S1 SUPPORTING INFORMATION

S1.1 Isolated subsystem: original TSL model

When $\delta = \lambda = 0$, the original TSL model[24] describes each subsystem. Dropping subsystem labels for brevity, we summarise its features as a benchmark for assessing effects of perturbations due to biophysical and social couplings in the coupled system. TSL calculated biophysical and social fixed points for a range of initial social conditions, $\phi(0)$ and parameters, μ. Setting $\dot{\rho}(\tau) = 0$, (4a) gives biophysical fixed points as a function of total community effort:

$$\rho_0^* = \frac{q}{2d}\left(-E^* R_m + S(E^*)\right), \tag{S1}$$

where $S(E^*) = \sqrt{4cd + E^{*2} R_m^2}$ and asterisks denote fixed-point values. Substitution into (6) gives socially optimal (cooperative) and individually optimal (Nash) effort levels (E_c^*, E_N^*), conditions for which are[39] respectively

$$w = \frac{df}{dE^*} \text{ and} \tag{S2a}$$

$$w = \frac{f}{E^*} - \frac{1}{n}\left(\frac{f}{E^*} - \frac{df}{dE^*}\right). \tag{S2b}$$

It is assumed that $1 < \mu \leq \mu_N$, where $\mu_N = E_N^*/E_c^*$. Using parameters from[24] ($n = 50$, $q = 1$, $c = 50$, $d = 50$, $R_m = 200$, $\alpha = 0.6$, $\beta = 0.2$, $\gamma = 10$) gives $E_c^* = 0.483$, $E_N^* = 1.83$ and corresponding resource fixed points $\rho_{0_c}^* = 0.425$, $\rho_{0_N}^* = 0.134$. Socially optimal and Nash efforts for individual agents are $e_c = E_c^*/n = 9.65 \times 10^{-3}$ and $e_N = E_N^*/n = 3.65 \times 10^{-2}$. The former quantifies the norm of socially optimal harvesting. It is assumed that $e_c < e_d \leq e_N$, so $1 < \mu \leq 3.78$, since $e_N/e_c \equiv \mu_N = 3.78$. Setting $\dot{\phi}(\tau) = 0$ in (9) yields monomorphic social fixed points ($\phi^* = 0, 1$) and mixed social fixed points on loci described by $\omega(\phi^*) = \pi_d(e_c, \mu, \phi^*, \rho_0^*)$.

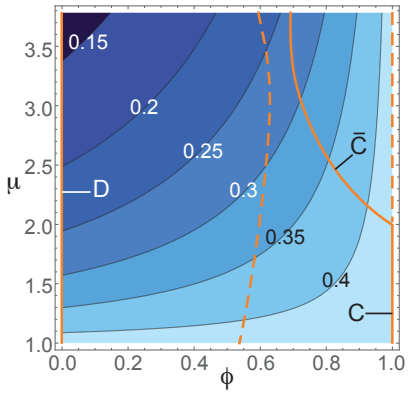

Figure S1. Social (ϕ^* – orange) and biophysical (ρ^* – blue) fixed points of an isolated TSL subsystem, parameterised by cooperative population fraction ($\phi(\tau)$) and effort factor (μ). Dashed and solid orange loci respectively comprise unstable social fixed points and stable social fixed points (equilibria). The equilibria are labelled by type: D- purely defective, C- purely cooperative, C̄- mixed, cooperator dominated. Parameters: $s = 6.8 \times 10^{-3}$, $T = -150$, $g = -10$. See figure 3 in[24] for more information.

Figure S1 overlays these on corresponding biophysical fixed points (S1). The stable fixed points (equilibria) are attractors toward which the system evolves. The μ parameter space is divided between two qualitatively different types of social stability landscape characterised by different combinations of attractors. When $\mu \gtrsim 2.0$, there are two attractors, one purely defective ($\phi^* = 0$) and one mixed and cooperator-dominated ($0.5 < \phi^* < 1$). We label these attractor types, respectively, D and C̄. When $\mu \lesssim 2.0$, D and a purely cooperative attractor (C; $\phi^* = 1$) are present. As each combination of attractors persists over a region of the parameter space, they may be thought of as different phases of the system's behaviour, analogous to familiar examples of phases from thermodynamics, in which substances such as water behave differently in the phases, solid, liquid and gas across different ranges of temperature and pressure.

S1.2 Assessment of perturbative approximation

We approximate the biophysical dynamics in the coupled subsystem model (2) with a perturbation expansion (3), truncated at first order. Here we assess this approximation by comparing it with the full numerical solution of the mean-field replicator dynamics (2). Figure S2 compares both the transient dynamics and equilibria for a range of coupling strengths. In each case, calculations were run until the system converged sufficiently close to equilibrium that the following condition was met:

$$|\dot{\rho}^{(1)}(\tau)|+|\dot{\rho}^{(2)}(\tau)|+|\dot{\phi}^{(1)}(\tau)|+|\dot{\phi}^{(2)}(\tau)|<\Gamma, \tag{S3}$$

where Γ is a tolerance which we set to $\Gamma = 10^{-5}$, since this gave a good balance between accuracy and practicality.

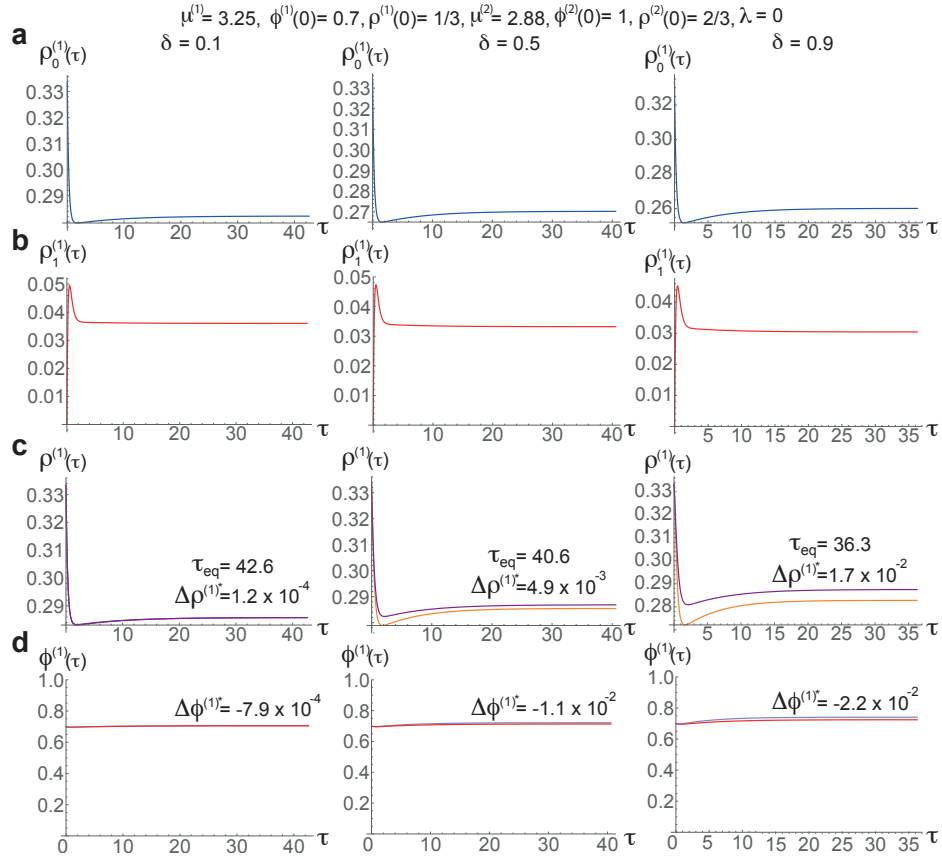

Figure S2. Comparison of biophysical and social dynamics of subsystem 1, in the presence of purely biophysical subsystem coupling, between full numerical solution and perturbative approximation of the biophysical dynamics. Cases of very weak (**left**), weak (**middle**) and strong (**right**) coupling are shown. **a**: zeroth-order resource dynamics, $\rho_0^{(1)}(\tau)$; **b**: first-order resource perturbation dynamics, $\rho_1^{(1)}(\tau)$; **c**: comparison of perturbative approximation $\left(\rho^{(1)}(\tau) \approx \rho_0^{(1)}(\tau) + \delta\rho_1^{(1)}(\tau),\ \text{purple}\right)$ with full numerical resource dynamics (orange); **d**: comparison of cooperative population fraction dynamics between full numerical solution (blue) and solution using perturbative approximation for the biophysical dynamics (red). Parameters and initial conditions shown at top.

