Joint Discussion Paper
Series in Economics

by the Universities of

Aachen · Gießen · Göttingen
Kassel · Marburg · Siegen

ISSN 1867-3678

No. 18-2024

Fabian Mankat

Norms as a solution to the tragedy of the commons: A co-evolutionary model

Coordination: Bernd Hayo • Philipps-University Marburg
School of Business and Economics • Universitätsstraße 24, D-35032 Marburg
Tel: +49-6421-2823091, Fax: +49-6421-2823088, e-mail: hayo@wiwi.uni-marburg.de

Norms as a solution to the tragedy of the commons: A co-evolutionary model

Fabian Mankat*

August 14, 2024

Abstract

This paper examines how societies can conserve common-pool resources (CPRs) through the cultural transmission of norms. To this end, we introduce an evolutionary model that endogenizes the formation of personal, social, and descriptive norms, thereby unifying existing economic theories on norm evolution. By studying this model in a binary CPR game, we also account for the dynamics of the resource stock and its interplay with norms and behavior. We find that the resource can persist through (1) asymptotically stable equilibrium points where moral perceptions and behavioral routines are either homogeneous or heterogeneous across individuals and (2) an asymptotically stable limit cycle in which moral perceptions remain constant, but herding causes alternating aggregate behavior and fluctuating resource stocks. We examine the degree of substitutability between two key factors — (a) the active promotion of norm adoption by institutions and (b) the impact of social recognition on the opinion formation of peers — for upholding norms and thus securing the CPR. Moreover, we find that while larger sanctions for norm violations and lower material benefits from resource exploitation generally favor resource-conserving behavior in the short run, they may, surprisingly, adversely affect resource conservation in the long run by interfering with cultural dynamics and thereby threatening norm persistence.

JEL Classification Codes: C73, D02, D62, Q20, Z13

*University of Kassel, Faculty of Economics and Management

1 Introduction

Governing common-pool resources (CPRs) is among societies' most critical challenges, with its failure often resulting in severe economic and ecological consequences (Ostrom, 2009). The underlying challenge arises from the individuals' short-term benefits of exploitation, leading to the collective depletion and eventual erosion of the resource (Hardin, 1968). Economists widely accept norms as a solution for societies to bridge the divergence between self- and social interests in such dilemmas (Elster, 1989; Ostrom, 2000). Scholars generally differentiate between different types of norms (see, e.g., Bicchieri and Dimant, 2019; Dannenberg et al., 2024). A personal (or moral) norm captures an individual's perception of what is morally right, guiding behavior through inner feelings such as guilt and self-disapproval (Nyborg, 2018). A social norm captures a society's shared understanding of what is morally right, while a descriptive norm captures what behavior is generally executed (Crawford and Ostrom, 1995; Thøgersen, 2006; Farrow, Grolleau, and Ibanez, 2017). Individuals follow social and descriptive norms to avoid social sanctions (Voss, 2001; Fehr and Fischbacher, 2004; te Velde and Louis, 2022). Clearly, norms introduce incentives to abstain from self-serving behavior and, thereby, may enable resource conservation.[1] However, rather than being exogenous to society, norms are endogenous constructs that evolve through social interactions and cultural transmission processes, exposing them to the potential threat of erosion themselves. Therefore, this paper addresses the following research questions: Under what conditions can a society rely on norms to secure a CPR? And, more particularly, how do different cultural and socio-economic factors affect society's possibilities to conserve a CPR?

We introduce a novel evolutionary game theoretical model to shed light on these questions. The main contribution to the existing literature is threefold. First, to the best of our knowledge, we present the first model that endogenizes personal, social, and descriptive norm formation while accounting for their joint influences on behavior. To this end, we merge existing ideas and present a general model of norms that can be applied to various

[1] For some theoretical contributions on how norms may affect economic outcomes see Bernheim (1994), Nyborg (2000), Brekke, Kverndokk, and Nyborg (2003), Nyborg and Rege (2003a), Andreoni and Bernheim (2009), Traxler (2010), Bénabou and Tirole (2011), Figuieres, Masclet, and Willinger (2013), and d'Adda, Dufwenberg, Passarelli, and Tabellini (2020) among others.

economic problems. Second, by studying this dynamic model of norms in a CPR game, we introduce an additional dynamic dimension and its interactions with norms and behavior, namely the resource stock. This allows us to address the research questions and enables the thematic insights of this paper. To this end, and this is our third contribution, we present the first model that integrates the *indirect evolutionary approach* as proposed by Güth and Yaari (1992) into a CPR game.

The model considers a continuum of individuals who recurrently interact in a binary CPR game. Following the indirect evolutionary approach, an individual's behavior results from utility maximization. The behaviors of all individuals jointly determine the descriptive norm. Resource dynamics depend on the share of individuals exploiting the resource and the resource's natural growth. Horizontal and oblique cultural transmission drive the adoption of personal norms. Horizontal cultural transmission occurs through peer-to-peer interactions, whereas oblique transmission occurs through socialization institutions. These cultural transmission processes are biased in that socially successful individuals more significantly impact the opinion-formation process of others. Social success derives from two factors, namely material payoff and social sanctions. Additionally, institutions promote norm adoption through oblique cultural transmission. The distribution of individuals' moral perceptions, their personal norms, defines what society generally considers morally appropriate, namely the social norm.

We elicit conditions for the existence and asymptotic stability of (1) a boundary equilibrium point, where all individuals agree one should not exploit the resource, behave accordingly, and the resource stock is at the maximum sustainable level, (2) an interior equilibrium point, where only some individuals agree one should not exploit the resource, only these individuals behave accordingly, and the resource is below the maximum sustainable capacity, and (3) a boundary limit cycle, where all individuals agree one should not exploit the resource, aggregate behavior alternates between everyone and no-one exploiting the resource, and the resource stock exhibits a cyclic evolution. These equilibria are generally compatible with empirical observations of heterogeneous behavior and moral perceptions across individuals (see, e.g., Kotchen and Moore, 2008; Sundt and Rehdanz, 2015; Minton, Spielmann, Kahle, and Kim, 2018) as well as resource conservation in ever-evolving systems (see, e.g., Folke, Colding, and Berkes, 2002; Olsson, Folke, and Berkes, 2004).

By analyzing the existence and asymptotic stability conditions of the identified equilibria, we find that institutional pressure on norm adoption and the weight of social sanctions on the cultural fitness that drives norm adoption play a crucial role in securing the CPR. In some instances, these two factors function as substitutes, making either one dispensable, whereas institutional pressure cannot be substituted in others. The latter may apply, for example, when personal norms play a subordinate role in determining behavioral outcomes, as this generally increases institutional pressure's relative effectiveness. Moreover, material costs of non-exploitation and sanctions for norm violation play ambiguous roles in securing the CPR: Changes that reduce exploitation incentives may lead society to reach a non-favorable behavioral equilibrium in terms of norm persistence. Consequently, seemingly favorable changes may negatively impact resource conservation in the long run.

The rest of the paper unfolds as follows. Section 2 discusses the existing literature and how this paper relates to it. Section 3 presents the static framework. We discuss equilibrium behavior for exogenous norms and resource stocks in Section 4. Section 5 presents the evolutionary framework, which we study in Section 6. The results are discussed in Section 7. Finally, Section 8 concludes.

2 Related literature

This paper contributes to the theoretical literature on norms using evolutionary game theory, particularly those dealing with CPR games. A significant contribution comes from Sethi and Somanathan (1996), who study a model where each agent in a population belongs to one of three strategy types: defectors, cooperators, and enforcers. Enforcers punish defectors, with the costs of punishing and being punished depending on the distribution of strategies. Over time, agents adapt their strategy by imitating materially successful others in the population. Sethi and Somanathan (1996) identify two potentially stable Nash equilibria: (1) a population of only defectors and (2) a population consisting of cooperators and enforcers. Noailly, van den Bergh, and Withagen (2003), Noailly, Withagen, and Van den Bergh (2007), and Tilman, Plotkin, and Akçay (2020) provide further rationalizations for resource conservation when material payoff alone determines the adopted behavioral strategies. In parallel, Young (1993, 1996, 2015), Binmore and Samuelson (1994), and Lindbeck, Nyberg, and Weibull (1999) study similar models in public goods (PG) games.

We differ from that literature by focusing on norm-based behavioral motivations. Hence, rather than explicit material sanctioning through peers, we focus on indirect enforcement mechanisms such as self- and social disapproval. Osés-Eraso and Viladrich-Grau (2007) study such behavioral mechanisms in a CPR game. Individuals either cooperate or defect and adapt their strategies over time. Defection yields material gains, whereas cooperation yields social approval. This social approval increases in the share of others who cooperate, which introduces concerns for behavioral conformity. However, material gains of defection also increase in the share of cooperators, which enables the potential existence of a stable equilibrium characterized by the coexistence of cooperators and defectors. In addition, Osés-Eraso and Viladrich-Grau (2007) show when stable homogeneous equilibria exist. Nyborg and Rege (2003b), Rege (2004), and Nyborg, Howarth, and Brekke (2006) incorporate similar concerns for behavioral conformity into PG games, which effectively turn the corresponding situations into coordination games with full cooperation being a potentially stable equilibrium.

These works provide insights into how descriptive norms may affect behavioral incentives in dynamic settings. We expand on these contributions by also endogenizing the formation of personal and social norms. Conceptually and methodologically, our model closely relates to the literature that deals with the cultural transmission of norms using the indirect evolutionary approach.[2]

Mengel (2008) offers a notable contribution to this literature. She examines the cultural

[2] Other theoretical studies of norm evolution in PG and CPR games encompass theories on rational socialization (PG: Bisin and Verdier, 1998, Bisin and Verdier, 2001, Bisin, Topa, and Verdier, 2004, Tabellini, 2008; CPR: Schumacher, 2009, Schumacher, 2015, Bezin, 2019), group selection (PG: Boyd and Richerson, 1990, Boyd and Richerson, 2005, Bowles and Gintis, 1998, Mitteldorf and Wilson, 2000, Henrich, 2004; CPR: Waring, Goff, and Smaldino, 2017), peer persuasion (PG: Panebianco, 2016), and contagious cooperation (CPR: Richter, van Soest, and Grasman, 2013, Richter and Grasman, 2013), among others. Methodological, our paper also closely relates to the literature that utilizes the indirect evolutionary approach in PG games to endogenize the formation of preferences for norm-adherence (see, e.g., Fershtman and Weiss, 1998; Traxler and Spichtig, 2011; Alger and Weibull, 2013; Alger and Weibull, 2016) and pro-social preferences in general (see, e.g., Bester and Güth, 1998, Guttman, 2003, Guttman, 2013, Poulsen and Poulsen, 2006, Müller and von Wangenheim, 2019).

transmission and internalization process of a pro-social norm in a partially integrated society. The norm spreads through the imitation of successful individuals and institutional pressure. Individuals are repeatedly matched into pairs, which then face the prisoner's dilemma. Having internalized the norm introduces an additional non-monetary incentive for cooperating. Moreover, individuals who have internalized the norm are more likely to interact with each other due to partial integration in society. In equilibrium, low levels of integration or high institutional pressure are necessary for strict norms to persist, while high levels of integration and low institutional pressure suffice to uphold norms of intermediate strength.

Since this paper examines a CPR game in which the entire society interacts, assortative matching does not influence the situation. Therefore, this paper emphasizes the role of social sanctions as a co-determinant of cultural fitness rather than that of assortative matching. In this respect, we closely relate to Mankat (in press), who studies the co-evolution of personal norms, social norms, and preferences for norm compliance in a binary public goods game. He shows that if the social success that determines norm adoption depends on material and social payoffs, an interplay of social disapproval mechanisms can explain the persistence of cooperation-prescribing personal and social norms. This paper's model is similar to that of Mankat (in press), but expands on it by incorporating descriptive norms and institutional pressure into the dynamic analysis. Moreover, we study norm evolution in a CPR rather than a PG game.

3 Static framework

The model consists of a continuum of individuals $i \in [0,1]$ who recurrently interact in a (binary) CPR game. An individual i's behavior $a_i \in \mathcal{A} = \{0,1\}$ is closely linked to the resource stock $r \in [0,1]$. In particular, each individual can exploit the resource, $a_i = 0$, or not exploit it, $a_i = 1$. If an individual i chooses not to exploit the resource, she incurs material costs $c(r) \geq 0$. Material costs $c(r)$ is a continuous and differentiable function of the resource stock r. Exploitation becomes increasingly profitable as the resource stock rises, $c'(r) > 0$. We write the share of individuals who do not exploit the resource as ψ. Hence, ψ indicates commonly executed behavior in society, the descriptive norm.

Since we study potential resource conservation, we normalize the resource stock to $r \in [0,1]$, where $r = 1$ indicates the maximum resource stock that can persist if no individual exhibits exploitative behavior and $r = 0$ indicates the point of no return so that the resource erodes irrespective of society's behavior. We discuss resource dynamics in more detail in Section 5.

Each individual i holds a personal norm $n_i \in \{0,1\}$ that indicates the behavior i considers morally appropriate. If $n_i = 1$, we say that individual i holds the *sustainability norm* and only considers non-exploitation to be morally appropriate. Alternatively, an individual considers all possible actions morally appropriate. We then say an individual does not hold the sustainability norm and write $n_i = 0$. If an individual believes resource exploitation is morally inappropriate but behaves contrarily, she experiences feelings such as guilt and loss of self-esteem. We capture these feelings by personal sanctions p, where $p > 0$. Hence, given behavior a_i and personal norm n_i, individual i experiences personal sanctions $(a_i - 1)pn_i$.

In line with Cooter (1998) and Mankat (in press), we define the social norm by the distribution of personal norms. We write the proportion of individuals i with the personal norm $n_i = 1$ as ϕ. If ϕ is large, the social norm is strong since many individuals agree that non-exploitation is the only morally appropriate behavior. If an individual exploits the resource, she is subject to social sanctions $s(\phi, \psi)$ in the form of disapproval. These social sanctions depend on the social norm ϕ and the descriptive norm ψ. Specifically, the stronger the social norm is, the more individuals disapprove of exploitation, and, thus, the greater the social consequences for exploiting the resource, $s'_\phi(\phi, \psi) > 0$. Second, the more individuals in society adhere to the social norm, the greater the social sanctions for non-adherence, $s'_\psi(\phi, \psi) > 0$. If the social and descriptive norms are absent, $\phi = \psi = 0$, then social sanctions are too, $s(0,0) = 0$. Moreover, we assume $s(\phi, \psi)$ is continuous and differentiable in both arguments.

4 Equilibrium behavior

This section analyzes equilibrium behavior for an exogenous social norm ϕ and resource stock r. An individual i's behavior depends on her utility function, which is, in turn, subject to

	Equilibrium behavior	Existence condition
1.	$\psi^* = 0$, $(\sigma_1^*, \sigma_0^*) = (0, 0)$	$s(\phi, 0) + p < c(r)$
2.	$\psi^* = \phi$, $(\sigma_1^*, \sigma_0^*) = (1, 0)$	$s(\phi, \phi) < c(r) < s(\phi, \phi) + p$
3.	$\psi^* = 1$, $(\sigma_1^*, \sigma_0^*) = (1, 1)$	$c(r) < s(\phi, 1)$

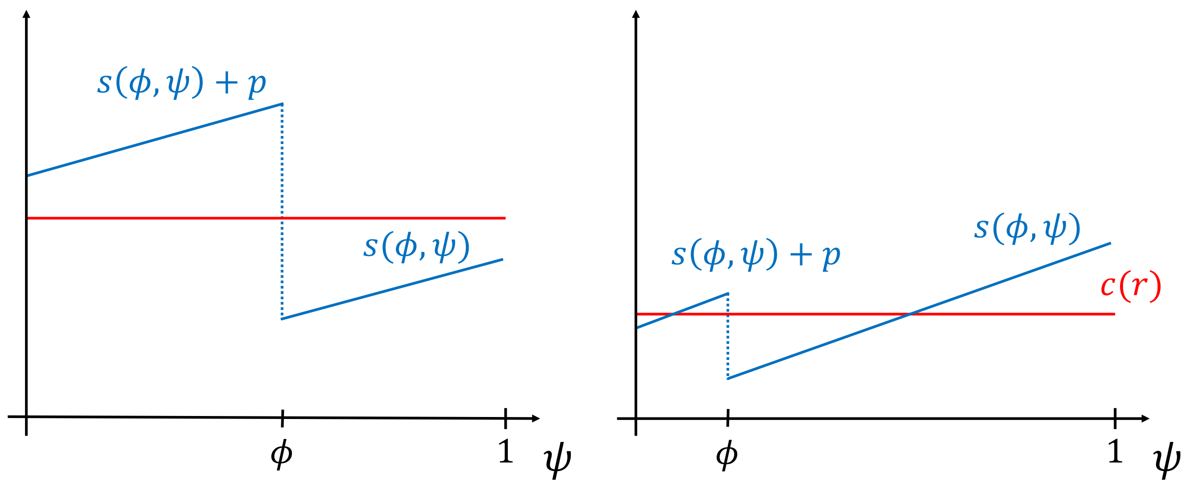
Table 1: Equilibrium behavior

material payoff, personal sanctions, and social sanctions.

Definition 4.1 (Utility)**.** $u_i = a_i(n_i p + s(\phi, \psi) - c(r))$.

Rather than employing the Nash equilibrium concept, we require a *behavioral equilibrium* to be an evolutionary stable state. The underlying idea is that behavior itself is subject to a (very fast) evolutionary process driven by utility improvements (i.e., best-response dynamics). We write the share of individuals i with personal norm $n_i = n$ who do not exploit the resource as σ_n. It follows that $\psi = \phi\sigma_1 + (1 - \phi)\sigma_0$. Throughout, we indicate an equilibrium share of non-exploitation by ψ^*. Moreover, the set $\Psi^*(r, \phi)$ contains all equilibrium shares ψ^* at some resource stock r and social norm ϕ. Table 1 presents all possible behavioral equilibria and when they exist. There are three potential behavioral equilibria: (1) everyone exploits the resource, (2) any individual exploits the resource if and only if she does not hold the sustainability norm, and (3) no individual exploits the resource. Since the existence conditions are not mutually exclusive, different behavioral equilibria potentially co-exist. Moreover, the existence conditions cannot all be violated simultaneously, implying that $\Psi^*(r, \phi)$ is always non-empty. Below, we discuss the results using Figure 1.

Figure 1 depicts two cases for which $s(\phi, \psi)$ is linear in ψ. The red curve $c(r)$ indicates the costs of non-exploitation at resource stock r. The blue curve indicates personal and social sanctions at any exploitation level ψ for the next individual willing to refrain from exploiting the resource. The first ϕ individuals hold the sustainability norm and are, thus, subject to personal and social sanctions, whereas the later $1-\phi$ individuals are solely subject to social sanctions. This generates the discontinuity at ϕ. At any ψ where the blue curve is



(a) $\Psi^*(r,\phi)=\{\phi\}$ (b) $\Psi^*(r,\phi)=\{0,\phi,1\}$

Figure 1: Equilibrium behavior

above (below) the red curve, an individual must be currently (not) exploiting the resource but prefers to behave otherwise. For example, consider Figure 1a and any $\psi<\phi$. Some individual who holds the sustainability norm exploits the resource (since $\psi<\phi$). She incurs personal and social sanctions of $s(\phi,\psi)+p$ while only avoiding material costs of $c(r)$. Consequently, she prefers to change her behavior, inducing ψ to rise. Following this line of argument, it becomes apparent why the behavioral equilibria in Table 1 exist under the stated conditions. Any intersection of the blue and red curve not listed in Table 1 must occur in one of the continuous segments of the blue curve. Figure 1b provides an example of such intersections. Following the above reasoning, such intersections constitute Nash equilibria but cannot be evolutionary stable.[3]

Lastly, note that lower levels of exploitation are more likely to constitute a behavioral equilibrium when material costs $c(\cdot)$ and the resource stock r are low, while personal sanctions p, social sanctions $s(\cdot,\cdot)$, and social norm ϕ are high.

[3]Consider, for example, a slight increase in the share of non-exploiting individuals. Social sanctions increase so that individuals previously indifferent between the two actions now prefer not to exploit the resource and change their behavior accordingly. Society coordinates into a behavioral equilibrium characterized by a smaller level of exploitation.

5 Evolutionary framework

Next, we turn to the evolutionary framework that enables us to study the co-evolution of resource stock r, social norm ϕ, and behavior ψ. Throughout, we write the solution of the dynamic system at time t with initial population state $\rho = (r, \phi, \psi)$ as $\xi(\rho, t)$.

