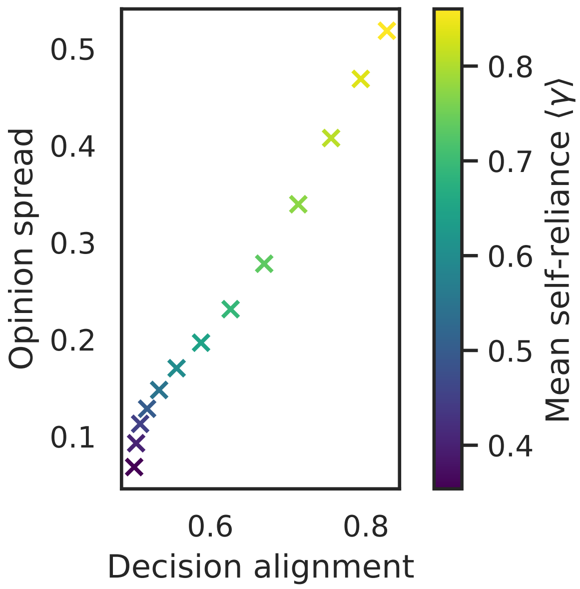
Supplementary Figure 13 Opinion spread increases in more self-reliant societies – standard deviation $\sigma = 0.15$. Panels show histograms of final attitude: **a** for uniform parameterised population, **b** for normal distributed γ with mean 0.5, **c** for normal distributed γ with mean 0.7 and **d** for normal distributed γ with mean 0.9. Red shows initial conviction < 0.5, blue shows initial conviction ≥ 0.5. Solid lines show medians, dashed black line 0.5.



Supplementary Figure 14 Highly self-reliant individuals have more polarised opinions – standard deviation $\sigma = 0.15$. Histograms of final attitude for bins of individual self-reliance γ based on ensemble of normal distributions of γ with means between 0.5 and 0.9.

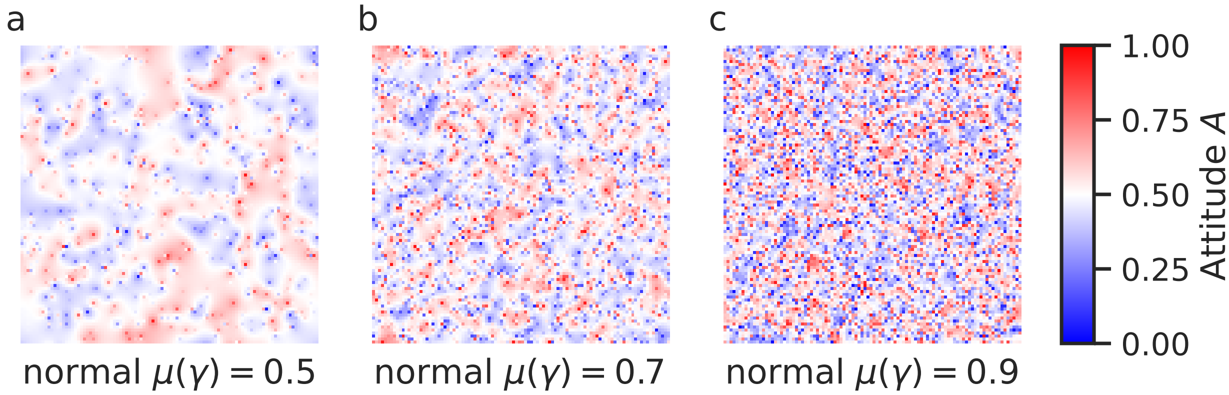


Supplementary Figure 15 Decision alignment and social cohesion for narrower opinion spread. – standard deviation $\sigma = 0.15$. a shows the decision alignment in dependence of the average number of self-reliant agents. The decision alignment is computed as the average over the differences between the agents' initial decision and their final decision after evolving the model. The grey dots show an analytical approximation based on a mean-field approximation (Assuming 25% of agents are in equilibrium with their neighbours, i. e. $\mathbb{P}(N_i = 0) = 0.25$). **b** shows the opinion spread in dependence of the average number of self-reliant agents measured as the difference between the 90th and 10th percentiles of the distribution of the final attitude. (The initial opinion spread is about 0.8 due to the uniform distribution of the initial attitude.) If the opinion spread is large societal opinions are drifting apart and social cohesion lowers. Hence, there is a trade-off between higher personal decision alignment with more self-reliant agents and more social cohesion with less self-reliant agents. In contrast to the main text, the average number of self-reliant agents here corresponds to the mean of a normal distribution with mean γ and standard deviation $\sigma = 0.15$ from which the self-reliance was sampled. Confidence bands show the [5,95] Confidence interval based on 100 simulations with varying initial conditions.

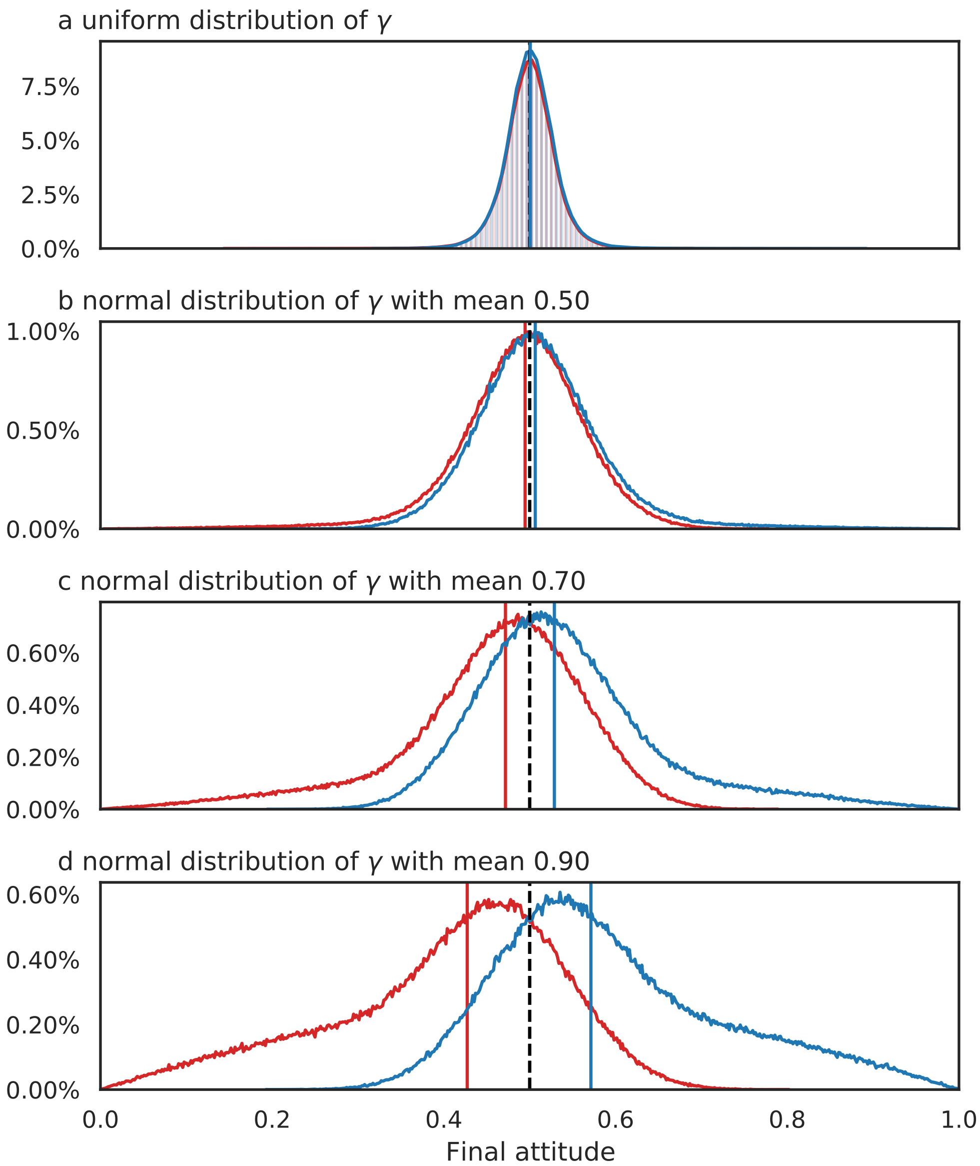


Supplementary Figure 16 Trade-off between opinion spread and alignment with inherent decision – standard deviation $\sigma = 0.15$. Decision alignment and opinion spread increase for higher mean self-reliance. Markers show mean values for colour-coded mean self-reliance.