We label the lowest time at which (S3) is satisfied τ_{eq}, and define the relative error in the resource equilibrium obtained from the perturbative approximation,

$$\Delta\rho^{(1)*} = \frac{\rho_0^{(1)}(\tau_{eq}) + \delta\rho_1^{(1)}(\tau_{eq}) - \rho^{(1)}(\tau_{eq})}{\rho^{(1)}(\tau_{eq})}, \tag{S4}$$

where $\rho^{(1)}(\tau_{eq})$ is the numerical solution of the full resource dynamics (2). Similarly, we define the relative error

in the equilibrium cooperative fraction of the subsystem 1 population,

$$\Delta\phi^{(1)^*} = \frac{\phi_p^{(1)}(\tau_{eq}) - \phi^{(1)}(\tau_{eq})}{\phi^{(1)}(\tau_{eq})}, \tag{S5}$$

where $\phi_p^{(1)}(\tau_{eq})$ is the cooperative fraction obtained when the perturbative approximation is used for the resource dynamics and $\phi^{(1)}(\tau_{eq})$ is the full numerical solution. These relative errors are assessed for each case in figure S2. The results show excellent agreement between the perturbative and full solutions in the cases of very weak ($\Delta\rho^{(1)^*} = 1.2 \times 10^{-4}$, $\Delta\phi^{(1)^*} = -7.9 \times 10^{-4}$) and weak coupling ($\Delta\rho^{(1)^*} = 4.9 \times 10^{-3}$, $\Delta\phi^{(1)^*} = -1.1 \times 10^{-2}$). Surprisingly, even though the strong coupling case (panel c) violates the weak coupling assumption underpinning our perturbative treatment ($0 < \delta \ll 1$), the relative errors in equilibrium resource and cooperative population levels are both only $\sim 2\%$ in this case. Although this varies across the parameter space, we found comparably small errors across a wide range of parameter values, suggesting that the perturbative treatment is more robust under strong biophysical coupling than might be expected. We leave elucidation of the reasons for this unexpected robustness to future work.

S1.3 Norm scenarios in coupled system

We study two scenarios in which harvesting norms respectively are and are not adjusted for biophysical coupling between subsystems. In each subsystem, the effort level satisfying (S2a) when the resource is at equilibrium quantifies the socially optimal harvesting norm. To compute subsystem 1's resource equilibria, we assume that both subsystems' zeroth- and first-order resource dynamics have equilibrated. Setting $\dot{\rho_1}^{(1)}(\tau) = \dot{\rho_1}^{(2)}(\tau) = 0$ yields

$$\rho_1^{(1)^*} = \frac{d(\rho_0^{(2)^*} - \rho_0^{(1)^*})}{R_m E^{(1)^*} + 2d\rho_0^{(1)^*}} \tag{S6a}$$

$$= \frac{R_m(E^{(1)^*} - E^{(2)^*}) - S(E^{(1)^*}) + S(E^{(2)^*})}{2S(E^{(1)^*})}, \tag{S6b}$$

When multiplied by δ, this gives the shift in subsystem 1's resource equilibrium due to transfer between subsystems (relative to $\rho_0^{(1)^*}$). Substituting (S6b) and (S1) into (3) gives

$$\begin{aligned}\rho^{(1)^*} &\approx \rho_0^{(1)^*} + \delta\rho_1^{(1)^*} \\ &\approx \frac{1}{2S(E^{(1)^*})d}\left(4cd + R_m E^{(1)^*}(R_m E^{(1)^*} - S(E^{(1)^*})) + \delta d(R_m(E^{(1)^*} - E^{(2)^*}) - S(E^{(1)^*}) + S(E^{(2)^*}))\right).\end{aligned} \tag{S7}$$

Empirical studies have shown that communication leading to norm-based collaboration between common pool resource harvesters can strongly improve resource sustainability compared with cases of no communication[6]. We consider two scenarios for norm setting, called 'uncollaborative' and 'collaborative'. In the former, the communities' weak social relationship prevents them from sharing information about their local resource pools and collaboratively managing the resource based on complete information. Accordingly, the communities set norms as if their subsystems were isolated ($E_c^* \equiv E_{c0}^* = 0.483$ and $E_N^* \equiv E_{N0}^* = 1.83$; $\mu_N = E_{N0}^*/E_{c0}^* \equiv \mu_{N0} = 3.78$, as calculated in section S1.1). In the collaborative scenario, communities proactively overcome their weak social relationship to communicate and set norms using complete resource information from both subsystems. Substituting (S7) into (6), and the result into (S2a) and (S2b), gives the shifted E_c^* and E_N^* values. Numerically solving for these values across the weak coupling range ($0 \leq \delta \leq 0.5$) yields distributions that are very well approximated by linear fits from least-squares regression (Pearson correlation, $r > 0.999$ for both):

$$\begin{aligned}E_c^*(\delta) &= E_{c0}^* + 7.70 \times 10^{-2}\delta \\ \mu_N(\delta) &= \mu_{N0} - 5.56 \times 10^{-1}\delta.\end{aligned} \tag{S8}$$

In the collaborative scenario, these shifts must be included in both $\rho_0^{(i)}$ and $\rho_1^{(i)}$ to calculate $\rho^{(i)}$.

S1.4 Supplementary results

Here we present detailed results to explain the deviations in the subsystem 1 fixed points under different cases of coupling (fig's 3, 4a), compared with an isolated subsystem (fig. S1). Evaluating equation (S6b) determines the shifts in biophysical equilibria across the parameter space (fig. S3). As equation (S6a) makes clear, $\rho_1^{(1)^*}$ varies proportionally to the difference in zeroth-order resource equilibria; when $\rho_0^{(1)^*}$ is higher than $\rho_0^{(2)^*}$, the first-order dynamics transfer resource from subsystem 1 to 2, and conversely.

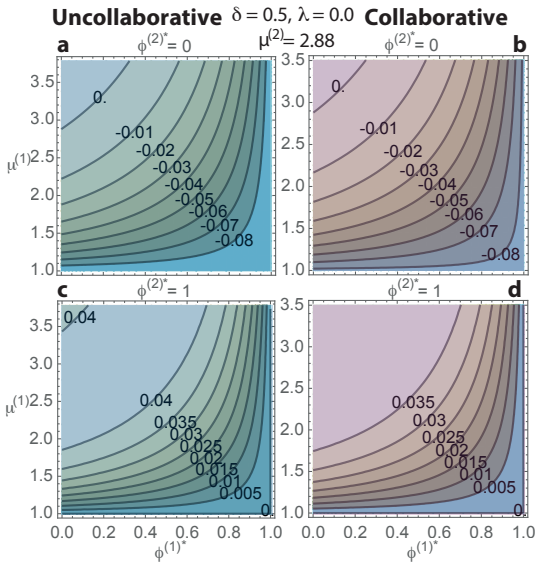

Figure S3. Fixed points of the first-order biophysical dynamics in subsystem 1 ($\rho_1^{(1)^*}$), in the uncollaborative and collaborative norm scenarios. Per equation (3), when multiplied by δ these give the biophysical coupling-induced shifts in the subsystem 1 fixed points compared with an isolated subsystem (see fig. S1). Results are shown for cases in which subsystem 2 equilibrates at each of its two monomorphic social fixed points. Values in panels a and b are negative because in those cases the equilibria correspond to ongoing transfer of resource from subsystem 1 to 2. The reverse is true for panels c and d.