Assumption 5.1. (Solution to the dynamic system) For any population profile $\rho = (r, \phi, \psi)$, the solution $\xi(\rho, t) : [0, 1]^3 \times \mathbb{R}_{\geq 0} \to [0, 1]^3$ satisfies:

1. $\xi(\rho, 0) = (r, \phi, \psi^*)$ and
2. $\xi(\rho, t) = \xi(\lim_{\tilde{t} \to t^-} \xi(\rho, \tilde{t}), 0)\ \forall t > 0$.

Condition 1 of Assumption 5.1 formalizes that behavior reaches an equilibrium before changes in the social norm and resource stock occur. The underlying intuition is that at any population state ρ for which behavior is (possibly) not in equilibrium, society moves into the nearest behavioral equilibrium at infinite speed. The initial solution $\xi(\rho, 0)$ specifies which behavioral equilibrium society coordinates into at any ρ. Section 5.1 discusses it in more detail. Condition 2 of Assumption 5.1 indicates that $\xi(\rho, t)$ can always be represented as an initial solution, implying that behavior is always in equilibrium along $\xi(\rho, t)$.

This introduces possible discontinuities of $\xi(\rho, t)$. If the solution approaches some state for which behavior is not in equilibrium, $\lim_{\tilde{t} \to t^-} \xi(\rho, \tilde{t}) = (\tilde{r}, \tilde{\phi}, \tilde{\psi})$ s.t. $\tilde{\psi} \notin \Psi^*(\tilde{r}, \tilde{\phi})$, then it must experience an abrupt change at t for behavior to be in equilibrium at t, $\xi(\rho, t) \neq (\tilde{r}, \tilde{\phi}, \tilde{\psi}) = \lim_{\tilde{t} \to t^-} \xi(\rho, \tilde{t})$. Condition 2 of Assumption 5.1 implies that if the solution approaches such a discontinuity, the dynamic system behaves as if society reaches the limit $(\tilde{r}, \tilde{\phi}, \tilde{\psi})$ and then immediately jumps into state $\xi((\tilde{r}, \tilde{\phi}, \tilde{\psi}), 0) = (\tilde{r}, \tilde{\phi}, \tilde{\psi}^*)$, for which behavior is in equilibrium $\tilde{\psi}^*$. Although conceptually, we can think of society being in $(\tilde{r}, \tilde{\phi}, \tilde{\psi})$ in the blink of a moment, formally, the solution $\xi(\rho, t)$ never actually reaches this point.

While $\xi(\rho, t)$ may exhibit discontinuities due to jumps in equilibrium behavior ψ^*, the resource stock r and social norm ϕ are continuous along $\xi(\rho, t)$. Their values at time t derive from their initial values and the cumulative changes up to t.

5.1 Behavioral evolution

We continue to describe the evolution of behavior along the solution $\xi(\rho, t)$. Section 4 already established which behavioral equilibria exist at any resource stock r and social norm ϕ. Hence, it remains for us to identify which behavioral equilibrium society reaches.

As a starting point, consider some population profile $\rho = (r, \phi, \psi^*)$ for which behavior is in equilibrium. If society experiences marginal changes in the social norm and resource stock, the behavioral equilibrium persists, and we expect society to remain in it. For example, suppose society is at population state $\rho = (r, \phi, \psi^*)$ for which $s(\phi, \psi^*) < c(r) < s(\phi, \psi^*) + p$. Therefore, equilibrium behavior corresponds to non-exploitation by all (and only) norm holders, $\psi^* = \phi$. For marginal changes in the resource and social norm to $\tilde{r}$ and $\tilde{\phi}$ respectively, the behavioral equilibrium $\tilde{\psi}^* = \tilde{\phi}$ still exists, $s(\tilde{\phi}, \tilde{\phi}) < c(\tilde{r}) < s(\tilde{\phi}, \tilde{\phi}) + p \Rightarrow \tilde{\phi} \in \Psi^*(\tilde{r}, \tilde{\phi})$. Since the previous level of non-exploitation $\psi^* = \phi$ is close to the new behavioral equilibrium $\tilde{\psi}^* = \tilde{\phi}$, it is in its basin of attraction. Consequently, we expect society to coordinate into the behavioral equilibrium $\tilde{\psi}^* = \tilde{\phi}$. Alternatively, suppose that either all individuals prefer to exploit the resource or not, $c(r) > s(\phi, \psi^*) + p$ or $c(r) < s(\phi, \psi^*)$. Hence, $\psi^* \in \{0, 1\}$. Marginal changes in the resource stock and social norm leave the behavioral equilibrium ψ^* undisturbed, and we expect it not to change.

Consequently, behavior remains in a particular equilibrium as long as that equilibrium persists, and the following dynamics describe it during such a period.

Definition 5.1 (Behavioral dynamics for equilibrium behavior)**.**

$$\dot{\psi}^* = \begin{cases} \dot{\phi} & \text{if } s(\phi, \psi^*) < c(r) < s(\phi, \psi^*) + p \\ 0 & \text{else} \end{cases}$$

Throughout, we understand all time derivatives as right-hand derivatives. Hence, they express how the variable changes as time proceeds.

Equilibrium behavior along the solution $\xi(\rho, t)$ is discontinuous if it approaches some $\tilde{\rho} = (\tilde{r}, \tilde{\phi}, \tilde{\psi})$ for which behavior is not in equilibrium, $\lim_{\tilde{t} \to t^-} \xi(\rho, \tilde{t}) = (\tilde{r}, \tilde{\phi}, \tilde{\psi})$ s.t. $\tilde{\psi} \notin \Psi^*(r, \phi)$. Recall that Condition 2 of Assumption 5.1 implies that at any such discontinuity, the system behaves as if it reaches the limit $\lim_{\tilde{t}\to t^-} \xi(\rho,\tilde{t})$ and then jumps into a population profile for which behavior is in equilibrium, $\xi(\lim_{\tilde{t}\to t^-} \xi(\rho,\tilde{t}),0) = (\tilde{r},\tilde{\phi},\tilde{\psi}^*)$.[4] To describe transitions between behavioral equilibria, we thus need to specify which behavioral equilibrium the dynamic system reaches at any ρ. For this purpose, we introduce the following assumptions on the initial solution $\xi(\rho,0)$ for any ρ.

Assumption 5.2 (Initial solution)**.** For any population profile $\rho = (r,\phi,\psi)$ and $\psi^* \in \Psi^*(r,\phi)$:

1. $\xi(\rho,0) = (r,\phi,\psi^*)$ if

   (a) $\psi^* \geq \psi$ and $\exists\epsilon > 0$ s.t. $s(x,\phi) + p\,1_{\leq\phi}(x) \geq c(r)$ $\forall x \in (\psi-\epsilon,\psi^*)\cap[0,1]$ or

   (b) $\psi^* \leq \psi$ and $\exists\epsilon > 0$ s.t. $s(x,\phi) + p\,1_{\leq\phi}(x) \leq c(r)$ $\forall x \in (\psi^*,\psi+\epsilon)\cap[0,1]$,

   where the function $1_{\leq\phi}(x)$ returns 1 if $x \leq \phi$ and 0 if $x > \phi$.

2. $\xi(\rho,0) \neq (r,\phi,\psi^*)$ if $\phi \in \Psi^*(r,\phi)$ and $\exists x \in (0,1)$ s.t. $x\psi + (1-x)\psi^* = \phi$.

The intuitive reasoning for Condition 1 of Assumption 5.2 closely follows that of the behavioral equilibria in Section 4. Recall Figure 1 and note that $c(r)$ and $s(x,\phi) + p\,1_{\leq\phi}(x)$ correspond to the red and blue curves, respectively. For illustration purposes, we focus on Condition 1a of Assumption 5.2. The underlying reasoning for Condition 1b is The analogous. Consider any starting point ψ so that for all levels of non-exploitation x below the equilibrium share ψ^*, $x < \psi^*$, and sufficiently close to ψ, $x > \psi - \epsilon$, the blue curve $s(x,\phi) + p\,1_{\leq\phi}(x)$ lies above the red curve $c(r)$. By similar reasoning as in Section 4, the share of non-exploiting individuals must increase at any such x. Hence, behavior reaches the behavioral equilibrium ψ^*. Note that for single points $x \in (\psi-\epsilon,\psi^*)$, we allow for $c(r) = s(x,\phi) + p\,1_{\leq\phi}(x)$. Although behavior is at rest at such x, x is not a behavioral equilibrium. Moreover, it must hold that the non-exploitation share increases for all points in some close neighborhood around x. In the presence of constant small mutations, behavior eventually passes x and reaches the behavioral equilibrium ψ^*. Note that Condition 1 of Assumption 5.2 implies that if ψ is a behavioral equilibrium, then society reaches it, $\psi^* = \psi$.

[4]Note that for all $t > 0$, $\lim_{x\to t^-} \xi(\rho,x)$ always exists since (1) r and ϕ are perfectly continuous along $\xi(\rho,t)$ and (2) ψ^* can be represented in a piecewise manner, with each segment coinciding with either 0, 1, or ϕ as described by $\xi(\rho,t)$, all of which are perfectly continuous.

Whenever ψ is not a behavioral equilibrium but a Nash equilibrium, $\xi(\rho, 0)$ is possibly not uniquely defined by Condition 1 of Assumption 5.2. If so, we remain relatively agnostic about which behavioral equilibrium society reaches. Condition 2 of Assumption 5.2 only requires that society does not pass through a behavioral equilibrium $\phi \in \Psi^*(r, \phi)$ to reach another one. For example, suppose $\psi \in (\phi, 1)$ is the Nash equilibrium of the right intersection of the blue and red curves in Figure 1b. Constant small mutations disrupt ψ, and society coordinates into one of the two neighboring equilibria, $\psi^* \in \{\phi, 1\}$.

For analytical purposes, we treat $\xi(\rho, 0)$ as uniquely defined and address the possible non-uniqueness by allowing $\xi(\rho, 0)$ to take any value consistent with Assumption 5.2.[5]

5.2 Resource evolution

Resource evolution depends on equilibrium behavior ψ^* and resource stock r.

Definition 5.2 (Resource dynamics)**.**

$$\dot{r} = \delta(r) - e(\psi^*),$$

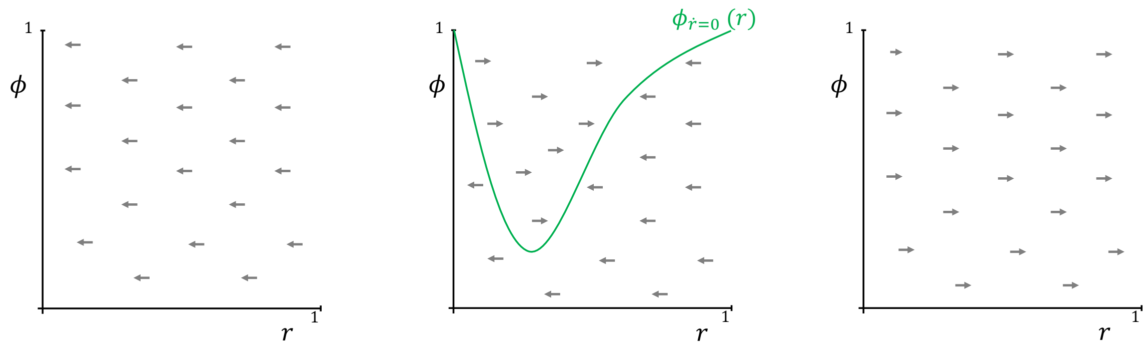
where $\delta(1) = 0$, $\delta(r) > 0\ \forall r \in (0, 1)$, $e(1) = 0$, $e'(\psi^*) < 0\ \forall \psi^* \in [0, 1]$, $\max_{r \in [0,1]} \delta(r) < e(0)$, and $e(\cdot)$ and $\delta(\cdot)$ are continuous and differentiable.

The first part of the resource dynamics captures the natural growth of the resource, $\delta(r)$. If the resource is at maximum capacity, it cannot grow further, $\delta(1) = 0$. For all other positive stocks, natural resource growth is positive, $\delta(r) > 0\ \forall r \in (0, 1)$.[6] The second part of the resource dynamics captures resource extraction by society, $e(\psi^*)$. The more individuals behave sustainably, the smaller is resource extraction, $e'(\psi^*) < 0\ \forall \psi^* \in [0, 1]$. We assume $\max_{r \in [0,1]} \delta(r) < e(0)$, implying that the resource always deteriorates if everyone exploits it.[7] In the following, we discuss resource dynamics at each behavioral equilibrium of Section

[5] Note that for any ρ, the non-uniqueness of $\xi(\rho, t)$ only arises at time $t = 0$. Hence, although Condition 1 of Assumption 5.2 sometimes does not specify where exactly $\xi(\rho, t)$ starts off (i.e., at which behavioral equilibrium), it unambiguously characterizes it (and all possible discontinuities) after the initial time point.

[6] One popular functional form corresponds to $\delta(r) = \tilde{\delta}(r) * r$, where $\tilde{\delta}$ constitutes the natural growth rate of the resource.

[7] Otherwise, the situation becomes inherently uninteresting, as there is a positive resource stock $\tilde{r} > 0$ below which the resource may never fall.



(a) $\psi^* = 0$ (b) $\psi^* = \phi$ (c) $\psi^* = 1$

Figure 2: Resource evolution

5.1. Figure 2 presents graphical illustrations for the exemplary case of $\delta(r)$ having a unique maximum and $\delta(0) = 0$.

First, if no individual exploits, $\psi^* = 1$, then resource extraction is minimal, $e(1) = 0$. If the resource stock is below maximum capacity and above the point of no return, it must be increasing, $\dot{r} > 0 \ \forall r \in (0, 1)$. Resource dynamics are at rest at maximum capacity, $\dot{r} = 0$ if $r = 1$. Analogously, if all individuals exploit the resource, $\psi^* = 0$, then resource extraction is at a maximum, and the resource stock must be diminishing, $\dot{r} < 0 \ \forall r \in [0, 1]$. Lastly, we look at partial exploitation, $\psi^* = \phi$. Depending on the share of individuals who hold the sustainability norm and, thus, do not exploit, resource extraction $e(\phi)$ is either larger or smaller than the natural growth of the resource $\delta(r)$. Given equilibrium behavior $\psi^* = \phi$, let the nullcline $\phi_{\dot{r}=0}(r)$ return a social norm ϕ for each resource stock r so that $\delta(r) = e(\psi^*)$. Hence, resource dynamics are at rest for any point on $\phi_{\dot{r}=0}(r)$. Suppose that at some resource stock r the social norm ϕ lies below $\phi_{\dot{r}=0}(r)$. So few individuals hold the sustainability norm and refrain from exploitation that resource extraction $e(\phi)$ outweighs natural resource growth $\delta(r)$. The resource stock diminishes. Analogously, the resource stock rises if the social norm exceeds $\phi_{\dot{r}=0}(r)$.

5.3 Cultural evolution

Following the existing literature, we focus on norm change through horizontal and oblique cultural transmission (Mengel, 2008; Mankat, in press). Interactions between peers and cultural imitation characterize horizontal transmission, while oblique transmission involves a learning process facilitated by socialization institutions such as schools and media. We assume that these transmission processes are biased regarding the cultural fitness of individuals. The underlying idea is that socially successful individuals have a more significant influence on the opinion formation of their peers as they are more likely to be imitated (Henrich and Gil-White, 2001) and have better prospects of occupying influential roles positions such as teachers or politicians (Bowles and Gintis, 1998). In line with Traxler and Spichtig (2011) and Mankat (in press), cultural fitness results from both material and social factors.

Definition 5.3 (Cultural fitness)**.** An individual i's cultural fitness is

$$f_i = a_i(\kappa s(\phi, \psi^*) - c(r)),$$

where $\kappa > 0$ is the weight of social sanctions on cultural fitness.

We model fitness-biased norm adoption using the well-known replicator dynamics. Hence, the resulting change in the social norm at any population profile ρ is proportional to $\phi(1-\phi)[(\sigma_1 - \sigma_0)(\kappa s(\phi, \psi^*) - c(r))$.

In addition, institutions can promote the adoption of the sustainability norm by structuring interactions, stigmatizing and promoting behavior through legal norms and policies, and influencing perceptions through (public) communication and socialization institutions such as schools, universities, and media (Gintis, 2003; Mengel, 2008). Following Gintis (2003) and Mengel (2008), we capture the degree to which institutions foster norm adoption by institutional pressure $\gamma \geq 0$. Institutional pressure induces a share of individuals to adopt the sustainability norm as their personal one. Moreover, the effectiveness of institutional pressure γ is proportional to the share of individuals in society who consider the behavior morally right, the social norm.[8] Formally, the change in the social norm from institutional pressure corresponds to $\phi(1-\phi)\gamma \geq 0$. Combining the above yields the cultural dynamics below.

[8] Gintis (2003) and Mengel (2008) point out that proportionality is a rather conservative assumption.

Definition 5.4 (Cultural dynamics)**.**

$$\dot{\phi} = v\phi(1-\phi)[(\sigma_1 - \sigma_0)(\kappa s(\phi, \psi^*) - c(r)) + \gamma],$$

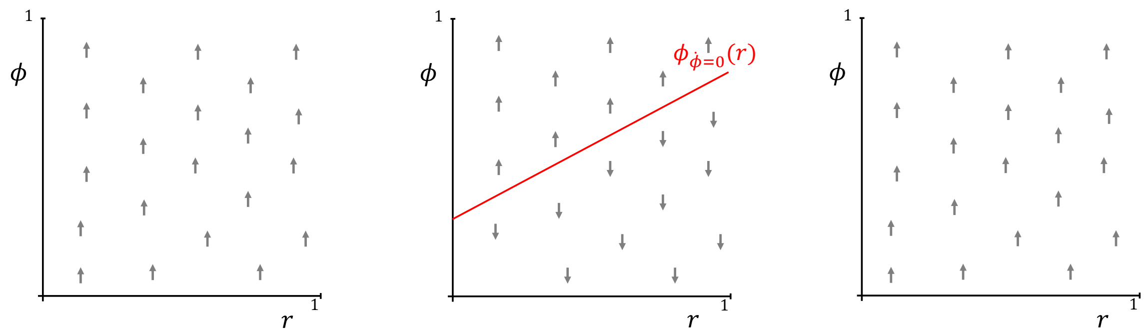
where v is the relative speed of cultural change.

As with resource dynamics, we discuss cultural dynamics for each possible behavioral equilibrium ψ^*. Figure 3 presents exemplary graphical illustrations. We start with the case of either all or no individual exploiting the resource, $\psi^* \in \{0, 1\}$. Since all individuals behave equally, all individuals incur the same material costs $c(r)$ and social sanctions $s(\phi, \psi^*)$. All individuals have the same cultural fitness, implying that cultural dynamics are solely driven by institutional pressure γ, $\dot{\phi} = v\phi(1-\phi)\gamma$.

Next, we turn to the behavioral equilibrium where an individual exploits the resource if and only if she does not hold the sustainability norm, $\psi^* = \phi$. From Definition 5.4, we can easily derive that cultural dynamics are at rest if either (1) all individuals hold the same personal norm, $\phi \in \{0, 1\}$ or (2) the difference in cultural fitness of both behavioral routines offsets institutional pressure, $\kappa s(\phi, \psi^*) + \gamma = c(r)$. Let $\phi_{\dot{\phi}=0}(r)$ be the nullcline, which for each resource stock $r \in (0, 1]$ returns a social norm ϕ solving $\kappa s(\phi, \psi^*) + \gamma = c(r)$. Hence, cultural dynamics are at rest at any point on the nullcline. Note that $\phi_{\dot{\phi}=0}(r)$ is strictly increasing. A larger resource stock r corresponds to greater costs of non-exploitation $c(r)$. This requires more social sanctions $\kappa s(\phi, \phi)$ and, thus, a stronger social norm ϕ for cultural dynamics to be at rest. Along these lines of argument, suppose that at some resource stock r, the social norm ϕ exceeds $\phi_{\dot{\phi}=0}(r)$. Social sanctions $\kappa s(\phi, \phi)$ and institutional pressure γ outweigh material costs of non-exploitation $c(r)$, and the social norm strengthens, $\dot{\phi} > 0$. The analogous holds if ϕ is below $\phi_{\dot{\phi}=0}(r)$.

6 Evolutionary analysis

In this section, we analyze the evolutionary model of Section 5. We identify different equilibrium sets, analyze their asymptotic stability, and discuss the roles of different cultural and socio-economic factors. Throughout, we adopt the following equilibrium notion.



(a) $\psi^* = 0$ (b) $\psi^* = \phi$ (c) $\psi^* = 1$

Figure 3: Cultural evolution

Definition 6.1 (Socio-ecological equilibrium)**.** A *socio-ecological equilibrium* is a closed set $\Omega \subset [0,1]^3$ s.t. $\forall \rho \in \Omega$,

- $\xi(\rho, t) \in \Omega$ for all $t \geq 0$ and
- $\nexists \hat{\Omega} \subset \Omega$ s.t. $\hat{\Omega}$ is a socio-ecological equilibrium.

