Contagious cooperation, temptation, and ecosystem collapse*

Andries Richter[1,2**], Daan van Soest[3,4], Johan Grasman[1]

[1] Department of Mathematical and Statistical Methods, Wageningen University, the Netherlands

[2] Centre for Ecological and Evolutionary Synthesis (CEES), University of Oslo, Norway

[3] Department of Spatial Economics and IVM, VU University Amsterdam, the Netherlands

[4] Department of Economics, Tilburg University, the Netherlands

Draft version, March 20, 2012

Abstract

Real world observations suggest that social norms of cooperation can be effective in overcoming social dilemmas such as the joint management of a common pool resource – but also that they can be subject to slow erosion and sudden collapse. We show that these patterns of erosion and collapse emerge endogenously in a model of a closed community harvesting a renewable natural resource in which individual agents face the temptation to overexploit the resource, while a cooperative harvesting norm spreads through the community via interpersonal relations. We analyze under what circumstances small changes in key parameters (including the size of the community, and the rate of technological progress) trigger catastrophic transitions from relatively high levels of cooperation to widespread norm violation – causing the social-ecological system to collapse.

Keywords: social norms, common pool resource, co-evolution, resilience, multiple equilibria.
JEL codes: C73, D62, D64, Q20.

* The authors acknowledge the Netherlands Organisation for Scientific Research for their financial support as part of the program on Evolution and Behavior. A.R. also acknowledges the European Commission through the Marie Curie Programme (PIEF-GA-2010-274356).

** Corresponding author: a.p.richter@bio.uio.no. Centre for Ecological and Evolutionary Synthesis (CEES), University of Oslo, P.O. Box 1066 Blindern, N-0316 Oslo, Norway.

1 Introduction

The history of mankind is one of gradual change in environmental quality and natural resource abundance, punctuated with sudden collapses of populations, species, ecosystems, and sometimes even of entire civilizations [1,2]. The most common example is the collapse of the human population on Easter Island following the depletion of forest resources [2,3]. To explain patterns of gradual change and sudden collapse the literature has focused on the existence of non-linear relationships in the dynamics of renewable natural resources. Examples of natural systems characterized by non-linearities are those that feature a minimum population size below which extinction is inevitable [4,5], but also those with complex interactions between the various components of the ecological system as is the case in, for example, shallow lakes and semi-arid ecosystems [6-8]. The strong non-linearities in the regeneration functions typically give rise to the prediction that continued overharvesting of the resource results in a gradual demise of the resource until a threshold – or tipping point – is reached, beyond which collapse is inevitable and subsequent system restoration is very costly, or even impossible [9].

In this paper we contribute to the literature on tipping points in social-ecological systems by analyzing how social interactions between the users of a natural system affect its resilience. Building on [10-13] we use evolutionary game theory to develop a model in which a finite number of community members have access to a commonly owned renewable resource. As is the case in the real world, we assume that the common property regime is such that community members are allowed to harvest the resource, but that they are not allowed to hire non-community members to engage in resource harvesting too if their own time constraint is binding [14]. Next, natural regeneration is captured by a standard logistic growth function, and community members can decide to act cooperatively by limiting their extraction, or not. Agents

are tempted to act non-cooperatively (also referred to as defecting) because of the higher associated profits, but we also allow for the possibility that whenever a cooperator and a defector meet, the cooperator may convince the defector of the social desirability of acting cooperatively. The diffusion of social norms regarding harvesting is thus assumed to take place via interpersonal relations, with cooperation being "contagious" [15-18]; see [19] for empirical evidence in the context of renewable resource use. This modeling approach is consistent with the experimental evidence that verbal expressions of discontent can induce and sustain cooperation in social dilemma situations [20], but the mechanism can also be interpreted as reflecting peer-to-peer sanctions or rewards [21-23].

Our paper generates tipping points without explicitly introducing (strong) non-linearities in the dynamics of either the ecological system or the social-economic system. The resource's logistic growth function implies that the percentage rate of resource regeneration increases linearly with resource scarcity, and the social-economic system is self-stabilizing as well. If, for whatever reason, the number of cooperators increases, the social pressure on defectors rises, but the benefits of defecting are larger too. Despite this apparent stability of its two components, the system can still generate positive feedbacks between them, giving rise to multiple equilibria. For some range of parameter values the "good equilibrium" can be very resilient to exogenous shocks or external developments (such as population growth or technological progress in harvesting), while the same exogenous changes cause the social-ecological system to collapse if the parameters are close enough to a critical threshold. The positive feedbacks (giving rise to tipping points) emerge because the property rights regime implies that each community member's harvesting time endowment is finite. If an exogenous shock causes a decline in the resource stock, the socially optimal individual harvesting effort level decreases. Cooperation thus

requires agents to decrease their effort levels, and hence the temptation to defect increases with resource scarcity. As a result, more cooperators decide to defect, putting even more pressure on the resource stock. This leads to a spiral of depletion and defection, and eventually, the system flips to the "bad" equilibrium. The societal consequences of such a flip can be substantial because the system exhibits hysteresis. Upon system collapse, moving back to the "good equilibrium" can be difficult and costly – if it is feasible at all.

We thus show that collapse can be caused by interpersonal interactions and economic constraints, rather than by the presence of inherently non-linear functional forms. In that sense, our model is related to models that generates tipping points in a general equilibrium framework because of interactions between economic sectors, with increased harvesting in the resource sector imposing a negative externality on another sector, resulting in even more intensive resource harvesting [24,25]. Our focus on the social dynamics at the community level is especially relevant because of the role of social norms in community governance of common pool resources such as fish, forests, or grazing lands [26-28]. Our paper identifies a mechanism why community resource management can be successful in some situations and not in others, and is even relevant for resources whose regeneration functions are not characterized by strong non-linearities. As such, the mechanism may have been one of the factors that contributed to social-ecological collapses in the past [29,30]. But the insights obtained by this paper may also be relevant for today's policy makers. If centralized enforcement is cheap and effective, community resource management is inefficient. But if the monitoring and policing costs of formal regulation are high (for example when it regards resources that are geographically remote), community management may be more efficient as long as the community's support for the social harvesting norms is sufficiently large, and this paper provides insights into the "early

warning signals" for the government to start intervening to prevent collapse. In that sense the paper also complements the literature in which a formal regulator aims to enforce property rights [31,32].

Our paper is, however, not the first in noting that coupled social-ecological systems can be inherently complex [33-35]. Iwasa et al. [36] analyze a system in which agents are more inclined to undertake pollution-mitigating activities when the environmental quality is poor, and also when social pressure is high. In their model, alternative stable states occur when social pressure increases strongly with the fraction of cooperators in the community. This framework has been extended to incorporate non-linear resource dynamics as well, leading to even richer dynamics [37]. Taylor [25] develops a minimum viable population model in which resource extraction has a negative effect on the profitability of a competing sector, rendering extraction even more attractive. Our paper is complementary to this research in that we do not use any functional forms that, by themselves, give rise to tipping points; in our model collapse can occur because of personal interactions, and the fact that individuals' time endowments are not infinite.

The setup of the paper is as follows. In section 2 we present the model, focusing on the mechanisms driving changes in the size of the resource stock and on those affecting the number of cooperating individuals in the community. The analysis is fairly complex, and hence we present the intuition behind the underlying mechanism in section 3, providing the proofs as well as a numerical robustness analysis in section 4. Section 5 concludes.

2 The model

We assume that there are $N > 1$ agents in a community who have access to a commonly-owned natural resource. The right to extract is exclusively associated with community membership;

community members are not allowed to employ outsiders to assist in harvesting [14]. The size of the resource stock at time *t* is denoted by $X(t)$. Each agent is endowed with a fixed effort level $\hat{e}$ which she can allocate to harvesting the common pool resource, or to an alternative economic activity. The amount of effort agent *i* (*i* = 1…*N*) allocates to resource harvesting at time *t* is denoted by $e_i(t)$. Assuming that the return to effort in the alternative economic activity is constant and equal to *w*, the income agent *i* derives from this activity at time *t* is equal to $w(\hat{e}-e_i(t))$, where $0 \le e_i(t) \le \hat{e}$.

The relationship between an individual agent's harvesting effort $e_i(t)$ and the quantity of resource goods harvested $h_i(t)$ is given by the Schaefer production function, $h_i(t) = qX(t)e_i(t)$, where q is a technology parameter. Assuming logistic growth, denoting the intrinsic growth rate by *r* and rescaling resource units such that the carrying capacity is equal to unity, net natural growth of the resource is equal to $G(X(t)) = rX(t)(1-X(t))$, and hence resource growth is given by

$$\mathrm{d}X(t)/\mathrm{d}t = rX(t)(1-X(t)) - qX(t)\sum_{i=1}^{N} e_i(t). \quad (1)$$

Regarding harvesting revenues, we assume that resource goods can be sold at a time-invariant unit price *P* so that agent *i*'s sales revenues are $Ph_i(t) = PqX(t)e_i(t)$. Harvesting gives rise to an intertemporal negative externality as excessive extraction today reduces the size of the available resource stock tomorrow. This intertemporal consequence of today's harvesting is sometimes referred to as the "Class I problem" [38]. "Class II problem" then refers to the problem caused by instantaneous externalities, where an agent's return to her harvesting effort in a specific period negatively depends on the total effort put in by the *N* – 1 agents in the community in that same period – think of congestion or crowding. In this paper, we do not just account for the

intertemporal externality; we also introduce an instantaneous one (see below). We do so for two reasons. First, while the intertemporal externality may be economically more severe in the real world than the instantaneous ones, the latter tend to be important too [39-41]. Second, analyzing the consequences of an instantaneous externality on (myopic) agents' propensity to cooperate, is much less complicated than in case of forward-looking agents trying to solve the intertemporal externality, while the underlying mechanism that is explored here is essentially the same.

We follow [42-44] by modeling the instantaneous negative externality as a cost component in the profit function. While gross harvesting income is equal to $PqX(t)e_i(t)$, we assume that there are congestion costs that are dependent on the total amount of effort ($E(t)$) the community puts into resource harvesting, where $E(t) \equiv \sum_{j=1}^{N} e_j(t)$. These congestion costs materialize because agents have to search longer for good spots, spend more money on fuel and transportation, or because fishing nets may interfere and tear, etc. [42]. We thus assume that if an agent puts in effort $e_i(t)$, she incurs congestion costs equal to $vE(t)e_i(t)$,where v reflects the marginal costs of congestion associated with a one unit increase in aggregate effort.[1] Hence, total income earned by agent i at time t is:

(2) $\pi_i(t) = PqX(t)e_i(t) + w\left(\hat{e} - e_i(t)\right) - vE(t)e_i(t)$.[2]

[1] Note that the crowding costs are assumed independent of the wage rate, implying that we focus on crowding resulting in increased expenditures on for example fuel or nets. We do so because in section 4.1 setting the wage rate equal to zero considerably facilitates the analytical solutions. If congestion costs were assumed to exclusively consist of forgone income from outside employment, there would be no externality in section 4.1. For consistency we thus need to assume that there are cost components other than foregone income as well. As the wage rate is assumed to be exogenous, making the congestion costs dependent on the wage rate increases complexity without yielding additional insights. Therefore we chose to capture the congestion costs by a specific constant, v.

[2] We model crowding as a cost component rather than via the production function (with crowding reducing harvesting productivity; cf. [41,45]) to keep the model analytically tractable. The latter modeling approach would imply that the harvesting production function equals $h_i(t) = \left(qX(t) - v'E(t)\right)e_i(t)$, and hence the dynamics of the

Using δ to denote the agents' discount rate, social welfare maximization requires

$$(3) \quad \max_E \int_0^\infty e^{-\delta t}\left[PqX(t)E(t)+w\left(N\hat{e}-E(t)\right)-vE^2(t)\right]\mathrm{d}t$$

subject to $0 \le e_i(t) \le \hat{e}$ for all *i*, as well as (1). Hence, (3) allows us to capture both the intertemporal and the instantaneous externalities.[3] When we focus on just the Class II problem (the instantaneous one), $\delta \to \infty$ (because agents are very impatient, or simply because they are unaware of the intertemporal externality); the case of Class I but no Class II problem can be analyzed by setting $\infty > \delta \ge 0$ and $v = 0$.