Figure S4 plots the sanctioning, $\omega^{(1)}(\phi^{(1)^*})$ and payoff, $\pi_d^{(1)}(e_c, \mu^{(1)}, \mu^{(2)}, \phi^{(1)^*}, \phi^{(2)^*})$ experienced by a subsystem 1 defector when coupling is purely biophysical. The payoff differs between the uncollaborative and collaborative scenarios, and is shown for both. Curve intersections correspond to mixed social fixed points because $\dot{\phi}^{(1)}(\tau) = 0$ when $\omega^{(1)}(\tau) = \pi_d{}^{(1)}(\tau)$ in (9); the sanctioning and payoff experienced by defectors balance under this condition. These fixed points are stable equilibria (dynamical attractors) when they satisfy the condition, $\frac{\partial \pi_d^{(1)}}{\partial \phi^{(1)}} > \frac{\partial \omega^{(1)}}{\partial \phi^{(1)}}$. When defectors harvest with effort only modestly higher than cooperators (e.g. 30% higher: $\mu^{(1)} = 0.30(\mu_N - 1) + 1 = 1.83$; panels c and d) there is only one mixed fixed point (unstable) and it is not significantly affected by either subsystem 2 cooperativity ($\phi^{(2)^*} = 0, 1$) or norm scenario. However, when defectors harvest with effort closer to the Nash level (e.g. $\mu^{(1)} = 0.7(\mu_N - 1) + 1 = 2.95$), both of these quantities strongly affect subsystem 1 behaviour. There are two mixed fixed points (one unstable and one $\bar{\mathrm{C}}$ attractor) in the collaborative scenario under both subsystem 2 cooperativity conditions, and also in the uncollaborative scenario when subsystem 2 is purely defective ($\phi^{(2)^*} = 0$). However, when $\phi^{(2)^*} = 1$ (panel b), subsystem 1 has no mixed fixed points because the defector payoff, $\pi_d^{(1)}$, is higher than the sanctioning, $\omega^{(1)}$ for all values of $\phi^{(1)}$. Intuitively, this is because biophysical coupling provides subsystem 1 agents access to additional resource from subsystem 2, which increases the defector payoff, while in the absence of social coupling, subsystem 2 cannot provide any extra sanctioning to balance this increased payoff for subsystem 1 defectors.

Figure S5 plots $\omega^{(1)}(\phi^{(1)^*}, \phi^{(2)^*})$ and $\pi_d^{(1)}(e_c, \mu^{(1)}, \phi^{(1)^*})$ in the case of purely social coupling. When subsystem 2 equilibrates at $\phi^{(2)^*} = 0$ (panels a, c), the sanctioning felt by defectors in subsystem 1 is significant only when $\phi^{(1)^*}$ reaches the threshold (~ 0.35) determined by the parameterisation of (8) (fig. 2). Subsystem 1 behaviour then shows no phase transition in stability landscape relative to the isolated case; when defectors harvest only modestly harder

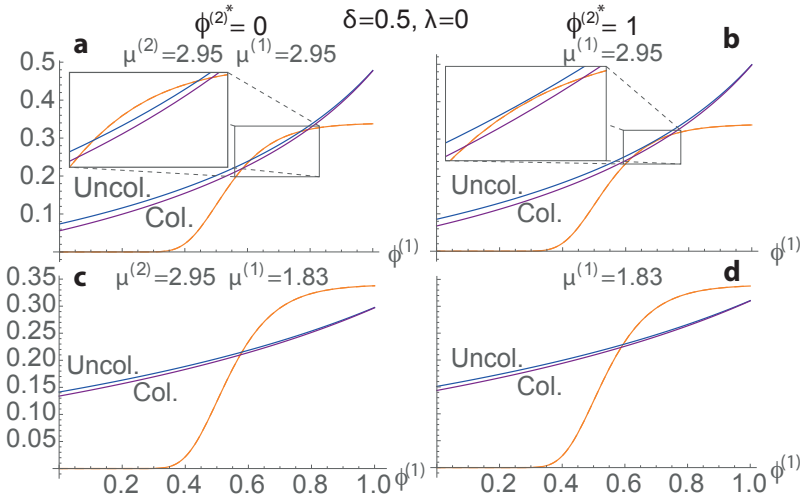

Figure S4. Determination of mixed social fixed points in subsystem 1 in the case of purely biophysical subsystem coupling, in the uncollaborative and collaborative norm scenarios. Fixed points are determined by intersections between the defector payoff (blue, purple) and sanctioning (orange). Results shown for both monomorphic social equilibria in subsystem 2. In the uncollaborative scenario in b, there is no intersection, resulting in the collapse of cooperation.

than cooperators (e.g. $\mu^{(1)} = 1.83$) there is a single mixed (unstable) fixed point (panel c), and when defectors harvest with effort closer to the Nash level ($\mu^{(1)} = 2.95$), a $\bar{\text{C}}$ attractor is also present (panel a). However, when subsystem 2 equilibrates at its purely cooperative fixed point ($\phi^{(2)^*} = 1$; panels b, d), subsystem 1 defectors experience nonzero sanctioning even when the cooperative population fraction in subsystem 1 is below its threshold (~ 0.35), due to the influence of subsystem 2 cooperators mediated by the social coupling. In this case, when $\mu^{(1)} = 2.95$ (panel b), there are no mixed fixed points and only the C attractor is present, so collapse of defection is inevitable. When $\mu^{(1)} = 1.83$ (panel d), however, two mixed social fixed points occur; an unstable one and the stable $\bar{\text{D}}$ attractor (defined in section 3.2), which is not seen in the case of purely biophysical coupling.

Figure S6 plots $\omega^{(1)}(\phi^{(1)^*}, \phi^{(2)^*})$ and $\pi_d^{(1)}(e_c, \mu^{(1)}, \mu^{(2)}, \phi^{(1)^*}, \phi^{(2)^*})$ when biophysical and social couplings are both moderate ($\delta = \lambda = 0.5$). The results can largely be understood as a simple combination of the effects seen above, in figures S4 and S5. Of particular note, however, is the prevention of collapse of cooperation in the uncollaborative scenario, previously seen for purely biophysical coupling (fig. S4b). Here (fig. S6b) it is instead replaced by collapse of defection, which also occurs in the collaborative scenario. This happens because the added sanctioning due to social coupling overwhelms the added defector payoff due to extra resource availability from biophysical coupling.