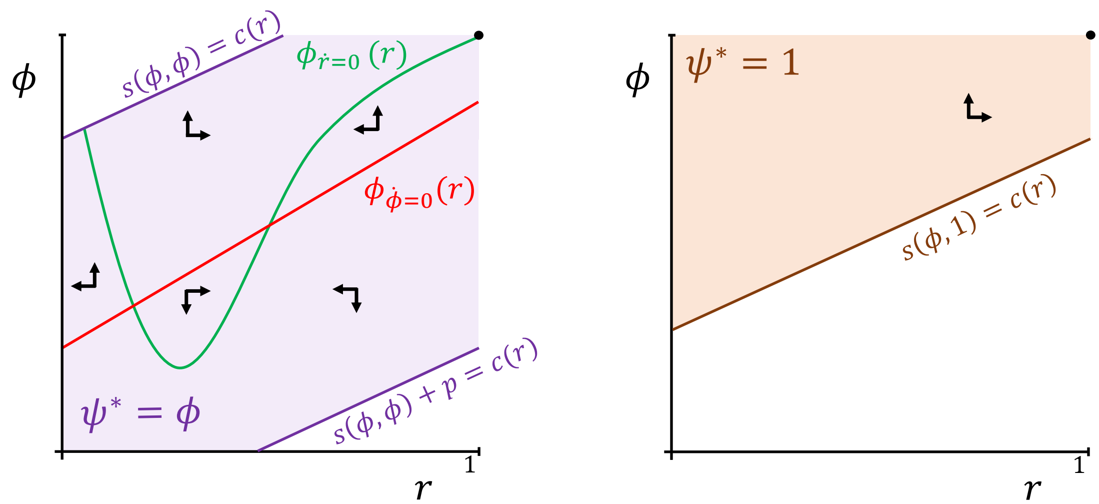
Hence, a socio-ecological equilibrium corresponds to a closed set $\Omega \subset [0,1]^3$, for which the dynamic system of Section 5 starting at any $\rho \in \Omega$ always remains in Ω. Moreover, Ω is minimal in the sense that it does not contain a strict subset $\hat{\Omega}$, which is itself a socio-ecological equilibrium. Our analysis identifies two different types of equilibria, namely: (1) equilibrium (or rest) points and (2) limit cycles.

6.1 Boundary equilibrium point

We first analyze a socio-ecological equilibrium for which the resource stock is at maximum capacity, $r = 1$, all individuals hold the sustainability norm, $\phi = 1$, and all individuals do not exploit the resource, $\psi^* = 1$.

Proposition 6.1 (Boundary equilibrium point)**.** $(1,1,1) \in (0,1]^3$ *is a socio-ecological equilibrium if* $c(1) < s(1,1) + p$.

Proof: See Appendix A.



(a) $s(1,1) < c(1) < \kappa s(1,1) + \gamma$ (b) $s(1,1) > c(1)$ and $\gamma > 0$

Figure 4: Boundary equilibrium point

Proposition 6.1 states that this point is a socio-ecological equilibrium if all individuals not exploiting the resource is indeed a behavioral equilibrium, $c(1) < s(1,1) + p \Rightarrow \psi^* = \phi = 1$. Given this equilibrium behavior, $\psi^* = 1$, resource dynamics are at rest if and only if the resource is at maximum capacity, $\dot{r} = 0 \Leftrightarrow r = 1$. Moreover, since all individuals hold the sustainability norm, cultural dynamics are at rest too, $\phi = 1 \Rightarrow \dot{\phi} = 0$.

Proposition 6.2 (Asymptotic stability of a boundary equilibrium point)**.** *Suppose* $(1,1,1)$ *is a socio-ecological equilibrium of Proposition 6.1.* $(1,1,1)$ *is asymptotically stable if*

1. $s(1,1) \leq c(1) < \kappa s(1,1) + \gamma$ *or*

2. $s(1,1) > c(1)$ *and* $\gamma > 0$.

Proof: The proposition follows from Lemma 6.1 and Lemma 6.2.

The above proposition summarizes the results on asymptotic stability of the boundary equilibrium point $(1,1,1)$. To discuss Proposition 6.2 in more detail, we differentiate two cases based on the material costs of non-exploitation. First, we look at the case of the material costs of non-exploitation at maximum resource capacity exceeding social sanctions at the perfect social norm and no exploitation, $s(1,1) \leq c(1)$.

Lemma 6.1. *Suppose* $s(1,1) \leq c(1) < s(1,1) + p$*. The socio-ecological equilibrium* $(1,1,1)$ *is asymptotically stable if* $\kappa s(1,1) + \gamma > c(1)$.

Proof: See Appendix A.

Lemma 6.1 states that a socio-ecological equilibrium $(1,1,1)$ is asymptotically stable if social sanctions on cultural fitness and institutional pressure outweigh material costs of non-exploitation, $\kappa s(1,1) + \gamma > c(1)$. The following discussion focuses on the case of social sanctions being strictly smaller than material costs of non-exploitation, $s(1,1) < c(1)$.[9] Hence, for all $\hat{\rho} = (\hat{r}, \hat{\phi}, \hat{\psi})$ in some neighborhood U of $(1,1,1)$, the behavioral equilibrium of partial exploitation exists, $s(\hat{\phi}, \hat{\phi}) < c(\hat{r}) < s(\hat{\phi}, \hat{\phi}) + p \Rightarrow \hat{\phi} \in \Psi^*(\hat{r}, \hat{\phi})$, and society always coordinates into it, $s(\hat{\phi}, \hat{\psi}) < c(\hat{r}) < s(\hat{\phi}, \hat{\psi}) + p \Rightarrow \xi((\hat{r}, \hat{\phi}, \hat{\psi}), 0) = (\hat{r}, \hat{\phi}, \hat{\phi})$. Thus, an individual does not exploit the resource if and only if she holds the sustainability norm, $\hat{\psi}^* = \hat{\phi}$. Moreover, if U is sufficiently small, then at any population state $\hat{\rho} = (\hat{r}, \hat{\phi}, \hat{\phi}) \in U$, social sanctions and institutional pressure outweigh material costs of non-exploitation, $\kappa s(\hat{\phi}, \hat{\phi}) + \gamma > c(\hat{r})$. Cultural transmission induces the norm non-holders to adopt the sustainability norm so that the social norm strengthens, $\dot{\hat{\phi}} > 0$, and returns to $\phi = 1$. Since no individual exploits the resource at the perfect social norm, $\psi^* = \phi = 1$, the resource stock recovers and reaches its maximum capacity, $r = 1$. Hence, society returns to the socio-ecological equilibrium $(1,1,1)$. Figure 4a provides a graphical illustration. Note that $\kappa s(1,1) + \gamma > c(1)$ implies that $\phi_{\dot{\phi}=0}(1) < 1$. Next, we turn to the case of non-exploitation being relatively cheap, $c(1) < s(1,1)$.

Lemma 6.2. *Suppose* $c(1) < s(1,1)$*. The socio-ecological equilibrium* $(1,1,1)$ *is asymptotically stable if and only if* $\gamma > 0$.

Proof: See Appendix A.

Lemma 6.2 states that in this case, the socio-ecological equilibrium $(1,1,1)$ is asymptotically stable if (and only if) institutional pressure exists, $\gamma > 0$. Given that material costs of non-exploitation are small, $c(1) < s(1,1)$, there is some neighborhood U of $(1,1,1)$ so

[9] The reasoning for the alternative case of $s(1,1) = c(1)$ is similar in intuition but somewhat more elaborate.

that at each population profile $\hat{\rho} = (\hat{r}, \hat{\phi}, \hat{\psi})$ in that neighborhood, the behavioral equilibrium of no exploitation exists, $c(\hat{r}) < s(\hat{\phi}, \hat{\psi}) \Rightarrow 1 \in \Psi^*(\hat{r}, \hat{\phi})$, and society coordinates into it, $c(\hat{r}) < s(\hat{\phi}, \hat{\psi}) \Rightarrow \xi((\hat{r}, \hat{\phi}, \hat{\psi}), 0) = (\hat{r}, \hat{\phi}, 1)$. Since all individuals refrain from resource exploitation, $\hat{\psi}^* = 1$, the resource stock increases, $\dot{\hat{r}} > 0$. Moreover, since all individuals behave alike, cultural evolution is solely driven by institutional pressure. If institutional pressure exists, $\gamma > 0$, the social norm strengthens, $\dot{\hat{\phi}} > 0$. Hence, society evolves towards the socio-ecological equilibrium $(1, 1, 1)$. Figure 4b presents a graphical illustration.

Note that positive institutional pressure, $\gamma > 0$, is necessary for asymptotic stability of the socio-ecological equilibrium $(1, 1, 1)$. This holds since in the absence of institutional pressure, $\gamma = 0$, the social norm is at rest if no individual exploits the resource, $\psi^* = 1 \Rightarrow \dot{\phi} = 0$. Hence, it is at rest at any population profile $\hat{\rho}$ in some neighborhood U of the socio-ecological equilibrium $(1, 1, 1)$. At any such $\hat{\rho}$, society only experiences an increase in the resource stock, $\dot{\hat{r}} > 0$. All population profiles for which the resource stock is at maximum capacity, $\hat{r} = 1$, and non-exploitation by all individuals is a behavioral equilibrium, $1 \in \Psi^*(1, \phi) \Leftrightarrow c(1) < s(\phi, 1)$, form a connected set of rest points, $\Omega = \{(1, \phi, 1)\}_{s(\phi,1)>c(1)}$. Each point in this set is Lyapunov stable but not asymptotically stable. It appears a natural next question to ask when Ω is (part of) a dynamically stable set. We can easily show that the closure of Ω is never asymptotically stable. Suppose society experiences random mutation across Ω, eventually reaching population profile $\tilde{\rho} = (1, \tilde{\phi}, 1)$ at the border of Ω, $s(\tilde{\phi}, 1) = c(\tilde{r})$. No-exploitation no longer constitutes a behavioral equilibrium, $1 \notin \Psi^*(1, \tilde{\phi})$. Society coordinates into a new behavioral equilibrium, $\psi^* < 1$, at which the resource stock must be decreasing, $r = 1 \wedge \psi^* < 1 \Rightarrow \dot{r} < 0$. The dynamic system evolves away from Ω.

The results of this section indicate that larger institutional pressure γ, weight of social sanctions on cultural fitness κ, and personal sanctions p unambiguously favor resource conservation through a boundary equilibrium point. Larger p favors its existence, whereas larger κ and γ favor asymptotic stability. The roles of social sanctions $s(\cdot, \cdot)$ and material costs of non-exploitation $c(\cdot)$ are more ambiguous. Although larger $s(1, 1)$ and smaller $c(1)$ favor the existence of the socio-ecological equilibrium, they may alter equilibrium behavior in its neighborhood. In the absence of institutional pressure, $\gamma = 0$, this may harm asymptotic stability as society moves from the case of Lemma 6.1 to that of Lemma 6.2. However, in the presence of institutional pressure, $\gamma > 0$, the changes in $s(1,1)$ and $c(1)$ favor asymptotic stability.

6.2 Interior equilibrium point

Next, we investigate equilibrium points for which the resource stock is (potentially) below maximum capacity, $r < 0$.

Proposition 6.3 (Interior equilibrium point). *Any* $(r, \phi, \phi) \in (0,1]^3$ *is a socio-ecological equilibrium if*

1. $s(\phi, \phi) < c(r) < s(\phi, \phi) + p$,
2. $\kappa s(\phi, \phi) + \gamma = c(r)$, *and*
3. $\delta(r) = e(\phi)$.

Proof: See Appendix A.

Consider any point $(r, \phi, \phi) \in (0,1]^3$ that satisfies the conditions of Proposition 6.3. The first condition implies that non-exploitation by all and only norm holders indeed corresponds to a behavioral equilibrium, $\psi^* = \phi$. Given this equilibrium behavior, Section 5.2 established that resource dynamics are at rest for any point on the nullcline $\phi_{\dot{r}=0}(r)$. Formally, social norm ϕ and resource stock r are on this nullcline if the resource growth equals resource extraction, $\delta(r) = e(\phi)$. Similarly, cultural dynamics are at rest on the nullcline $\phi_{\dot{\phi}=0}(r)$, which is formally equivalent to $\kappa s(\phi, \phi) + \gamma = c(r)$. Social sanctions on cultural fitness and institutional pressure equal the material costs of non-exploitation. If both conditions hold, then (r, ϕ) corresponds to an intersection of the two nullclines, and the dynamic system is at rest. Hence, (r, ϕ, ϕ) is a socio-ecological equilibrium. Figure 5 provides multiple graphical examples of interior equilibrium points.

Note that such an interior equilibrium point can only exist at some social norm ϕ if social and personal sanctions exceed social sanctions on cultural fitness and institutional pressure, which exceed social sanctions, $s(\phi, \phi) < \kappa s(\phi, \phi) + \gamma < s(\phi, \phi) + p$. This is necessary for the behavioral equilibrium of partial exploitation to exist at ϕ and cultural dynamics to be at rest. Graphically, the condition ensures that the nullcline $\phi_{\dot{\phi}=0}(r)$ lies between the behavioral boundaries where $s(\phi,\phi) = c(r)$ and $s(\phi,\phi) + p = c(r)$ in Figure 5. For any social norm ϕ, this condition is more likely to hold if personal sanctions p are large, as this increases the domain of points (r,ϕ) in Figure 5 for which partial exploitation corresponds to equilibrium behavior. Moreover, the condition is more likely satisfied if the weight of social sanctions on cultural fitness κ and institutional pressure γ jointly aggregate to an intermediate value. If κ and γ are either very large (e.g., $\gamma > p \wedge \kappa > 1$ or $\gamma > p + s(1,1)$) or very small (e.g., $\gamma = 0 \wedge \kappa < 1$), the condition may not hold for any social norm. Suppose the condition does hold at some social norm ϕ. In that case, we can easily show that if the social norm lies above some lower bound (i.e., $\phi \geq \min(\phi_{\dot{r}=0}(r))$), then there are cost curves $c(\cdot)$ for which an interior equilibrium point with social norm ϕ exists.

Proposition 6.4 (Asymptotic stability of an interior equilibrium point)**.** *Any socio-ecological equilibrium* $(r,\phi,\phi) \in (0,1]^3$ *of Proposition 6.3 is asymptotically stable if*

1. $0 < \frac{\delta'(r)}{e'(\phi)} < \frac{c'(r)}{\kappa(s'_\phi(\phi,\phi)+s'_\psi(\phi,\phi))}$ *and*

2. $v\phi(1-\phi)\kappa\left(s'_\phi(\phi,\phi) + s'_\psi(\phi,\phi)\right) < -\delta'(r)$.

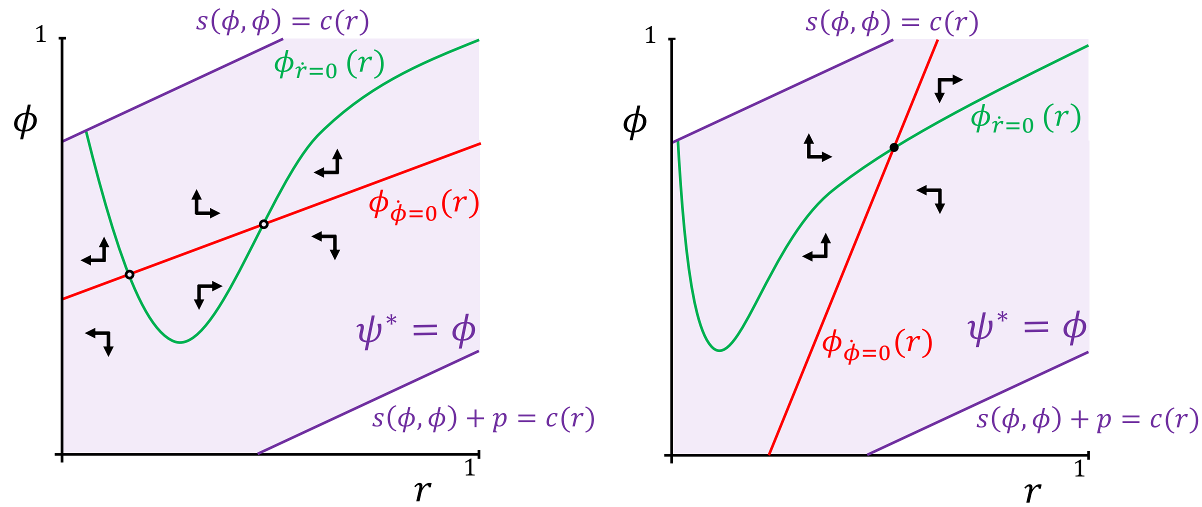
Proof: See Appendix A.

Proposition 6.4 introduces conditions that ensure asymptotic stability of an interior equilibrium point of Proposition 6.3. Recall that at such an equilibrium point (r,ϕ,ϕ), individuals exploit the resource if and only if they do not hold the sustainability norm, $\psi^* = \phi \Leftrightarrow s(\phi,\phi) < c(r) < s(\phi,\phi) + p$. For all $\hat{\rho} = (\hat{r},\hat{\phi},\hat{\psi})$ in some neighborhood U of a socio-ecological equilibrium (r,ϕ,ϕ), the behavioral equilibrium of partial exploitation exists, $s(\hat{\phi},\hat{\phi}) < c(\hat{r}) < s(\hat{\phi},\hat{\phi}) + p \Rightarrow \hat{\phi} \in \Psi^*(\hat{r},\hat{\phi})$, and society coordinates into it, $s(\hat{\phi},\hat{\psi}) < c(\hat{r}) < s(\hat{\phi},\hat{\psi}) + p \Rightarrow \xi((\hat{r},\hat{\phi},\hat{\psi}),0) = (\hat{r},\hat{\phi},\hat{\phi})$. Below, we focus on the dynamic system in that neighborhood U. Hence, throughout, all individuals exploit the resource if and only if they do not hold the sustainability norm, $\hat{\psi}^* = \hat{\phi}$. Moreover, recall that the interior equilibrium point corresponds to an intersection of the two nullclines $\phi_{\dot{\phi}=0}(r)$ and $\phi_{\dot{r}=0}(r)$. Their respective derivatives are $\phi'_{\dot{r}=0}(r) = \frac{\delta'(r)}{e'(\phi)}$ and $\phi'_{\dot{\phi}=0}(r) = \frac{c'(r)}{\kappa\left(s'_\phi(\phi,\phi)+s'_\psi(\phi,\phi)\right)}$.

Condition 1 of Proposition 6.4 requires that the socio-ecological equilibrium constitutes an intersection of the two nullclines in the increasing segment of $\phi_{\dot{\phi}=0}(r)$, $0 < \frac{\delta'(r)}{e'(\phi)}$. Recall from Section 5.3 that the social norm evolves away from the nullcline $\phi_{\dot{\phi}=0}(r)$. Moreover, Section 5.2 shows that the resource evolves away from the decreasing segment of the nullcline $\phi_{\dot{r}=0}(r)$. Since aggregate resource extraction decreases in the share of non-exploiting individuals, $e'(\phi) < 0$, the nullcline $\phi_{\dot{r}=0}(r)$ is decreasing at resource stock r if and only if an increase (decrease) in the resource stock leads to greater (lower) resource growth, $\phi'_{\dot{r}=0}(r) < 0 \Leftrightarrow \delta'(r) > 0$. If the socio-ecological equilibrium was not an intersection of the two nullclines in the increasing segment of $\phi_{\dot{r}=0}(r)$, resource and cultural evolution would favor a move away from it at any $\hat{\rho}$ in some close neighborhood U, implying that (r, ϕ, ϕ) is asymptotically unstable. The left intersection in Figure 5a presents a corresponding example.

Furthermore, the first condition also requires that $\phi_{\dot{\phi}=0}(r)$ intersects $\phi_{\dot{r}=0}(r)$ from below, $\frac{\delta'(r)}{e'(\phi)} < \frac{c'(r)}{\kappa\,(s'_{\phi}(\phi,\phi)+s'_{\psi}(\phi,\phi))}$. For illustration purposes, suppose that this was not the case. The right intersection in Figure 5a constitutes a graphical example. Large values of $-\delta'(r)$ imply that an increase (decrease) in the resource stock drastically decreases (increases) resource growth below (above) zero. Moreover, if $e'(\phi)$ is close to zero, then marginal changes in the social norm, and thus the share of individuals not exploiting the resource ψ^*, only affect resource extraction slightly. Hence, at any population profile $(\hat{r}, \hat{\phi}, \hat{\psi})$ close to (r, ϕ, ϕ), resource evolution drives society towards some stock relatively close to r. Graphically, $\phi_{\dot{r}=0}(r)$ is steep. In conjunction with the costs of non-exploitation being relatively insensitive to changes in the resource stock, so $c'(r)$ is small, material costs of non-exploitation remain relatively steady. However, a large value of $\kappa s'(\phi)$ corresponds to social sanctions being very responsive to changes in the social norm. Hence, at any $(\hat{r}, \hat{\phi}, \hat{\phi})$ close to (r, ϕ, ϕ), there is a strong cultural push away from the social norm ϕ. Combining these insights implies that (r, ϕ, ϕ) is a saddle point.