We assume that agents can choose between two types of behavior: to act cooperatively, or to "defect". All agents are aware of the social benefits of internalizing the (instantaneous and/or the intertemporal) negative externalities, and some of them decide to act cooperatively. Those who do, are assumed to put in their fair share (i.e., 1/*N*) of the socially optimal aggregate harvesting effort (that is, the one that solves (3) given the current size of *X*). Others, however, decide to act non-cooperatively because of the higher profits associated with defection. Each agent that defects is assumed to choose the effort level that maximizes his private welfare level given the aggregate amount of effort put in by the $N-1$ other agents.[4] Using superscripts *C* and

---

resource stock (see (1)) would be specified as $\mathrm{d}X(t)/\mathrm{d}t = rX(t)\left(1-X(t)\right)-\left(qX(t)-v'E(t)\right)E(t)$. The severity of the crowding externality would then not just affect profits but it would also shift the nullcline of the resource. This would complicate the analytical solution substantially without yielding any new insights, and hence we decided to model the instantaneous externality via crowding costs rather than via decreased harvesting productivity. What is essential for our model is that agents are tempted to defect because of differences in profits between acting cooperatively and selfishly, and this is the case if one agent's effort decreases the returns other agents receive on their harvesting effort, but also if it increases the per-unit harvesting costs of the other agents.

[3] From here onwards we omit time arguments, unless omitting them may cause confusion.

[4] Most evolutionary game theory models on cooperation and defection assume that effort levels chosen only depend on the behavioral mode chosen (cooperation, or defection), but not on the size of the resource stock [10-13]. In other words, effort is either "high" or "low", depending on whether an individual is a cooperator or not. In our model, the

D to respectively denote cooperators and defectors, the above assumptions imply that $\pi^D \geq \pi^C$. The prospect of having higher payoffs is what tempts agents to start acting selfishly, and we assume that agents are more likely to defect the larger is π^D as compared to π^C. More specifically we assume that the fraction of cooperators that decide to defect at time *t* is equal to $\frac{\mathrm{d}C/\mathrm{d}t}{C} = -\beta\left(1-\frac{\pi^C(X)}{\pi^D(X,C)}\right)$, where $C(t)$ denotes the number of cooperators at time *t*, and β is the percentage decrease in $C(t)$ associated with a one unit decrease in π^C/π^D.[5]

Next, we assume that whenever a cooperator meets a defector, there is a probability μ that the former succeeds in convincing the latter to act cooperatively. Assuming that social encounters occur randomly, the probability of a cooperator meeting a defector can be modeled as a Poisson process. Using $D(t) \equiv N - C(t)$ to denote the current number of defectors in the community, the probability of an encounter taking place in time interval $(t, t+\Delta t)$ is equal to $\lambda C(t)D(t)\Delta t / N$, where λ is the Poisson parameter. Social pressure thus increases the number of cooperators by $C(t+\Delta t) - C(t) = \alpha C(t)D(t)\Delta t / N$, where $\alpha \equiv \lambda\mu$. Using the continuous-time equivalent and combining the effects of social pressure and temptation, we have:

allocation of effort is endogenous. This assumption is not just realistic, it also is crucial for explaining under what circumstances a community is able to maintain cooperation, and when cooperation collapses.

[5] Agents are not divisible, and denoting the change in cooperators by d*C*/d*t* implies a slight abuse of notation as this suggests that we do not explicitly constrain the change to be an integer number. Similarly, the equilibrium number of cooperators is not exactly at the point where d*C*/d*t* = 0 (cf. equation (A1) in the Appendix), but the first integer number above the *C* that solves d*C*/d*t* = 0. Explicitly acknowledging indivisibilities complicates the notation without affecting the essence of the results as long as the number of agents is sufficiently large (as assumed in this paper).

$$(4) \quad \mathrm{d}C/\mathrm{d}t = \frac{\alpha}{N} C(N-C) - \beta C\left(1 - \frac{\pi_C(X)}{\pi_D(X,C)}\right).^{6}$$

Equation (4) thus captures what we label "contagious cooperation with the temptation to defect", and relies essentially on three assumptions. First, some agents are willing to uphold a social extraction norm (doing what is optimal for the group as a whole), and try to impose social pressure on non-cooperators to also start adhering to the norm [20,46,47]. Second, the propensity to (dis)obey a cooperative norm depends on the temptation to defect, but also on whether individuals have recently been exposed to cooperatively-minded agents. There is a vast literature on the role of personal encounters in spreading social norms (see for example [48-51]). Face-to-face communication is found to be very effective in inducing cooperation in laboratory experiments, and much more so than alternative modes of communication [46]. Being confronted with other people's behavior conveys information and induces people to update their "best mode of behavior" cf. [18,49]. Indeed, even subtle cues of peer pressure are often enough to induce rule-obeying behavior [48]. The idea of being watched (even if this is induced by mere photographs of human eyes) tends to improve rule-compliance [52]. And if social interactions are repetitive, the resulting behavior can become a social norm [50] which may subsequently be internalized [51]. The mechanisms by which cooperation spreads include moral persuasion, social pressure and feelings of guilt [20,47,53-56].[7] Third, the probability of a cooperator meeting a defector follows a random Poisson process. This last assumption is more likely to be

[6] If $\alpha \leq \beta(N-1)/(2N)$, the system's collapse results in $C = 0$ – see Proposition 1 in section 4.1 and equation (A6) in Appendix A3. If $C = 0$, we have $\mathrm{d}C/\mathrm{d}t = 0$ independent of whatever policy the regulator may want to undertake (see (4)). This is neither plausible nor very interesting, and hence we assume that $\alpha > \beta(N-1)/(2N)$.

[7] Another mechanism often cited are punishments [23,57,58]. Although our model does not cover the option for costly punishment, it is straightforward to see that both gradual change and sudden collapse can be generated by a model with costly punishment. All what we need for these to materialize is a countervailing force for defection.

met in some circumstances (e.g., when defectors can hide their harvests so that cooperators can only indentify defectors "in the field") but not in all – allowing for targeted encounters would be an interesting extension of the model.

3 Gradual changes in cooperation and resource conservation, and sudden collapse

The analysis of why the social-ecological system (1)-(4) is characterized by multiple equilibria is complicated. Because of this, we first provide the intuition behind the mechanism in this section, and present all the proofs and robustness checks in the next.

The mechanism giving rise to positive feedbacks and multiple equilibria is as follows. To maintain cooperation, social pressure should be sufficiently large, and the temptation to defect should be sufficiently small; see (4). The strength of social pressure is a function of the number of cooperators: the larger is C, the larger the pressure on defectors to change their behavior. The temptation to defect is also a function of the number of cooperators. For given X the temptation to defect (weakly) increases with C because of the following. Solving (3), cooperators aim to maximize the (net present value of the) total amount of resource rents accruing to the community. They take into account the (instantaneous and/or intertemporal) negative externalities associated with their harvesting activities, and hence they put in less effort into harvesting than defectors ($e^C \leq e^D$).[8] The larger C, the larger the available amount of resource rents, and hence the more tempting it is to defect and appropriate these rents.

[8] Recall that we assume that all agents know the current resource stock and the socially optimal effort level is public knowledge. Furthermore, defectors can infer the number of cooperators, for example by observing their own net returns to harvesting (from which they can derive the total amount of harvesting effort put in by the rest of the

So, the temptation to defect depends on *C*, but it is a function of the size of the remaining stock as well. Perhaps surprisingly, the smaller *X*, the larger is the temptation to defect. Cooperators take into account the negative harvesting externalities and switch to interior effort levels at an earlier stage of resource depletion than defectors. Thus, e^C / e^D falls with resource depletion, and so does π^C / π^D (because $\partial\pi/\partial e > 0$ unless all rents have been dissipated). Hence, the temptation to defect is larger the larger is *C* and the smaller is *X*.

With these mechanisms in mind, we now address the question why the social-ecological system may experience periods of gradual change, punctuated by sudden collapse. Suppose that, for a given set of parameters, the net social marginal harvesting productivity is larger than the wage rate even if all agents harvest as much as they can ($e_i = \hat{e}$ for all i = 1…*N*). Then it is socially optimal for each agent to put in $\hat{e}$. This would be the case if the community is fairly small (small *N*), if the outside wage rate is not too high (small *w*), if the rate of regeneration is fairly high (*r* large), etc. Let us use $\hat{e}_0(r,P,q,N,v,w,\delta)$ to denote the critical time (or effort) endowment at which the socially optimal individual effort level (and hence e^C) is exactly equal to $\hat{e}$ in steady state.[9] That means that there is no social dilemma as long as $\hat{e} \le \hat{e}_0(\bullet)$: $e^C = e^D = \hat{e}$ so that $\pi^C = \pi^D$ (see (2)), and hence, in steady state, *C* = *N* (cf. (4)).

So how do exogenous developments such as technical progress or population growth affect the critical effort endowment $\hat{e}_0(\bullet)$ for which $e^C = \hat{e}$? An increase in *q* or *N* amplifies the –

community). These assumptions are fairly standard in economics but not necessarily very realistic. However, note that our results hold as long as cooperators choose lower effort levels than defectors. Hence, all we need is that defectors try to appropriate (part of) the resource rents created by other community members acting cooperatively, and this assumption is likely to be met in the real world too.

[9] Hence, $\hat{e}_0$ is a function of all system parameters except α and β. As long as $e^C \ge \hat{e}$ there is no temptation to defect, and hence the probability of a cooperator convincing a defector to become cooperative, is immaterial too.

effective – labor input in harvesting: $\partial \hat{e}_0(\bullet)/\partial q < 0$[10] and $\partial \hat{e}_0(\bullet)/\partial N < 0$. The larger *N* or *q*, the lower the socially optimal individual effort level and hence the smaller the critical effort endowment level for which a social dilemma materializes.

As long as the parameter set is such that $\hat{e} < \hat{e}_0(\bullet)$ we have full cooperation, and increases in *N* or *q* result in a gradual decrease in the size of the resource stock, because (effective) labor input in harvesting increases. If exogenous developments in *N*, *q*, or any other parameter cause $\hat{e}_0(\bullet)$ to fall below $\hat{e}$, the social dilemma emerges and cooperators choose interior effort levels. We then have $e^C < e^D = \hat{e}$, and we also have $\pi^C < \pi^D$ (cf. (2)). As long as $\hat{e} - \hat{e}_0(\bullet)$ is positive but sufficiently close to zero, the decrease in cooperation and resource conservation is small, because temptation to defect is small (as *X* is large) while the probability of a defector meeting a cooperator, is high (as *C* is large).

However, we can identify a tipping point, $\hat{e}_2(r, P, q, N, v, w, \delta, \alpha, \beta)$, where the same gradual changes cause the social-ecological system to collapse. For the system to be in steady state, we need d*X*/d*t* = 0 and d*C*/d*t* = 0 (cf. (1) and (4)). At $\hat{e}_2(\bullet)$, the steady state is stable, but the nullclines of *X* and *C* are just tangent. That means that any change in *N*, *q*, or any other parameter can cause the equilibrium to disappear, triggering a spiral of defection and resource depletion. As before, changes in *N* or *q* cause cooperators to reduce their harvesting effort levels, while defectors continue to allocate their entire time endowment to harvesting. As a result π^C / π^D decreases, some cooperators defect, aggregate harvesting effort increases, and the resource stock is reduced further. This induces cooperators to reduce their harvesting effort even more, thus

[10] At least, for $q > \sqrt{2rv/P}$; see Appendix 3.

resulting in an even stronger decrease in the payoff ratio, and a spiral of defection and resource depletion unfolds. This positive feedback mechanism gives rise to a rapid deterioration of both cooperation and the resource stock, and the negative spiral is stopped only when (almost) all rents have been dissipated. That is, if the average net private return on harvesting effort is equal to the wage rate (possibly zero), such that defectors are indifferent between putting an extra unit of effort into harvesting, or not. Hence, if $\hat{e}_2(\bullet)$ falls below $\hat{e}$, the system moves from an equilibrium with reasonably high levels of cooperation and resource conservation (the "good equilibrium"), to one characterized by little cooperation and near-complete rent dissipation (the "bad equilibrium").[11]

The typical pattern is depicted in Figure 1. On the horizontal axis we plot $\hat{e}/\hat{e}_0(\bullet)$, which portrays the strength of the social dilemma and on the vertical axes *C* and *X* (in Panels A and B, respectively). We plot $\hat{e}/\hat{e}_0(\bullet)$ rather than just $\hat{e}$ to emphasize that a change in *any* parameter can cause the number of equilibria to change – not just changes in $\hat{e}$ itself. The straight lines connect the system's stable steady states that emerge under various parameter constellations, while the dashed lines indicate unstable steady states. There are two branches of stable equilibria (for both *C* and *X*), an upper and a lower branch. An equilibrium located on the lower branch is characterized by very low levels of cooperation and with relatively small resource stocks – all resource rents have been dissipated, and hence we refer to this equilibrium as the "bad

[11] Note that, unlike $\hat{e}_0$, $\hat{e}_2$ is a function of α and β – as well as of all other system parameters $(r, P, q, N, v, w, \delta)$. The larger the steady-state number of cooperators, the larger the defectors' optimal amount of effort, and hence the more likely it is that a given time endowment $\hat{e}$ is binding. Hence, the probability of a defector facing a binding time constraint is smaller the larger is α and the smaller is β (cf. (4) and also (A12) in Appendix A3).

equilibrium". The upper branch connects all the "good equilibria" – those characterized by relatively high levels of *C* and *X*.