S1.5 Mechanisms underlying stability landscapes

Figure 4b presents social stability landscape phase plots for subsystem 1, showing up to five different phases across the weak coupling parameter space. In this section, we briefly summarise mechanisms underlying the behaviour of each phase. In the purely defective D phase, additional sanctioning due to very weak social coupling is inadequate to balance the extra defector payoff due to added resource availability from moderate biophysical coupling. Defector payoff exceeds sanctioning for all subsystem 1 population compositions, making collapse of cooperation inevitable. In the $\text{D}\bar{\text{C}}$ phase, added sanctioning from subsystem 2 enables the sanctioning experienced by subsystem 1 defectors to balance the defector payoff but only when sanctioning from subsystem 1 is also high, which occurs only with a cooperative majority. This phase includes the isolated subsystem case ($\delta = \lambda = 0$) and region of what may be considered very weak couplings ($\delta \lesssim 0.1$, $\lambda \lesssim 0.1$). In the $\bar{\text{D}}\bar{\text{C}}$ phase, added sanctioning from subsystem 2 is sufficient to balance the defector payoff even when subsystem 1 has a defector majority, giving rise to the mixed, defector-dominated attractor. The mixed, cooperator-dominated attractor persists in this phase because the defector sanctioning and payoff can also balance when subsystem 1 has a cooperator majority (e.g. fig. S6a, b). In the $\bar{\text{D}}\text{C}$ phase, defector sanctioning and payoff balance only when both are low because subsystem 1 is dominated by defectors. When it is instead dominated by cooperators, sanctioning exceeds defector payoff due to the moderate social coupling. Finally, in the purely cooperative C phase, sanctioning exceeds defector payoff for all subsystem 1 population compositions, necessitating collapse of defection. In most regions of the parameter space,

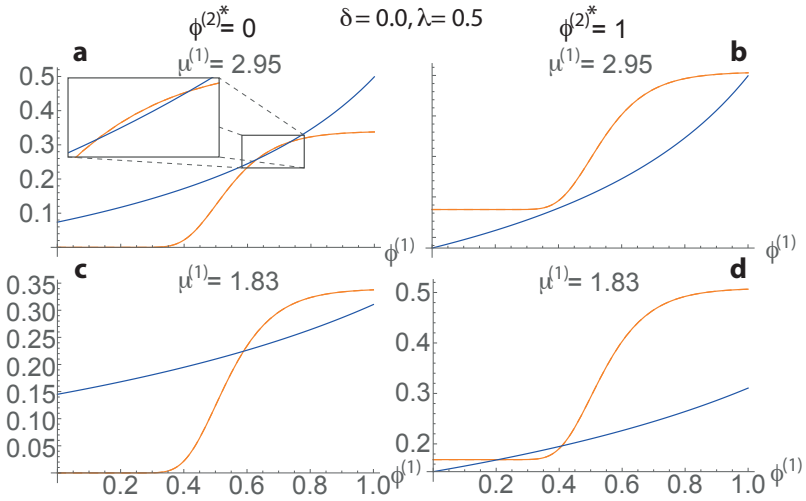

Figure S5. Determination of mixed social fixed points in subsystem 1 in the case of purely social subsystem coupling, in the uncollaborative and collaborative norm scenarios (which are identical in this case). Results shown for both monomorphic social equilibria in subsystem 2. Sanctioning is nonzero even for small $\phi^{(1)}$ in panels b and d because $\phi^{(2)^*} = 1$ and the social coupling conveys sanctioning to subsystem 1 even when its own population is largely or completely defective. This enables the appearance of a fixed point in d, not seen in the absence of social coupling. This is a stable, mixed fixed point with a large defector majority.

increasing the social coupling, even within its weak range, causes successive phase transitions in the subsystem 1 stability landscape. In general, this gradually increases its cooperativity, though not always monotonically. Note that the different phases' cooperativities overlap and are not strictly ordered, since the mixed attractors are defined by ranges of $\phi^{(1)}$ rather than specific values.

S1.6 Subsystem parameter sensitivity

In this section we explore the robustness of our main findings to variations in the subsystem parameters, $\mu^{(1)}$, $\mu^{(2)}$, n, s, g, and T. Figures S8–S13 show social stability landscape plots for subsystem 1 under different parameter combinations. In each figure, we hold most parameters at the values used in the main text (adopted from the original TSL model[24]), while varying others to study how they affect the phase plot. In all cases, we vary the effort factor, $\mu^{(1)}(\psi) = \psi(\mu_N - 1) + 1$, choosing low, medium and high values corresponding respectively to $\psi = 0.3, 0.5, 0.7$. Figure S8 shows the effect of varying $\mu^{(1)}$ alone. Figures S9–S13 respectively show effects of varying $\mu^{(2)}$, n, s, g and T in combination with $\mu^{(1)}$. For each parameter combination, we present phase plots for both the uncollaborative and collaborative scenarios, and also plot $\omega^{(1)}(\phi^{(1)^*}, \phi^{(2)^*})$ and $\pi_d^{(1)}(e_c, \mu^{(1)}, \mu^{(2)}, \phi^{(1)^*}, \phi^{(2)^*})$ for biophysical and social couplings both at the centres of their weak ranges ($\delta = \lambda = 0.25$). This point is chosen arbitrarily as an example of how the given parameter combination affects how $\omega^{(1)}$ and $\pi_d^{(1)}$ depend on $\phi^{(1)}$, since this determines their intersections. These correspond to mixed social fixed points, which are stable equilibria (dynamical attractors) when $\frac{\partial \pi_d^{(1)}}{\partial \phi^{(1)}} > \frac{\partial \omega^{(1)}}{\partial \phi^{(1)}}$.

Overall, the results reveal a rich diversity of model behaviour across the parameter subspace analysed. While we do not give a detailed interpretation of results specific to every parameter combination, overall we observed that model behaviour is strongly sensitive to $\phi^{(2)}$, $\mu^{(1)}$, n and s, moderately sensitive to g and less sensitive to $\mu^{(2)}$ and T. We also note that multiple features seen in figure 4b are consistent across the larger parameter subspace analysed here. First, stability landscape depends much more strongly on social coupling than biophysical coupling; whereas varying the biophysical coupling over a constant social coupling causes at most one phase transition, varying social coupling over a constant biophysical coupling causes up to four phase transitions. Second, in general, stronger social coupling is associated with more cooperative stability landscapes. Third, whereas the phase boundary gradients over the biophysical coupling are always non-negative in the uncollaborative scenario, they are always non-positive in the collaborative scenario. This means that collaborative norm setting consistently increases the power of social sanctioning as a resource-sustaining mechanism, by allowing it to cause phase transitions in the subsystem stability landscape at lower levels of social coupling for a given biophysical coupling. The consistency of

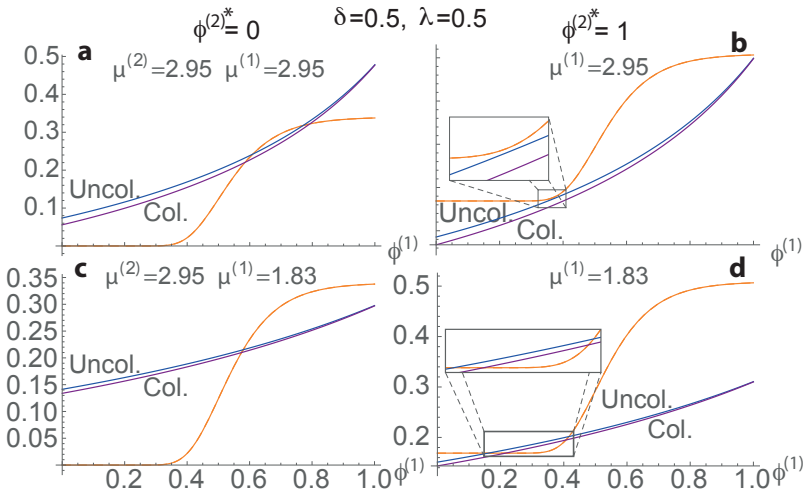

Figure S6. Determination of mixed social fixed points in subsystem 1 in the case of combined biophysical and social subsystem coupling, in the uncollaborative and collaborative norm scenarios. Results shown for both monomorphic social equilibria in subsystem 2. The norm scenarios differ only through small shifts in the fixed points and have no qualitative differences in the cases shown.

these trends across the parameter subspace analysed here suggests that our findings described in the main text, based on a narrower range of parameters, hold more generally. We invite interested readers to further explore the parameter space for themselves, using the Supplementary Code supplied in a Mathematica notebook.