Lastly, suppose Condition 1 of Proposition 6.4 holds. Figure 5b illustrates this case graphically. If Condition 1 of Proposition 6.4 holds, then the socio-ecological equilibrium is asymptotically stable if resource dynamics change more drastically for marginal changes in the resource than cultural dynamics change for marginal changes in the social norm. Condition 2 of Proposition 6.4 captures this. Recall that cultural evolution favors a move



(a) Unstable equilibria

(b) Potentially stable equilibrium

Figure 5: Interior equilibrium points

away, whereas resource evolution favors a move towards the socio-ecological equilibrium. If Condition 2 of Proposition 6.4 holds, then, after slight changes in the resource stock and social norm, the stabilizing resource dynamics dominate the destabilizing cultural dynamics. The system spirals inwards. Alternatively, if the second condition in the lemma is violated, the evolutionary force driving society away from the socio-ecological equilibrium dominates, and the dynamic system spirals outwards.

The above results offer insights into when society may conserve the CPR through an interior equilibrium point. An interior equilibrium point more likely exists for large personal sanctions p. Moreover, the weight of social sanctions on cultural fitness κ and institutional pressure γ must balance the costs of non-exploitation $c(r)$. Proposition 6.4 sheds light onto when an existing interior equilibrium point is likely asymptotically stable, namely if culture evolves relatively slow (v is small), social sanctions are relatively non-responsive to changes in the social and descriptive norms (s'_ϕ and s'_ψ are small), material costs are relatively responsive to changes in the resource stock ($c'(r)$ is large), and the weight of social sanctions on cultural fitness is small (κ is small). The last bit implies that an interior equilibrium point is more likely asymptotically stable if cultural transmission mainly occurs through institutional pressure γ. For illustration purposes, suppose that social sanctions do not affect cultural fitness, $\kappa = 0$. In that case, changes in the social norm do not impact cultural evolution. The social norm spreads due to institutional pressure γ and erodes due to cultural fitness differences, which entirely derive from material costs $c(r)$. These two forces balance at an interior equilibrium point, $\gamma = c(r)$. Graphically, $\phi_{\dot{\phi}=0}(r)$ corresponds to a vertical line passing through the socio-ecological equilibrium. For resource stocks $\hat{r}$ above (below) the equilibrium value r, the material costs exceed (fall short of) institutional pressure, $c(\hat{r}) > \gamma$ ($c(\hat{r}) < \gamma$), implying a norm decrease (increase). As long as the socio-ecological equilibrium occurs in the increasing segment of $\phi_{\dot{r}=0}(r)$, it is asymptotically stable.

6.3 Boundary limit cycle

Finally, we discuss a socio-ecological equilibrium resulting from the existence of a limit cycle. Proposition 6.5 describes such a limit cycle with orbit Ω_{BLC}. Consequently, the minimal closed set containing Ω_{BLC}, $cl(\Omega_{BLC})$, constitutes a socio-ecological equilibrium.[10]

Proposition 6.5 (Boundary limit cycle)**.** *If* $c(0) < s(1,0) + p < s(1,1) + p \leq c(1)$*, then there is some* $\hat{r}, \check{r} \in (0,1]$ *s.t.*

- $\forall \rho \in \Omega_{BLC} := \{(r,1,1)\}_{\check{r} \leq r < \hat{r}} \cup \{(r,1,0)\}_{\check{r} < r \leq \hat{r}}$*,* $\exists t > 0$ *s.t.* $\xi(\rho, 0) = \xi(\rho, t)$ *and*
- $cl(\Omega_{BLC}) = \{(r,1,1)\}_{\check{r} \leq r \leq \hat{r}} \cup \{(r,1,0)\}_{\check{r} \leq r \leq \hat{r}}$ *is a socio-ecological equilibrium.*

Proof: See Appendix A.

Proposition 6.5 specifies that the limit cycle exists in the case of (1) the costs of non-exploitation at the full resource stock outweighing personal and social sanctions at the perfect social norm and no exploitation, $s(1,1) + p \leq c(1)$, and (2) personal and social sanctions at the perfect social norm and full exploitation outweighing the costs of non-exploitation at the minimum resource stock, $c(0) < p + s(1,0)$. Throughout the course of evolution, all individuals hold the sustainability norm, $\phi = 1$, implying that cultural dynamics are always

[10]Note that further limit cycles are possible, depending on the exact specifications of the dynamic system. We require extensive information about the dynamic system to ensure the existence and stability of such limit cycles. Due to the general approach taken, the detailed investigation of these limit cycles is beyond the scope of this paper.

at rest, $\dot{\phi} = 0$. Since all individuals hold the same personal norm, $\phi = 1$, society can only reach behavioral equilibria for which all individuals behave alike, $\psi^* \in \{0, 1\}$. Whenever no individual exploits, $\psi^* = 1$, the resource stock grows, $\dot{r} > 0$. Eventually, the resource stock reaches some level $\hat{r}$ for which the costs of non-exploitation equal personal and social sanctions when no individual exploits, $s(1, 1) + p = c(\hat{r}) \leq c(1)$. Non-exploitation by all individuals is no longer a behavioral equilibrium, $1 \notin \Psi^*(\hat{r}, 1)$. Society transits into the behavioral equilibrium where all individuals exploit the resource, $\psi^* = 0$. The resource stock diminishes, $\dot{r} < 0$, until it reaches some level $\check{r}$ for which the costs of non-exploitation equal personal and social sanctions when all individuals exploit, $c(0) < c(\check{r}) = s(1, 0) + p$. Full exploitation is no longer a behavioral equilibrium, $0 \notin \Psi^*(\check{r}, 1)$, and society transits into the behavioral equilibrium where no individual exploits the resource, $\psi^* = 1$. The above process repeats.

Figure 6 presents a graphical illustration of such a limit cycle in the case of personal sanctions exceeding differences in social sanctions for all and no individuals exploiting the resource, $s(1, 1) - s(0, 1) < p$. Figures 6a and 6b depict norm and resource evolution at each behavioral equilibrium $\psi^* \in \{0, 1\}$. Figure 6c depicts the whole dynamic system in a three-dimensional graph.

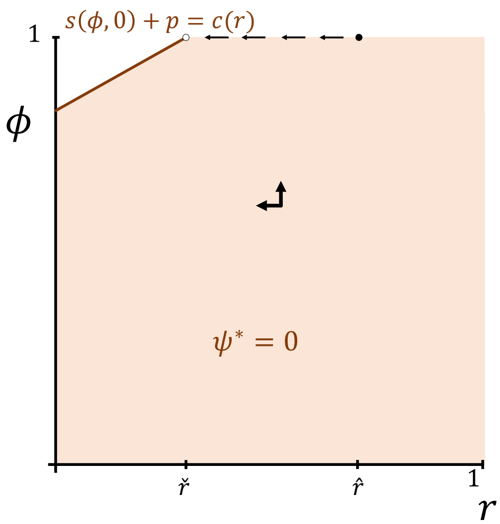
We can rephrase Proposition 6.5 to state that the limit cycle exists only if material costs of non-exploitation are more responsive to extreme changes in the resource stock than social sanctions are to extreme changes in behavior, $s(1, 1) - s(1, 0) < c(1) - c(0)$. Another interesting observation is that the limit cycle can only exist if the boundary equilibrium point of Proposition 6.1 does not exist and vice versa.

Proposition 6.6 (Asymptotic stability of a boundary limit cycle)**.** *For any specification of the model:*

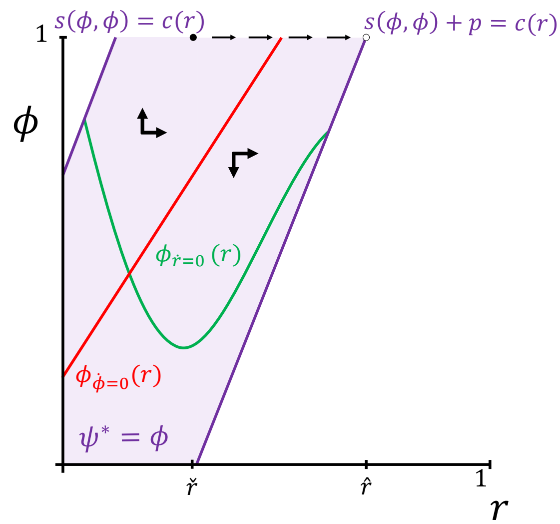
1. *If $\gamma + \kappa s(1, 1) > s(1, 1) + p$, then any $cl(\Omega_{BLC})$ of Proposition 6.5 is asymptotically stable, provided it exists.*

2. *If $\gamma > 0$, then $\exists \bar{p} > 0$ s.t. ($p < \bar{p}$ and $cl(\Omega_{BLC})$ of Proposition 6.5 exists) $\Rightarrow cl(\Omega_{BLC})$ is asymptotically stable.*

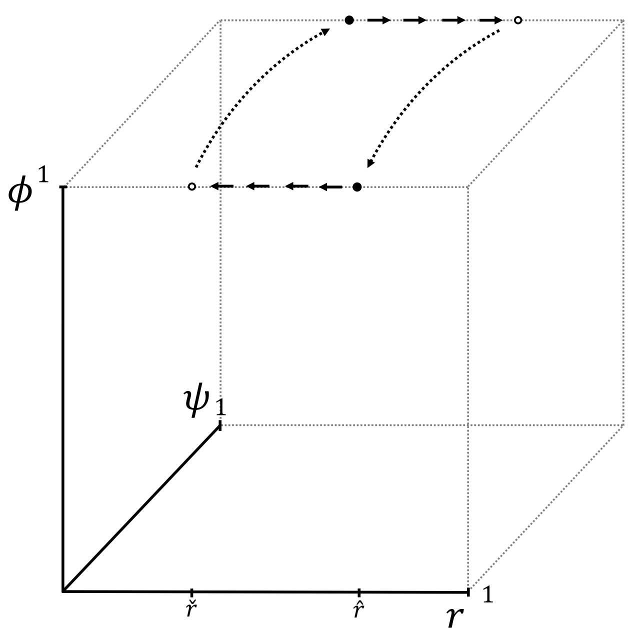
Proof: The Proposition follows from Lemmas A.5 and A.8 in Appendix A.



(a) $\psi^* = 0$



(b) $\psi^* = \phi$



(c) Three dimensional illustration

Figure 6: Boundary limit cycle

Proposition 6.6 above provides insights into when a boundary limit cycle of proposition 6.5 is asymptotic stable.[11] Below, we discuss each condition of the proposition individually.

Condition 1 states that the boundary limit cycle is asymptotically stable if at the perfect social norm $\phi = 1$ and no exploitation $\psi = 1$, social sanctions on cultural fitness $\kappa s(1,1)$ and institutional pressure γ jointly exceed social and personal sanctions $s(1,1) + p$. Hence, we can infer from the condition that (a) the boundary limit cycle is asymptotically stable if the weight of social sanctions on cultural fitness κ and institutional pressure γ are sufficiently large and (b) lower personal sanctions p might positively affect asymptotic stability.

Reconsider figure 6 and recall that in the specific example, it is the case that $s(1,1) < s(1,0) + p$. Hence, at any population profile $\tilde{\rho} = (\tilde{r}, \tilde{\phi}, \tilde{\psi})$ in a sufficiently close neighborhood U of the limit cycle, society is either in the behavioral equilibrium of full or partial exploitation, $\tilde{\psi}^* \in \{0, \tilde{\phi}\}$. When all individuals exploit the resource, $\tilde{\psi}^* = 0$, the resource stock diminishes, $\dot{\tilde{r}} < 0$. If a decrease in the resource stock renders full exploitation no longer a behavioral equilibrium, society transits to the behavioral equilibrium of partial exploitation. Similarly, at any population profile $\tilde{\rho}$ with equilibrium behavior of partial exploitation, $\tilde{\psi}^* = \tilde{\phi}$, we experience an increase in the resource stock, $\dot{\tilde{r}} > 0$. When an increase in the resource stock renders partial exploitation no longer a behavioral equilibrium, equilibrium behavior transits. Hence, the cyclic tendencies of the dynamic system remain present in the boundary limit cycle's neighborhood.

The boundary limit cycle is asymptotically stable if the dynamic system approaches it in these cyclic movements. When all individuals exploit the resource, cultural transmission solely derives from institutional pressure γ and, hence, the social norm unambiguously strengthens. When equilibrium behavior corresponds to partial exploitation, the social norm strengthens above the nullcline $\phi_{\dot{\phi}=0}(r)$ and diminishes below it. Greater institutional pressure γ and weight of social sanctions on cultural fitness κ lower the nullcline and, hence, increase the area for which there is a move towards the boundary limit cycle, thereby favoring asymptotic stability.

A decrease in personal sanctions p shifts the behavioral boundaries by moving the

[11] For ease of writing, we speak of asymptotic stability of the limit cycle, although technically, we consider asymptotic stability of the minimal closed set that contains it.

nullclines solving $s(\phi, 0) + p = c(r)$ and $s(\phi, \phi) + p = c(r)$ to the left. Hence, the limit cycle moves to the left, and society transits between behavioral equilibria at smaller resource stocks. The nullcline $\phi_{\dot{\phi}=0}(r)$, however, is unaffected by such changes. Moving the behavioral boundaries to the left increases the area for which society is above the nullcline $\phi_{\dot{\phi}=0}(r)$ at any neighborhood of the limit cycle and, hence, experiences a strengthening of the social norm — consequently, the decrease in personal sanctions p favors asymptotic stability.

Condition 2 of Proposition 6.6 adds to the insights on lower personal sanctions favoring asymptotic stability of the limit cycle. It states that in the mere presence of institutional pressure, $\gamma > 0$, sufficiently small personal sanctions p render the limit cycle asymptotically stable. The underlying intuition is that as we decrease personal sanctions p, the set of points for which the dynamic system is in the behavioral equilibrium $\psi^* = \phi$ decreases. If the system exhibits this behavioral equilibrium less often, it more often exhibits a behavioral equilibrium where institutional pressure γ alone determines cultural evolution. Hence, institutional pressure becomes more effective in stabilizing the limit cycle. Consider, for example, the limit case where personal sanctions are absent, $p = 0$. The limit cycle exists if $c(0) < s(1, 0) < s(1, 1) < c(1)$. In the neighborhood of the limit cycle, all individuals always behave alike, $\psi^* \in \{0, 1\}$, and cultural dynamics derive fully from institutional pressure, $\dot{\phi} = \phi(1 - \phi)\gamma$. The mere existence of institutional pressure, $\gamma > 0$, drives society back to the perfect social norm $\phi = 1$, thereby ensuring asymptotic stability of the limit cycle. Note that in this limit case of $p = 0$, the limit cycle is asymptotically stable if and only if institutional pressure exists.

Finally, we briefly discuss how changes in the costs of non-exploitation $c(\cdot)$ and social sanctions $s(\cdot, \cdot)$ might affect asymptotic stability of a boundary limit cycle. The following observation captures two opposing effects of such changes verbally, while Lemma A.13 in Appendix A provides a more formal description.

Observation 6.1. *The consequence of raising (lowering) material costs $c(\cdot)$ and lowering (raising) social sanctions $s(\cdot, \cdot)$ for asymptotic stability of a boundary limit cycle of Proposition 6.5 is ambiguous. In particular, two opposing effects exist that affect cultural dynamics in some neighborhood U of the limit cycle:*

1. *The changes may have a positive (negative) effect on cultural dynamics $\dot{\phi}$ at some population profiles $\rho \in U$ by changing equilibrium behavior.*

2. *The changes may have a negative (positive) effect on cultural dynamics $\dot{\phi}$ at some population profiles $\rho \in U$ by decreasing (increasing) the cultural fitness of non-exploiting norm holders.*

First, similar to a decrease in personal sanctions p in Condition 1 of Proposition 6.2, raising the costs $c(\cdot)$ or lowering social sanctions $s(\cdot, \cdot)$ may shift the boundaries of the behavioral equilibria and, hence, the whole limit cycle to the left in Figure 6. Suppose cultural dynamics and, thus, the nullcine $\phi_{\dot{\phi}=0}(r)$ were unaffected by these changes. Consequently, by the same reasoning as for personal sanctions, the shifts of the behavioral boundaries favor asymptotic stability of the boundary limit cycle. Second, and contrary to personal sanctions, the described changes in $c(\cdot)$ and $s(\cdot, \cdot)$ also impact cultural evolution. The higher $c(\cdot)$ and lower $s(\cdot, \cdot)$ affect differences in cultural fitness in the norm non-holders favor whenever behavior is in the partial exploitation equilibrium $\psi^* = \phi$, and thereby weakens norm dynamics. The changes move the nullcline $\phi_{\dot{\phi}=0}(r)$ to the left. This effect generally harms asymptotic stability of the limit cycle. Whether the changes in $c(\cdot)$ and $s(\cdot, \cdot)$ are beneficial for asymptotic stability depends on which of these two effects dominates, which is, in turn, subject to the exact changes that occur.

In summary, the results of this section show that resource conservation through a boundary limit cycle is possible if the costs of non-exploitation at the maximum resource stock $c(1)$ are too high for a boundary equilibrium point of Proposition 6.1 to exist, $s(1, 1) + p \leq c(1)$, but relatively responsive to changes in the resource stock as compared to social sanctions for behavioral non-conformity, $s(1, 1) - s(1, 0) < c(1) - c(0)$. Although larger personal sanctions p may aid the existence of a limit cycle (by ensuring that full exploitation is not a behavioral equilibrium as society approaches the point of no return $r = 0$), smaller personal sanctions p favor asymptotic stability of an existing one. This adverse effect of personal sanctions on asymptotic stability becomes even more prominent when observing that institutional pressure γ is especially effective in stabilizing the limit cycle when personal sanctions are low. Social sanctions $s(\cdot, \cdot)$ and material costs $c(\cdot)$ play similar ambiguous roles. Changes in these factors shift the boundaries of behavioral equilibria and alter cultural dynamics at the behavioral equilibrium of partial exploitation. Finally, greater institutional pressure γ and weight of cultural fitness κ generally positively affect asymptotic stability.

7 Discussion

Our analysis suggests that if the costs of non-exploitation at the maximum resource stock are small, $c(1) < s(1,1)$, society can conserve the CPR through an asymptotically stable boundary equilibrium point with minimal levels of institutional pressure, $\gamma > 0$. For intermediate costs of non-exploitation, $s(1,1) \leq c(1) < s(1,1) + p$, society requires jointly sufficiently large institutional pressure and weight of social sanctions on cultural fitness to do so, $\kappa s(1,1) + \gamma > c(1)$. However, contrary to the previous case, institutional pressure is no longer necessary for asymptotic stability. If non-exploitation is very costly at the maximum resource stock, $s(1,1) + p \leq c(1)$, but relatively cheap as the resource approaches the point of no return and behavioral conformity plays a relatively minor role in determining sanctions for norm violation, $c(0) < s(1,0) + p$, then society may conserve the CPR through an asymptotically stable boundary limit cycle if the weight of social sanctions on cultural fitness κ and institutional pressure γ are again sufficiently large. Alternatively, there may exist an interior equilibrium point at social norm ϕ and resource stock r if only all norm holders not exploiting the resource is a behavioral equilibrium, $s(\phi, \phi) < c(r) < s(\phi, \phi) + p$, and institutional pressure and the weight of social sanctions on cultural fitness balance material costs of non-exploitation, $\kappa s(\phi, \phi) + \gamma = c(r)$. Society is more likely to conserve the resource through such an asymptotically stable interior equilibrium point if institutional pressure γ plays a more significant role than the weight κ in balancing material costs $c(r)$.

The discussion above indicates that, at times, institutional pressure γ and the weight of social sanctions on cultural fitness κ constitute substitutes for securing the CPR. In contrast, sometimes γ cannot be substituted by κ. In these latter cases, a (relatively) large κ either does not affect cultural evolution (at the boundary equilibrium point if $s(1,1) > c(1)$) or harms asymptotic stability (at the interior equilibrium point).

Moreover, our results suggest that the substitutability for upholding a perfect social norm $\phi = 1$ weakens if personal sanctions p are relatively unimportant in the sense that either social sanctions $s(\phi, \psi)$ suffice to motivate non-exploitative behavior (at the boundary equilibrium point) or personal sanctions p are small (at the boundary limit cycle and in general). In these instances, institutional pressure is especially effective in promoting norm adoption. Moreover, the weight of social sanctions on cultural fitness becomes less effective: If personal sanctions play a subordinate role, norm holders and non-holders behave alike, in which case no difference in social sanctions that could affect cultural evolution exists. Hence, our findings suggest that, in these instances, policies that foster norm adoption through institutions prove especially effective and may, therefore, be desirable.