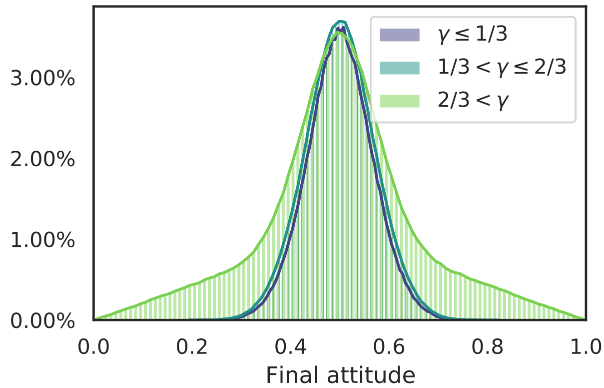
Standard deviation $\sigma = 0.20$



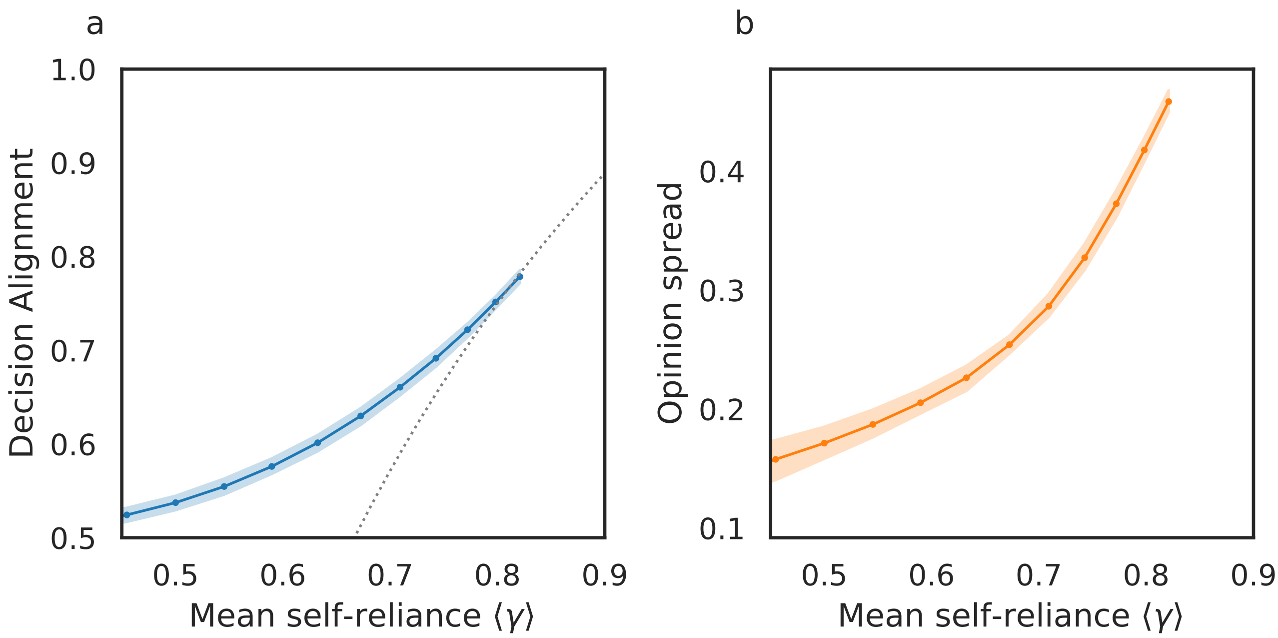
Supplementary Figure 17 Size and number of opinion clusters changes with mean self reliance; $\sigma = 0.20$. Panels show the final attitude after evolving the model for 1000 time steps for normally distributed self-reliance γ with varying means and fixed standard deviation $\sigma = 0.1$ as in the main text. Means of 0.5 (panel **a**), 0.7 (panel **b**) and 0.9 (panel **c**) are considered. The higher the mean of the normal distribution, the more fine-grained the clustering.



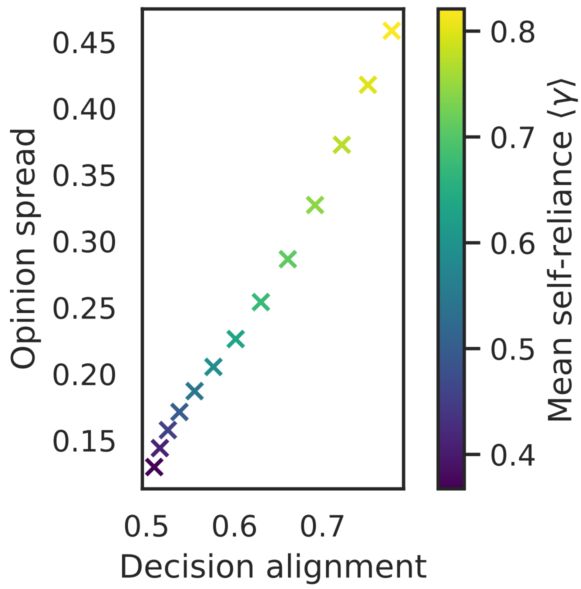
Supplementary Figure 18 Opinion spread increases in more self-reliant societies – standard deviation $\sigma = 0.20$. Panels show histograms of final attitude: **a** for uniform parameterised population, **b** for normal distributed γ with mean 0.5, **c** for normal distributed γ with mean 0.7 and **d** for normal distributed γ with mean 0.9. Red shows initial conviction < 0.5, blue shows initial conviction ≥ 0.5. Solid lines show medians, dashed black line 0.5.



Supplementary Figure 19 Highly self-reliant individuals have more polarised opinions – standard deviation $\sigma = 0.20$. Histograms of final attitude for bins of individual self-reliance γ based on ensemble of normal distributions of γ with means between 0.5 and 0.9.



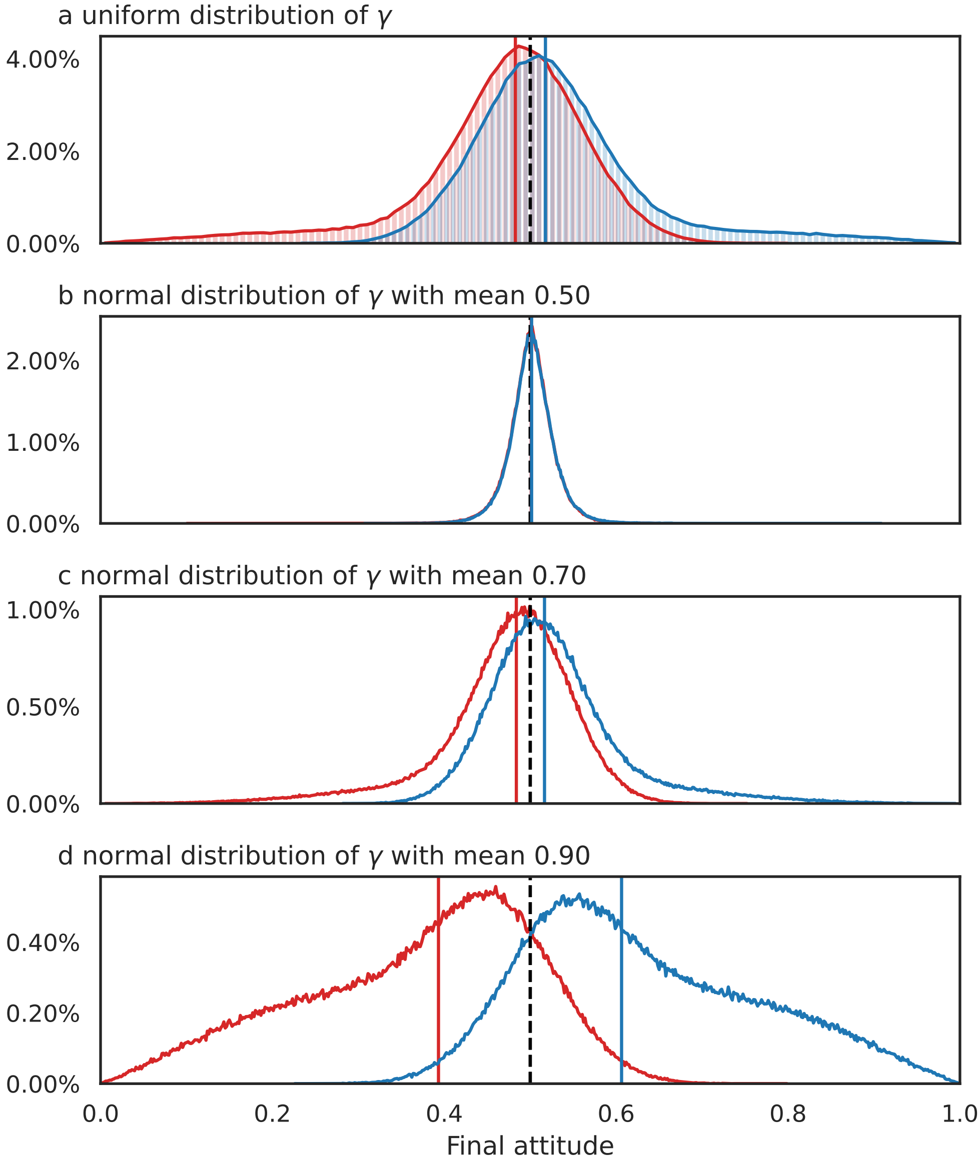
Supplementary Figure 20 Decision alignment and social cohesion for narrower opinion spread. – standard deviation $\sigma = 0.20$. a shows the decision alignment in dependence of the average number of self-reliant agents. The decision alignment is computed as the average over the differences between the agents' initial decision and their final decision after evolving the model. The grey dots show an analytical approximation based on a mean-field approximation (Assuming 25% of agents are in equilibrium with their neighbours, i. e. $\mathbb{P}(N_i = 0) = 0.25$). **b** shows the opinion spread in dependence of the average number of self-reliant agents measured as the difference between the 90th and 10th percentiles of the distribution of the final attitude. (The initial opinion spread is about 0.8 due to the uniform distribution of the initial attitude.) If the opinion spread is large societal opinions are drifting apart and social cohesion lowers. Hence, there is a trade-off between higher personal decision alignment with more self-reliant agents and more social cohesion with less self-reliant agents. In contrast to the main text, the average number of self-reliant agents here corresponds to the mean of a normal distribution with mean γ and standard deviation $\sigma = 0.20$ from which the self-reliance was sampled. Confidence bands show the [5,95] Confidence interval based on 100 simulations with varying initial conditions.