<Insert Figure 1 about here>

Because we plot the ratio $\hat{e}/\hat{e}_0(\bullet)$ on the horizontal axis, a move to the right can be interpreted as an intensification of the social dilemma, possibly the result of an exogenous increase in $\hat{e}$, but it may also be the result of an exogenous increase in for example *N* or *q*. If $\hat{e}/\hat{e}_0(\bullet)<1$, the social dilemma is absent and we have $e^C=\hat{e}$ and hence *C* = *N*. Any increases in $\hat{e}$, *N*, *q*, etc. just result in a gradual decrease in the size of the resource stock – the system is always in the good equilibrium; see Panels A and B in Figure 1. If $\hat{e}/\hat{e}_0(\bullet)>1$, a social dilemma materializes, because $e^C<\hat{e}$ becomes socially optimal while it is privately optimal to continue putting in $\hat{e}$ (that is, $e^D=\hat{e}$). Cooperation then decreases as the temptation to defect increases, but the resource stock itself does not fall by much. This is because cooperators compensate for the extra effort put in by the new defectors by choosing lower effort levels themselves, thus limiting the increase in aggregate harvesting effort. As long as $\hat{e}$ is below a second threshold level (or tipping point), $\hat{e}_1(r,P,q,N,v,w,\delta,\alpha,\beta)$, there is just one stable equilibrium, the good one. If $\hat{e}>\hat{e}_2(\bullet)$ there is also just one equilibrium – the bad one. If $\hat{e}_1(\bullet)<\hat{e}<\hat{e}_2(\bullet)$ the system is in either the bad or the good equilibrium, depending on the history of parameter changes – the system is located on the upper (lower) branch if the system approaches $\hat{e}_2(\bullet)$ from below (above).

Starting from a situation in which $\hat{e} < \hat{e}_2(\bullet)$, small changes in *q*, *N* or any other parameter result in small changes in *C* and *X* as the system moves along the upper branch of stable equilibria – until $\hat{e} = \hat{e}_2(\bullet)$. When the system moves beyond $\hat{e} = \hat{e}_2(\bullet)$, the same small changes result in the social-ecological system collapsing to the bad equilibrium. Upon collapse the social-ecological system is in the locally stable bad equilibrium, and the system can only flip back to the good one if effective labor time becomes scarce again (that is, if *q*, *N* etc. fall such that $\hat{e} / \hat{e}_0$ decreases towards unity). For the system to flip back to the good equilibrium on the upper branch it is insufficient to restore $\hat{e} < \hat{e}_2(\bullet)$. Only if $\hat{e}$ falls below the second tipping point, $\hat{e}_1(\bullet)$, defectors are sufficiently constrained in their harvesting efforts that they are unable to appropriate all the extra rents accruing from additional cooperation. Cooperation increases, the resource is exploited less intensively, and the stock recovers. That means that the difference between π^C and π^D decreases, while the subsequent increase in the number of cooperators causes the social pressure on defectors to increase too. As a result, a positive spiral of cooperation and resource restoration pushes the system back to the good equilibrium on the upper branches of Figure 1.

4 Analysis

Having provided the intuition why the system is characterized by a positive feedback, we now proceed as follows. In section 4.1 we analyze the case where (i) agents are assumed to be aware of the instantaneous externality (the Class II problem) but not of the intertemporal one (the Class I problem), and (ii) there is no outside employment opportunity. That means that we assume $v > 0, w = 0$ and $\delta \to \infty$. These assumptions enable us to present the full analytical solution, and the results correspond perfectly with Figure 1 presented above.

The assumptions of agents being myopic and fully dependent on the resource are analytically convenient but maybe not always equally realistic. Therefore, we relax the assumption of no external labor market in section 4.2. Analytically solving the case of $w > 0$ is cumbersome, and hence we rely on numerical methods (together with a robustness analysis testing whether the mechanism is the same for all possible drivers of change – population growth, technical progress, etc.). In section 4.3 we drop the assumption of agents being ignorant of the intertemporal externality (that is, we then assume $\delta \geq 0$).

4.1 Cooperation and collapse when agents are myopic and dependent on the resource

In this subsection we assume that community members are aware of the instantaneous crowding externality (the Class II problem) but that they do not take the intertemporal externality (the Class I problem) into account – because they are not fully informed about the dynamics of resource regeneration, or simply because they are myopic. Letting $\delta \to \infty$ in (3), the relevant benchmark for cooperation is the aggregate effort level $\tilde{E}$ that maximizes the community's instantaneous aggregate income while taking into account the Class II problem:

$$\tilde{E}(X) = \max_E \left\{ PqXE + w\left(N\hat{e} - E\right) - vE^2 \,\middle|\, 0 \leq E \leq N\hat{e} \right\}. \quad (5)$$

Following Bischi et al. [59], we assume that cooperators always put in their fair share of the aggregate effort, $e^C(X) = \tilde{E}(X)/N$. Solving (5) and dividing by N, we have:

$$e^C(X) = \begin{cases} \hat{e} & \text{if} \quad X \geq \left(w + 2vN\hat{e}\right)/(Pq), \\ \dfrac{PXq - w}{2vN} & \text{if} \quad w/(Pq) \leq X < \left(w + 2vN\hat{e}\right)/(Pq), \\ 0 & \text{if} \quad 0 \leq X < w/(Pq). \end{cases} \quad (6)$$

Defectors choose e_i $\left(0 \le e_i \le \hat{e}\right)$ to maximize individual profits, as given in (2). Using $E_{-i} \equiv \sum_{j \neq i} e_j$ to denote the total amount of effort put in by the $N-1$ other agents, a defector's

350 best response (BR) function is

$$(7) \quad e^{\mathrm{BR}}(X, E_{-i}) = \min\left\{\frac{PXq - w}{2v} - \frac{1}{2}E_{-i}, \hat{e}\right\}.$$

Noting that $E_{-i} = Ce^{\mathrm{C}} + \left(D-1\right)e^{\mathrm{BR}}$ and using (6), the equilibrium effort of defectors is:

$$(8) \quad e^{D}(X, C) = \begin{cases} \hat{e} & \text{if} \quad X \ge \dfrac{w}{Pq} + \dfrac{2vN(N-C+1)\hat{e}}{Pq\left(2N-C\right)}, \\ \dfrac{\left(PXq - w\right)\left(2N-C\right)}{2vN(N-C+1)} & \text{if} \quad \dfrac{w}{Pq} \le X < \dfrac{w}{Pq} + \dfrac{2vN(N-C+1)\hat{e}}{Pq\left(2N-C\right)}, \\ 0 & \text{if} \quad 0 \le X < \dfrac{w}{Pq}. \end{cases}$$

Note that the harvesting effort of defectors depends on both *X* and *C*.

355 Having derived the effort levels of cooperators and defectors, we now analyze under what circumstances the social-ecological system is characterized by multiple steady states. To maintain analytical tractability we set $w = 0$ in the rest of this subsection – but see section 4.2 for the case where $w > 0$. Setting $w = 0$ is mathematically convenient because it substantially facilitates the analysis of the social dynamics as embodied in equation (4).[12]

[12] The temptation to defect is a (decreasing) function of π^C / π^D. Substituting e^C and e^D into (2) and taking the ratio, we have $\dfrac{\pi^C}{\pi^D} = \dfrac{\left(PqX - w - vE\right)e^C + w\hat{e}}{\left(PqX - w - vE\right)e^D + w\hat{e}}$, and hence $\pi^C / \pi^D = e^C / e^D$ if $w = 0$, where e^C and e^D are identified in (6) and (8).

In steady state we have dX/dt = 0 and dC/dt = 0 (cf. (1) and (4)), and the (relative) levels of effort chosen by cooperators and defectors crucially affect both the location and slope of these two nullclines. Hence we first state the following Lemma:

Lemma 1: Effort levels of cooperators and defectors in (C,X) space if $w = 0$ and $\delta \to \infty$

In (C,X) space we can identify three regions (denoted R1, R2 and R3) that differ in the effort levels chosen by the cooperators and defectors:

(9.R1) If $X \geq \frac{2\nu N\hat{e}}{Pq}$, we have $e^C(X) = e^D(X,C) = \hat{e}$;

(9.R2) If $\frac{2\nu N(N-C+1)\hat{e}}{Pq(2N-C)} \leq X < \frac{2\nu N\hat{e}}{Pq}$, we have $e^C(X) < e^D(X,C) = \hat{e}$, and $\partial e^C / \partial X > 0$;

(9.R3) If $X < \frac{2\nu N(N-C+1)\hat{e}}{Pq(2N-C)}$, we have $e^C(X) < e^D(X,C) < \hat{e}$, with $\frac{e^C}{e^D} = \frac{N-C+1}{2N-C} \equiv \eta(C)$

and $\partial e^C / \partial X = \eta(C)\, \partial e^D / \partial X > 0$.

Proof: This follows immediately from inserting w = 0 into (6) and (8) and noting that $\eta(C) = \frac{N-C+1}{2N-C} < 1$ for all $N > 1$. ■

In region 1 (R1) the resource is sufficiently abundant so that there is no social dilemma (yet). In R2 defectors still allocate all their available effort to harvesting, but cooperators choose interior effort levels – and the latter type's effort level is smaller, the lower the remaining resource stock. In R3 both types choose interior effort levels (with $e^C < e^D < \hat{e}$). The three regions are crucial when drawing the phase planes of the system; see Figures 2A and 2B. The two boundaries between the three regimes are depicted using thin, uninterrupted lines. The horizontal one is the

boundary between R1 and R2 $\left(X = \frac{2\nu N\hat{e}}{Pq} \right)$, and the downward-sloping concave line is the boundary between R2 and R3 $\left(X = \frac{2\nu N\left(N - C + 1\right)\hat{e}}{Pq\left(2N - C\right)} \right)$.

The nullclines of C and X are also depicted in Figures 2A and 2B, and their locations and slopes are derived in Lemmas 2 and 3 below. Before doing that, two things should be noted about regions R1-R3. First, because defectors choose interior effort levels in R3, we can conclude that (almost) all resource rents are dissipated in this region – if not, it would pay for selfish agents to put in extra effort.[13] This implies that the aggregate payoff of the community decreases when the system moves from R1 to R3: the smaller the resource stock, the lower the aggregate welfare, and the higher the need for cooperation. Second, the payoff ratio π^C / π^D decreases when the system moves from R1 to R3. Second, the payoff ratio π^C / π^D decreases when the system moves from R1 to R3. The instantaneous net marginal benefits of harvesting ($PqX - \nu E$; cf. (2)) are the same for all agents, and hence $\pi^C / \pi^D = e^C / e^D$. From (9.R1)-(9.R3) we infer that π^C / π^D is equal to unity in R1, that it decreases when the stock is being depleted in R2, and that it reaches its minimum (and remains constant) as soon as the system is in R3. Hence, while the need for cooperation increases when the system moves down from R1 via R2 to R3, the temptation to defect increases too. Having identified the three regions, let us now have a closer look at the dC/dt = 0 isocline; see also Figure 2A.

[13] Indeed, for all $w \geq 0$ and using (6) and (8) we have $E \approx \left(PqX - w\right)/\nu$ in R3 because $\left(N - C\right)/\left(N - C + 1\right) \approx 1$. Inserting this into (2) we have $\pi_i = \left(PqX - w - \nu\left(PqX - w\right)/\nu\right)e_i + w\hat{e} = w\hat{e}$ for all i = 1…N.

Lemma 2: The dynamics of cooperation and defection if $w = 0$ and $\delta \to \infty$

Lemma 2.1: In (C,X) space, the nullcline of the number of cooperators, denoted as $C(X)\big|_{\mathrm{d}C/\mathrm{d}t=0}$, consists of three segments:

- In R1, $C(X)\big|_{\mathrm{d}C/\mathrm{d}t=0} = N$,
- In R2, $C(X)\big|_{\mathrm{d}C/\mathrm{d}t=0} = \kappa(X) < N$ with $\mathrm{d}\kappa/\mathrm{d}X > 0$ and $\mathrm{d}^2\kappa/\mathrm{d}X^2 = 0$,
- In R3, $C(X)\big|_{\mathrm{d}C/\mathrm{d}t=0} = \underline{\kappa} << N$.