Despite the above discussion, there also exist instances for which too much institutional pressure γ might be non-favorable in terms of resource conservation, as it may hinder the existence of an interior equilibrium point (e.g., if $\gamma > p + s(1, 1)$). If material costs at the maximum resource stock are too large for a boundary equilibrium point to exist, $s(1, 1) + p < c(1)$, and the descriptive social norm plays a relatively large role in determining total sanctions for exploitation, so that a boundary limit cycle does also not exist, $s(1, 0)+p < c(0)$, then society could benefit from sufficiently low levels of institutional pressure that enable an interior equilibrium point. This line of argument easily extends to the weight of social sanctions on cultural fitness κ.

Beyond institutional pressure γ and the weight of social sanctions on cultural fitness κ, we can draw conclusions regarding the further variables. We have already touched upon personal sanctions p, for which small values may favor asymptotic stability of a boundary limit cycle. However, larger values of p increase the non-exploitation incentives of norm holders, which, in turn, favors the existence of such a limit cycle as well as that of boundary and interior equilibrium points. Hence, personal sanctions play a somewhat ambiguous role. The ambiguity arises because, although an increase in personal sanctions favors norm holders refraining from exploitation, it also does so if this behavior is sub-optimal from a cultural fitness perspective, potentially harming the persistence of the social norm.

We can accredit a similar ambiguous role to social sanctions $s(\cdot, \cdot)$ and material costs of non-exploitation $c(\cdot)$. Larger social sanctions $s(\cdot, \cdot)$ and smaller material costs $c(\cdot)$ generally favor the existence of a boundary equilibrium point or limit cycle as they correspond to greater incentives to refrain from exploitation. However, for an existing boundary equilibrium, the magnitudes of these variables have an ambiguous effect on stability. For the boundary equilibrium point, for example, high social sanctions $s(\cdot,\cdot)$ and low material costs $c(\cdot)$ are favorable in the presence of institutional pressure and potentially non-favorable in its absence. Similar to personal sanctions, the adverse effects arise due to shifts in equilibrium behavior that are sub-optimal in terms of cultural dynamics.

These insights highlight the complexity of the dynamic system and the importance of understanding norms in a dynamic context, accounting for the interplay of behavioral incentives and cultural dynamics. For example, when designing and implementing policies targeting behavioral incentives, these policies must account for potentially adverse effects on norms.

Our findings also suggest that a strong sensitivity of material costs to variations in the resource stock is favorable for securing the CPR. The existence of a boundary limit cycle is contingent upon the material costs being responsive to extreme changes in the resource stock, and an existing interior equilibrium point is asymptotically stable when marginal changes in the resource stock induce a significant alteration in material costs. In both cases, highly responsive material costs induce a decrease in the resource stock to be met by a significant decrease in exploitation, either directly through all individuals changing their behavior when exploiting is no longer worth it (at the boundary limit cycle) or indirectly through individuals adopting a new personal norm and adapting their behavior in response (at the interior equilibrium point). Similarly, weakly responsive social sanctions favor resource conservation through an interior equilibrium point or boundary limit cycle. This holds since weak responsiveness ensures that even after a decrease in the social and descriptive norms, social sanctions still significantly impact behavioral incentives (at the boundary equilibrium limit cycle) and cultural evolution (at the interior equilibrium point).

Finally, note that, in addition to the exploratory insights that the analysis provides, the equilibrium findings are compatible with several real-world observations: (1) The interior point aligns with heterogeneous behavior and moral perceptions regarding resource conservation across individuals (see, e.g., Kotchen and Moore, 2008; Sundt and Rehdanz, 2015), (2) the multitude of socio-ecological equilibria matches the variety of outcomes in different (but possibly structurally similar) situations and societies (see, e.g., Minton et al., 2018; Dannenberg et al., 2024), and (3) the boundary limit cycle is consistent with the conservation of resources in constantly evolving systems through resource-stock-responsive adaptions (see, e.g., Folke et al., 2002; Olsson et al., 2004). These alignments suggest that the model provides some theoretical rationales for the observed phenomena and, to some extent, contributes to a better understanding of them.

8 Concluding remarks

This paper introduces and studies an evolutionary model that endogenizes the formation of behavior, personal, social, and descriptive norms, and the resource stock. We find that, under certain conditions, society can secure a positive resource stock through (a) an asymptotically stable equilibrium point where the social and descriptive norms align, and personal norms and behavior are either homogeneous or heterogeneous across individuals or (b) an asymptotically stable limit cycle in which personal and social norms remain constant, but herding causes alternating descriptive norms and a fluctuating resource stock. These results are consistent with real-world observations (e.g., heterogeneous behavior and moral perceptions across individuals, varying outcomes across situations and societies, time-varying norms and resource stocks), which suggests that the model provides some theoretical rationales.

The analysis highlights the importance of institutional pressure and the weight of social sanctions on cultural fitness for upholding the social norm and, consequently, conserving the CPR. In some cases, these two factors function as substitutes, whereas, in other cases, institutional pressure cannot be substituted. In the latter cases, the active promotion of norm adoption by institutions becomes particularly effective in upholding the social norm and is thus likely to be desirable.

Additionally, behavioral incentives to avoid exploitation must be in place to secure the resource. However, we find that overly large incentives can interfere with cultural dynamics, leading to adverse outcomes that harm resource conservation in the long run. These adverse effects have significant policy implications, highlighting the importance of well-tailored policies that account for the complex dynamics at play.

Future research may complement this paper by expanding on the equilibrium analyses and discussing how a society can transition from one equilibrium point to a socially preferred one. Such analyses would require, among others, social welfare comparisons of the different socio-ecological equilibria. Closely related, further research could analyze how different policies perform when accounting for the complex dynamic interactions of behavior, norms, and resources. Lastly, the analysis should be expanded to account for other mechanisms influencing norm evolution (e.g., cognitive dissonance, normative conformity concerns, network effects) to draw a more complete picture of the dynamic system.

References

Alger, I., and Weibull, J. W. (2013). Homo moralis—preference evolution under incomplete information and assortative matching. *Econometrica*, *81*(6), 2269–2302. doi: 10.3982/ECTA10637

Alger, I., and Weibull, J. W. (2016). Evolution and kantian morality. *Games and Economic Behavior*, *98*, 56–67. doi: 10.1016/j.geb.2016.05.006

Andreoni, J., and Bernheim, B. D. (2009). Social image and the 50–50 norm: A theoretical and experimental analysis of audience effects. *Econometrica*, *77*(5), 1607–1636. doi: 10.3982/ECTA7384

Bénabou, R., and Tirole, J. (2011). Identity, morals, and taboos: Beliefs as assets. *The Quarterly Journal of Economics*, *126*(2), 805–855. doi: 10.1093/qje/qjr002

Bernheim, B. D. (1994). A theory of conformity. *Journal of Political Economy*, *102*(5), 841–877. doi: 10.1086/261957

Bester, H., and Güth, W. (1998). Is altruism evolutionarily stable? *Journal of Economic Behavior and Organization*, *34*(2), 193–209. doi: 10.1016/S0167-2681(97)00060-7

Bezin, E. (2019). The economics of green consumption, cultural transmission and sustainable technological change. *Journal of Economic Theory*, *181*, 497–546. doi: 10.1016/j.jet.2019.03.005

Bicchieri, C., and Dimant, E. (2019). Nudging with care: The risks and benefits of social information. *Public Choice*, *191*(3), 443–464. doi: 10.1007/s11127-019-00684-6

Binmore, K., and Samuelson, L. (1994). An economist's perspective on the evolution of norms. *Journal of Institutional and Theoretical Economics (JITE)/Zeitschrift für die gesamte Staatswissenschaft*, *150*(1), 45–63. Retrieved from `https://www.jstor.org/stable/40753015`

Bisin, A., Topa, G., and Verdier, T. (2004). Cooperation as a transmitted cultural trait. *Rationality and Society*, *16*(4), 477–507. doi: 10.1177/1043463104046695

Bisin, A., and Verdier, T. (1998). On the cultural transmission of preferences for social status. *Journal of Public Economics*, *70*(1), 75–97. doi: 10.1016/S0047-2727(98)00061-9

Bisin, A., and Verdier, T. (2001). The economics of cultural transmission and the dynamics of preferences. *Journal of Economic Theory*, *97*(2), 298–319. doi: 10.1006/jeth.2000.2678

Bowles, S., and Gintis, H. (1998). The moral economy of communities: Structured populations and the evolution of pro-social norms. *Evolution and Human Behavior*, *19*(1), 3–25. doi: 10.1016/S1090-5138(98)00015-4

Boyd, R., and Richerson, P. J. (1990). Group selection among alternative evolutionarily stable strategies. *Journal of Theoretical Biology*, *145*(3), 331–342. doi: 10.1016/S0022-5193(05)80113-4

Boyd, R., and Richerson, P. J. (2005). *The origin and evolution of cultures.* Oxford University Press.

Brekke, K. A., Kverndokk, S., and Nyborg, K. (2003). An economic model of moral motivation. *Journal of Public Economics*, *87*(9-10), 1967–1983. doi: 10.1016/S0047-2727(01)00222-5

Cooter, R. (1998). Expressive law and economics. *The Journal of Legal Studies*, *27*(S2), 585–607. doi: 10.1086/468036

Crawford, S. E. S., and Ostrom, E. (1995). A grammar of institutions. *American Political Science Review*, *89*(3), 582–600. doi: 10.2307/2082975

d'Adda, G., Dufwenberg, M., Passarelli, F., and Tabellini, G. (2020). Social norms with private values: Theory and experiments. *Games and Economic Behavior*, *124*, 288–304. doi: 10.1016/j.geb.2020.08.012

Dannenberg, A., Gutsche, G., Batzke, M. C. L., Christens, S., Engler, D., Mankat, F., … Ziegler, A. (2024). The effects of norms on environmental behavior. *Review of Environmental Economics and Policy*, *18*(1), 165–187. doi: 10.1086/727588

Elster, J. (1989). Social norms and economic theory. *Journal of Economic Perspectives*, *3*(4), 99–117. doi: 10.1257/jep.3.4.99

Farrow, K., Grolleau, G., and Ibanez, L. (2017). Social norms and pro-environmental behavior: a review of the evidence. *Ecological Economics*, *140*, 1–13. doi: 10.1016/j.ecolecon.2017.04.017

Fehr, E., and Fischbacher, U. (2004). Social norms and human cooperation. *Trends in Cognitive Sciences*, *8*(4), 185–190. doi: 10.1016/j.tics.2004.02.007

Fershtman, C., and Weiss, Y. (1998). Social rewards, externalities and stable preferences. *Journal of Public Economics*, *70*(1), 53–73. doi: 10.1016/S0047-2727(98)00060-7

Figuieres, C., Masclet, D., and Willinger, M. (2013). Weak moral motivation leads to the decline of voluntary contributions. *Journal of Public Economic Theory*, *15*(5), 745–772. doi: 10.1111/jpet.12036

Folke, C., Colding, J., and Berkes, F. (2002). Synthesis: building resilience and adaptive capacity in social–ecological systems. In *Navigating social-ecological systems: Building resilience for complexity and change* (p. 352–387). Cambridge University Press. doi: 10.1017/CBO9780511541957.020

Gintis, H. (2003). Solving the puzzle of prosociality. *Rationality and Society*, *15*(2), 155–187. doi: 10.1177/1043463103015002001

Güth, W., and Yaari, M. E. (1992). An evolutionary approach to explain reciprocal behavior in a simple strategic game. In U. Witt (Ed.), *Explaining process and change: Approaches to evolutionary economics* (pp. 23–34). University of Michigan Press.

Guttman, J. M. (2003). Repeated interaction and the evolution of preferences for reciprocity. *The Economic Journal*, *113*(489), 631–656. doi: 10.1111/1468-0297.t01-1-00144

Guttman, J. M. (2013). On the evolution of conditional cooperation. *European Journal of Political Economy*, *30*, 15–34. doi: 10.1016/j.ejpoleco.2012.11.003

Hardin, G. (1968). The tragedy of the commons: the population problem has no technical solution; it requires a fundamental extension in morality. *Science*, *162*(3859), 1243–1248. doi: 10.1126/science.162.3859.124

Henrich, J. (2004). Cultural group selection, coevolutionary processes and large-scale cooperation. *Journal of Economic Behavior and Organization*, *53*(1), 3–35. doi: 10.1016/S0167-2681(03)00094-5

Henrich, J., and Gil-White, F. J. (2001). The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior*, *22*(3), 165–196. doi: 10.1016/S1090-5138(00)00071-4

Kotchen, M. J., and Moore, M. R. (2008). Conservation: From voluntary restraint to a voluntary price premium. *Environmental and Resource Economics*, *40*, 195–215. doi: https://doi.org/10.1007/s10640-007-9148-x

Lindbeck, A., Nyberg, S., and Weibull, J. W. (1999). Social norms and economic incentives in the welfare state. *The Quarterly Journal of Economics*, *114*(1), 1–35. doi: 10.1162/003355399555936

Mankat, F. (in press). Cooperation, norms, and gene-culture coevolution. *Games and Economic Behavior*. doi: 10.1016/j.geb.2024.07.006

Mengel, F. (2008). Matching structure and the cultural transmission of social norms. *Journal of Economic Behavior and Organization*, *67*(3-4), 608–623. doi: 10.1016/j.jebo.2008.01.001

Minton, E. A., Spielmann, N., Kahle, L. R., and Kim, C.-H. (2018). The subjective norms of sustainable consumption: A cross-cultural exploration. *Journal of Business Research*, *82*, 400–408. doi: 10.1016/j.jbusres.2016.12.031

Mitteldorf, J., and Wilson, D. S. (2000). Population viscosity and the evolution of altruism. *Journal of Theoretical Biology*, *204*(4), 481–496. doi: 10.1006/jtbi.2000.2007

Müller, S., and von Wangenheim, G. (2019). *Coevolution of cooperation, preferences, and cooperative signals in social dilemmas* (Discussion Paper No. 221). Center for European, Governance, and Economic Development Research.

Noailly, J., van den Bergh, J. C., and Withagen, C. A. (2003). Evolution of harvesting strategies: replicator and resource dynamics. *Journal of Evolutionary Economics*, *13*, 183–200. doi: 10.1007/s00191-003-0146-z

Noailly, J., Withagen, C. A., and Van den Bergh, J. C. (2007). Spatial evolution of social norms in a common-pool resource game. *Environmental and Resource Economics*, *36*, 113–141. doi: 10.1007/s10640-006-9046-7

Nyborg, K. (2000). Homo economicus and homo politicus: interpretation and aggregation of environmental values. *Journal of Economic Behavior and Organization*, *42*(3), 305–322. doi: 10.1016/S0167-2681(00)00091-3

Nyborg, K. (2018). Social norms and the environment. *Annual Review of Resource Economics*, *10*, 405–423. doi: 10.1146/annurev-resource-100517-023232

Nyborg, K., Howarth, R. B., and Brekke, K. A. (2006). Green consumers and public policy: On socially contingent moral motivation. *Resource and Energy Economics*, *28*(4), 351–366. doi: 10.1016/j.reseneeco.2006.03.001

Nyborg, K., and Rege, M. (2003a). Does public policy crowd out private contributions to public goods. *Public Choice*, *115*(3), 397–418. doi: 10.1023/A:1024245522958

Nyborg, K., and Rege, M. (2003b). On social norms: the evolution of considerate smoking behavior. *Journal of Economic Behavior and Organization*, *52*(3), 323–340. doi: 10.1016/S0167-2681(03)00031-3

Olsson, P., Folke, C., and Berkes, F. (2004). Adaptive comanagement for building resilience in social–ecological systems. *Environmental management*, *34*, 75–90. doi: 10.1007/s00267-003-0101-7

Osés-Eraso, N., and Viladrich-Grau, M. (2007). On the sustainability of common property resources. *Journal of Environmental Economics and Management*, *53*(3), 393–410. doi: 10.1016/j.jeem.2006.10.006

Ostrom, E. (2000). Collective action and the evolution of social norms. *Journal of Economic Perspectives*, *14*(3), 137–158. doi: 10.1257/jep.14.3.137

Ostrom, E. (2009). A general framework for analyzing sustainability of social-ecological systems. *Science*, *325*(5939), 419–422. doi: 10.1126/science.1172133

Panebianco, F. (2016). The role of persuasion in cultural evolution dynamics. *International Review of Economics*, *63*(3), 233–258. doi: 10.1007/s12232-016-0253-4

Poulsen, A., and Poulsen, O. (2006). Endogenous preferences and social-dilemma institutions. *Journal of Institutional and Theoretical Economics (JITE)/Zeitschrift für die gesamte Staatswissenschaft*, *162*(4), 627–660. doi: 10.1628/093245606779252742

Rege, M. (2004). Social norms and private provision of public goods. *Journal of Public Economic Theory*, *6*(1), 65–77. doi: 10.1111/j.1467-9779.2004.00157.x

Richter, A., and Grasman, J. (2013). The transmission of sustainable harvesting norms when agents are conditionally cooperative. *Ecological Economics*, *93*, 202–209. doi: 10.1016/j.ecolecon.2013.05.013

Richter, A., van Soest, D., and Grasman, J. (2013). Contagious cooperation, temptation, and ecosystem collapse. *Journal of Environmental Economics and Management*, *66*(1), 141–158. doi: 10.1016/j.jeem.2013.04.004

Schumacher, I. (2009). The dynamics of environmentalism and the environment. *Ecological Economics*, *68*(11), 2842–2849. doi: 10.1016/j.ecolecon.2009.06.004

Schumacher, I. (2015). The endogenous formation of an environmental culture. *European Economic Review*, *76*, 200–221. doi: 10.1016/j.euroecorev.2015.03.002

Sethi, R., and Somanathan, E. (1996). The evolution of social norms in common property resource use. *American Economic Review*, *86*(4), 766–788. Retrieved from https://www.jstor.org/stable/2118304

Sundt, S., and Rehdanz, K. (2015). Consumers' willingness to pay for green electricity: A meta-analysis of the literature. *Energy Economics*, *51*, 1–8. doi: https://doi.org/10.1016/j.eneco.2015.06.005

Tabellini, G. (2008). The scope of cooperation: Values and incentives. *The Quarterly Journal of Economics*, *123*(3), 905–950. doi: 10.1162/qjec.2008.123.3.905

te Velde, V. L., and Louis, W. (2022). Conformity to descriptive norms. *Journal of Economic Behavior and Organization*, *200*, 204–222. doi: 10.1016/j.jebo.2022.05.017

Thøgersen, J. (2006). Norms for environmentally responsible behaviour: An extended taxonomy. *Journal of Environmental Psychology*, *26*(4), 247–261. doi: 10.1016/j.jenvp.2006.09.004

Tilman, A. R., Plotkin, J. B., and Akçay, E. (2020). Evolutionary games with environmental feedbacks. *Nature Communications*, *11*(1), 915. doi: 10.1038/s41467-020-14531-6

Traxler, C. (2010). Social norms and conditional cooperative taxpayers. *European Journal of Political Economy*, *26*(1), 89–103. doi: 10.1016/j.ejpoleco.2009.11.001

Traxler, C., and Spichtig, M. (2011). Social norms and the indirect evolution of conditional cooperation. *Journal of Economics*, *102*(3), 237–262. doi: 10.1007/s00712-010-0173-9

Voss, T. (2001). Game theoretical perspectives on the emergence of social norms. In M. Hechter and K. D. Opp (Eds.), *Social norms* (p. 105-138). Russell Sage Foundation.

Waring, T. M., Goff, S. H., and Smaldino, P. E. (2017). The coevolution of economic institutions and sustainable consumption via cultural group selection. *Ecological Economics*, *131*, 524–532. doi: 10.1016/j.ecolecon.2016.09.022

Young, H. P. (1993). The evolution of conventions. *Econometrica*, *61*(1), 57–84. doi: 10.2307/2951778

Young, H. P. (1996). The economics of convention. *Journal of Economic Perspectives*, *10*(2), 105–122. doi: 10.1257/jep.10.2.105

Young, H. P. (2015). The evolution of social norms. *Annual Review of Economics*, *7*(1), 359–387. doi: 10.1146/annurev-economics-080614-115322

A Proofs

This section presents the formal analysis. The variables $r(\rho, t)$, $\phi(\rho, t)$, and $\psi^*(\rho, t)$ represent solution values as described by $\xi(\rho, t)$. We denote their respective (right-hand) time derivatives as $\dot{r}(\rho, t)$, $\dot{\phi}(\rho, t)$, and $\dot{\psi}^*(\rho, t)$. In line with the main text of the paper, we often use abbreviated notation for simplicity (e.g., $\dot{\phi}$). Throughout the analysis, we use the Chebyshev distance as the metric on $[0, 1]^3$.