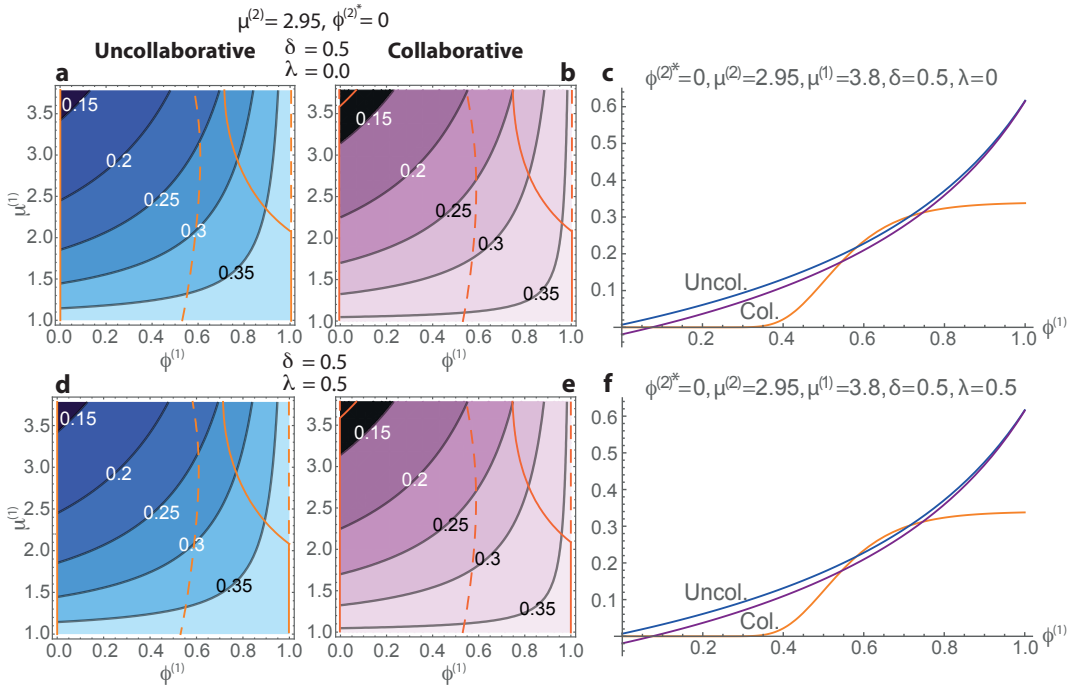

Figure S7. Shifted biophysical and social fixed points in subsystem 1 under purely biophysical subsystem coupling (panels a,b,c) and mixed social and biophysical coupling (d,e,f) when subsystem 2 equilibrates at its purely defective fixed point ($\phi^{(2)^*} = 0$). Results shown for uncollaborative (a,c,d,f) and collaborative (b,c,e,f) norm scenarios and moderate coupling ($\delta = \lambda = 0.5$). Dashed and solid orange loci respectively comprise unstable and stable social fixed points. The results show only small deviations and no qualitative change in the stability landscape, compared with an isolated subsystem (fig. S1), in both scenarios, except in the small region where $\mu^{(1)} \gtrsim 3.6$ in the collaborative scenario. Here, the $\bar{\mathrm{D}}$ attractor appears over a range of small $\phi^{(1)}$ values, in addition to the $\bar{\mathrm{C}}$ attractor. This is because, in this region, the resource level drops sufficiently low that the cost of production exceeds the production, so the defector payoff is negative (${\pi_d}^{(1)} < 0$, per equation (7)). This is shown in panels c and f, where the blue and purple curves represent the defector payoff in the uncollaborative and collaborative scenarios respectively, and orange curves represent the sanctioning. For $\phi^{(1)^*} \lesssim 0.075$, ${\pi_d}^{(1)} < 0$ in the collaborative scenario. The $\bar{\mathrm{D}}$ attractor then appears where the cost of production and production balance (${\pi_d}^{(1)} \approx 0$), since $\omega^{(1)} \approx 0$ for such small values of $\phi^{(1)}$. The effect of subsystem 2 long-time behaviour ($\phi^{(2)^*} = 0$ vs. $\phi^{(2)^*} = 1$) on subsystem 1 long-time behaviour under these couplings can be seen by comparing panels a,b of this figure with a-i and a-ii in figure 3, and panels d,e here with a-i and a-ii in figure 4.

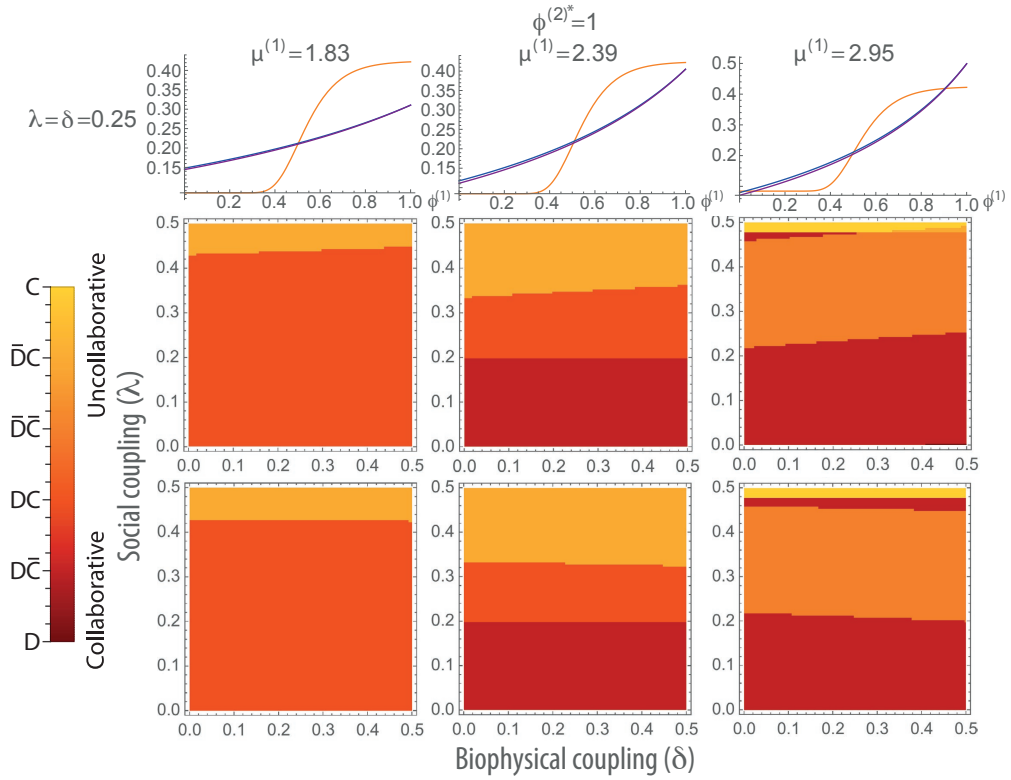

Figure S8. Dependence of subsystem 1 social stability landscape plot on effort factor, $\mu^{(1)}$ when $\phi^{(2)^*} = 1$. Results shown for low (left column), medium (centre) and high (right) $\mu^{(1)}$. **Top:** Defector sanctioning ($\omega^{(1)}$, orange) and payoff ($\pi_d^{(1)}$) in the uncollaborative (blue) and collaborative (purple) scenarios when $\delta = \lambda = 0.25$. **Middle and bottom:** Phase plots for the uncollaborative and collaborative scenarios over the space of weak social and biophysical couplings. Phase plots in the right column are identical to those in figure 4, though calculated at 5x lower resolution for practicality.