Lemma 2.2: $\forall\ X \geq 0,\ \dfrac{\mathrm{d}C(X)}{\mathrm{d}\hat{e}}\Big|_{\mathrm{d}C/\mathrm{d}t=0} \leq 0$ with the inequality being strict in R2.

Lemma 2.3: For given X, $\mathrm{d}C/\mathrm{d}t > (<)\ 0$ if $C < (>) C(X)\big|_{\mathrm{d}C/\mathrm{d}t=0}$.

Proof: See appendix A1. ■

<Insert Figure 2 about here>

Lemma 2.1 indicates that the $\mathrm{d}C/\mathrm{d}t = 0$ isocline consists of three segments. In R1 and R3 this nullcline is vertical in (C,X) space, while it is an upward-sloping linear function in R2; see also Figure 2A. The intuition is straightforward. Lemma 1 implies that the temptation to defect in R2 is high when X is small, and hence the equilibrium number of cooperators is smaller the lower is X. The nullcline of cooperation is vertical at $C = N$ in R1 because the temptation to defect is zero, while it is vertical in R3 because here e^C/e^D is a function of C but not of X; see (9.R3).

Next, Lemma 2.2 states that the larger is the effort endowment $\hat{e}$, the more the $\mathrm{d}C/\mathrm{d}t = 0$ isocline is located to the left in (C,X) space in R2. The larger is $\hat{e}$, the less constrained defectors are in their harvesting activities, the larger the temptation to defect and hence the smaller the equilibrium number of cooperators that can be sustained at any X.

Finally, Lemma 2.3 states that the nullcline of C is an attractor. For any given X, the larger is C, the larger the temptation to defect (because E is smaller), and the smaller the number of defectors becoming cooperators (as there are relatively few defectors). Hence, for a given X the strength of social pressure is larger (smaller) than the temptation to defect if C is small (large).

Let us now derive the dX/dt = 0 isocline, which we denote by $X(C)\big|_{\mathrm{d}X/\mathrm{d}t=0}$.

Lemma 3: The dynamics of the resource stock if $w = 0$ and $\delta \to \infty$

Lemma 3.1: If $\hat{e} \le \dfrac{rPq/N}{2rv + Pq^2} \equiv \hat{e}_0$, $X(C)\big|_{\mathrm{d}X/\mathrm{d}t=0} = \bar{\psi} > 2vN\hat{e}/Pq$.

Lemma 3.2: If $\hat{e} > \hat{e}_0$, $X(C)\big|_{\mathrm{d}X/\mathrm{d}t=0}$ does not exist in R1 and hence consists of just two segments:

- In R2, $X(C)\big|_{\mathrm{d}X/\mathrm{d}t=0} = \psi(C) < 2vN\hat{e}/(Pq)$ for all C, with $\mathrm{d}\psi/\mathrm{d}C > 0, \mathrm{d}^2\psi/\mathrm{d}C^2 < 0$,
- In R3, $X(C)\big|_{\mathrm{d}X/\mathrm{d}t=0} \approx \underline{\psi} << 2vN\hat{e}/(Pq)$.

Lemma 3.3: If $\hat{e} > \hat{e}_0$, $\left.\dfrac{\mathrm{d}X(C)}{\mathrm{d}\hat{e}}\right|_{\mathrm{d}X/\mathrm{d}t=0} \le 0$.

Lemma 3.4: For any C, dX/dt > (<) 0 for all $X < (>) X(C)\big|_{\mathrm{d}X/\mathrm{d}t=0}$.

Proof: See appendix A2. ■

Lemma 3.1 states that if $\hat{e} \le \hat{e}_0(\bullet)$ – see section 3 – the steady state must be located in R1: the total amount of effort available $(N\hat{e})$ is too small for the community to be able to draw down the resource stock to a level below the one where harvesting becomes a social dilemma. The case of $\hat{e} > \hat{e}_0(\bullet)$ (as described in Lemma 3.2) is more interesting, and is depicted in Figure 2B. In that case, the equilibria are located in R2 or even R3, and never in R1. That means that the nullcline

of X then consists of two segments, one in each region. In R2 it is upward-sloping (and concave). The larger is C, the lower is E, and hence the larger the resource stock that can be sustained in equilibrium. However, in R3 the X nullcline is (almost) horizontal because here all agents choose interior harvesting effort levels: if one defector decides to start acting cooperatively, the decrease in E is negligible because all other defectors increase their effort levels in response.

Regarding the location of $dX/dt = 0$, Lemma 3.3 states that it is located farther to the South in (C,X) space the larger is $\hat{e}$. The larger the effort endowment, the less agents are constrained in their harvesting, and hence (for every C and keeping everything else constant) the smaller the equilibrium size of the resource stock. And regarding the dynamics of resource regeneration, Lemma 3.4 states that the nullcline of X is an attractor. For any C, the lower is X, the smaller the aggregate quantity harvested (because of lower aggregate effort, and because of lower marginal productivity of effort), and the higher percentage resource growth rate; cf. (1). That means that for given C, regeneration is larger (smaller) than the quantity harvested if X is small (large).

Having derived the shape and location of the two nullclines, we can determine the number of steady states of the system. Visual inspection of Figure 2 suggests that the nullclines can intersect once, twice or three times. Proposition 1 proves the existence of multiple equilibria.

Proposition 1: For any set of parameters (r,P,q,N,v,α,β), three critical effort levels can be identified, $\hat{e}_0(\bullet)$, $\hat{e}_1(\bullet)$, and $\hat{e}_2(\bullet)$ (with $\hat{e}_0(\bullet)<\hat{e}_1(\bullet)<\hat{e}_2(\bullet)$), for which the following holds:

Proposition 1.1: The social-ecological system has just one non-trivial, globally stable steady state (C,X) for each $\hat{e}\notin\left[\hat{e}_1(\bullet),\hat{e}_2(\bullet)\right]$, where

- $(C, X) = \left(N, \left(r - \hat{e}Nq\right)/r\right) \equiv (C_1, X_1)$ if $\hat{e} \leq \hat{e}_0(\bullet)$,
- $(C, X) = (C_2, X_2)$ if $\hat{e}_0(\bullet) < \hat{e} < \hat{e}_1(\bullet)$, where $C_2 < C_1$ and $X_2 < X_1$, and
- $\left(C, X\right) = (C_3, X_3)$ if $\hat{e} > \hat{e}_2(\bullet)$, where $C_3 < C_2$ and $X_3 < X_2$.

Proposition 1.2: The social-ecological system is characterized by three non-trivial steady states (two stable ones and one unstable steady state) if $\hat{e} \in \left\langle \hat{e}_1(\bullet), \hat{e}_2(\bullet) \right\rangle$;

Proof: See appendix A3. ■

<Insert Figure 3 about here>

Depending on the values of $\hat{e}$, r, P, q, N, α, β and v, there may be one or two stable steady states (and zero or one unstable ones). For simplicity, let us vary just $\hat{e}$ and keep all other parameters constant, so that $\hat{e}_0(\bullet)$, $\hat{e}_1(\bullet)$ and $\hat{e}_2(\bullet)$ remain constant too. If $\hat{e} \leq \hat{e}_0(\bullet)$ there is no social dilemma (see Lemma 3.1) so that $C = N$ in steady state, and the resource stock is drawn down to the level where resource regeneration is equal to the maximum quantity the community can harvest.[14]

The analysis is more interesting when the community faces a social dilemma. Figure 3 shows the nullclines for the different qualitative cases of $\hat{e} > \hat{e}_0(\bullet)$. Lemmas 2.2 and 3.3 imply that the higher is $\hat{e}$ (for a given set of parameters), the more the nullcline of the social system (dC/dt = 0) is located to the North-West and the more the nullcline of the resource stock (dX/dt =

[14] Just substitute $\sum_{i=1}^{N} e_i = N\hat{e}$ into (1), set dX/dt = 0, and solve.

0) is located to the South-East. If $\hat{e}_0(\bullet) < \hat{e} < \hat{e}_1(\bullet)$, the nullclines intersect once in R2 giving rise to just one equilibrium (C_2,X_2); see Figure 3A.

An increase in $\hat{e}$ beyond $\hat{e}_1(\bullet)$ causes alternative stable states to emerge in the range $\hat{e}_1(\bullet) < \hat{e} < \hat{e}_2(\bullet)$; see Figure 3B.[15] As stated in Proposition 1.2, there are then three equilibria, two of which are located in R2 (denoted (C_2,X_2) and (C_{2B},X_{2B})), and one in R3, (C_3,X_3). Of these, (C_2,X_2) and C_3,X_3) are locally stable, while (C_{2B},X_{2B}) is unstable. When $\hat{e}$ increases in the range $\hat{e}_1(\bullet) < \hat{e} < \hat{e}_2(\bullet)$, the nullclines shift as indicated by Lemmas 2.2. and 3.3, (C_2,X_2) and (C_{2B},X_{2B}) move towards each other, coincide (when $\hat{e} = \hat{e}_2(\bullet)$), and then disappear (when $\hat{e} > \hat{e}_2(\bullet)$), implying that ($C_3$,$X_3$) is the only remaining equilibrium – as depicted in Figure 3C. [16, 17]

Figure 3 reveals the exact mechanics giving rise to the bifurcation diagrams presented in Figure 1.[18] For $\hat{e} < \hat{e}_0(\bullet)$ there is no social dilemma and just one stable equilibrium, (C_1,X_1), located in R1. The system is located on the upper branches of Panels A and B in Figure 1, with $C_1 = N$. If $\hat{e}_0(\bullet) < \hat{e} < \hat{e}_1(\bullet)$, Figure 3A applies, and there is a unique equilibrium: (C_2,X_2) in the North-East of R2 (implying $C_2 < N$). In Figures 1A and 1B this equilibrium is located on the

[15] In fact, $\hat{e}_1(\bullet)$ is defined as a fold bifurcation at which the "bad equilibrium" (C_3,X_3) is located on the boundary between R2 and R3, where cooperators choose interior effort levels but where the defectors' effort constraint is weakly binding; see equation (A7) in appendix A3.

[16] Here, $\hat{e}_2(\bullet)$ is defined as a fold bifurcation at which two nullclines are tangent in R2 (implying that the cooperators choose interior effort levels whereas the effort constrain of defectors is strictly binding); see equation (A12) in appendix A3.

[17] The relevance of the time constraint (relative to the rest of the system's parameters) is immediately clear from Lemma 1 and Proposition 1. If agents have unlimited amounts of effort at their disposal, the effort levels chosen are, by definition, interior, and then Lemma 1 indicates that the system is always in R3. In Figure 2 R1 and R2 are no longer relevant, and the isoclines in R3 just intersect just once – see also Proposition 1 and Appendix A3. With unlimited time endowments (that is, if agents can hire outside labor), the system is de facto open access, and hence the bad equilibrium is its unique steady state.

[18] Indeed, Figure 1 is the numerical solution to the analytical results obtained in section 4.1 using $\hat{e}$=0.71, N=100, P=50,000, q=0.01, v=1, r = 0.8, α=0.1 and β=0.2.

upper branches. If $\hat{e} > \hat{e}_2(\bullet)$ Figure 3C applies, (C_3,X_3) in R3 is the unique equilibrium (implying that X_3 is very close to the Nash equilibrium steady state stock), and it is located on the lower branches of Panels A and B in Figure 1. And path-dependency emerges in the system because of the fact that there are two stable equilibria in case $\hat{e}_1(\bullet) < \hat{e} < \hat{e}_2(\bullet)$; see Figure 3B. Whether the system is in equilibrium (C_2,X_2) or rather in (C_3,X_3), depends on whether the system approaches the threshold from a situation in which $\hat{e} < \hat{e}_2(\bullet)$, or rather $\hat{e} > \hat{e}_2(\bullet)$. In the first case, the system is in the good equilibrium (C_2,X_2) – on the upper branches in Figure 1 – until it collapses when exogenous changes move the system beyond tipping point $\hat{e}_2(\bullet)$, when the positive feedback identified in section 3 brings the system down to the bad equilibrium (C_3,X_3).[19] Having passed $\hat{e}_2(\bullet)$, reversion of the exogenous changes does not automatically move the system back to the good equilibrium (C_2,X_2) because (C_3,X_3) is stable. The jump back to (C_2,X_2) only occurs if parameters change so much that (C_3,X_3) disappears (which happens when (C_3,X_3) hits the boundary between R2 and R3 – that is, when $\hat{e} = \hat{e}_1(\bullet)$); compare Figures 3C and 3A.