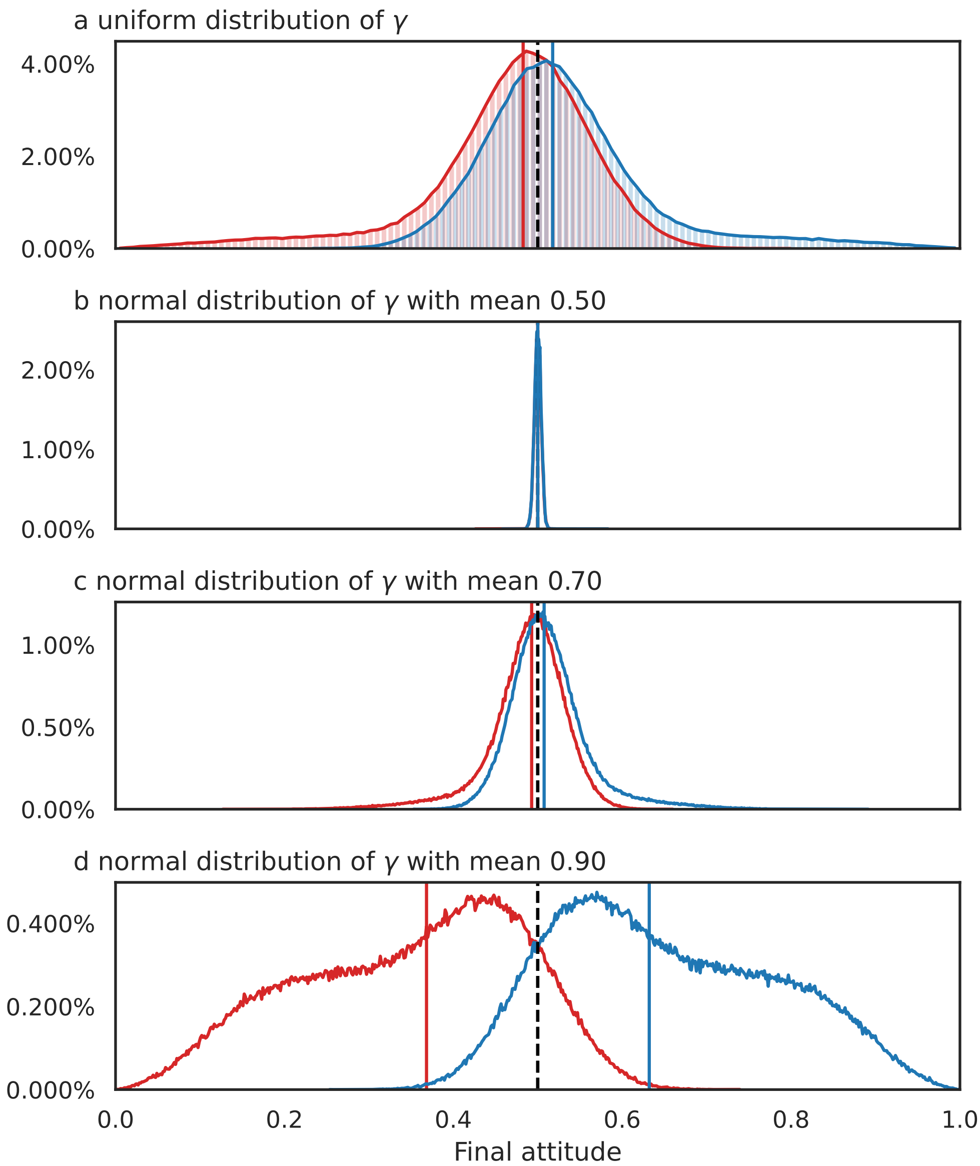
Supplementary Figure 21 Trade-off between opinion spread and alignment with inherent decision – standard deviation $\sigma = $ 0.20. Decision alignment and opinion spread increase for higher mean self-reliance. Markers show mean values for colour-coded mean self-reliance.

Assuming independence of initial conviction and initial attitude

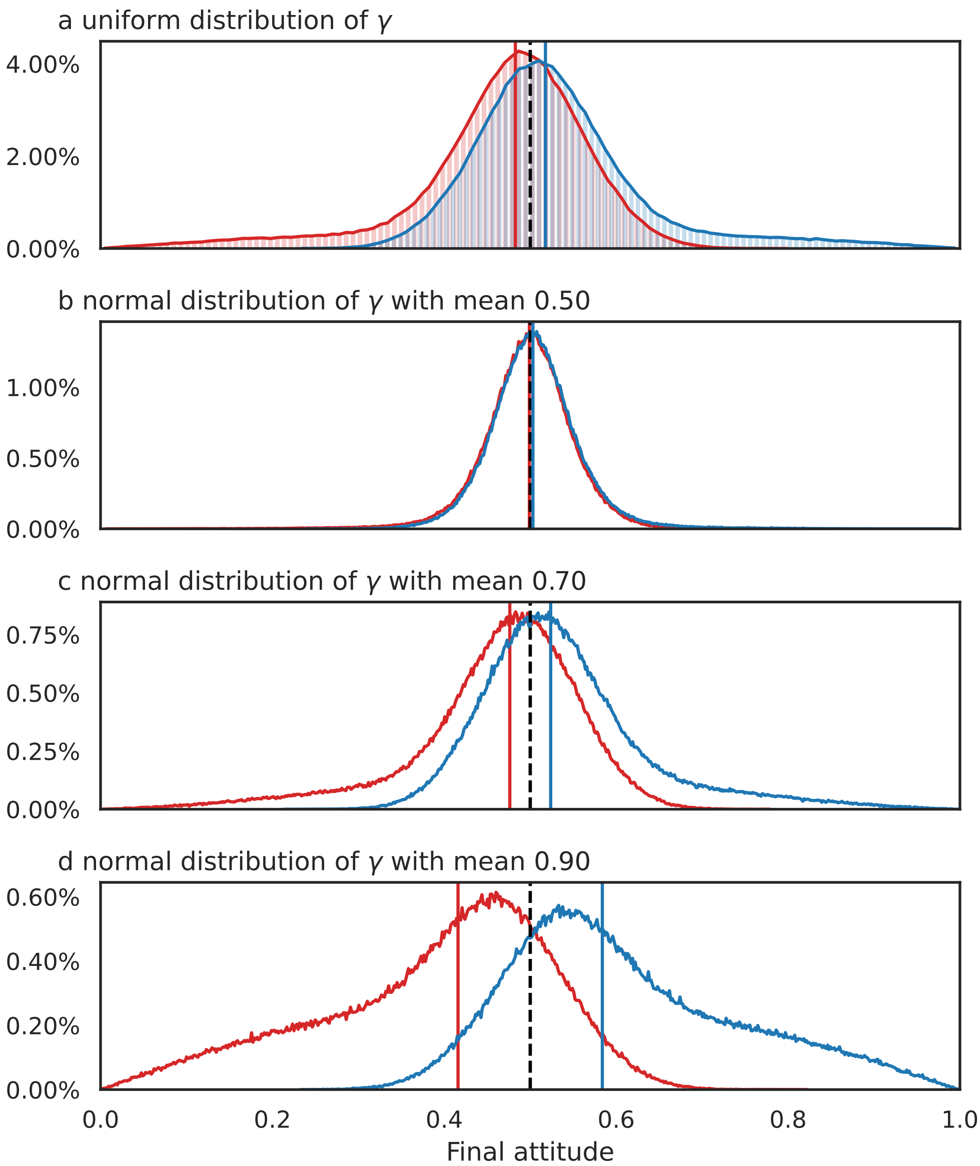
In the main analysis, we assume that the initial conviction A_i^* of an agent is equal to their initial attitude $A_i(0)$. Here, we relax this assumption as a robustness check and draw both parameters independently. To evaluate the impact on the results, we reproduce Fig. 3 from the main text with independently drawn A_i^* and $A_i(0)$ for different standard deviations of the self-reliance γ. Overall, we find similar distributions of the final attitude as shown in Supplementary Figs. 22–25.