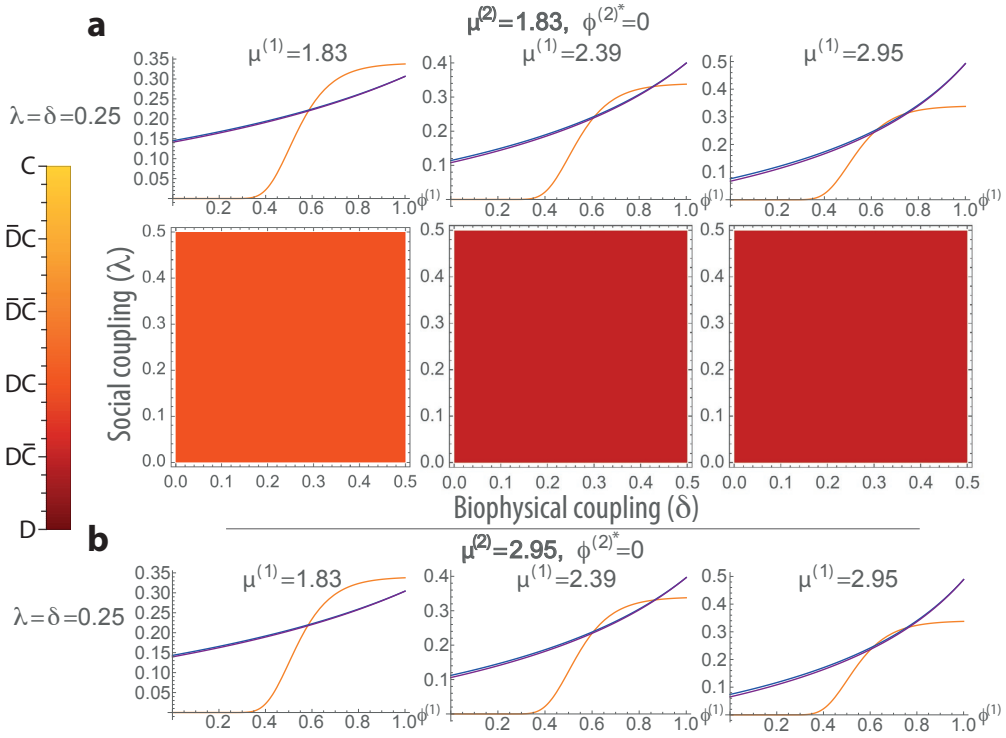

Figure S9. Dependence of subsystem 1 social stability landscape plot on effort factors, $\mu^{(2)}$ and $\mu^{(1)}$ when $\phi^{(2)*} = 0$. (Results are independent of $\mu^{(2)}$ when $\phi^{(2)*} = 1$). Results shown for low (left column), medium (centre) and high (right) $\mu^{(1)}$, and low and high $\mu^{(2)}$ (labelled). **a (Low $\mu^{(2)}$): Top-** Defector sanctioning ($\omega^{(1)}$, orange) and payoff ($\pi_d^{(1)}$) in the uncollaborative (blue) and collaborative (purple) scenarios when $\delta = \lambda = 0.25$. **Bottom:** Phase plots for the uncollaborative and collaborative scenarios over the space of weak social and biophysical subsystem couplings. stability landscape is independent of both couplings for all parameter combinations. However, some dependence on $\mu^{(1)}$ is observed. **b (High $\mu^{(2)}$):** Defector sanctioning ($\omega^{(1)}$, orange) and payoff ($\pi_d^{(1)}$) in the uncollaborative (blue) and collaborative (purple) scenarios when $\delta = \lambda = 0.25$. Phase plots for this condition are identical to those in (a), so were omitted for brevity. Overall, the results reveal that subsystem 1's social stability landscape depends only very weakly on (though, strictly, is not independent of) $\mu^{(2)}$.

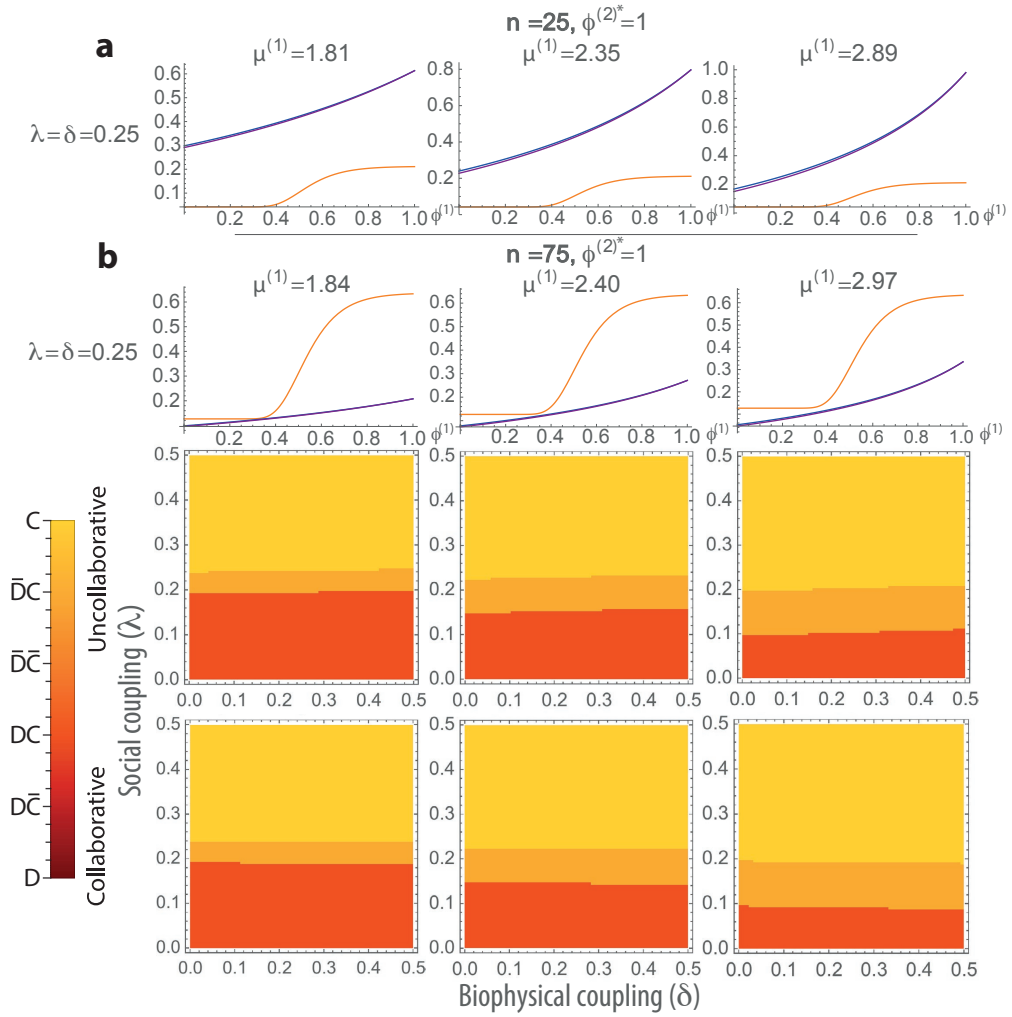

Figure S10. Dependence of subsystem 1 social stability landscape plot on subsystem population, n and effort factor, $\mu^{(1)}$ when $\phi^{(2)^*} = 1$. Results shown for low (left column), medium (centre) and high (right) $\mu^{(1)}$, and low and high n (labelled). The latter were chosen arbitrarily to be 50% lower and higher, respectively, than the TSL n value. **a (Low n):** Defector sanctioning ($\omega^{(1)}$, orange) and payoff ($\pi_d^{(1)}$) in the uncollaborative (blue) and collaborative (purple) scenarios when $\delta = \lambda = 0.25$. Due to the defector payoff exceeding the sanctioning by a large margin, the stability landscape was found to be purely defective under all couplings and parameter combinations for this n condition. Phase portraits were accordingly omitted for brevity. **b (High n): Top-** Defector sanctioning ($\omega^{(1)}$, orange) and payoff ($\pi_d^{(1)}$) in the uncollaborative (blue) and collaborative (purple) scenarios when $\delta = \lambda = 0.25$. **Middle and bottom:** Phase plots for the uncollaborative and collaborative scenarios over the space of weak social and biophysical subsystem couplings. The results reveal that subsystem 1's social stability landscape depends strongly on n.