4.2 Collapse triggered by various external changes if $w > 0$ and $\delta \to \infty$

In this subsection we relax the assumption of $w = 0$, and we also explore whether the system's properties are dependent on the underlying causes of change – increases in time endowments, technical progress, population growth, etc. With $w > 0$ we need to resort to a numerical analysis,

[19] Note that because $v > 0$ collapse does not result in the complete depletion of the resource. We have $X_3 > 0$ because all rents have disappeared before the resource is depleted. However, the fall to the bad equilibrium still constitutes a crisis, as defined by Taylor [25]: "a dramatic, unexpected, and [largely] irreversible worsening of the environment leading to significant welfare losses". Even if erosion of social capital does not necessarily lead to complete resource depletion, the welfare consequences can still be dramatic for some or even all stakeholders involved [60]. For cases in which the model does result in complete exhaustion of the stock, see section 4.3.

but we show that the presence of labor markets leads to results that are qualitatively very similar to those obtained in sections 3 and 4.1; see Figure 4.

<Insert Figure 4 about here>

Figures 4A and 4B show the internal equilibria of C and X for different values of the effort endowment $\hat{e}$ if $w > 0$ and $\delta \to \infty$. As stated before, Figure 1 plots the case of $w = 0$ and $\delta \to \infty$, and comparing Figures 1A and 1B to Figures 4A and 4B reveals that with $w > 0$ the social-ecological system behaves in qualitatively the same way as with $w = 0$ as analyzed in sections 3 and 4.1. A positive feedback emerges because a reduction in the size of the resource stock induces cooperators to spend less time harvesting, and the subsequent decrease in the payoff ratio π^C / π^D causes the number of cooperators to decrease. In turn, the social pressure to act cooperatively falls, the number of cooperators falls, and then the resource stock falls even more – triggering even more defection.

The main novel insight obtained from this analysis using $w > 0$ is that cooperation increases if $\hat{e} - \hat{e}_2(\bullet)$ continues to increase after collapse; see Figure 4A. This (small) increase in cooperation materializes because $\lim_{\hat{e} \to \infty} \pi^C / \pi^D = 1$ if $w > 0$. If $\hat{e} > \hat{e}_2$, the social–ecological system is in the bad equilibrium (where all agents choose interior effort levels), and hence increases in $\hat{e}$ or decreases in $\hat{e}_2(\bullet)$ only increase the amount of money earned at the external labor markets, where the same wage rate applies to cooperators and defectors alike. Hence, the larger $\hat{e} - \hat{e}_2(\bullet)$, the larger the income share of wages earned at the external labor market, and hence the closer the payoff ratio π^C / π^D is to unity. That means that the increase in cooperation following environmental collapse should not be interpreted as a sign that the system is moving

back to a better equilibrium. Similarly, if the system has collapsed, the regulator should not be concerned about the fact that policies aimed at reducing $\hat{e}-\hat{e}_2(\bullet)$ below zero (for example by decreasing q or by subsidizing outside employment) actually results in a *decrease* in cooperation – reducing $\hat{e}-\hat{e}_2(\bullet)$ reduces the wage share in total income and hence π^C / π^D increases, so that it becomes more tempting to defect. As was the case in Figure 1, the system only flips back to the good equilibrium if $\hat{e}$ falls below $\hat{e}_1(\bullet)$.

Having established that qualitatively the same patterns emerge for $w > 0$ as for $w = 0$ (with $\delta \to \infty$), we probe further into the robustness of our results and interpretations by numerically solving the system when changing the various key parameters, and then especially q, N, v and α. Figures 4C-4F indicate that increases in the harvesting technology parameter (q) and in the size of the population (N) yield qualitatively similar patterns as when $\hat{e}$ increases – not surprisingly, the only difference is that X continues to fall when q increases (Figure 4D). Next, if the instantaneous externality becomes more severe (that is, if v is larger), the steady-state resource stock tends to be larger (Figure 4H) while the equilibrium number of cooperators tends to be smaller (Figure 4G). The larger is v, the higher the need for cooperation, but also the more costly it is to cooperate. So the increase in X in Figure 4H does not occur because of an increase in C, but in spite of a decrease thereof.[20] The consequences of changes in α, the social pressure parameter, are presented in Figures 4I and 4J. These figures show that for low initial levels of α,

[20] Interestingly, the welfare effects of higher crowding costs v are ambiguous. While an increase in v always decreases welfare in regions 1 and 3, it may increase welfare in region 2. The presence of crowding costs reduces aggregate effort, and hence higher v always attenuate the stock externality (the Class I problem). We find that in region 2 the welfare increases resulting from the reduced intertemporal externality can dominate the welfare costs associated with larger instantaneous crowding costs, but only for intermediate levels of v. Hence, our results are similar to those established empirically by [61] in case of the shrimp fishery in North Carolina.

increases in the strength of persuasion do not have much impact on either C or X – until α reaches a tipping point. After crossing this threshold the system jumps to a much higher level of both cooperation and resource conservation, and the system is also quite robust against possible weakening of social pressure. If α has increased sufficiently for the system to flip to the good equilibrium, α can fall substantially before the system flips back to the bad equilibrium. Again, this is a direct result of the system exhibiting alternative stable states.

Exogenous developments in the system parameters all give rise to the same dynamics as do changes in $\hat{e}$ itself – see section 4.1. All qualitative results obtained analytically assuming $w = 0$ carry over to the case of $w > 0$ (with $\delta \to \infty$), and also the policy implications remain unchanged. If X is observed to stabilize at an intermediately high level this is no guarantee that the system is resilient against shocks. And if the system has collapsed, restoring the system to the good equilibrium requires changing the system parameters such that $\hat{e} << \hat{e}_2(\bullet)$.

4.3 Ecosystem collapse when agents are aware of both the Class I and II problems ($\delta \geq 0$)

Let us now consider the case where the community members are aware of both the instantaneous and intertemporal externality. Each defector takes into account the instantaneous externality vE_{-i} caused by the effort decisions of all community members, but ignores both the instantaneous and intertemporal consequences of his actions on the welfare of others. Hence, the best response function of defectors is still given by (7). Cooperators put in their fair share of the socially optimal aggregate harvesting effort, taking into account the two types of externalities. Hence, they solve (3) where $\delta \geq 0$. Let us use $\tilde{\tilde{z}}$ to denote the socially optimal steady state level of variable z when agents take into account both the Class I and Class II problem of resource harvesting. It is fairly straightforward to determine $\tilde{\tilde{X}}(\bullet)$, $\tilde{\tilde{E}}(\bullet)$ and $\tilde{\tilde{e}}^C(\bullet) = \tilde{\tilde{E}}(\bullet)/N$ (see

appendix A4). But it is very difficult to derive the socially optimal trajectories towards $\tilde{\tilde{X}}(\bullet)$, and it is also not very likely that communities in the real world are able to do that. Therefore, we assume that cooperators adopt a simple stock-dependent harvest strategy [62,63] by using a linear feedback control rule: $e^C = \max\{a + bX, 0\}$.[21] Figure 5 presents the numerical results.

<Insert Figure 5 about here>

The bifurcation diagrams presented in Figure 5 are qualitatively identical to those in Figures 1 and 4.[22] Indeed, the underlying mechanism is the same: the instantaneous profits earned by defectors are always at least as high as those earned by cooperators, socially optimal harvesting decreases with stock size, and the time constraint ceases to be binding for defectors at a stock size that is lower than the stock size at which cooperators start choosing interior effort levels.

The main difference is that maintaining cooperation is even more difficult than in the case in which agents are unaware of the intertemporal (or Class I) externality. Compared to the case of agents being myopic, the cooperators reduce their effort levels even more for every *X* because they now take both externalities into account. That means that for given *C* the payoff ratio π^C / π^D is even lower, and hence (i) collapse occurs sooner (i.e., all else equal, at lower levels of $\hat{e}$, *q*, *N*, etc.), and (ii) the number of cooperators in the bad equilibrium is even smaller. This may

[21] We assume cooperators adopt adaptive management (so that $a < 0$ and $b > 0$) aimed at steering the system towards the optimal steady state. Furthermore, *a* and *b* are set such that each cooperator invests $\tilde{\tilde{e}}^C(\bullet)$ when $X = \tilde{\tilde{X}}(\bullet)$. Our results carry over to the more realistic case where the optimal steady state is not exactly known by cooperators.

[22] Parameter values are $\delta = 0.05$, $a = -0.3$, $b = 1.2$. All other parameters are as before.

explain why communities are better able at overcoming the crowding externality than at solving the intertemporal one [64,65].

Finally, note that while the collapse of the social-ecological system did not result in the total demise of the resource, setting $w = v = 0$ in the intertemporal problem the resource is fully exhausted if the system falls to its bad equilibrium. While the static externality (or a positive wage rate) makes it uneconomical to actually deplete the resource, the intertemporal one does not: as long as the instantaneous benefits of harvesting are positive, defectors continue to extract, and even the last unit will be harvested if the social system collapses.

5. Conclusions

We developed a model of renewable resource use in which agents can decide to act cooperatively with respect to resource harvesting or behave selfishly. Adherence to social harvesting norms can spread through the community because of interpersonal relationships between cooperators and defectors (because the former try to convince the latter of the social desirability of acting cooperatively), but community members also always face the temptation to act non-cooperatively – because of the higher profits. The resulting social-ecological system is characterized by multiple equilibria, so that small changes in key parameters (such as population growth and technological progress) can trigger catastrophic transitions from relatively high levels of cooperation to widespread norm violation – causing the demise of the resource. Our setup is unique in that tipping points emerge even though both the ecological and the social-economic systems, by themselves, are inherently stable.

Positive feedback relationships occur in our model because of the fact that, in closed communities, the amount of labor a community member can allocate to resource harvesting is

necessarily finite because the property right system usually does not allow members to hire external labor. If the resource becomes scarcer, for example due to unfavorable climatic conditions, the cooperators in the community decrease their harvesting effort while defectors continue to allocate all their available time to harvesting – if the net private marginal benefits of harvesting are strictly positive. A decrease in the size of the resource thus increases the relative profitability of defecting, and makes cooperation even more costly. Fewer cooperators are unable to maintain sufficient social pressure, thus triggering even stronger defection and resource depletion. Thus, a positive feedback between the resource stock and the number of cooperators emerges endogenously – possibly resulting in the collapse of the social-ecological system.

Our model is purely theoretical in nature, but it does yield an important policy implication. If a regulator is monitoring the ecological system, she must be aware of potential collapses of cooperation in communities having access to renewable resources even if there are no non-linearities in the resource dynamics themselves. Monitoring the system is important because it is characterized by hysteresis. Upon collapse, it is not sufficient to reverse the small exogenous change that caused the system to collapse. More draconic measures are needed to generate a spiral of increasing cooperation and resource regeneration – possibly at very high cost.

Appendix A1: Proof of Lemma 2

Proof of Lemma 2.1: From (4) we have $\mathrm{d}C/\mathrm{d}t \gtreqless 0$ if $\alpha(N-C)/N \gtreqless \beta\left(1-\frac{\pi^C}{\pi^D}\right)$. Because $w = 0$, we have $\pi^C/\pi^D = e^C/e^D$; see (2). Using (6) and (8) and setting $w = 0$, we have $\left(e^C, e^D\right) = \left(\hat{e}, \hat{e}\right)$

if $X \geq \frac{2\nu N\hat{e}}{Pq}$, $\left(e^C, e^D\right) = \left(\frac{PqX}{2\nu N}, \hat{e}\right)$ if $\frac{2\nu N\left(N-C+1\right)\hat{e}}{Pq\left(2N-C\right)} < X < \frac{2\nu N\hat{e}}{Pq}$, and

$$\left(e^C, e^D\right) = \left(\frac{PqX}{2\nu N}, \left(\frac{2N-C}{N-C+1}\right)\frac{PqX}{2\nu N}\right) \text{ if } 0 \leq X \leq \frac{2\nu N\left(N-C+1\right)\hat{e}}{Pq\left(2N-C\right)}.$$

Inserting $\pi^C / \pi^D = e^C / e^D$ into (4) and setting dC/dt = 0, we have

640 (A1) $$C(X)\big|_{\mathrm{d}C/\mathrm{d}t=0} = \begin{cases} N & \text{if} \quad X \geq \frac{2\nu N\hat{e}}{Pq}, \\ \frac{\left(2\nu\left(\alpha-\beta\right)N\hat{e}+\beta PqX\right)}{2\nu\hat{e}\alpha} \equiv \kappa(X) < N & \text{if} \quad \frac{2\nu N\left(N-C+1\right)\hat{e}}{Pq\left(2N-C\right)} \leq X < \frac{2\nu N\hat{e}}{Pq}, \\ \frac{3}{2}N - \frac{1}{2}\sqrt{N^2 + \frac{4\beta N}{\alpha}(N-1)} \equiv \underline{\kappa} < \kappa(X) & \text{if} \quad 0 \leq X < \frac{2\nu N\left(N-C+1\right)\hat{e}}{Pq\left(2N-C\right)}. \end{cases}$$

Proof of Lemma 2.2: This can trivially be inferred from (A1).