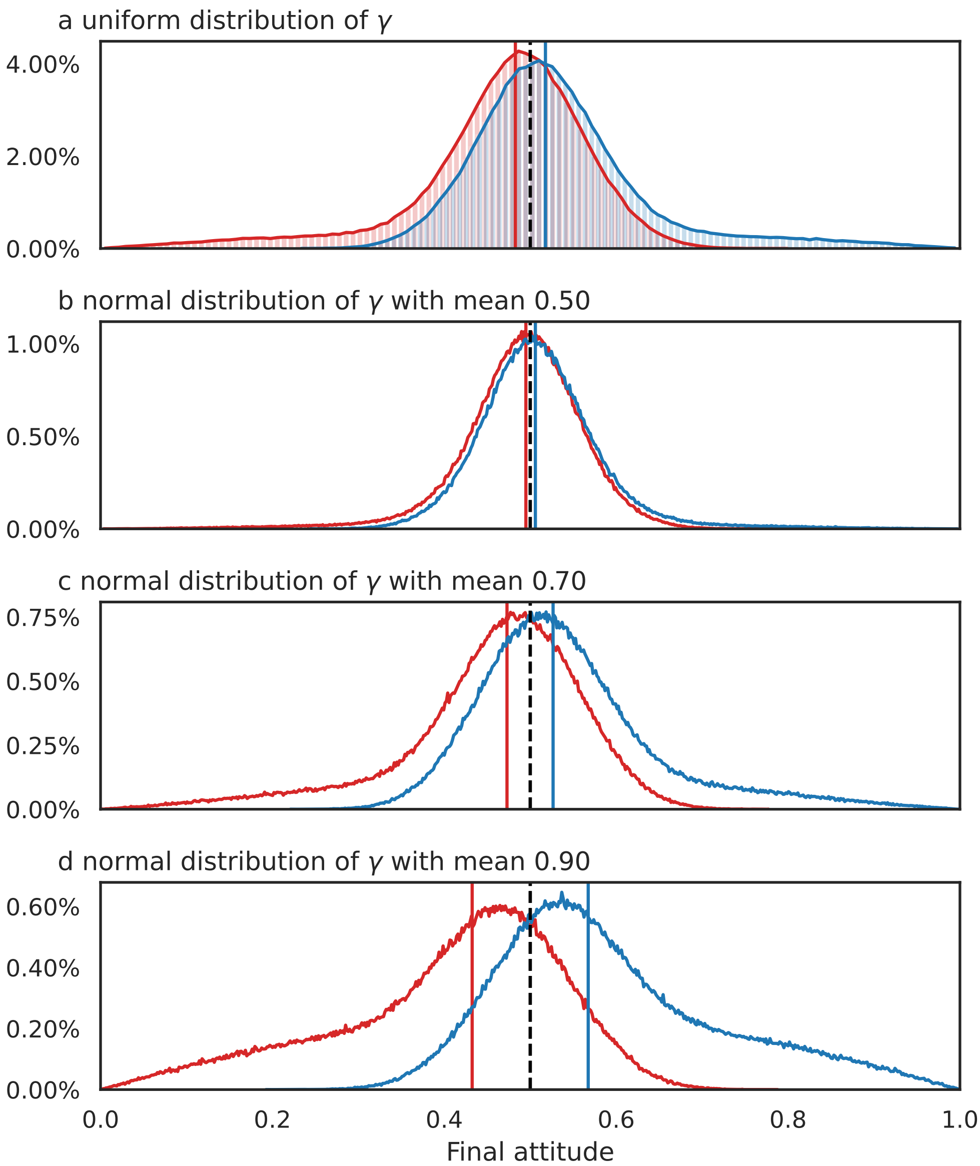
Supplementary Figure 22 Opinion spread increases in more self-reliant societies for independently sampling of A_i^* and $A_i(0)$ – SD of self-reliance $\sigma = 0.10$. Panels show histograms of final attitude: **a** for uniform parameterised population, **b** for normal distributed γ with mean 0.5, **c** for normal distributed γ with mean 0.7 and **d** for normal distributed γ with mean 0.9. Red shows initial conviction < 0.5, blue shows initial conviction ≥ 0.5. Solid lines show medians, the dashed black line 0.5. In contrast to the main text, A_i^* and $A_i(0)$ are not equal but sampled independently.



Supplementary Figure 23 Opinion spread increases in more self-reliant societies for independently sampling of A_i^* and $A_i(0)$ – SD of self-reliance $\sigma = 0.05$. Panels show histograms of final attitude: **a** for uniform parameterised population, **b** for normal distributed γ with mean 0.5, **c** for normal distributed γ with mean 0.7 and **d** for normal distributed γ with mean 0.9. Red shows initial conviction < 0.5, blue shows initial conviction ≥ 0.5. Solid lines show medians, the dashed black line 0.5. In contrast to the main text, A_i^* and $A_i(0)$ are not equal but sampled independently.



Supplementary Figure 24 Opinion spread increases in more self-reliant societies for independently sampling of A_i^* and $A_i(0)$ – SD of self-reliance $\sigma = 0.15$. Panels show histograms of final attitude: **a** for uniform parameterised population, **b** for normal distributed γ with mean 0.5, **c** for normal distributed γ with mean 0.7 and **d** for normal distributed γ with mean 0.9. Red shows initial conviction < 0.5, blue shows initial conviction ≥ 0.5. Solid lines show medians, the dashed black line 0.5. In contrast to the main text, A_i^* and $A_i(0)$ are not equal but sampled independently.



Supplementary Figure 25 Opinion spread increases in more self-reliant societies for independently sampling of A_i^* and $A_i(0)$ – SD of self-reliance $\sigma = 0.20$. Panels show histograms of final attitude: **a** for uniform parameterised population, **b** for normal distributed γ with mean 0.5, **c** for normal distributed γ with mean 0.7 and **d** for normal distributed γ with mean 0.9. Red shows initial conviction < 0.5, blue shows initial conviction ≥ 0.5. Solid lines show medians, the dashed black line 0.5. In contrast to the main text, A_i^* and $A_i(0)$ are not equal but sampled independently.

Randomisation of placement of agents on the grid

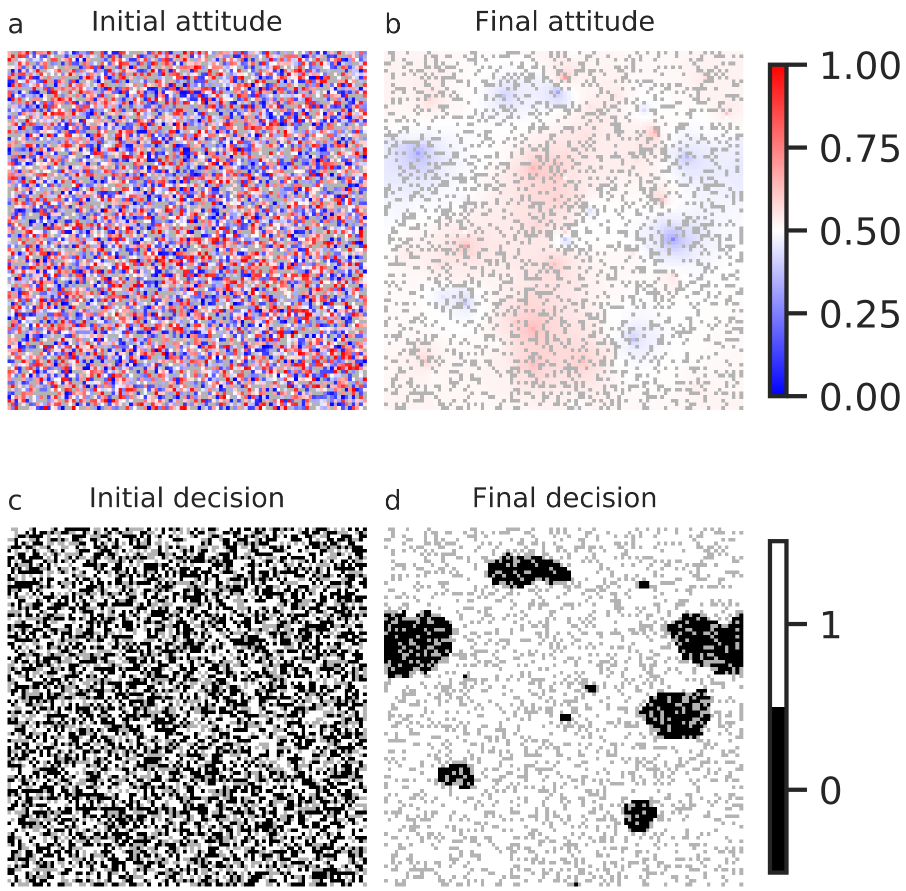
To check for robustness with respect to variation of the number of neighbours, we repeat the simulations after an random initialisation of the agents on the grid: Instead of having one agent in each grid position, we set a placement probability $p = 0.75$, with which agents are placed at each grid position.