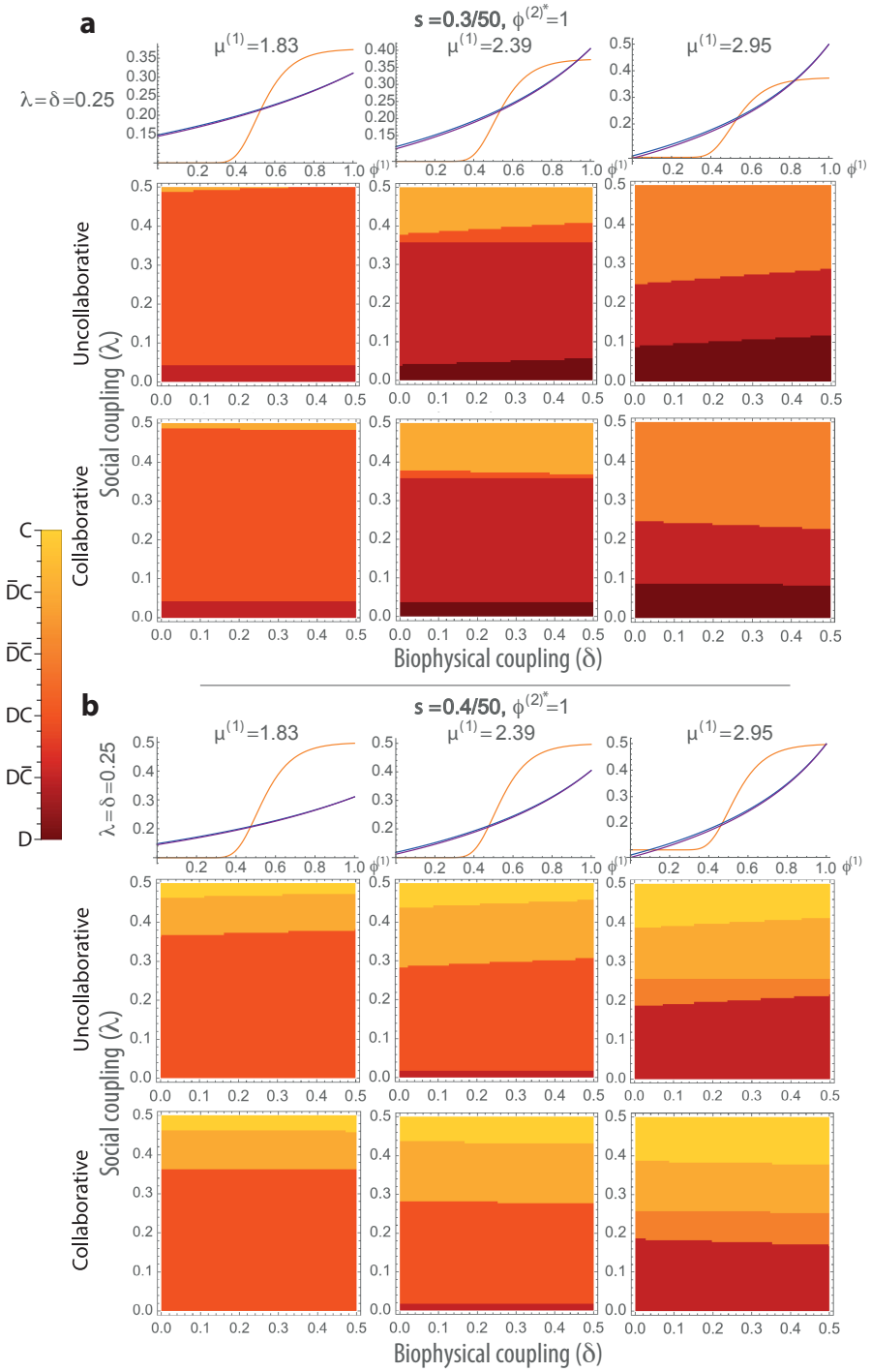

Figure S11. Dependence of subsystem 1 social stability landscape plot on the maximal individual sanctioning capacity, s and effort factor, $\mu^{(1)}$ when $\phi^{(2)^*} = 1$. Results shown for low (left column), medium (centre) and high (right) $\mu^{(1)}$, and low and high s (labelled). Prior analysis in[24] showed that the original TSL model is sensitive to maximal sanctioning capacity, so we chose low and high s values which vary from the TSL value by less than 20%. **a (Low s): Top-** Defector sanctioning ($\omega^{(1)}$, orange) and payoff ($\pi_d^{(1)}$) in the uncollaborative (blue) and collaborative (purple) scenarios when $\delta = \lambda = 0.25$. **Bottom:** Phase plots for the uncollaborative and collaborative scenarios over the space of weak social and biophysical subsystem couplings. **b (High s): Top-** Defector sanctioning ($\omega^{(1)}$, orange) and payoff ($\pi_d^{(1)}$) in the uncollaborative (blue) and collaborative (purple) scenarios when $\delta = \lambda = 0.25$. **Middle and bottom:** Phase plots for the uncollaborative and collaborative scenarios over the space of weak social and biophysical subsystem couplings. The results reveal that subsystem 1's social stability landscape depends strongly on s.