Proof of Lemma 2.3: Defining $V \equiv \alpha\left(N-C\right)/N - \beta\left(1-e^C\left(X\right)/e^D\left(C,X\right)\right)$ and using (8), we have

$\mathrm{d}V/\mathrm{d}C < 0$ for all X. Therefore, for any X, $\mathrm{d}C/\mathrm{d}t \overset{>}{\underset{<}{=}} 0$ if $C \overset{<}{\underset{>}{=}} C(X)\big|_{\mathrm{d}C/\mathrm{d}t=0}$. ■

645 **Appendix A2: Proof of Lemma 3**

Inserting (6) and (8) into (1) and setting w = 0, we have

(A2) $$\frac{\mathrm{d}X/\mathrm{d}t}{X} = \begin{cases} r\left(1-X\right) - qN\hat{e} & \text{if } X \geq \frac{2\nu N\hat{e}}{Pq}, \\ r\left(1-X\right) - q\left(\frac{CPqX}{2\nu N} + \left(N-C\right)\hat{e}\right) & \text{if } \frac{2\nu N\left(N-C+1\right)\hat{e}}{Pq\left(2N-C\right)} \leq X < \frac{2\nu N\hat{e}}{Pq}, \\ r\left(1-X\right) - Pq^2XZ(C)/\nu & \text{if } 0 \leq X < \frac{2\nu N\left(N-C+1\right)\hat{e}}{Pq\left(2N-C\right)}. \end{cases}$$

where $Z(C) = \frac{N-C+C/(2N)}{N-C+1} \approx 1$.

Proof of Lemma 3.1: Combining $X \geq 2vN\hat{e}/(Pq)$ and $\mathrm{d}X/\mathrm{d}t = \left[r(1-X) - qN\hat{e}\right]X = 0$, we have

$X(C)\big|_{\mathrm{d}X/\mathrm{d}t=0} = (r - Nq\hat{e})/r \equiv \bar{\psi}$ if and only if $\hat{e} \leq \dfrac{rPq/N}{2rv + Pq^2} \equiv \hat{e}_0$.

Proof of Lemma 3.2: If $\hat{e} > \hat{e}_0$, a corollary of Lemma 3.1 is that $\mathrm{d}X/\mathrm{d}t < 0$ for all $X \geq 2vN\hat{e}/(Pq)$. Using (A2) we have

$$\text{(A3)}\quad X(C)\big|_{\mathrm{d}X/\mathrm{d}t=0} = \begin{cases} \text{does not exist} & \text{if} \quad X \geq \dfrac{2vN\hat{e}}{Pq}, \\ \dfrac{2vN\left[r - q\hat{e}(N-C)\right]}{2rvN + Pq^2C} \equiv \psi(\mathrm{C}) & \text{if} \quad \dfrac{2vN(N-C+1)\hat{e}}{Pq(2N-C)} \leq X < \dfrac{2vN\hat{e}}{Pq}, \\ \dfrac{vr}{Pq^2Z(C) + vr} & \text{if} \quad 0 \leq X < \dfrac{2vN(N-C+1)\hat{e}}{Pq(2N-C)}. \end{cases}$$

The nullcline is concave in R2 for $\hat{e} > \hat{e}_0$ because $\dfrac{\partial X(C)_{\mathrm{d}X/\mathrm{d}t=0}}{\partial C} > 0, \dfrac{\partial^2 X(C)_{\mathrm{d}X/\mathrm{d}t=0}}{\partial C^2} < 0.$ And because $Z(C) \approx 1$ the nullcline is (almost) horizontal in R3, with $X(C)_{\mathrm{d}X/\mathrm{d}t=0} \approx \dfrac{vr}{Pq^2 + vr} \equiv \underline{\psi}$.

Proof of Lemma 3.3: From (A3) we have $\dfrac{\mathrm{d}X(C)_{\mathrm{d}X/\mathrm{d}t=0}}{\mathrm{d}\hat{e}} \leq 0$ in R2, and $\dfrac{\mathrm{d}X(C)_{\mathrm{d}X/\mathrm{d}t=0}}{\mathrm{d}\hat{e}} = 0$ in R3.

Proof of Lemma 3.4: Defining $W \equiv r(1-X) - q\left(Ce^C(X) + (N-C)e^D(C,X)\right)$ and using (6) and (8), we have $\mathrm{d}W/\mathrm{d}X < 0$ for all C. Therefore, for any C, $\mathrm{d}X/\mathrm{d}t \overset{>}{=} 0$ for all $X \overset{<}{=} X(C)\big|_{dX/dt=0}$. ■

Appendix A3: Proof of Proposition 1

(i) For $X \geq 2vN\hat{e}/(Pq)$ we have $e^C = e^D = \hat{e}$; see (6) and (8). Using (A1) and (A2) we have

(A4) $(C_1, X_1) = (N, (r - \hat{e}Nq)/r)$,

and this is an equilibrium in R1 if and only if $X_1 = (r - \hat{e}Nq)/r \geq 2vN\hat{e}/(Pq)$. Solving for $\hat{e}$, (C_1, X_1) is an equilibrium if and only if $\hat{e} \leq \dfrac{rPq/N}{2rv + Pq^2} \equiv \hat{e}_0$.

665

(ii) For $X < 2vN\left(N - C + 1\right)\hat{e}/\left(Pq\left(2N - C\right)\right)$ we have, from (A1) and (A3) respectively,

$C(X)\big|_{\mathrm{d}C/\mathrm{d}t=0} = \dfrac{3}{2}N - \dfrac{1}{2}\sqrt{N^2 + \dfrac{4\beta N(N-1)}{\alpha}}$ and $X(C)\big|_{\mathrm{d}X/\mathrm{d}t=0} = \dfrac{vr}{Pq^2 Z(C) + vr}$. Combining, we have

(A5) $X_3 = \dfrac{2rv\sqrt{N}\left(\sqrt{N}\sqrt{\alpha N + 4\beta\left(N-1\right)} + \sqrt{\alpha}\left(2 - N\right)\right)}{\left(2rNv + Pq^2\left(2N-1\right)\right)\sqrt{\alpha N + 4\beta\left(N-1\right)} - \sqrt{\alpha}\sqrt{N}\left(2rv\left(N-2\right) + Pq^2\left(2N-3\right)\right)}$,

(A6) $C_3 = \dfrac{3}{2}N - \dfrac{1}{2}\sqrt{N^2 + \dfrac{4\beta}{\alpha}(N^2 - N)} > 0$,

670 and this is an equilibrium if $X_3 \leq \dfrac{2vN\left(N - C_3 + 1\right)\hat{e}}{Pq\left(2N - C_3\right)}$. Solving, (C_3, X_3) is an equilibrium in R3 iff

(A7) $\hat{e} \geq \dfrac{\beta rPq}{\alpha N\left(rv + Pq^2\right) + \beta\left(2rNv + Pq^2\left(2N-1\right)\right) - \sqrt{\alpha}\sqrt{N}\sqrt{\alpha N + 4\beta\left(N-1\right)}\left(rv + Pq^2\right)} \equiv \hat{e}_1$.

(iii) For $\dfrac{2vN\left(N - C + 1\right)\hat{e}}{Pq\left(2N - C\right)} \leq X < \dfrac{2vN\hat{e}}{Pq}$, we have $C(X)\big|_{\mathrm{d}C/\mathrm{d}t=0} = \dfrac{2v\left(\alpha - \beta\right)N\hat{e} + \beta PqX}{2v\hat{e}\alpha}$ (see (A1)) and

$X(C)\big|_{\mathrm{d}X/\mathrm{d}t=0} = \dfrac{2vN\left[r - q\hat{e}(N - C)\right]}{2rvN + Pq^2 C}$ (see (A3)). Combining and defining $Q \equiv Pq^2 + 2rv$, we have

675 (A8) $X_2 = \dfrac{\sqrt{\hat{e}}\sqrt{N}\cdot v\left(\sqrt{\alpha}\sqrt{\alpha\hat{e}N\left(4r^2v^2 + 4rPvq^2 + P^2q^4\right) + 4\beta Pq^2\left(rPq - \hat{e}NQ\right)} - \sqrt{\hat{e}}\sqrt{N}\left(\alpha Q - 2\beta Pq^2\right)\right)}{\beta P^2 q^3}$,

(A9) $C_2 = \dfrac{\sqrt{N}\left(\sqrt{\alpha\hat{e}N\left(4r^2v^2 + 4rPvq^2 + P^2q^4\right) + 4\beta Pq^2\left(rPq - \hat{e}NQ\right)} + \sqrt{\alpha}\sqrt{\hat{e}}\sqrt{N}\left(Pq^2 - 2rv\right)\right)}{2\sqrt{\alpha}\sqrt{\hat{e}}Pq^2}$,

(A10) $X_{2B} = -\dfrac{\sqrt{\hat{e}}\sqrt{N}\cdot v\left(\sqrt{\alpha}\sqrt{\alpha\hat{e}N\left(4r^2v^2+4rPvq^2+P^2q^4\right)+4\beta Pq^2\left(rPq-\hat{e}NQ\right)}+\sqrt{\hat{e}}\sqrt{N}\left(\alpha Q-2\beta Pq^2\right)\right)}{\beta P^2q^3}$,

(A11) $C_{2B}=-\dfrac{\sqrt{N}\left(\sqrt{\alpha\hat{e}N\left(4r^2v^2+4rPvq^2+P^2q^4\right)+4\beta Pq^2\left(rPq-\hat{e}NQ\right)}+\sqrt{\alpha}\sqrt{\hat{e}}\sqrt{N}\left(2rv-Pq^2\right)\right)}{2\sqrt{\alpha}\sqrt{\hat{e}}Pq^2}$.

We have $X_2 = X_{2B}, C_2 = C_{2B}$ (and hence just one equilibrium in R2) iff

(A12) $\hat{e}=\dfrac{4\beta rP^2q^3}{N\left(4\beta Pq^2\left(2rv+Pq^2\right)-\alpha\left(4rv\left(rv+Pq^2\right)+P^2q^4\right)\right)}\equiv\hat{e}_2$.

If $\hat{e}>\hat{e}_2$ (A8)-(A11) do not have real roots; in that case there are no equilibria in R2. If $\hat{e}<\hat{e}_1$ we have $\dfrac{2vN\left(N-C_{2B}+1\right)\hat{e}}{Pq\left(2N-C_{2B}\right)}>X_{2B}$, and then (C_2, X_2) is the only equilibrium in R2. If $\hat{e}_1<\hat{e}<\hat{e}_2$ there are two equilibria ((C_2, X_2) and (C_{2B}, X_{2B})) in R2.

(iv) Combining **(i)-(iii)**, if $\hat{e}\le\hat{e}_0$ the system's equilibrium is (C_1, X_1) as defined in (A4), if $\hat{e}_0<\hat{e}<\hat{e}_1$ there is just one equilibrium ((C_2, X_2) as defined in (A8) and (A9)), if $\hat{e}>\hat{e}_2$ there is just one equilibrium ((C_3, X_3) as defined in (A5) and (A6)), and if $\hat{e}_1<\hat{e}<\hat{e}_2$ there are three equilibria ((C_2, X_2), (C_{2B}, X_{2B}) and (C_3, X_3) as defined in (A5)-(A6) and (A8)-(A11).■

Appendix A4: The socially optimal steady state in the presence of both externalities

Writing down the current value Hamiltonian of (3), taking the appropriate first derivates, setting all time derivatives equal to zero and solving, the socially optimal steady state levels of *E* and *X* (denoted by $\tilde{\tilde{E}}$ and $\tilde{\tilde{X}}$) are implicitly determined by the following two equations:

(A13) $$\delta - r = \frac{Pq^2 \tilde{\tilde{X}} \tilde{\tilde{E}}}{Pq\tilde{\tilde{X}} - w - 2v\tilde{\tilde{E}}} - 2r\tilde{\tilde{X}} - q\tilde{\tilde{E}}$$

(A14) $$\tilde{\tilde{E}} = r\left(1 - \tilde{\tilde{X}}\right)/q,$$

and where, in the optimum, each cooperator sets $\tilde{\tilde{e}}^C = \tilde{\tilde{E}} / N$. $\tilde{\tilde{X}}$ and $\tilde{\tilde{E}}$ are then equal to

(A15) $$\tilde{\tilde{X}} = \frac{\sqrt{\delta^2 Q^2 - 2\delta r\left(P^2 q^4 - Pq^2\left(3qw + 4rv\right) - 2rv\left(qw + 2rv\right)\right) + r^2\left(Q + qw\right)^2} - \delta Q - r\left(Q + qw\right)}{4r\left(Pq^2 + rv\right)}$$

(A16) $$\tilde{\tilde{E}} = \frac{N\sqrt{\delta^2 Q^2 - 2\delta r\left(P^2 q^4 - Pq^2\left(3qw + 4rv\right) - 2rv\left(qw + 2rv\right)\right) + r^2\left(Q + qw\right)^2} - \delta NQ - Nr\left(3Pq^2 - qw + 2rv\right)}{4q\left(Pq^2 + rv\right)}$$

where $Q \equiv Pq^2 + 2rv$.