Lemma A.1.

1. *If* $\psi^*(\rho, t_0) = \phi(\rho, t_0) \in (0, 1)$*, then (a)* $\psi^*(\rho, t_0)$ *is right-continuous, and if* $\bar{t} := \inf\{t > t_0 : \psi^*(\rho, t_0) \neq \phi(\rho, t_0)\}$ *exists, then (b)* $\bar{t} = \min\{t > t_0 : \psi^*(\rho, t_0) \neq \phi(\rho, t_0)\}$.

2. *If* $\psi^*(\rho, t_0) = a \in \{0, 1\}$*, then (a)* $\psi^*(\rho, t_0)$ *is right-continuous, and if* $\bar{t} := \inf\{t > t_0 : \psi^*(\rho, t_0) \neq \phi(\rho, t_0)\}$ *exists, then (b)* $\bar{t} = \min\{t > t_0 : \psi^*(\rho, t_0) \neq a\}$.

Proof. We start with the first condition. Suppose $\psi^*(\rho, t_0) = \phi(\rho, t_0) \in (0, 1)$. $\phi(\rho, t_0) \in (0, 1)$ and the asymptotic nature of $\dot{\phi}(\rho, t)$ as $\phi(\rho, t_0)$ approaches 0 or 1 imply $\phi(\rho, t) \in (0, 1)\ \forall t \geq t_0$. $\psi^*(\rho, t_0) = \phi(\rho, t_0) \in (0, 1) \Rightarrow s(\phi(\rho, t_0), \phi(\rho, t_0)) < c(r(\rho, t_0)) < s(\phi(\rho, t_0), \phi(\rho, t_0)) + p$. Continuity of $r(\rho, t_0)$ and $\phi(\rho, t_0)$ imply $\exists \epsilon > 0$ s.t. $s(\phi(\rho, t), \phi(\rho, t)) < c(r(\rho, t)) < s(\phi(\rho, t), \phi(\rho, t)) + p\ \forall t \in [t_0, t_0 + \epsilon)$. $s(\phi(\rho, t), \phi(\rho, t)) < c(r(\rho, t)) < s(\phi(\rho, t), \phi(\rho, t)) + p\ \forall t \in [t_0, t_0 + \epsilon) \Rightarrow \phi(\rho, t) \in \Psi^*(r(\rho, t), \phi(\rho, t))\ \forall t \in [t_0, t_0 + \epsilon)$. $(\psi^*(\rho, t_0) = \phi(\rho, t_0) \wedge \phi(\rho, t) \in \Psi^*(r(\rho, t), \phi(\rho, t))\ \forall t \in [t_0, t_0 + \epsilon)) \Rightarrow \psi^*(\rho, t) = \phi(\rho, t)\ \forall t \in [t_0, t_0 + \epsilon)$. Since $\phi(\rho, t)$ is continuous, $\psi^*(\rho, t)$ is right-continuous at t_0.

Suppose $\exists t > t_0$ s.t. $\psi^*(\rho, t) \neq \phi(\rho, t)$. $\exists t > t_0$ s.t. $\psi^*(\rho, t) \neq \phi(\rho, t) \Rightarrow \exists t > t_0$ s.t. $\phi(\rho, t) \notin \Psi^*(r(\rho, t), \phi(\rho, t)) \Rightarrow \exists t > t_0$ s.t. $s(\phi(\rho, t), \phi(\rho, t)) \geq c(r(\rho, t)) \vee s(\phi(\rho, t), \phi(\rho, t)) + p \leq c(r(\rho, t))$. Let $\bar{t} := \min\{t > t_0 : s(\phi(\rho, t), \phi(\rho, t)) = c(r(\rho, t)) \vee s(\phi(\rho, t), \phi(\rho, t)) + p = c(r(\rho, t))\}$, which must exist by continuity. Note that $s(\phi(\rho, t), \phi(\rho, t)) < c(r(\rho, t)) < s(\phi(\rho, t), \phi(\rho, t)) + p\ \forall t \in [t_0, \bar{t})$ and $\psi^*(\rho, \bar{t}) \neq \phi(\rho, \bar{t})$. $s(\phi(\rho, t), \phi(\rho, t)) < c(r(\rho, t)) < s(\phi(\rho, t), \phi(\rho, t)) + p\ \forall t \in [t_0, \bar{t}) \Rightarrow \phi(\rho, t) \in \Psi^*(r(\rho, t), \phi(\rho, t))\ \forall t \in [t_0, \bar{t})$. $(\psi^*(\rho, t_0) = \phi(\rho, t_0) \wedge \phi(\rho, t) \in \Psi^*(r(\rho, t), \phi(\rho, t))\ \forall t \in [t_0, \bar{t})) \Rightarrow \psi^*(\rho, t) = \phi(\rho, t)\ \forall t \in [t_0, \bar{t}) \Rightarrow \bar{t} = \min\{t > t_0 : \psi^*(\rho, t) \neq \phi(\rho, t)\}$.

We continue with Condition 2. For the case where $\psi^*(\rho,t_0)=a\in\{0,1\}$, we need to differentiate whether (a) $\phi(\rho,t_0)\in(0,1)$ or (b) $\phi(\rho,t_0)\in\{0,1\}$. (a): In the former, $\phi(\rho,t)\in(0,1)\ \forall t\geq 0$ by similar reasoning as above. For $\psi^*(\rho,t_0)=1$ we can proceed analogously to the above to show that $\psi(\rho,t)$ is right-continuous at t_0, and if $\bar{t}$ exists, $\bar{t}=\min\{t>t_0:\psi(\rho,t_0)\neq 1\}=\min\{t>t_0:s(\phi(\rho,t),1)=c(r(\rho,t))\}$. For $\psi^*(\rho,t_0)=0$ we can proceed analogously to the above to show that $\psi(\rho,t)$ is right-continuous at t_0, and if $\bar{t}$ exists, $\bar{t}=\min\{t>t_0:\psi(\rho,t_0)\neq 0\}=\min\{t>t_0:s(\phi(\rho,t),0)+p=c(r(\rho,t))\}$. (b): In the latter, $\phi(\rho,t_0)=b\in\{0,1\}$. $\phi(\rho,t_0)=b\in\{0,1\}\Rightarrow\dot{\phi}(\rho,t_0)=0\Rightarrow\phi(\rho,t_0)=b\ \forall t\geq t_0$. For $\phi(\rho,t_0)=1$ and $\psi^*(\rho,t_0)=a\in\{0,1\}$ we can proceed analogously to the above to show that $\psi(\rho,t)$ is right-continuous at t_0, and if $\bar{t}$ exists, $\bar{t}=\min\{t>t_0:\psi(\rho,t_0)\neq 1\}=\min\{t>t_0:s(1,a)+p=c(r(\rho,t))\}$. For $\phi(\rho,t_0)=0$ and $\psi^*(\rho,t_0)=a\in\{0,1\}$ we can proceed analogously to the above to show that $\psi(\rho,t)$ is right-continuous at t_0, and if $\bar{t}$ exists, $\bar{t}=\min\{t>t_0:\psi(\rho,t_0)\neq 0\}=\min\{t>t_0:s(0,a)=c(r(\rho,t))\}$. □

Lemma A.2. *Consider any $\rho=(r,\phi,\psi)$ s.t. $\xi(\rho,t)=\rho\ \forall t\geq 0$ and there is a neighborhood U of ρ s.t. (a) $\psi^*(\hat{\rho},0)=1\ \forall\hat{\rho}=(\hat{r},\hat{\phi},\hat{\psi})\in U$ or (b) $\psi^*(\hat{\rho},0)=\hat{\phi}\ \forall\hat{\rho}=(\hat{r},\hat{\phi},\hat{\psi})\in U$.*

- *If (a) holds, then asymptotic stability of ρ in the reduced dynamic system consisting of (1) $\dot{r}_r(\rho,t)=\delta(r_r(\rho,t))-e(1)$, (2) $\dot{\phi}_r(\rho,t)=\phi_r(\rho,t)(1-\phi_r(\rho,t))\upsilon\gamma$ and (3) $\psi_r^*(\rho,t)=1\ \forall t\geq 0$ implies asymptotic stability of ρ in the original (non-reduced) dynamic system.*
- *If (b) holds, then asymptotic stability of ρ in the reduced dynamic system consisting of (1) $\dot{r}_r(\rho,t)=\delta(r_r(\rho,t))-e(\phi)$, (2) $\dot{\phi}_r(\rho,t)=\phi_r(\rho,t)(1-\phi_r(\rho,t))\upsilon(\kappa s(\phi,\phi)+\gamma-c(r))$, and (3) $\psi_r^*(\rho,t)=\phi_r(\rho,t)\ \forall t\geq 0$ implies asymptotic stability of ρ in the original (non-reduced) dynamic system.*

Proof. Consider $\check{\rho}=(\check{r},\check{\phi},\check{\psi})$ s.t. $\xi(\check{\rho},t)=\check{\rho}\ \forall t\geq 0$ and suppose (a) holds. Consider any U of $\check{\rho}$. Let $Z\subset U$ be s.t. $\psi^*(\hat{\rho},0)=1\ \forall\hat{\rho}=(\hat{r},\hat{\phi},\hat{\psi})\in Z$. Let $\xi_r(\rho,t)$ be the solution of the reduced dynamic system, described by $r_r(\hat{\rho},t)$, $\phi_r(\hat{\rho},t)$, and $\psi_r^*(\rho,t)=1\ \forall t\geq 0$. Suppose $\check{\rho}$ is asymptotically stable in the reduced dynamic system. Let $\tilde{U}\subset Z$ be s.t. for all $\hat{\rho}\in\tilde{U}$, $\xi_r(\hat{\rho},t)\in Z\subset U\ \forall t\geq 0$ and $\lim_{t\to\infty}\xi_r(\hat{\rho},t)=\check{\rho}$. Consider any $\hat{\rho}=(\hat{r},\hat{\phi},\hat{\psi})\in\tilde{U}$. For the original (non-reduced) dynamic system, $\xi(\hat{\rho},0)=(\hat{r},\hat{\phi},1)$. For all $\tilde{t}>0$ s.t. $\psi^*(\hat{\rho},t)=1\ \forall t\in[0,\tilde{t}]$, $\xi(\hat{\rho},t)=\xi_r(\hat{\rho},t)\ \forall t\in[0,\tilde{t}]$ as both solutions are described by the same system of equations on this interval (and no discontinuities occur). $\xi_r(\hat{\rho},t) \in Z\ \forall t \geq 0 \Rightarrow s(\phi_r(\hat{\rho},t),1) > c(r_r(\rho,t))\ \forall t \geq 0 \Rightarrow \nexists \tilde{t} > 0$ s.t. $1 \notin \Psi^*(\lim_{t\to\tilde{t}} r_r(\hat{\rho},t), \lim_{t\to\tilde{t}} \phi_r(\hat{\rho},\tilde{t})) = \Psi^*(\lim_{t\to\tilde{t}} r(\hat{\rho},t), \lim_{t\to\tilde{t}} \phi(\hat{\rho},\tilde{t})) \Rightarrow \psi^*(\hat{\rho},t) = 1\ \forall t \geq 0 \Rightarrow \xi(\hat{\rho},t) = \xi_r(\hat{\rho},t)\ \forall t \geq 0 \Rightarrow (\xi(\hat{\rho},t) \in Z \subset U\ \forall t \geq 0 \wedge \lim_{t\to\infty} \xi(\hat{\rho},t) = \check{\rho})$. Since U and $\hat{\rho} \in \tilde{U}$ were chosen arbitrarily, $\check{\rho}$ is asymptotically stable in the original (non-reduced) dynamic system. The proof of the second condition works analogously, so we refrain from writing it out. □

Proof of Proposition 6.1

Proof. $(r,\phi,\psi) = (1,1,1)$ is a rest point if $\xi((1,1,1),0) = (1,1,1)$, $\dot{\psi}^* = 0$, $\dot{\phi} = 0$, and $\dot{r} = 0$. $(s(1,1) + p > c(1) \wedge \phi = 1) \Rightarrow 1 \in \Psi^*(1,1)$. $(1 \in \Psi^*(1,1) \wedge$ Condition 1 of Assumption 5.2) $\Rightarrow \xi((1,1,1),0) = (1,1,1)$, $\phi = 1 \Rightarrow \dot{\phi} = 0$, $(\dot{\psi}^* \in \{0,\dot{\phi}\} \wedge \dot{\phi} = 0) \Rightarrow \dot{\psi}^* = 0$, and $(\psi^* = 1 \wedge r = 1) \Rightarrow \dot{r} = 0$. □

Proof of Lemma 6.1

Proof. Suppose $s(1,1) \leq c(1) < \min\{s(1,1)+p, \kappa s(1,1)+\gamma\}$. Consider any neighborhood U of $(1,1,1)$ with distance $\epsilon > 0$: $U := [1-\epsilon]^3$. Let $\tilde{\epsilon}_\psi \in (0,\epsilon)$ be s.t. $\forall \rho \in [1-\tilde{\epsilon}_\psi]^3$, $s(\phi,\psi)+p > c(1)$. Such $\tilde{\epsilon}_\psi$ exists since $s(1,1)+p > c(1)$ and $s(\cdot,\cdot)$ is continuous. Let $\tilde{\epsilon}_\phi \in (0,\tilde{\epsilon}_\psi)$ be s.t. $\forall x \in [1-\tilde{\epsilon}_\phi, 1]$, (a) $\kappa s(x,x)+\gamma > c(1)$ and (b) $e(x) < \max_{r\in[0,1]} \delta(r)$. Such $\tilde{\epsilon}_\phi$ exists since $\kappa s(1,1)+\gamma > c(1)$, $e(1) = 0 < \delta(y)\ \forall y \in (0,1)$, and all involved functions are continuous. Let $\tilde{\epsilon}_r > 0$ be s.t. $e(1-\tilde{\epsilon}_\phi) < \delta(1-\tilde{\epsilon}_r)$. Such $\tilde{\epsilon}_r$ exists since $e(1-\tilde{\epsilon}_\phi) < \max_{r[0,1]} \delta(r)$. Since $e(\cdot)$ is strictly decreasing, $e(x) < \delta(1-\tilde{\epsilon}_r)\ \forall x \in [1-\tilde{\epsilon}_\phi, 1]$. Let $Z := [1-\tilde{\epsilon}_r] \times [1-\tilde{\epsilon}_\phi]^2$. Throughout the following analysis, we consider a neighborhood $\tilde{U} \subset Z$ of $(1,1,1)$ and $\rho = (r,\phi,\psi) \in \tilde{U}$. Note that for any $t \geq 0$, $(\psi^*(\rho,t) = \phi(\rho,t) \wedge \phi(\rho,t) \in [1-\tilde{\epsilon}_\phi, 1)) \Rightarrow \kappa s(\phi(\rho,t),\phi(\rho,t)) + \gamma > c(1) \geq c(r(\rho,t)) \Rightarrow \dot{\phi}(\rho,t) > 0$.

<u>Lyapunov stability:</u> We continue to show that $\xi(\rho,t) \in Z \subset U\ \forall t \geq 0$. First, we establish that $\phi(\rho,t) \geq 1-\tilde{\epsilon}_\phi\ \forall t \geq 0$. Suppose by contradiction this was not the case: $\exists \check{t} \geq 0$ s.t. $\phi(\rho,\check{t}) < 1-\tilde{\epsilon}_\phi$. By continuity of $\phi(\rho,t)$ w.r.t. t, $\exists \hat{t} < \check{t}$ s.t. (a) $\phi(\rho,\hat{t}) = 1-\tilde{\epsilon}_\phi$ and (b) $\phi(\rho,t) < 1-\tilde{\epsilon}_\phi\ \forall t \in (\hat{t},\check{t})$. $\dot{\phi}(\rho,\hat{t}) \geq 0$, since $\psi^*(\rho,\hat{t}) \in \{0,1\} \Rightarrow \dot{\phi}(\rho,\hat{t}) \geq 0$ and $(\psi^*(\rho,\hat{t}) = \phi(\rho,\hat{t}) \wedge \phi(\rho,\hat{t}) = 1-\tilde{\epsilon}_\phi) \Rightarrow \dot{\phi}(\rho,\hat{t}) > 0$. $\dot{\phi}(\rho,\hat{t}) \geq 0 \Rightarrow \exists h > 0$ s.t. $\phi(\rho,t) \geq \phi(\rho,\hat{t}) = 1-\tilde{\epsilon}_\phi\ \forall t \in (\hat{t},\hat{t}+h)$, which is a contradiction. Hence, $\phi(\rho,t) \geq 1-\tilde{\epsilon}_\phi\ \forall t \geq 0$.

Second, we show $\psi^*(\rho,t) \geq 1-\tilde{\epsilon}_\phi\ \forall t \geq 0$. Suppose by contradiction that $\exists \check{t} > 0$ s.t. $\psi^*(\rho,\check{t}) < 1-\tilde{\epsilon}_\phi$. $(\psi^*(\rho,\check{t}) < 1-\tilde{\epsilon}_\phi \wedge \psi^*(\rho,\check{t}) \in \{0, \phi(\rho,\check{t}), 1\} \wedge \phi(\rho,\check{t}) \geq 1-\tilde{\epsilon}_\phi) \Rightarrow \psi^*(\rho,\check{t}) = 0$. Let $\bar{t} := \min\{t \geq 0 : \psi^*(\rho,t) = 0\}$. $\rho \in \tilde{U} \Rightarrow (s(\phi,\phi) + p > c(r) \wedge s(\phi,\psi) + p > c(r))$.

Lemma A.3. $\forall \rho = (r,\phi,\psi)$, $(s(\phi,\phi) + p > c(r) \wedge s(\phi,\psi) + p > c(r)) \Rightarrow \psi^*(\rho,0) \neq 0$.

Proof. First, suppose $s(\phi,\phi) < c(r)$. Hence, $\phi \in \Psi^*(r,\phi)$. If $\psi \geq \phi$, then Condition 2 of Assumption 5.2 implies $\psi^*(\rho,0) \neq 0$. If $\psi < \phi$, then $(\psi < \phi \wedge s(\phi,\psi)+p > c(r)) \Rightarrow \exists \epsilon > 0$ s.t. $s(\phi,x)+p > c(r)\ \forall x \in (\psi-\epsilon,\phi)\cap[0,1] \Rightarrow \psi^*(\rho,0) = \phi$, where the last implication follows from Condition 1 of Assumption 5.2. Second, suppose $s(\phi,\phi) \geq c(r)$. If $\phi = 1$, then $\phi \in \Psi^*(r,\phi)$ and $\psi^*(\rho,0) \neq 0$ by analogous reasoning to the above. If $\phi < 1$, then $s(\phi,\phi) \geq c(r) \Rightarrow s(\phi,1) > c(r) \Rightarrow 1 \in \Psi^*(r,\phi)$. $(s(\phi,\psi) + p > c(r) \wedge s(\phi,\phi) \geq c(r) \wedge s(\phi,1) > c(r)) \Rightarrow \exists \epsilon > 0$ s.t. $s(\phi,x)+p1_{\leq\phi} > c(r)\ \forall x \in (\psi-\epsilon,1)\cap[0,1] \Rightarrow \psi^*(\rho,0) = 1$, where the last implication follows from Condition 1 of Assumption 5.2. □

$(s(\phi,\psi) + p > c(r) \wedge s(\phi,\phi) + p > c(r)) \Rightarrow \psi^*(\rho,0) \neq 0 \Rightarrow \bar{t} \neq 0$. $\psi^*(\rho,t) \neq 0\ \forall t < \bar{t} \Rightarrow \psi^*(\rho,t) \geq \phi(\rho,t) \geq 1-\tilde{\epsilon}_\phi\ \forall t < \bar{t} \Rightarrow s(\lim_{t\to\bar{t}^-}\phi(\rho,t), \lim_{t\to\bar{t}^-}\psi^*(\rho,t)) \geq s(\lim_{t\to\bar{t}^-}\phi(\rho,t), \lim_{t\to\bar{t}^-}\phi(\rho,t)) + p \geq s(1-\tilde{\epsilon}_\phi, 1-\tilde{\epsilon}_\phi) + p > c(1) \geq c(\lim_{t\to\bar{t}^-} r(\rho,t)) \Rightarrow s(\lim_{t\to\bar{t}^-}\phi(\rho,t), \lim_{t\to\bar{t}^-}\psi^*(\rho,t)) + p \geq s(\lim_{t\to\bar{t}^-}\phi(\rho,t), \lim_{t\to\bar{t}^-}\phi(\rho,t)) + p > c(\lim_{t\to\bar{t}^-} r(\rho,t)) \Rightarrow \psi^*(\rho,\bar{t}) \neq 0$ (by analogous reasoning as Lemma A.3). We have reached a contradiction. Hence, $\psi^*(\rho,t) \geq 1-\tilde{\epsilon}_\phi\ \forall t \geq 0$.