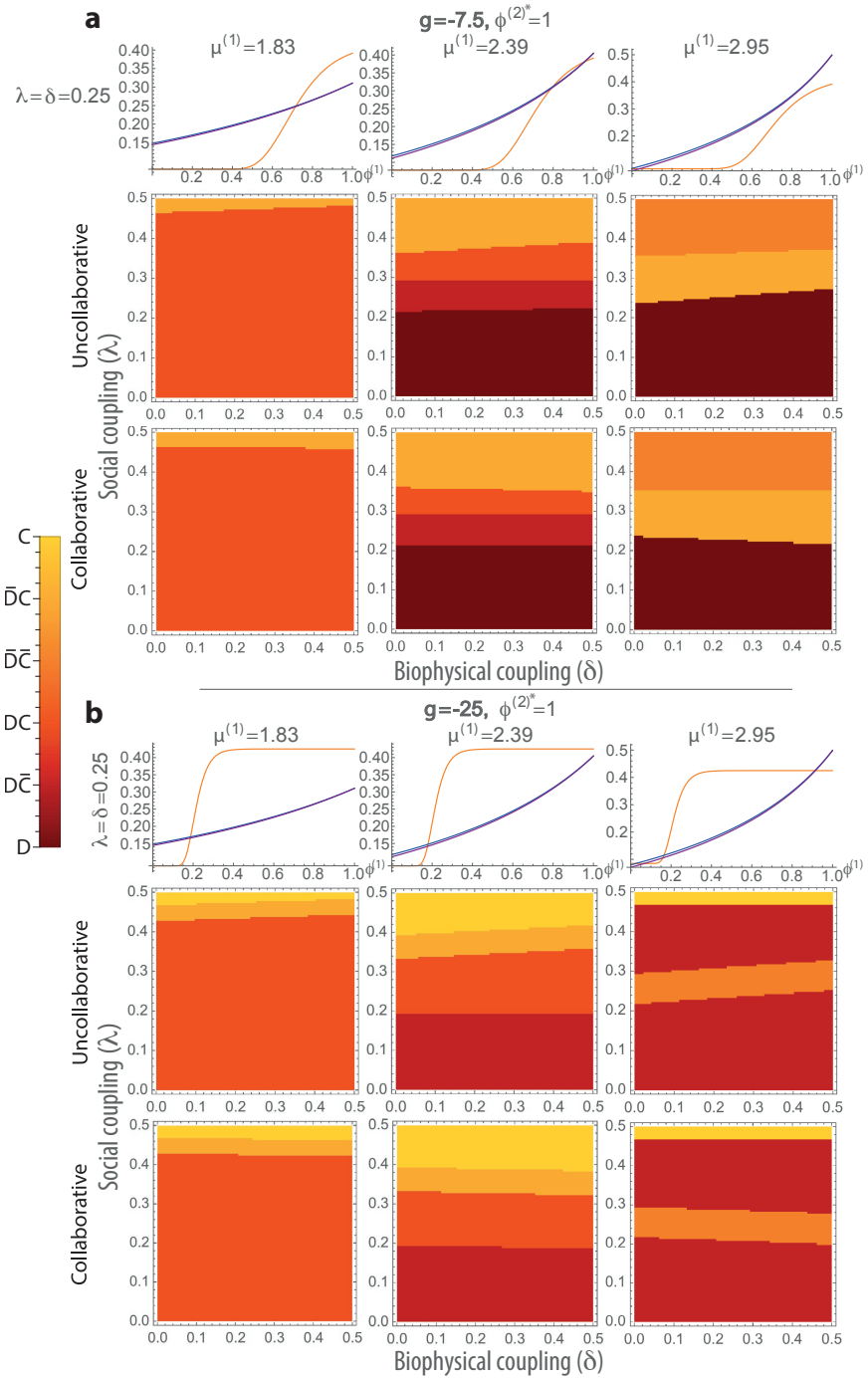

Figure S12. Dependence of subsystem 1 social stability landscape plot on the sanctioning curvature parameter, g and effort factor, $\mu^{(1)}$ when $\phi^{(2)^*} = 1$. Results shown for low (left column), medium (centre) and high (right) $\mu^{(1)}$, and low and high g (labelled). The latter values were chosen as the most extreme which were deemed (by eye) to adequately preserve a sigmoidal shape in $\omega^{(1)}$; these varied from the TSL value by 25% and 150% respectively. **a (Low g): Top-** Defector sanctioning ($\omega^{(1)}$, orange) and payoff ($\pi_d^{(1)}$) in the uncollaborative (blue) and collaborative (purple) scenarios when $\delta = \lambda = 0.25$. **Bottom:** Phase plots for the uncollaborative and collaborative scenarios over the space of weak social and biophysical subsystem couplings. **b (High g): Top-** Defector sanctioning ($\omega^{(1)}$, orange) and payoff ($\pi_d^{(1)}$) in the uncollaborative (blue) and collaborative (purple) scenarios when $\delta = \lambda = 0.25$. **Middle and bottom:** Phase plots for the uncollaborative and collaborative scenarios over the space of weak social and biophysical subsystem couplings. The results reveal that subsystem 1's social stability landscape depends strongly on g over the tested range.

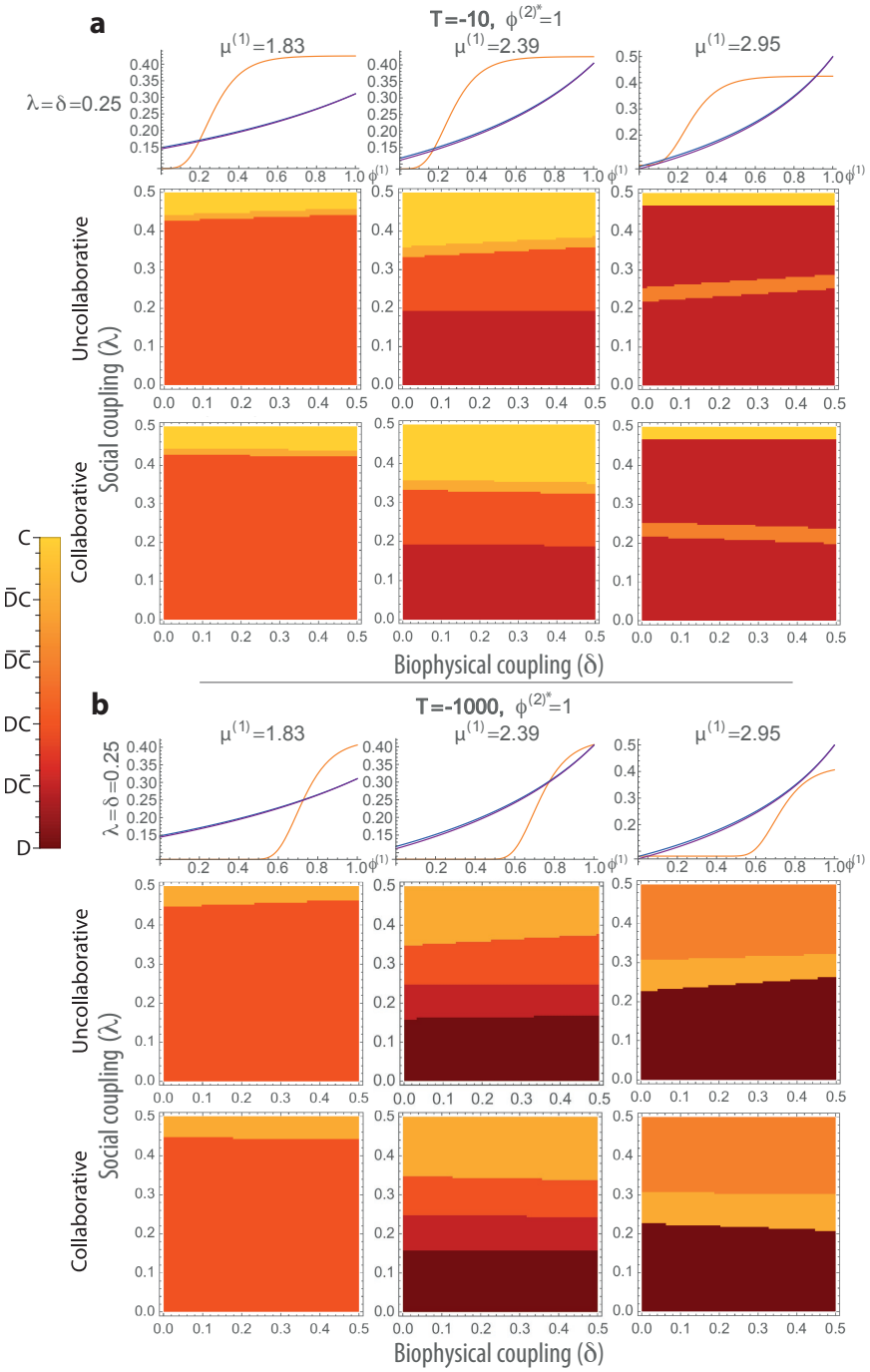

Figure S13. Dependence of subsystem 1 social stability landscape plot on the sanctioning curvature parameter, T and effort factor, $\mu^{(1)}$ when $\phi^{(2)*} = 1$. Results shown for low (left column), medium (centre) and high (right) $\mu^{(1)}$, and low and high T (labelled). The latter values were chosen as the most extreme which were deemed (by eye) to adequately preserve a sigmoidal shape in $\omega^{(1)}$; these varied from the TSL value by 1400% and 10000% respectively. **a (Low T): Top-** Defector sanctioning ($\omega^{(1)}$, orange) and payoff ($\pi_d^{(1)}$) in the uncollaborative (blue) and collaborative (purple) scenarios when $\delta = \lambda = 0.25$. **Bottom:** Phase plots for the uncollaborative and collaborative scenarios over the space of weak social and biophysical subsystem couplings. **b (High T): Top-** Defector sanctioning ($\omega^{(1)}$, orange) and payoff ($\pi_d^{(1)}$) in the uncollaborative (blue) and collaborative (purple) scenarios when $\delta = \lambda = 0.25$. **Middle and bottom:** Phase plots for the uncollaborative and collaborative scenarios over the space of weak social and biophysical subsystem couplings. The results reveal that subsystem 1's social stability landscape depends strongly on T over the tested range, although this range is very large when compared with the other parameters tested, indicating lower sensitivity to T.