References

[1] J.M. Diamond, Collapse: how societies choose to fail or succeed, Blackwell Synergy, 2005.
[2] M. Scheffer, Critical transitions in nature and society, Princeton University Press, 2009.
[3] J.A. Brander, M.S. Taylor, The Simple Economics of Easter Island: A Ricardo-Malthus Model of Renewable Resource Use, *American Economic Review*, **88** (1998) 119-138.
[4] P. Berck, Open Access and Extinction, *Econometrica*, **47** (1979) 877-882.
[5] J.R. Gould, Extinction of a Fishery by Commercial Exploitation: A Note, *The Journal of Political Economy*, **80** (1972) 1031-1038.
[6] J.M. Anderies, M.A. Janssen, B.H. Walker, Grazing Management, Resilience, and the Dynamics of a Fire-driven Rangeland System, *Ecosystems*, **5** (2002) 23-44.
[7] K.G. Mäler, A. Xepapadeas, A. de Zeeuw, The Economics of Shallow Lakes, *Environmental and Resource Economics*, **26** (2003) 603-624.
[8] S. Polasky, A. de Zeeuw, F. Wagener, Optimal management with potential regime shifts, *Journal of Environmental Economics and Management*, **62** (2011) 229-240.
[9] M. Scheffer, S. Carpenter, J.A. Foley, C. Folke, B. Walker, Catastrophic shifts in ecosystems, *Nature*, **413** (2001) 591-596.
[10] E.H. Bulte, R.D. Horan, Identities in the Commons: The Dynamics of Norms and Social Capital, *The BE Journal of Economic Analysis & Policy*, **10** (2010) Art.13.
[11] N. Osés-Eraso, M. Viladrich-Grau, On the sustainability of common property resources, *Journal of Environmental Economics and Management*, **53** (2007) 393-410.
[12] R. Sethi, E. Somanathan, The Evolution of Social Norms in Common Property Resource Use, *American Economic Review*, **86** (1996) 766-788.
[13] J. Noailly, J. van den Bergh, C.A. Withagen, Evolution of harvesting strategies: replicator and resource dynamics, *Journal of Evolutionary Economics*, **13** (2003) 183-200.
[14] N. McCarthy, E. Sadoulet, A. de Janvry, Common Pool Resource Appropriation under Costly Cooperation, *Journal of Environmental Economics and Management*, **42** (2001) 297-309.

[15] L. Cavalli-Sforza, M.W. Feldman, Cultural Transmission and Evolution: A Quantitative Approach.(MPB-16), Princeton University Press, 1981.
[16] P.S. Dodds, D.J. Watts, A generalized model of social and biological contagion, *Journal of Theoretical Biology*, **232** (2005) 587-604.
[17] G. Heal, H. Kunreuther, Social Reinforcement: Cascades, Entrapment, and Tipping, *American Economic Journal: Microeconomics*, **2** (2010) 86-99.
[18] H.P. Young, Innovation Diffusion in Heterogeneous Populations: Contagion, Social Influence, and Social Learning, *American Economic Review*, **99** (2009) 1899-1924.
[19] J.A. Robalino, A. Pfaff, Contagious development: Neighbor interactions in deforestation, *Journal of Development Economics*, **In Press** (2011).
[20] D. Masclet, C. Noussair, S. Tucker, M.-C. Villeval, Monetary and Nonmonetary Punishment in the Voluntary Contributions Mechanism, *American Economic Review*, **93** (2003) 366-380.
[21] E. Fehr, S. Gächter, Cooperation and punishment in public goods experiments, *American Economic Review*, **90** (2000) 980-994.
[22] D.G. Rand, A. Dreber, T. Ellingsen, D. Fudenberg, M.A. Nowak, Positive Interactions Promote Public Cooperation, *Science*, **325** (2009) 1272-1275.
[23] S. Gächter, E. Renner, M. Sefton, The Long-Run Benefits of Punishment, *Science*, **322** (2008) 1510.
[24] L. Karp, S. Sacheti, J. Zhao, Common Ground Between Free-Traders and Environmentalists, *International Economic Review*, **42** (2001) 617-648.
[25] M.S. Taylor, Environmental crises: past, present, and future, *Canadian Journal of Economics/Revue canadienne d'économique*, **42** (2009) 1240-1275.
[26] S. Bowles, H. Gintis, Social Capital and Community Governance, *The Economic Journal*, **112** (2002) 419.
[27] J.M. Baland, J.P. Platteau, Halting degradation of natural resources, Clarendon Press for FAO, Oxford., 1996.
[28] E. Ostrom, Governing the Commons: The Evolution of Institutions for Collective Action, Cambridge University Press, 1990.
[29] J. Aimers, D. Hodell, Societal collapse: Drought and the Maya, *Nature*, **479** (2011) 44-45.
[30] T. Hunt, C. Lipo, The statues that walked: unraveling the mystery of Easter Island, Free Press, 2011.
[31] B.R. Copeland, M.S. Taylor, Trade, Tragedy, and the Commons, *American Economic Review*, **99** (2009) 725-749.
[32] L. Hotte, N.V. Long, H. Tian, International trade with endogenous enforcement of property rights, *Journal of Development Economics*, **62** (2000) 25-54.
[33] J. Barkley Rosser, Complex ecologic-economic dynamics and environmental policy, *Ecological Economics*, **37** (2001) 23-37.
[34] J. Liu, T. Dietz, S.R. Carpenter, M. Alberti, C. Folke, E. Moran, A.N. Pell, P. Deadman, T. Kratz, J. Lubchenco, Complexity of Coupled Human and Natural Systems, *Science*, **317** (2007) 1513.
[35] M. Scheffer, F. Westley, W. Brock, Slow Response of Societies to New Problems: Causes and Costs, *Ecosystems*, **6** (2003) 493-502.
[36] Y. Iwasa, T. Uchida, H. Yokomizo, Nonlinear behavior of the socio-economic dynamics for lake eutrophication control, *Ecological Economics*, **63** (2007) 219-229.
[37] Y. Suzuki, Y. Iwasa, The coupled dynamics of human socio-economic choice and lake water system: the interaction of two sources of nonlinearity, *Ecological Research*, **24** (2009) 479-489.
[38] G.R. Munro, A.D. Scott, The Economics of Fisheries Management, in: A.V. Kneese, J.L. Sweeney (Eds.) Handbook of Natural Resource and Energy Economics, Elsevier, 1985, pp. 623-676.
[39] J.A. Wilson, The Economical Management of Multispecies Fisheries, *Land Economics*, **58** (1982) 417-434.
[40] C.W. Clark, Mathematical bioeconomics: The optimal management of renewable resources, 2 ed., John Willy & Sons, Inc. , New York 1990.
[41] J.R. Boyce, Individual Transferable Quotas and Production Externalities in a Fishery, *Natural Resource Modeling*, **6** (1992) 385-408.

[42] V.L. Smith, Economics of production from natural resources, *The American Economic Review*, **58** (1968) 409-431.
[43] V.L. Smith, On models of commercial fishing, *The Journal of Political Economy*, **77** (1969) 181-198.
[44] C.W. Clark, Towards a predictive model for the economic regulation of commercial fisheries, *Canadian Journal of Fisheries and Aquatic Science*, **37** (1980) 1111-1129.
[45] J.M. Karpoff, Suboptimal Controls in Common Resource Management: The Case of the Fishery, *The Journal of Political Economy*, **95** (1987) 179-194.
[46] D. Balliet, Communication and Cooperation in Social Dilemmas: A Meta-Analytic Review, *Journal of Conflict Resolution*, **54** (2010) 39-57.
[47] E. Ostrom, J. Walker, R. Gardner, Covenants with and without a Sword - Self-Governance Is Possible, *American Political Science Review*, **86** (1992) 404-417.
[48] K. Keizer, S. Lindenberg, L. Steg, The Spreading of Disorder, *Science*, **322** (2008) 1681-1685.
[49] C.F. Manski, Economic Analysis of Social Interactions, *The Journal of Economic Perspectives*, **14** (2000) 115-136.
[50] C. Bicchieri, The Grammar of Society: The Nature And Dynamics Of Social Norms, Cambridge University Press, 2006.
[51] J.F. Scott, Internalization of norms : A Sociological Theory of Moral Commitment, Prentice Hall, 1971.
[52] M. Bateson, D. Nettle, G. Roberts, Cues of being watched enhance cooperation in a real-world setting, *Biology Letters*, **2** (2006) 412-414.
[53] P. Keefer, S. Knack, Social Capital, Social Norms and the New Institutional Economics, in: C. Ménard, M.M. Shirley (Eds.) Handbook of New Institutional Economics, Kluwer Academic Pub, 2005.
[54] M. Servátka, S. Tucker, R. Vadovič, Words speak louder than money, *Journal of Economic Psychology*, **32** (2011) 700-709.
[55] M.A. Janssen, R. Holahan, A. Lee, E. Ostrom, Lab Experiments for the Study of Social-Ecological Systems, *Science*, **328** (2010) 613-617.
[56] E. Ostrom, A Behavioral Approach to the Rational Choice Theory of Collective Action: Presidential Address, American Political Science Association, 1997, *The American Political Science Review*, **92** (1998) 1-22.
[57] E. Fehr, S. Gächter, Cooperation and Punishment in Public Goods Experiments, *The American Economic Review*, **90** (2000) 980-994.
[58] R. Boyd, H. Gintis, S. Bowles, P.J. Richerson, The evolution of altruistic punishment, *Proceedings of the National Academy of Sciences*, **100** (2003) 3531-3535.
[59] G.I. Bischi, F. Lamantia, L. Sbragia, Competition and cooperation in natural resources exploitation: an evolutionary game approach, in: C. Carraro, V. Fragnelli (Eds.) Game Practice and the Environment, Edward Elgar Publishing, 2004, pp. 187–121.
[60] L.H. Gunderson, C.S. Holling, Panarchy: Understanding Transformations in Human and Natural Systems, Island Press, Washington D.C. , 2002.
[61] L. Huang, M.D. Smith, The Dynamic Efficiency Costs of Common-Pool Resource Exploitation, manuscript, (2011).
[62] R. Hilborn, C.J. Walters, Quantitative fisheries stock assessment: choice, dynamics, and uncertainty, *Reviews in Fish Biology and Fisheries*, **2** (1992) 177-178.
[63] C. Walters, Adaptative management of renewable resource, Macmillan Publishing, 1986.
[64] P. Holm, B. Hersoug, S.A. Rånes, Revisiting Lofoten: Co-Managing Fish Stocks or Fishing Space?, *Human Organization*, **59** (2000) 353-364.
[65] L. Taylor, The river would run red with blood: Community and common property in Irish fishing settlement, in: B.J. McCay, J.M. Acheson (Eds.) The Question of the Commons: The Culture and Ecology of Communal Resources., University of Arizona Press, 1987, pp. 290-307.