Third, we show $r(\rho,t) \geq 1-\tilde{\epsilon}_r\ \forall t \geq 0$. Suppose by contradiction this was not the case: $\exists \check{t} > 0$ s.t $r(\rho,\check{t}) < 1-\tilde{\epsilon}_r$. Since $r(\rho,t)$ is continuous w.r.t t, $\exists \hat{t} \in (0,\check{t})$ s.t. $r(\rho,\hat{t}) = 1-\tilde{\epsilon}_r$ and $r(\rho,t) < 1-\tilde{\epsilon}_r\ \forall t \in (\hat{t},\check{t})$. $\psi^*(\rho,\hat{t}) \geq 1-\tilde{\epsilon}_\phi \Rightarrow \dot{r}(\rho,\hat{t}) > 0 \Rightarrow \exists h > 0$ s.t. $r(\rho,t) \geq r(\rho,\hat{t}) = 1-\tilde{\epsilon}_\phi\ \forall t \in (\hat{t},\hat{t}+h)$, which is a contradiction. Hence, $r(\rho,t) \geq 1-\tilde{\epsilon}_r\ \forall t \geq 0$. Combining the above yields that $\xi(\rho,t) \in Z \subset U\ \forall t \geq 0$. Since U and $\rho \in \tilde{U}$ were arbitrarily chosen, $(1,1,1)$ is Lyapunov stable.

Convergence: Next, we show that $\lim_{t\to\infty}\xi(\rho,t) = (1,1,1)$. First, we show that $\lim_{t\to\infty}\phi(\rho,t) = 1$. Note that $\xi(\rho,t) \in Z \subset U\ \forall t \geq 0 \Rightarrow \dot{\phi}(\rho,t) \geq 0\ \forall t \geq 0 \Rightarrow \lim_{t\to\infty}\phi(\rho,t)$ exists. Before proceeding, we establish the following lemma.

Lemma A.4. $\forall t \geq 0$, $\phi(\rho,t) < 1 \Rightarrow\ \exists \check{t} \geq t$ *s.t.* $\dot{\phi}(\rho,\check{t}) > 0$.

Proof. Consider any $\hat{t} \geq 0$ and suppose $\phi(\rho, \hat{t}) < 1$. Assume by contradiction that the lemma does not hold. Thus, $\phi(\rho, t) = \phi < 1\ \forall t \geq \hat{t}$. For all $t \geq \hat{t}$, $(\psi^*(\rho,t) = \phi(\rho,t) \Rightarrow \dot{\phi}(\rho,t) > 0) \Rightarrow \psi^*(\rho,t) \neq \phi(\rho,t) = \phi \Rightarrow \psi^*(\rho,t) = 1 \Rightarrow \dot{r}(\rho,t) > 0$. Let z be s.t. $s(\phi, 1) = c(z)$ and note that $z < 1$ since $s(\phi,1) < c(1)$ and $c(\cdot)$ is strictly increasing. For all $t \geq \hat{t}$ s.t. $r(\rho,t) < z$, $\dot{r}(\rho,t) > b_l > 0$, where $b_l = \min_{y \in [1-\tilde{\epsilon}_r, z]} \delta(y) > 0$. Since $\dot{r}(\rho,t)$ is bounded below by a positive number if $r(\rho,t) < z$, $\exists \bar{t} \geq \hat{t}$ s.t. $r(\rho,\bar{t}) \geq z$. $r(\rho,\bar{t}) \geq z \Rightarrow s(\phi,1) \leq c(r(\rho,\bar{t})) \Rightarrow \psi^*(\rho,\bar{t}) \neq 1 \Rightarrow \psi^*(\rho,\bar{t}) \in \{0,\phi\}$, which is a contradiction. □

Suppose by contradiction that $\lim_{t\to\infty} \phi(\rho,t) = x$ for some $x < 1$. $\dot{\phi}(\rho,t) \geq 0\ \forall t \geq 0 \Rightarrow \phi(\rho,t) \in [\phi, x]\ \forall t \geq 0$. $\lim_{t\to\infty} \phi(\rho,t) = x$ only if $\lim_{t\to\infty} \dot{\phi}(\rho,t) = 0$. However, $\lim_{t\to\infty} \dot{\phi}(\rho,t) \neq 0$, since Lemma A.4 holds and $\dot{\phi}(\rho,t)$ has a strictly positive lower bound $b_n > 0$ for all t s.t. $\dot{\phi}(\rho,t) > 0$ and $\phi(\rho,t) \in [\phi, x]$. Particularly: If $\gamma = 0$, then $\dot{\phi}(\rho,t) > 0 \Rightarrow \psi^*(\rho,t) = \phi(\rho,t)$. $(\psi^*(\rho,t) = \phi(\rho,t) \wedge \phi(\rho,t) \geq 1 - \tilde{\epsilon}_\phi) \Rightarrow \kappa s(\phi(\rho,t),\phi(\rho,t)) > c(1) \Rightarrow \dot{\phi}(\rho,t) = v\phi(\rho,t)(1-\phi(\rho,t))(\kappa s(\phi(\rho,t),\phi(\rho,t)) - c(r(\rho,t))) > v\phi(\rho,t)(1-\phi(\rho,t))(\kappa s(\phi(\rho,t),\phi(\rho,t)) - c(1)) > 0 \Rightarrow b_n = v \min_{y\in[\phi,x]}[y(1-y)(\kappa s(y,y) - c(1))] > 0$. If $\gamma > 0$, then $((\psi^*(\rho,t) = \phi(\rho,t) \Rightarrow \dot{\phi}(\rho,t) > v\phi(\rho,t)(1-\phi(\rho,t))(\kappa s(\phi(\rho,t),\phi(\rho,t)) + \gamma - c(1)) > 0) \wedge (\psi^*(\rho,t) \in \{0,1\} \Rightarrow \dot{\phi}(\rho,t) > v\phi(\rho,t)(1-\phi(\rho,t))\gamma > 0)) \Rightarrow b_n = v\min\{\min_{y\in[\phi,x]}[y(1-y)(\gamma + \kappa s(y,y) - c(1))], \min_{y\in[\phi,x]}[y(1-y)\gamma]\} > 0$. Hence, $\lim_{t\to\infty}\phi(\rho,t) \neq x$ for $x < 1$, implying that $\lim_{t\to\infty}\phi(\rho,t) = 1$. Note that $(\lim_{t\to\infty}\phi(\rho,t) = 1 \wedge \psi^*(\rho,t) \in \{\phi(\rho,t), 1\}\ \forall t \geq 0) \Rightarrow \lim_{t\to\infty}\psi^*(\rho,t) = 1$.

It remains for us to argue that the solution also converges to $(1,1,1)$ on the r-dimension: $\lim_{t\to\infty} r(\rho,t) = 1$. This holds if $\forall \alpha > 0\ \exists \bar{t} > 0$ s.t. $t > \bar{t} \Rightarrow 1 - \alpha < r(\rho,t)$. Consider any arbitrary $\alpha > 0$. Let $\lambda_r \in (0, \min\{\alpha, \tilde{\epsilon}_r\})$. Let $\lambda_\phi > 0$ be s.t. $\delta(y) > e(1-\lambda_\phi) \forall y \in [1-\tilde{\epsilon}_r, 1-\lambda_r]$. Such λ_ϕ exists since $\delta(y) > 0\ \forall y \in [1-\tilde{\epsilon}_r, 1-\lambda_r]$, $e(1) = 0$, and $e(\cdot)$ is continuous. Let $\hat{t} > 0$ be s.t. $\phi(\rho,t) > 1 - \lambda_\phi\ \forall t > \hat{t}$. Such $\hat{t}$ exists since $\lim_{t\to\infty}\phi(\rho,t) = 1$. $(\psi^*(\rho,t) \in \{\phi(\rho,t),1\} \wedge r(\rho,t) \leq 1-\lambda_r) \Rightarrow \dot{r}(\rho,t) > \min_{y\in[1-\tilde{\epsilon}_r, 1-\lambda_r]}\delta(y) - e(1-\lambda_\phi) > 0$. $\dot{r}(\rho,t)$ is bounded below by a positive number if $r(\rho,t) \leq 1-\lambda_r$ and $t > \hat{t}$. Hence, $\exists \bar{t} \geq \hat{t}$ s.t. $r(\rho,\bar{t}) \geq 1-\lambda_r$. Moreover, for any $\bar{t} \geq \hat{t}$ s.t. $r(\rho,\bar{t}) \geq 1-\lambda_r$, continuity of $r(\rho,t)$ and $(\forall t \geq \bar{t} \geq \hat{t},\ r(\rho,t) = 1-\lambda_r \Rightarrow \dot{r}(\rho,t) > 0)$ imply $\nexists t > \bar{t}$ s.t. $r(\rho,t) < 1-\lambda_r$. Thus, $r(\rho,t) \geq 1-\lambda_r > 1-\alpha\ \forall t > \bar{t}$, which proves $\lim_{t\to\infty} r(\rho,t) = 1$. Consequently,

$\lim_{t\to\infty} \xi(\rho, t) = (1, 1, 1)$, which, coupled with Lyapunov stability, proves that $(1, 1, 1)$ is asymptotically stable. □

Proof of Lemma 6.2

Proof. Suppose $s(1, 1) > c(1)$. Note that $s(1, 1) > c(1)$ implies $(1, 1, 1)$ is a socio-ecological equilibrium of Proposition 6.1. Consider any neighborhood U of $(1, 1, 1)$ s.t. $s(\phi, \psi) > c(r)\ \forall \rho = (r, \phi, \psi) \in U$. Such U exists due to $s(1, 1) > c(1)$ and continuity of all involved functions. For all $\rho = (r, \phi, \psi) \in U$, $s(\phi, \psi) > c(r) \Rightarrow \xi(\rho, 0) = (r, \phi, 1)$ (by Condition 1 of Assumption 5.2).

Only if: Suppose $\gamma = 0$ and consider some $\rho = (1, \phi, 1) \in U \setminus \{(1, 1, 1)\}$. $s(\phi, \psi) > c(r) \Rightarrow \xi(\rho, 0) = (1, \phi, 1) \Rightarrow (\psi^* = 1 \wedge \dot{\psi}^* = 0)$. $(r = 1 \wedge \psi^* = 1) \Rightarrow \dot{r} = 0$. $(\gamma = 0 \wedge \psi^* = 1) \Rightarrow \dot{\phi} = v\phi(1 - \phi)\gamma = 0$. Hence, the dynamic system remains at rest, and the solution $\xi(\rho, t)$ does not converge to $(1, 1, 1)$.

If: Suppose $\gamma > 0$. Since $\xi(\rho, 0) = (r, \phi, 1)\ \forall \rho = (r, \phi, \psi) \in U$, we can restrict our attention to the reduced dynamic system consisting of (1) $\dot{r} = \delta(r) - e(\phi)$ and (2) $\dot{\phi} = v\phi(1-\phi)(\kappa s(\phi, \phi) + \gamma - c(r))$, while $\psi^* = 1$ throughout (see Lemma A.2).

$\psi^* = 1 \Rightarrow \dot{r} \geq 0$. $(\psi^* = 1 \wedge r < 1) \Rightarrow \dot{r} > 0$. Moreover, $\phi = 1 \Rightarrow \dot{\phi} = 0$ and $(\psi^* = 1 \wedge \phi < 1) \Rightarrow \dot{\phi} = v\phi(1 - \phi)\gamma > 0$. Hence, $\xi(\rho, t)$ moves towards $(1, 1, 1)$ at any $\rho \in U \setminus \{(1, 1, 1)\}$. Moreover, dynamics (1) and (2) above are continuous. It follows that $(1, 1, 1)$ is asymptotically stable in the reduced dynamic system and, hence, in the original (non-reduced) one. □

Proof of Proposition 6.3

Proof. (r, ϕ, ϕ) is a rest point if $\xi((r, \phi, \phi), 0) = (r, \phi, \phi)$, $\dot{\psi}^* = 0$, $\dot{\phi} = 0$, and $\dot{r} = 0$. $s(\phi, \phi) < c(r) < s(\phi, \phi) + p \Rightarrow \xi((r, \phi, \phi), 0) = (r, \phi, \phi)$ (by Condition 1 of Assumption 5.2), $\xi((r, \phi, \phi), 0) = (r, \phi, \phi) \Rightarrow \psi^* = \phi \Rightarrow \dot{\psi}^* = \dot{\phi}$, $(\kappa s(\phi, \phi) + \gamma = c(r) \wedge \psi^* = \phi) \Rightarrow \dot{\phi} = 0 = \dot{\psi}^*$, and $(\delta(r) = e(\phi) \wedge \psi^* = \phi) \Rightarrow \dot{r} = 0$. □

Proof of Proposition 6.4

Proof. Consider any $(r,\phi,\phi) \in (0,1]^3$ of Proposition 6.3. Consider some neighborhood U of (r,ϕ,ϕ) s.t. for all $\hat{\rho} = (\hat{r},\hat{\phi},\hat{\psi}) \in U$, $s(\hat{\phi},\hat{\psi}) < c(\hat{r}) < s(\hat{\phi},\hat{\psi}) + p$. Such U exists due to the continuity of all involved functions. Hence, $\xi(\hat{\rho},0) = (\hat{r},\hat{\phi},\hat{\phi})\ \forall \hat{\rho} \in U$. Since $\xi(\hat{\rho},0) = (\hat{r},\hat{\phi},\hat{\phi})\ \forall \hat{\rho} \in U$, we can restrict our attention to the reduced dynamic system consisting of (1) $\dot{\hat{r}} = \delta(\hat{r}) - e(\hat{\phi})$ and (2) $\dot{\hat{\phi}} = v\hat{\phi}(1-\hat{\phi})(\kappa\, s(\phi,\phi) + \gamma - c(\hat{r}))$ to analyze asymptotic stability of (r,ϕ,ϕ) (recall Lemma A.2). Linearization of this reduced system around equilibrium point (r,ϕ,ϕ) yields the Jacobian matrix $J = \begin{pmatrix} \delta'(r) & -e'(\phi) \\ -v\phi(1-\phi)c'(r) & v\phi(1-\phi)\kappa\,(s'_\phi(\phi,\phi)+s'_\psi(\phi,\phi)) \end{pmatrix}$. We obtain determinant $det(J) = \delta'(r) * v\phi(1-\phi)\kappa\,(s'_\phi(\phi,\phi) + s'_\psi(\phi,\phi)) - e'(\phi) * v\phi(1-\phi)c'(r)$ and trace $tr(J) = \delta'(r) + v\phi(1-\phi)\kappa\,(s'_\phi(\phi,\phi) + s'_\psi(\phi,\phi))$. The equilibrium point is asymptotically stable if $det(J) > 0$ and $tr(J) < 0$, which is equivalent to $\frac{\delta'(r)}{e'(\phi^*)} < \frac{c'(r)}{\kappa\,(s'_\phi(\phi,\phi)+s'_\psi(\phi,\phi))}$ and $v\phi(1-\phi)\kappa\,(s'_\phi(\phi,\phi) + s'_\psi(\phi,\phi)) < -\delta'(r)$. The conditions of Proposition 6.4 imply this holds. It follows that (r,ϕ,ϕ) is asymptotically stable in the reduced dynamic system and, hence, in the original (non-reduced) one, which proves the proposition. □

Proof of Proposition 6.5

Proof. Suppose $c(0) < s(1,0) + p < s(1,1) + p \leq c(1)$ and let $0 < r_{min} < r_{max} \leq 1$ s.t. $c(r_{min}) = s(1,0) + p < s(1,1) + p = c(r_{max})$. First, consider any $\rho = (r,\phi,\psi) \in \{(r,1,1)\}_{r_{min} \leq r < r_{max}}$. $\phi = 1 \Rightarrow \dot{\phi} = 0$. $s(1,1) + p > c(r) \Rightarrow \psi^* = 1$. $(\psi^* = 1 \wedge r < 1) \Rightarrow \dot{r} \geq \min_{x \in [r_{min}, r_{max}]} \delta(x) > 0$. $(\dot{\psi}^* \in \{0, \dot{\phi}\} \wedge \dot{\phi} = 0) \Rightarrow \dot{\psi}^* = 0$. Hence, $\dot{\phi} = 0$, $\dot{\psi}^* = 0$, and $\dot{r} \geq \min_{x \in [r_{min}, r_{max}]} \delta(x) > 0$. It follows that when starting at any $\rho \in \{(r,1,1)\}_{r_{min} \leq r < r_{max}}$, the dynamic system moves across $\{(r,1,1)\}_{r_{min} \leq r < r_{max}}$, with r moving towards r_{max}, while $\phi = 1$ and $\psi^* = 1$ remain unaltered. Since $\dot{r}$ is bounded below by a strictly positive number, $\min_{x \in [r_{min}, r_{max}]} \delta(x) > 0$, the system eventually reaches r_{max}. However, $s(1,1) + p = c(r_{max}) \Rightarrow 1 \notin \Psi^*(r_{max}, 1) \Rightarrow \xi((r_{max},1,1),0) \neq (r_{max},1,1)$. $(\xi((r_{max},1,1),0) \neq (r_{max},1,1) \wedge \phi = 1) \Rightarrow \xi((r_{max},1,1),0) = (r_{max},1,0)$. When the resource reaches r_{max}, equilibrium behavior transits to $\psi^* = 0$.

Second, consider any $\rho = (r,\phi,\psi) \in \{(r,1,0)\}_{r_{min} < r \leq r_{max}}$. $\phi = 1 \Rightarrow \dot{\phi} = 0$. $s(1,0) + p < c(r) \Rightarrow \psi^* = 0$, $(\psi^* = 0 \wedge r > 0) \Rightarrow \dot{r} \leq \max_{x \in [r_{min}, r_{max}]} \delta(x) - e(0) < 0$, and $(\dot{\psi}^* \in \{0, \dot{\phi}\} \wedge \dot{\phi} = 0) \Rightarrow \dot{\psi}^* = 0$. Hence, $\dot{\phi} = 0$, $\dot{\psi}^* = 0$, and $\dot{r} < \max_{x \in [r_{min}, r_{max}]} \delta(x) - e(0) < 0$. Hence, starting at any $\rho \in \{(r,1,0)\}_{r_{min} < r \leq r_{max}}$, the dynamic system moves across $\rho \in \{(r,1,0)\}_{r_{min} < r \leq r_{max}}$, with r moving towards r_{min}, while $\phi = 1$ and $\psi^* = 0$ remain unaltered. Since $\dot{r}$ is bounded above by a strictly negative number, $\min_{x\in[r_{min},r_{max}]} \delta(x) - e(0) < 0$, the system eventually reaches r_{min}. However, $s(1,0) + p = c(r_{min}) \Rightarrow 0 \notin \Psi^*(r_{min}, 1) \Rightarrow \xi((r_{min}, 1, 0), 0) \neq (r_{min}, 1, 0)$. $\xi((r_{min}, 1, 0), 0) \neq (r_{min}, 1, 0) \Rightarrow \xi((r_{min}, 1, 0), 0) \in \{(r_{min}, 1, \phi), (r_{min}, 1, 1)\}$. $\phi = 1 \Rightarrow \xi((r_{min}, 1, 0), 0) = (r_{min}, 1, 1)$. When the resource stock reaches r_{min}, equilibrium behavior transits to $\psi^* = 1$.