During initialisation of the model, for each position in the grid we now draw a value x from an uniform distribution between 0 and 1, $\mathcal{U}(0, 1)$. If $x \leq p$, an agent is placed, otherwise the position remains empty.

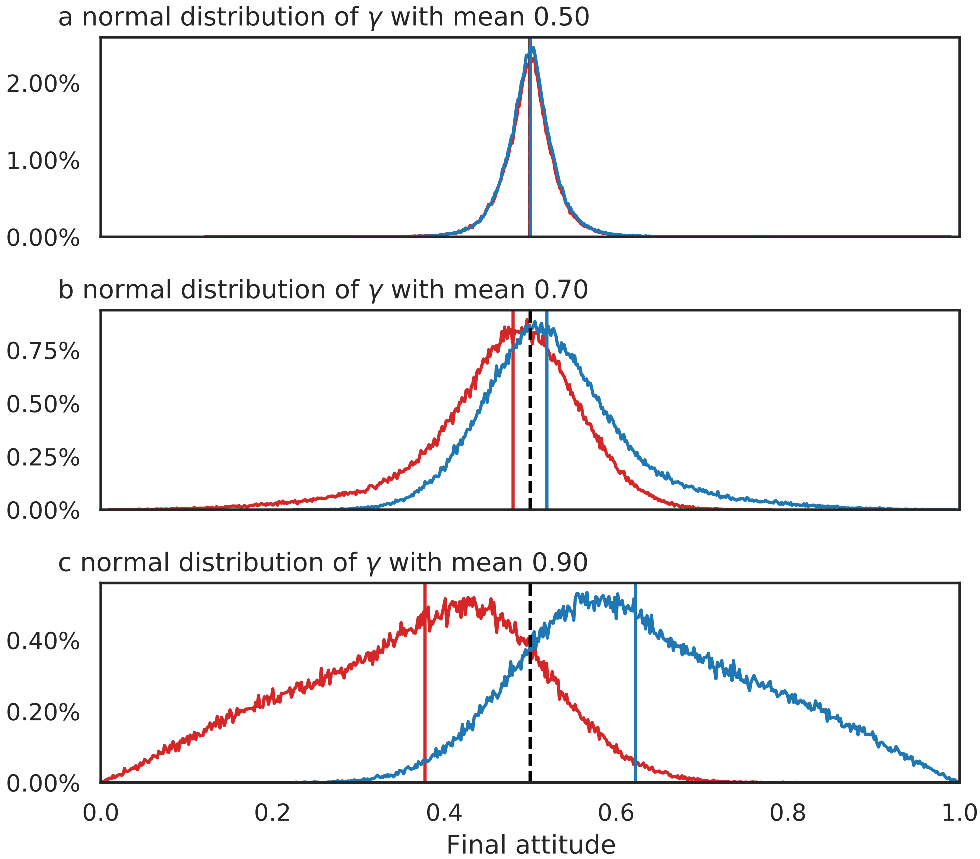
As can be seen in Supplementary Figs. 26–29, all qualitative results remain, despite the less regular structure of the partly populated grid.

Varying neighbourhood radius

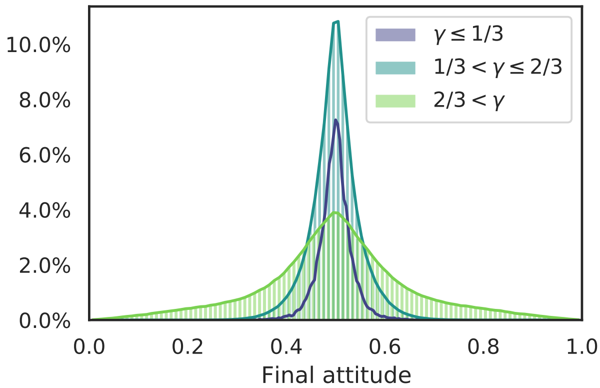
To check for the efffect of increasing neighbourdhood size, we run simulations increasing the neighbourhood radius to 2, i.e. all agents with Chebyshev distance <=2 are considered part of the neighbourhood of an agent.



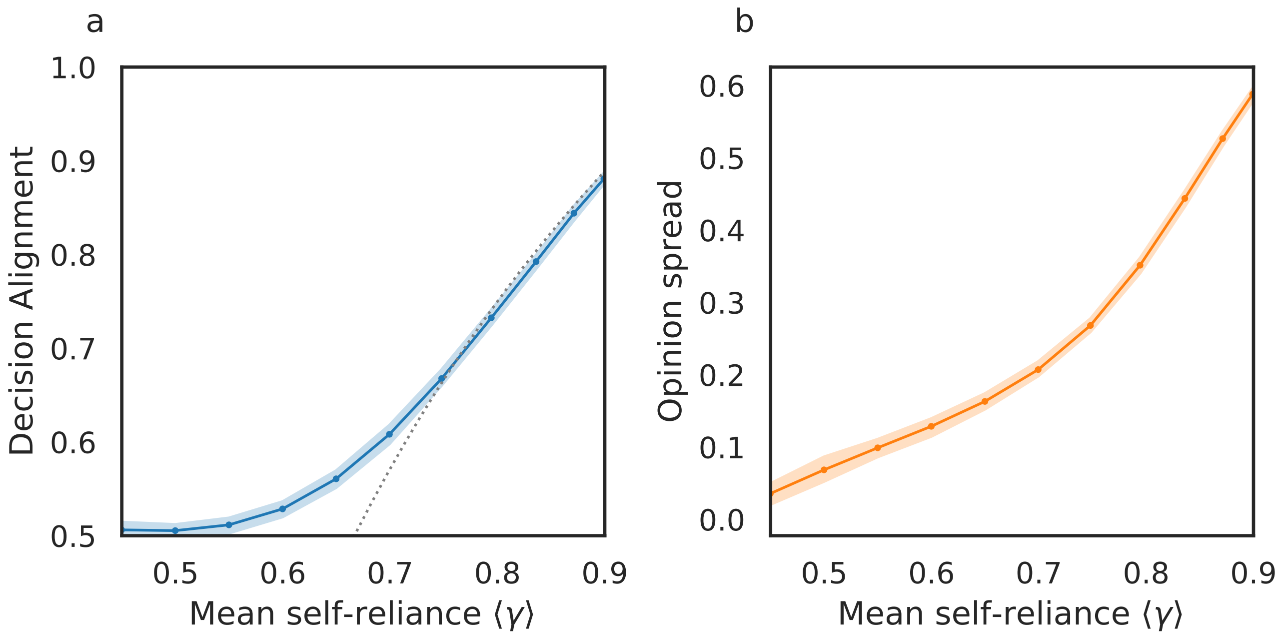
Supplementary Figure 26 Partly populated grid conserves trade-off between belonging and individualism leads to emergence of stable, opposing opinion clusters. a shows the uniformly distributed initial attitude for each agent. **b** shows the attitude after evolving the model for 1000 time steps. Opinion clusters emerge around few agents with strong opinions and many agents with more moderate views. **c** visualises the binary initial decision, which is based on the initial conviction (threshold of 0.5). No clusters are visible. In contrast, the final decision (**d**), which is based on the final attitude, shows clear opinion clusters. Grey show grid points without an agent.