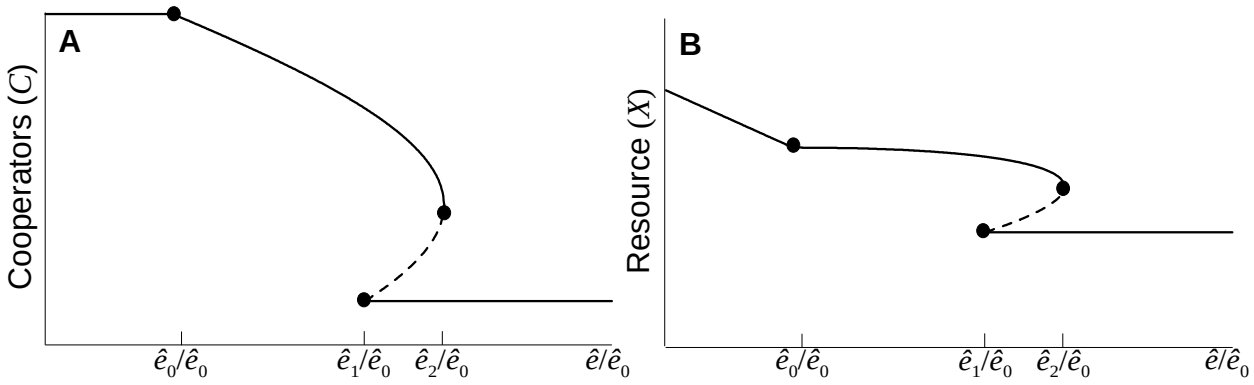


Figure 1. Bifurcation diagram showing internal equilibria of the number of cooperators C (panel A) and the resource stock X (panel B) for different values of $\hat{e}/\hat{e}_0$, reflecting the strength of the social dilemma. Stable equilibria are connected by solid lines, unstable equilibria are connected by dashed lines. Dots denote the two tipping points $\hat{e}_1/\hat{e}_0$ and $\hat{e}_2/\hat{e}_0$, and the point $\hat{e}_0/\hat{e}_0 = 1$, where the social dilemma materializes.

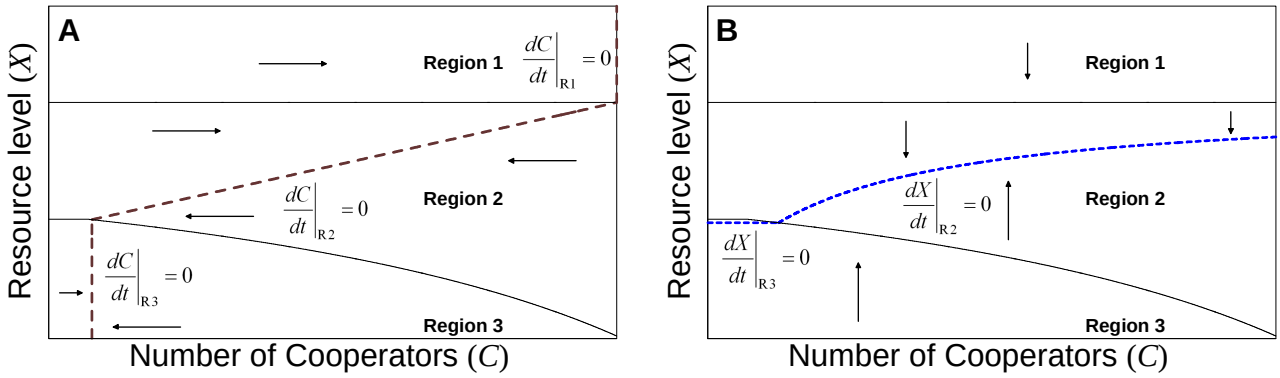
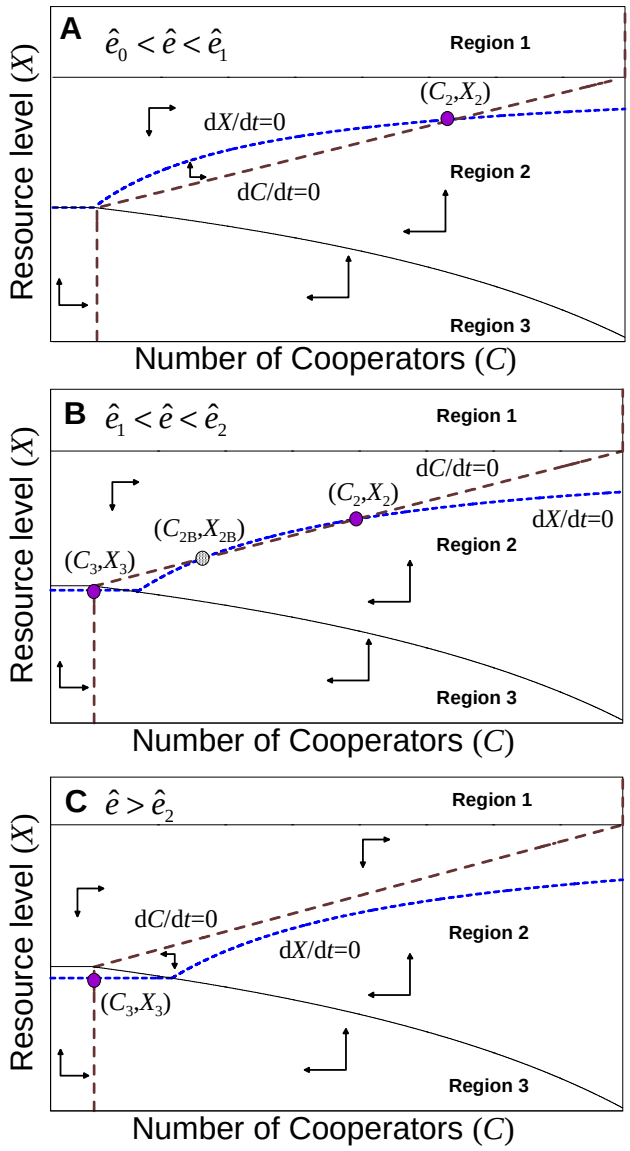


Figure 2. The nullclines (dashed lines) and direction vectors of the number of cooperators (panel A) and of the resource stock (panel B). The three different regions are separated by solid lines.[23]

[23] Trivial nullclines ($dC/dt = 0$ at $C = 0$ and $dX/dt = 0$ at $X = 0$) are not shown.



845

Figure 3. Phase planes showing the nullclines (dashed lines), the region boundaries (solid lines) and the vector fields for different values of $\hat{e}$. In panel A $(\hat{e}_0 < \hat{e} < \hat{e}_1)$ and C $(\hat{e} > \hat{e}_2)$ there is only one stable equilibrium, while panel B $(\hat{e}_1 < \hat{e} < \hat{e}_2)$ exhibits alternative stable states.[24]

[24] Parameter values as in footnote 18 and the focal parameter $\hat{e}$ equals 0.689 in panel A, 0.713 in panel B, and 0.75 (panel C). Again, the trivial nullclines are not shown. Note that the intersection point of $\mathrm{d}C/\mathrm{d}t = 0$ and the horizontal axis (X = 0) and the intersection point of $\mathrm{d}X/\mathrm{d}t = 0$ and the vertical axis (C = 0) are equilibria too, and so is the origin of the system; cf. (1) and (4). As these three equilibria are unstable, we omit them in this figure.

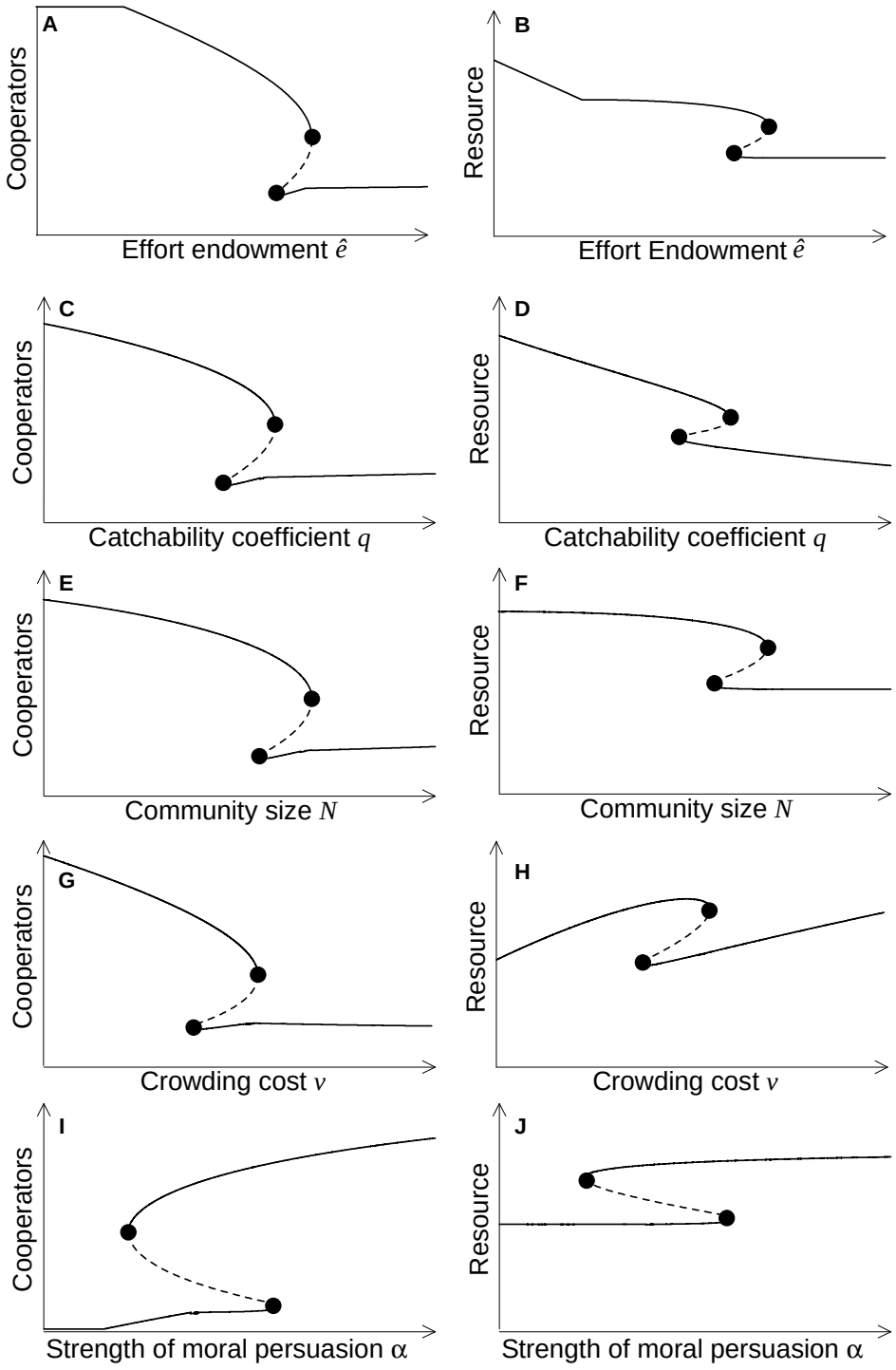


Figure 4. Bifurcation diagrams showing internal equilibria of the number of cooperators C and the resource stock X for w=0.1. Stable equilibria are connected by solid lines, unstable equilibria are connected by dashed lines, and dots denote the tipping points.

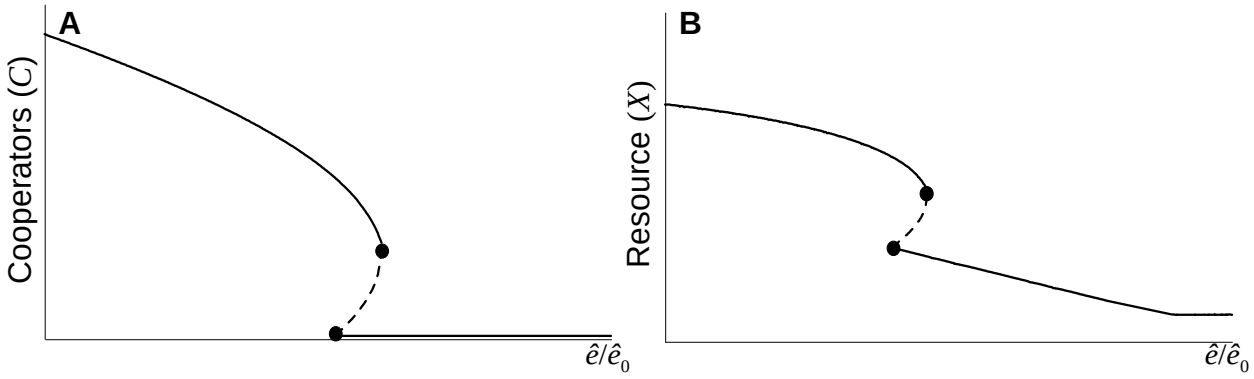


Figure 5. Bifurcation diagrams showing internal equilibria of the number of cooperators C and the resource stock X for different values of the effort endowment $\hat{e}$ and with w=0.1 when agents are aware of both the Class I and Class II externalities $(\delta = 0.05)$. Stable equilibria are connected by solid lines, unstable equilibria are connected by dashed lines, and dots denote the tipping points.