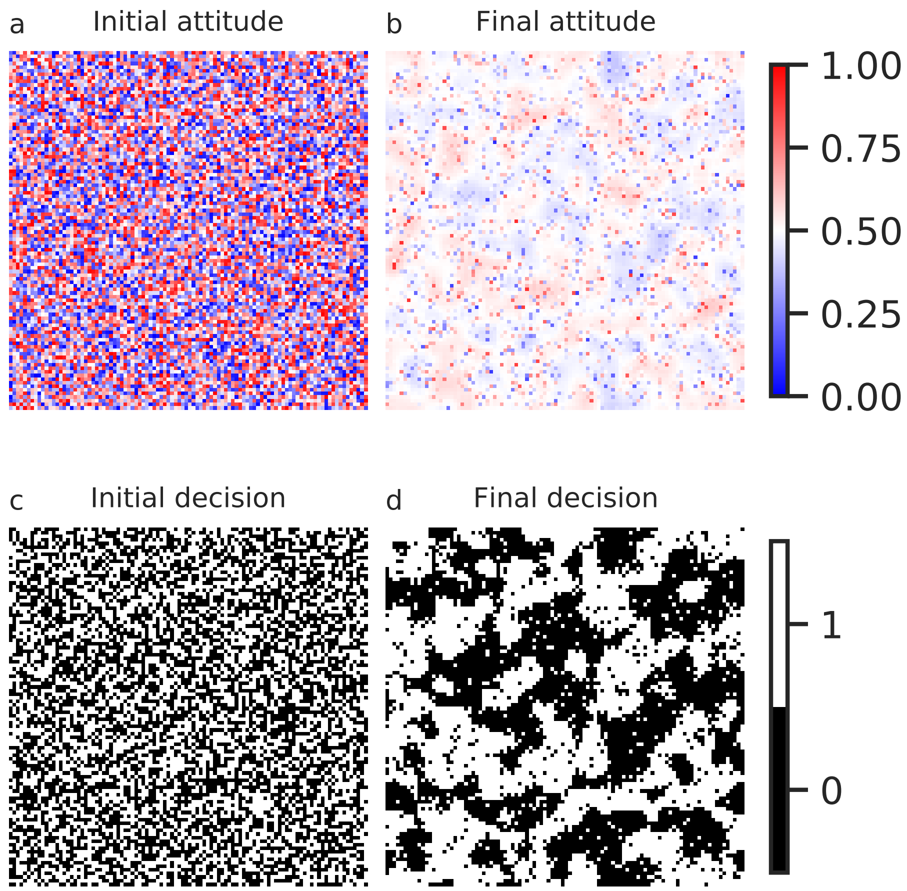
Supplementary Figure 27 Partly populated grid: Opinion spread increases in more self-reliant societies. Panels show histograms of final attitude: **a** for normal distributed γ with mean 0.5, **c** for normal distributed γ with mean 0.7 and **d** for normal distributed γ with mean 0.9. Red shows initial conviction $<$ 0.5, blue shows initial conviction $\geq$ 0.5. Solid lines show medians, dashed black line 0.5.



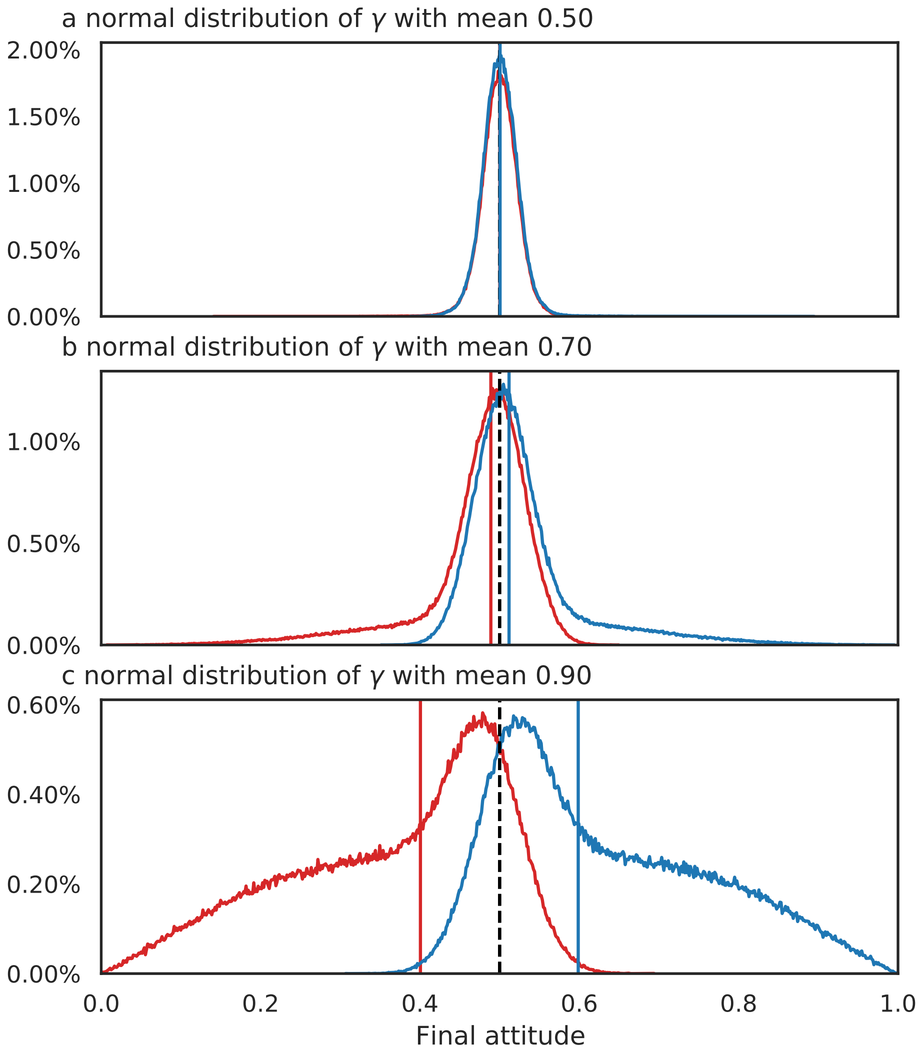
Supplementary Figure 28 Partly populated grid: Highly self-reliant individuals have more polarised opinions. Histograms of final attitude for bins of individual self-reliance γ based on ensemble of normal distributions of γ with means between 0.5 and 0.9



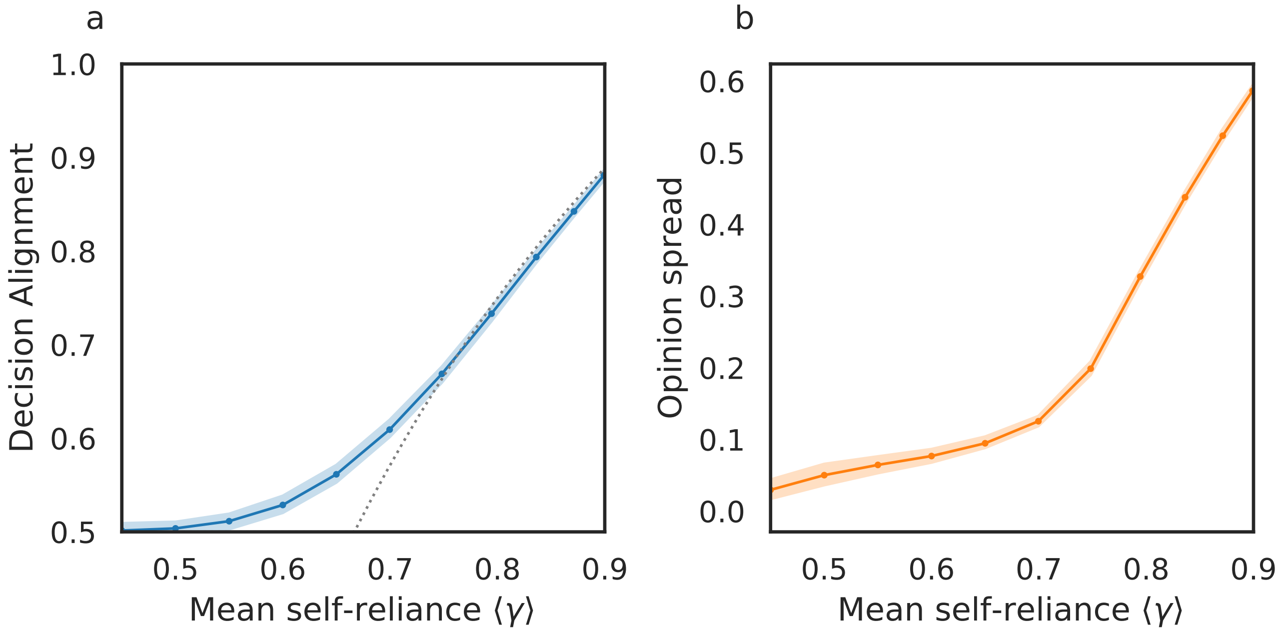
Supplementary Figure 29 Partly populated grid: Trade-off between decision alignment and social cohesion. a shows the decision alignment in dependence on the average number of self-reliant agents. The decision alignment is computed as the average over the differences between the agents' initial decision and their final decision after evolving the model. The grey dots show an analytical approximation based on a mean-field approximation (Assuming 25% of agents are in equilibrium with their neighbours, i. e. $\mathbb{P}(N_i = 0) = 0.25$). **b** shows the opinion spread in dependence on the average number of self-reliant agents measured as the difference between the 90th and 10th percentiles of the distribution of the final attitude. The initial opinion spread is about 0.8 due to the uniform distribution of the initial conviction. If the opinion spread is large, societal opinions are drifting apart and social cohesion lowers. Hence, there is a trade-off between higher personal decision alignment with more self-reliant agents and more social cohesion with less self-reliant agents. The average number of self-reliant agents corresponds to the mean of a normal distribution with mean γ and standard deviation σ = 0.1 from which the self-reliance was sampled. Confidence bands show the [5,95] confidence interval based on 50 simulations with varying initial conditions.



Supplementary Figure 30 Increasing neighbourhood radius conserves trade-off between belonging and individualism leads to emergence of stable, opposing opinion clusters. a shows the uniformly distributed initial attitude for each agent. **b** shows the attitude after evolving the model for 1000 time steps. Opinion clusters emerge around few agents with strong opinions and many agents with more moderate views. **c** visualises the binary initial decision, which is based on the initial conviction (threshold of 0.5). No clusters are visible. In contrast, the final decision (**d**), which is based on the final attitude, shows clear opinion clusters.



Supplementary Figure 31 Increased neighbourhood radius to 2: Opinion spread increases in more self-reliant societies, but slightly more central tendencies due to the increased radius leading to more connection between agents. Panels show histograms of final attitude: **a** for normal distributed γ with mean 0.5, **c** for normal distributed γ with mean 0.7 and **d** for normal distributed γ with mean 0.9. Red shows initial conviction < 0.5, blue shows initial conviction ≥ 0.5. Solid lines show medians, dashed black line 0.5.



Supplementary Figure 32 Increased neighbourhood radius to 2: Trade-off between decision alignment and social cohesion. a shows the decision alignment in dependence on the average number of self-reliant agents. The decision alignment is computed as the average over the differences between the agents' initial decision and their final decision after evolving the model. The grey dots show an analytical approximation based on a mean-field approximation (Assuming 25% of agents are in equilibrium with their neighbours, i. e. $\mathbb{P}(N_i = 0) = 0.25$). **b** shows the opinion spread in dependence on the average number of self-reliant agents measured as the difference between the 90th and 10th percentiles of the distribution of the final attitude. The initial opinion spread is about 0.8 due to the uniform distribution of the initial conviction. If the opinion spread is large, societal opinions are drifting apart and social cohesion lowers. Hence, there is a trade-off between higher personal decision alignment with more self-reliant agents and more social cohesion with less self-reliant agents. The average number of self-reliant agents corresponds to the mean of a normal distribution with mean γ and standard deviation σ = 0.1 from which the self-reliance was sampled. Confidence bands show the [5,95] confidence interval based on 100 simulations with varying initial conditions.

1 Equilibrium conditions

The model equilibrium is given by the concurrent individual agent equilibria defined by the condition

$$0 = \frac{1}{\tau}\left((1-\gamma_i)N_i + \gamma_i|N_i|(A_i^* - A_i)\right) \ . \tag{8}$$

This equation has two qualitatively different solutions:

$$N_i = 0 \text{ and } A_i^* - A_i = \frac{1-\gamma_i}{\gamma_i}\text{sign}\,(N_i) \tag{9}$$

The first solution describes the situation when an agents attitude equals the average of their neighbours attitudes. The second equilibrium with $N_i \neq 0$ is reached if the dissonance between initial conviction and actual attitude is equal to the strength of individuality and direction of the disagreement with the neighbours. From the parameter ranges of γ_i, A_i^* and A_i follows that this equilibrium can only be met by agents with $\gamma_i \geq 0.5$, since

$$1 \geq |A_i^* - A_i| = \frac{1-\gamma_i}{\gamma_i} \Rightarrow \gamma_i \geq 0.5 \tag{10}$$

2 Detailed derivation of eq. (6)

We derive eq. (6),first we decompose the expected value by the sign of N_i:

$$\Delta = \mathbb{E}\left[\delta_i\right] \approx \mathbb{P}(N_i < 0)\mathbb{P}(\delta_i = 1|N_i < 0) + \mathbb{P}(N_i > 0)\mathbb{P}(\delta_i = 1|N_i > 0) + \mathbb{P}(N_i = 0)\mathbb{P}(\delta_i = 1|N_i = 0). \tag{11}$$

Here we assume independence of N_i from A_i, which can be considered similar to a mean-field approximation, i.e. the assumption of independence does not hold on the individual level, but for the overall society at large and thus this style of approximation is instrumental to predict system wide metrics like the average decision alignment.

Assuming equilibrium with $N_i \neq 0$, we show in detail in the next paragraph, that

$$\mathbb{P}(\delta_i = 1|N_i < 0) = \mathbb{P}(\delta_i = 1|N_i > 0) = \frac{1}{2} + \frac{1}{2}\mathbb{P}(\mathcal{U}(0, 0.5) > \frac{1-\gamma_i}{\gamma_i}). \tag{12}$$

For $N_i = 0$ it follows due to the uniform distribution of A_i and A_i^* that

$$\mathbb{P}(\delta_i = 1|N_i = 0) = \mathbb{P}(\delta_i = 1|A_i < 0.5) + \mathbb{P}(\delta_i = 1|A_i \geq 0.5) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

Thus, we find using $\mathbb{P}(N_i = 0) = 1 - \mathbb{P}(N_i \neq 0)$

$$\mathbb{E}\left[\delta_i\right] = \left[\mathbb{P}(N_i < 0) + \mathbb{P}(N_i > 0)\right]\left[\frac{1}{2} + \mathbb{P}(\mathcal{U}(0, 0.5) > \frac{1-\gamma_i}{\gamma_i})\right] + \frac{1}{2}\mathbb{P}(N_i = 0) \quad (13)$$

$$= \frac{1}{2} + \left[\mathbb{P}(N_i \neq 0)\right]\mathbb{P}(\mathcal{U}(0, 0.5) > \frac{1-\gamma_i}{\gamma_i}) \quad (14)$$

Derivation of eq. (12) We consider the case $N_i < 0$ and $N_i > 0$ separately. We use the uniform distribution of $A_i^* \in [0, 1]$. We assume that N_i is independent of A_i or A_i^*. This assumption is simplifying and may be used to estimate the average over all agents, not for individual agents i. Decision alignment $\delta_i = 1$, if $A_i^* > \frac{1}{2}$ and $A_i > \frac{1}{2}$ or $A_i^* < \frac{1}{2}$ and $A_i < \frac{1}{2}$.

If $N_i < 0$, the equilibrium is given by

$$A_i = A_i^* + \frac{1-\gamma_i}{\gamma_i}.$$

For $A_i^* > \frac{1}{2}$, $A_i > \frac{1}{2}$ if $A_i^* + \frac{1-\gamma_i}{\gamma_i} > \frac{1}{2}$, which holds since $\frac{1-\gamma_i}{\gamma_i} \geq 0$. For $A_i^* < \frac{1}{2}$, $A_i < \frac{1}{2}$ if $A_i^* + \frac{1-\gamma_i}{\gamma_i} < \frac{1}{2}$, which holds if $A_i^* - \frac{1}{2} < -\frac{1-\gamma_i}{\gamma_i}$. Since $A_i^* < \frac{1}{2}$, this is equivalent to $\mathbb{P}(\mathcal{U}(-0.5, 0) < -\frac{1-\gamma_i}{\gamma_i})$.

If $N_i > 0$, the equilibrium is given by

$$A_i = A_i^* - \frac{1-\gamma_i}{\gamma_i}.$$

For $A_i^* > \frac{1}{2}$, $A_i > \frac{1}{2}$ if $A_i^* + \frac{1-\gamma_i}{\gamma_i} > \frac{1}{2}$, which holds if $A_i^* - \frac{1}{2} > \frac{1-\gamma_i}{\gamma_i}$. Since $A_i^* > \frac{1}{2}$, this is equivalent to $\mathbb{P}(\mathcal{U}(0, 0.5) > \frac{1-\gamma_i}{\gamma_i})$. For $A_i^* < \frac{1}{2}$, $A_i < \frac{1}{2}$ if $A_i^* - \frac{1-\gamma_i}{\gamma_i} < \frac{1}{2}$, which holds since $\frac{1-\gamma_i}{\gamma_i} \geq 0$.

Due to symmetry $\mathbb{P}(A_i^*) < 0.5) = \mathbb{P}(A_i^*) > 0.5) = \frac{1}{2}$ and $\mathbb{P}(\mathcal{U}(0, 0.5) > \frac{1-\gamma_i}{\gamma_i}) = \mathbb{P}(\mathcal{U}(-0.5, 0) < -\frac{1-\gamma_i}{\gamma_i})$, thus eq. (12) follows.

3 Example: Analytical considerations for a two-agent model

For the most general two-agents model with $\gamma_1 \neq \gamma_2$ we analyse conditions for the existence of a disagreeing equilibrium. Assuming the time-scaling factor τ to be 1, the dynamic is described by the map

$$\mathbf{f}(x,y) = \begin{pmatrix} \gamma_1 x + (1-\gamma_1)y + \gamma_1(\hat{x}-x)|y-x| \\ \gamma_2 y + (1-\gamma_2)x + \gamma_2(\hat{y}-y)|y-x| \end{pmatrix} \tag{15}$$

and has the fixed points

$$x = y \qquad \text{and} \qquad x = \hat{x} + \sigma\frac{1-\gamma_1}{\gamma_1}, \quad y = \hat{y} - \sigma\frac{1-\gamma_2}{\gamma_2}\,. \tag{16}$$

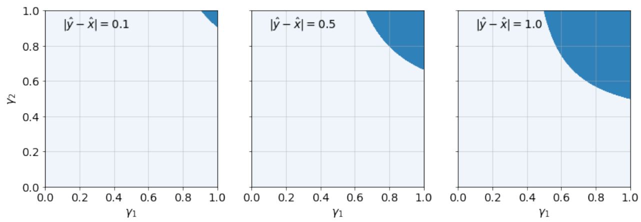
For the sign of the attitude difference $\sigma \in \{-1, 1\}$ we can assume w.l.o.g. $\sigma = 1$ due to symmetry. Thus $2^3 \cdot \frac{1}{2} = 4$ qualitatively different parameter configurations are possible.

Case 1	Case 2	Case 3	Case 4
$\hat{y} > \hat{x}$, $\gamma_2 > \gamma_1$	$\hat{y} > \hat{x}$, $\gamma_2 < \gamma_1$	$\hat{y} < \hat{x}$, $\gamma_2 > \gamma_1$	$\hat{y} < \hat{x}$, $\gamma_2 < \gamma_1$

The second fixed point exists (and is stable as shown in the following) if

$$y > x \quad \leftrightarrow \quad \hat{y} - \frac{1-\gamma_2}{\gamma_2} \geq \hat{x} + \frac{1-\gamma_1}{\gamma_1} \quad \rightarrow \quad 2 + \hat{y} - \hat{x} \geq \frac{\gamma_1 + \gamma_2}{\gamma_1\gamma_2}\,. \tag{17}$$

The area of the parameter space spanned by γ_1 and γ_2 where this condition is fulfilled is shown for different values of the initial convictions difference in Supplementary Fig. 33. For $\hat{x} > \hat{y}$ this is never the case so that case 3 and 4 of the configurations only allow the stable equilibrium $x = y$.



Supplementary Figure 33 Parameter space for γ of the two-agents model. Points inside the dark area meet condition eq. (17) which is required for the existence of a disagreeing equilibrium. Each panel shows the parameter space for a different gap of initial conviction in the system.

The condition for the initial parameters that allow a change of σ ($\hat{y} > \hat{x}$) becomes:

$$(\gamma_2 + \gamma_1 - 1)\,\sigma + \gamma_2\hat{y} - \gamma_1\hat{x} > \gamma_2 y - \gamma_1 x \,. \tag{18}$$

The Jacobian matrix is given by eq. (19).

$$\mathbf{Df}(x,y) = \begin{pmatrix} \gamma_1(1-\hat{x}-y+2x) & 1-\gamma_1(1-\hat{x}+x) \\ 1-\gamma_2(1+\hat{y}-y) & \gamma_2(1+\hat{y}-2y+x) \end{pmatrix} \tag{19}$$

Evaluated at the fixed point defined by $x = y \equiv \eta$ this becomes:

$$\mathbf{Df}(\eta,\eta) = \begin{pmatrix} q_1 & 1-q_1 \\ 1-q_2 & q_2 \end{pmatrix} \quad q_1 \equiv \gamma_1(1-\hat{x}+\eta) \quad q_2 \equiv \gamma_2(1+\hat{y}-\eta)\,. \tag{20}$$

The Eigenvalues are given by

$$\begin{aligned} 0 =& (q_1-\lambda)(q_2-\lambda) - (1-q_1)(1-q_2) & (21)\\ & [\hat{q} \equiv q_1 + q_2] & (22)\\ =& \lambda^2 - \hat{q}\lambda + \hat{q} - 1 \quad \rightarrow \quad \lambda_+ = \hat{q} - 1 \quad \lambda_- = 1\,. & (23) \end{aligned}$$

Since every point along the line $x = y$ is a fixed point of the dynamics λ_- is equal to 1.

$$\mathbf{0} = \begin{pmatrix} (q_1 - \hat{q} + 1)v_+^1 + (1-q_1)v_+^2 \\ (1-q_2)v_+^1 + (q_2 - \hat{q} + 1)v_+^2 \end{pmatrix} \quad \rightarrow \quad \mathbf{v}_+ = \begin{pmatrix} \frac{q_1-1}{1-q_2}\nu \\ \nu \end{pmatrix} \tag{24}$$

The fixed points are attracting along $\mathbf{v}_+$ if $|\lambda_+| < 1$. If $\gamma_1 > 0 \vee \gamma_2 > 0$ it holds

$$\begin{aligned} & |\gamma_1(1+\eta-\hat{x}) + \gamma_2(1+\hat{y}-\eta) - 1| < 1 \\ \leftrightarrow \quad & 0 < \gamma_1(1+\eta-\hat{x}) + \gamma_2(1+\hat{y}-\eta) < 2. \end{aligned} \tag{25}$$

In general, it is not trivial to determine if the condition is satisfied as it depends strongly on the specific values of the parameters and η. The left side of the inequality is always true for the parameters lying inside the open unit interval since $\min(1-a-b) = \epsilon > 0$ for $a, b \in [0,1]$.

$$\gamma_1(1+\eta-\hat{x}) + \gamma_2(1+\hat{y}-\eta) > \gamma_1\epsilon_1 + \gamma_2\epsilon_2 > 0 \tag{26}$$

For $\hat{x} > \hat{y}$ the right side is also met because

$$\gamma_1(1+\eta-\hat{x}) + \gamma_2(1+\hat{y}-\eta) \leq 2 + \eta - \eta + \hat{y} - \hat{x} < 2\,. \tag{27}$$

It follows that the fixed point is semi-stable. For $\hat{y} > \hat{x}$ it is possible to verify numerically using the appropriate long-term limit for η that the inequality is satisfied if and only if eq. (17) is not satisfied. In fact, it can be shown that the second fixed point given in eq. (16) is stable if it exists. Therefore the Jacobian matrix evaluates to

$$\mathbf{Df}(\eta_1, \eta_2) = \begin{pmatrix} q_1 & q_2 \\ q_3 & q_4 \end{pmatrix} \tag{28}$$

$$\begin{aligned}
q_1 &\equiv \gamma_1(1 - \hat{x} - y + 2x) = \gamma_1(1 - \hat{y} + \hat{x}) + 2\sigma(1 - \gamma_1) + \sigma\gamma_1\frac{1-\gamma_2}{\gamma_2} \\
q_2 &\equiv 1 - \gamma_1(1 - \hat{x} + x) = 1 - \gamma_1 - \sigma(1 - \gamma_1) \\
q_3 &\equiv 1 - \gamma_2(1 + \hat{y} - y) = 1 - \gamma_2 - \sigma(1 - \gamma_2) \\
q_4 &\equiv \gamma(1 + \hat{y} - 2y + x) = \gamma_2(1 - \hat{y} + \hat{x}) + 2\sigma(1 - \gamma) + \sigma\gamma_2\frac{1-\gamma_1}{\gamma_1}
\end{aligned} \tag{29}$$

Using $\sigma = 1$ it follows that $q_2 = 0$ and $q_3 = 0$. The resulting matrix is diagonal with Eigenvalues q_1, q_4. Stability is guaranteed if $|\lambda| < 1$.

$$\begin{aligned}
& 1 > |\lambda| = |q_1| \\
& \Leftrightarrow -1 > \gamma_1(1 - \hat{y} + \hat{x}) - 2\gamma_1 + \gamma_1\frac{1-\gamma_2}{\gamma_2} \\
& \Leftrightarrow 1 + \hat{y} - \hat{x} > \frac{\gamma_2}{\gamma_1\gamma_2} + \gamma_1\frac{1-\gamma_2}{\gamma_1\gamma_2} \\
& \Leftrightarrow 1 + \hat{y} - \hat{x} > \frac{\gamma_1 + \gamma_2 - \gamma_1\gamma_2}{\gamma_1\gamma_2} \\
& \Leftrightarrow 2 + \hat{y} - \hat{x} > \frac{\gamma_1 + \gamma_2}{\gamma_1\gamma_2} \qquad \text{(for } q_2 \text{ analogously)}
\end{aligned} \tag{30}$$

The resulting equation matches the condition for the existence of the fixed point, such that if the fixed point exists, it is stable. Thus the disagreeing equilibrium is stable if it can be reached, which depends on the parameters of the model.