The above implies a cyclic movement across Ω_{BLC}, which, in turn, implies that (a) for all $\rho \in \Omega_{BLC}$ and $t \geq 0$, $\xi(\rho, t) \in \Omega_{BLC}$ and (b) $\forall \hat{\rho}, \check{\rho} \in \Omega_{BLC}$, $\exists t > 0$ s.t. $\xi(\hat{\rho}, t) = \check{\rho}$. (a) and (b) jointly imply that $cl(\Omega_{BLC})$ is a socio-ecological equilibrium, and (a) implies the first bullet point of Proposition 6.5. □

Lemma A.5. *Suppose $cl(\Omega_{BLC})$ of Proposition 6.5 exists. $\kappa s(1,1) + \gamma > s(1,1) + p \Rightarrow cl(\Omega_{BLC})$ is asymptotically stable.*

Proof. Assume $\gamma + \kappa s(1,1) > s(1,1) + p$ and Ω_{BLC} of Proposition 6.5 exists. Let r_{max} and r_{min} be the maximum and minimum resource levels of Ω_{BLC}, respectively. Consider any neighborhood U of $cl(\Omega_{BLC})$ with distance $\epsilon > 0$. Let $\tilde{\epsilon}_r \in (0, \epsilon)$ be s.t. (1) $0 < r_{min} - \tilde{\epsilon}_r$ and (2) $c(y) \in (s(1,1) + p, \kappa s(1,1) + \gamma)\ \forall y \in (r_{max}, r_{max} + \tilde{\epsilon}_r] \cap [0,1]$. Such $\tilde{\epsilon}_r$ exists due to continuity of all involved functions as well as (1) $r_{min} > 0$ and (2) $c(r_{max}) = s(1,1) + p < \kappa s(1,1) + \gamma$. Let $\tilde{\epsilon}_\phi \in (0, \epsilon)$ be s.t. for all $x \in [1 - \tilde{\epsilon}_\phi, 1]$, (1) $e(x) < \delta(z)\ \forall z \in [r_{min} - \tilde{\epsilon}_r, r_{min}]$, (2) $c(r_{min} - \tilde{\epsilon}_r) < p + s(x, 0)$, and (3) $\kappa s(x,x) + \gamma > s(x,x) + p > s(x,1)$. Such $\tilde{\epsilon}_\phi$ exists due to continuity of all involved functions and (1) $e(1) < \delta(r)\ \forall r \in (0,1)$, (2) $c(r_{min} - \tilde{\epsilon}_r) < c(r_{min}) = p + s(1,0)$, and (3) $\kappa s(1,1) + \gamma > s(1,1) + p > s(1,1)$. Since $\max\{\tilde{\epsilon}_\phi, \tilde{\epsilon}_r\} < \epsilon$, $Z := [r_{min} - \tilde{\epsilon}_r, \min\{r_{max} + \tilde{\epsilon}_r, 1\}] \times [1 - \tilde{\epsilon}_\phi, 1] \times ([0, \tilde{\epsilon}_\phi] \cup [1 - \tilde{\epsilon}_\phi, 1]) \subset U$. Note, $[r_{min} - \tilde{\epsilon}_r, \min\{r_{max} + \tilde{\epsilon}_r, 1\}] = [r_{min} - \tilde{\epsilon}_r, r_{max} + \tilde{\epsilon}_r] \cap [0,1]$. Throughout the following analysis, we consider any neighborhood $\tilde{U} \subset Z$ of $cl(\Omega_{BLC})$ and $\rho = (r, \phi, \psi) \in \tilde{U}$. Before proceeding, we establish results we use repeatedly at later stages of the proof.

Lemma A.6. *$\forall t > 0$ s.t. $\xi(\rho, t) \in Z$: (a) $\psi^*(\rho, t) = \phi(\rho, t) < 1 \Rightarrow \dot{\phi}(\rho, t) > 0$, (b) $\dot{\phi}(\rho, t) \geq 0$, (c) $(\psi^*(\rho, t) \in \{\phi(\rho, t), 1\} \wedge r(\rho, t) \in [r_{min} - \tilde{\epsilon}_r, r_{min}]) \Rightarrow \dot{r}(\rho, t) \geq b_l$ for some $b_l > 0$, and (d) $\psi^*(\rho, t) = 0 \Rightarrow \dot{r}(\rho, t) \leq b_u$ for some $b_u < 0$.*

Proof. Consider any $\xi(\rho,t) \in Z$. (a) follows from: $\xi(\rho,t) \in Z \Rightarrow \phi(\rho,t) \geq 1 - \tilde{\epsilon}_\phi \Rightarrow p + s(\phi(\rho,t),\phi(\rho,t)) < \gamma + \kappa s(\phi(\rho,t),\phi(\rho,t)) \Rightarrow (\psi^*(\rho,t) = \phi(\rho,t) < 1 \Rightarrow c(r(\rho,t)) < p + s(\phi(\rho,t),\phi(\rho,t)) < \gamma + \kappa s(\phi(\rho,t),\phi(\rho,t)) \Rightarrow \dot{\phi}(\rho,t) > 0)$. (b) follows from (a) and $\psi^*(\rho,t) \neq \phi(\rho,t) \Rightarrow \dot{\phi}(\rho,t) \geq 0$. (c) holds since $\forall x \in [1-\tilde{\epsilon}_\phi, 1]$ and $z \in [r_{min} - \tilde{\epsilon}_r, r_{min}]$, $e(x) < \delta(z)$, with an obvious candidate for the lower bound being $b_l = \min_{(z,x)\in[r_{min}-\tilde{\epsilon}_r, r_{min}]\times[1-\tilde{\epsilon}_\phi,1]}[\delta(z) - e(x)] > 0$. (d) holds due to the specifications of $e(0)$ and $\delta(\cdot)$, with an obvious example of the upper bound being $b_u = \max_{y\in[0,1]}[\delta(y) - e(0)] < 0$. □

Lyapunov stability: We show that $\rho \in \tilde{U} \Rightarrow \xi(\rho,t) \in Z \subset U \; \forall t \geq 0$. Suppose by contradiction that it is not true: $\exists \rho \in \tilde{U}$ and $\hat{t} > 0$ s.t. $\xi(\rho,\hat{t}) = (\hat{r},\hat{\phi},\hat{\psi}^*) \notin Z$. $((\hat{r},\hat{\phi},\hat{\psi}^*) \notin Z \wedge \hat{\psi}^* \in \{0,\hat{\phi},1\}) \Rightarrow (\hat{\phi} < 1 - \tilde{\epsilon}_\phi \vee \hat{r} \notin [r_{min} - \tilde{\epsilon}_r, \min\{r_{max} + \tilde{\epsilon}_r, 1\}])$. Since $r(\rho,t)$ and $\phi(\rho,t)$ are continuous w.r.t. t, there is $\bar{t} \in (0,\hat{t})$ s.t.: (1) for all $t \in (\bar{t},\hat{t}]$, $\phi(\rho,t) < 1 - \tilde{\epsilon}_\phi \vee r(\rho,t) \notin [r_{min} - \tilde{\epsilon}_r, \min\{r_{max} + \tilde{\epsilon}_r, 1\}]$ and (2) (a) $\phi(\rho,\bar{t}) = 1 - \tilde{\epsilon}_\phi \vee r(\rho,\bar{t}) \in \{r_{min} - \tilde{\epsilon}_r, r_{max} + \tilde{\epsilon}_r\}$ and (b) $\phi(\rho,\bar{t}) \geq 1 - \tilde{\epsilon}_\phi \wedge r(\rho,\bar{t}) \in [r_{min} - \tilde{\epsilon}_r, \min\{r_{max} + \tilde{\epsilon}_r, 1\}]$. Consider such $\bar{t}$. Note that (2) implies $\xi(\rho,\bar{t}) \in Z$. We can derive the following: (A) $(\xi(\rho,\bar{t}) \in Z \wedge$ (b) of Lemma A.6 $) \Rightarrow \dot{\phi}(\rho,\bar{t}) \geq 0 \Rightarrow (\exists h_\phi > 0$ s.t. $\forall t \in (\bar{t},\bar{t}+h_\phi)$, $\phi(\rho,t) \geq 1 - \tilde{\epsilon}_\phi \wedge \psi^*(\rho,t) \in \{0\} \cup [1-\tilde{\epsilon}_\phi, 1])$. Consider such h_ϕ. And (B) (i) $(r(\rho,\bar{t}) = r_{max} + \tilde{\epsilon}_r \wedge \phi(\rho,\bar{t}) \geq 1 - \tilde{\epsilon}_\phi) \Rightarrow c(r_{max} + \tilde{\epsilon}_r) > c(r_{max}) \geq s(1,1) + p \geq s(\phi(\rho,\bar{t}),\phi(\rho,\bar{t})) + p > s(\phi(\rho,\bar{t}),1) \Rightarrow \psi^*(\rho,\bar{t}) = 0 \Rightarrow \dot{r}(\rho,\bar{t}) < 0$ and (ii) $(r(\rho,\bar{t}) = r_{min} - \tilde{\epsilon}_r \wedge \phi(\rho,\bar{t}) \geq 1 - \tilde{\epsilon}_\phi) \Rightarrow c(r_{min} - \tilde{\epsilon}_r) \leq s(\phi(\rho,\bar{t}),0) + p \Rightarrow \psi^*(\rho,\bar{t}) \neq 0$. $(\psi^*(\rho,\bar{t}) \neq 0 \wedge$ (c) of Lemma A.6 $) \Rightarrow \dot{r}(\rho,\bar{t}) > 0$. (i) and (ii) imply $r(\rho,\bar{t}) \in \{r_{min} - \tilde{\epsilon}_r, r_{max} + \tilde{\epsilon}_r\} \Rightarrow (h_r > 0$ s.t. $\forall t \in (\bar{t},\bar{t}+h_r)$, $r(\rho,t) \in (r_{min} - \tilde{\epsilon}_r, \min\{r_{max} + \tilde{\epsilon}_r, 1\}))$. Consider such h_r. (A) and (B) imply that for $h := \min\{h_r, h_\phi\}$ and all $t \in (\bar{t},\bar{t}+h)$, $\xi(\rho,t) \in Z$. This is a contradiction to (1), implying that $\forall \rho \in \tilde{U}$, $\xi(\rho,t) \in Z \subset U \; \forall t \geq 0$. Since U was arbitrarily chosen, this proves that $cl(\Omega_{BLC})$ is Lyapunov stable.

Convergence: Next, we show that for any $\rho \in \tilde{U}$, the solution $\xi(\rho,t)$ converges to Ω_{BLC}. Consider any $\rho = (r,\phi,\psi) \in \tilde{U}$.

Lemma A.7. *$\forall t > 0$, $\phi(\rho,t) < 1 \Rightarrow \exists \hat{t} > t$ s.t. $\dot{\phi}(\rho,\hat{t}) > 0$.*

Proof. To see why the lemma holds, suppose by contradiction that it does not and $\phi < 1$. Thus, $\phi(\rho,t) = \phi < 1 \; \forall t \geq 0$. This is possible only if $\nexists \tilde{t}$ s.t. $\psi^*(\rho,\tilde{t}) = \phi$, since $(\psi^*(\rho,\tilde{t}) = \phi(\rho,\tilde{t}) \wedge \xi(\rho,\tilde{t}) \in Z) \Rightarrow \dot{\phi}(\rho,\tilde{t}) > 0$ (see (a) of Lemma A.6). $\forall t \geq 0$ s.t. $\psi^*(\rho,t) = 0$, $\exists \tilde{t} > t$ s.t. $\psi^*(\rho,\tilde{t}) \neq 0$, due to the following contradiction: $\psi^*(\rho,t) = 0\ \forall t > 0 \Rightarrow \dot{r}(\rho,t) < b_u$ for some $b_u < 0\ \forall t > 0$ (i.e. $\dot{r}(\rho,t)$ is bounded above by a negative value for all $r(\rho,t) \in [r_{min} - \tilde{\epsilon}_r, r_{max} + \tilde{\epsilon}_r]$ due to (d) of Lemma A.6) $\Rightarrow \exists \tilde{t} > 0$ s.t. $r(\rho,\tilde{t}) = r_{min} - \tilde{\epsilon}_r \Rightarrow \exists \tilde{t} > 0$ s.t. $s(\phi,0) + p = c(r_{min}) > c(r(\rho,\tilde{t})) = c(r_{min} - \tilde{\epsilon}_r) \Rightarrow \exists \tilde{t} > 0$ s.t. $\psi^*(\rho,\tilde{t}) \neq 0$. Thus, $\forall t > 0\ \exists \tilde{t} > t$ s.t. $\psi^*(\rho,\tilde{t}) \neq 0$. Clearly, this implies $\exists \tilde{t} > 0$ s.t. $\psi^*(\rho,\tilde{t}) \neq 0$. Consider any such $\tilde{t}$. Since $\psi^*(\rho,\tilde{t}) \neq 0$ and $\psi^*(\rho,\tilde{t}) \neq \phi$, $\psi^*(\rho,\tilde{t}) = 1$. Let us now investigate what happens after $\tilde{t}$. By similar reasoning as above and considering $r_{max} + \tilde{\epsilon}_r$ instead of $r_{min} - \tilde{\epsilon}_r$, we can show: $\forall \tilde{t} \geq 0$ s.t. $\psi^*(\rho,\tilde{t}) = 1$, $\exists t > \tilde{t}$ s.t. $\psi^*(\rho,t) \neq 1$. Let $\bar{t} := \min\{t > 0 : \psi^*(\rho,t) \neq 1\}$. Note, $\lim_{t\to\bar{t}^-} \psi^*(\rho,t) = 1$. $\psi^*(\rho,t) = 1\ \forall t \in (\tilde{t},\bar{t}) \Rightarrow (\psi^*(\rho,\bar{t}) \neq 1 \Rightarrow 1 \notin \Psi^*(r(\rho,\bar{t}),\phi) \Rightarrow s(\phi,1) \leq c(r(\rho,\bar{t}))$. $\psi^*(\rho,t) = 1 \forall t \in (\tilde{t},\bar{t}) \Rightarrow s(\phi,1) > c(r(\rho,t))\ \forall t \in (\tilde{t},\bar{t})$. By continuity of $r(\rho,t)$ and $c(\cdot)$, $s(\phi,1) = c(r(\rho,\bar{t}))$. $(s(\phi,1) = c(r(\rho,\bar{t})) \wedge s(\phi,\phi) < s(\phi,1) \wedge s(\phi,1) < s(\phi,\phi) + p) \Rightarrow s(\phi,\phi) < c(r(\rho,\bar{t})) < s(\phi,\phi) + p \Rightarrow \phi \in \Psi^*(r(\rho,\bar{t}),\phi)$. $(\phi \in \Psi^*(r(\rho,\bar{t}),\phi) \wedge \lim_{t\to\bar{t}^-} \psi^*(\rho,t) = 1 > \phi \wedge$ condition 2 of Assumption 5.2) $\Rightarrow \xi(\rho,\bar{t}) = \xi(\lim_{t\to\bar{t}^-} \xi(\rho,t), 0) = \xi((r(\rho,\bar{t}),\phi,1),0) \neq (r(\rho,\bar{t}),\phi,0) \Rightarrow \psi^*(\rho,\bar{t}) = \phi$, which is a contradiction. Hence, if $\phi < 1$, then $\exists \bar{t} > 0$ s.t. $\dot{\phi}(\rho,\bar{t}) > 0$. From the above follows that $\forall t > 0$, $\phi(\rho,t) < 1 \Rightarrow \exists \bar{t} > t$ s.t. $\dot{\phi}(\rho,\bar{t}) > 0$. □

We now show that $\lim_{t\to\infty} \phi(\rho,t) = 1$. Since $\dot{\phi}(\rho,t) \geq 0\ \forall t \geq 0$ (recall Lemma A.6), $\lim_{t\to\infty} \phi(\rho,t)$ exists. Suppose by contradiction that $\lim_{t\to\infty} \phi(\rho,t) = x$ for some $x < 1$. $\dot{\phi}(\rho,t) \geq 0\ \forall t \geq 0$ implies $\phi(\rho,t) \leq x\ \forall t > 0$. $\lim_{t\to\infty} \phi(\rho,t) = x$ only if $\lim_{t\to\infty} \dot{\phi}(\rho,t) = 0$. However, $\lim_{t\to\infty} \dot{\phi}(\rho,t) \neq 0$, since (1) $\dot{\phi}(\rho,t)$ has a strictly positive lower bound $b_n > 0$ for all t s.t. $\dot{\phi}(\rho,t) > 0$ and $\phi(\rho,t) \in [\phi,x]$ and (2) Lemma A.7. Particularly: If $\gamma = 0$, then $\dot{\phi}(\rho,t) > 0 \Rightarrow \psi^*(\rho,t) = \phi(\rho,t)$. $(\psi^*(\rho,t) = \phi(\rho,t) \wedge \phi(\rho,t) \geq 1 - \tilde{\epsilon}_\phi) \Rightarrow \kappa s(\phi(\rho,t),\phi(\rho,t)) > s(\phi(\rho,t),\phi(\rho,t)) + p > c(r(\rho,t)) \Rightarrow \dot{\phi}(\rho,t) = v\phi(\rho,t)(1-\phi(\rho,t))(\kappa s(\phi(\rho,t),\phi(\rho,t)) - c(r(\rho,t))) > v\phi(\rho,t)(1-\phi(\rho,t))(\kappa s(\phi(\rho,t),\phi(\rho,t)) - s(\phi(\rho,t),\phi(\rho,t)) - p) > 0 \Rightarrow b_n = v\min_{y\in[\phi,x]}[y(1-y)(\kappa s(y,y) - s(y,y) - p)] > 0$. If $\gamma > 0$, then $((\psi^*(\rho,t) = \phi(\rho,t) \Rightarrow \dot{\phi}(\rho,t) > v\phi(\rho,t)(1-\phi(\rho,t))(\kappa s(\phi(\rho,t),\phi(\rho,t)) + \gamma - s(\phi(\rho,t),\phi(\rho,t)) - p) > 0) \wedge (\psi^*(\rho,t) \in \{0,1\} \Rightarrow \dot{\phi}(\rho,t) > v\phi(\rho,t)(1-\phi(\rho,t))\gamma > 0)) \Rightarrow b_n = v\min\{\min_{y\in[\phi,x]}[y(1-y)(\gamma + \kappa s(y,y) - s(y,y) - p)], \min_{y\in[\phi,x]}[y(1-y)\gamma]\} > 0$. Hence, $\nexists x < 1$

s.t. $\lim_{t\to\infty}\phi(\rho,t) = x \Rightarrow \lim_{t\to\infty}\phi(\rho,t) = 1$. Note that $(\lim_{t\to\infty}\phi(\rho,t) = 1 \wedge \psi^*(\rho,t) \in \{0,\phi(\rho,t),1\}) \Rightarrow \psi^*(\rho,t)$ converges to $\{0,1\}$.

It remains for us to argue that the solution also converges to $cl(\Omega_{BLC})$ on the r-dimension: $\forall\alpha > 0\ \exists\bar{t} > 0$ s.t. $t > \bar{t} \Rightarrow r(\rho,t) \in [r_{min} - \alpha, r_{max} + \alpha]$. Consider any $\alpha > 0$. $r(\rho,t) \geq r_{max} \Rightarrow c(r(\rho,t)) \geq c(r_{max}) = s(1,1) + p \geq s(\phi(\rho,t),\phi(\rho,t)) + p > s(\phi(\rho,t),1) \Rightarrow \psi^*(\rho,t) = 0$. In combination with (d) of Lemma A.6, it follows that $\dot{r}(\rho,t)$ has an upper bound $b_u < 0$ whenever $r(\rho,t) \geq r_{max}$. Hence, $\exists\bar{t}_1$ s.t. $r(\rho,\bar{t}_1) \leq r_{max}$. Consider such $\bar{t}_1$. Continuity of $r(\rho,t)$ and $(r(\rho,t) = r_{max} \Rightarrow \dot{r}(\rho,t) < 0)$ imply $\nexists t > \bar{t}_1$ s.t. $r(\rho,t) > r_{max}$. Hence, $r(\rho,t) \leq r_{max} < r_{max} + \alpha$, $\forall t \geq \bar{t}_1$.

Next, consider any $\lambda \in (0, \min\{\tilde{\epsilon}_r, \alpha\})$ Let z be s.t. $c(r_{min} - \lambda) < s(z,0) + p$, which exists since $c(r_{min} - \lambda) < c(r_{min}) \leq s(1,0) + p$ and all involved functions are continuous. Note that for all $x \geq z$ and $y \leq r_{min} - \lambda$, $c(y) < s(x,0) + p$. Let $\hat{t}$ be s.t. $\phi(\rho,t) > z\ \forall t \geq \hat{t}$. Such $\hat{t}$ exists since $\lim_{t\to\infty}\phi(\rho,t) = 1$. $(t \geq \hat{t} \wedge r(\rho,t) \leq r_{min} - \lambda) \Rightarrow s(\phi(\rho,t),0) + p \geq s(\phi(\rho,\hat{t}),0) + p > c(r_{min} - \lambda) \geq c(r(\rho,t)) \Rightarrow \psi^*(\rho,t) \neq 0 \Rightarrow \psi^*(\rho,t) \in \{\phi(\rho,t),1\} \Rightarrow \dot{r}(\rho,t) > \min_{y\in[r_{min}-\tilde{\epsilon}_r, r_{min}]}\delta(y) - e(\phi(\rho,\hat{t})) > 0$. $\dot{r}(\rho,t)$ is bounded below by a positive number if $r(\rho,t) \leq 1 - \lambda_r$ and $t > \hat{t}$. Hence, $\exists\bar{t}_2 \geq \hat{t}$ s.t. $r(\rho,\bar{t}_2) \geq 1 - \lambda_r$. Moreover, for any $\bar{t}_2 \geq \hat{t}$ s.t. $r(\rho,\bar{t}_2) \geq 1 - \lambda_r$, continuity of $r(\rho,t)$, and $(\forall t \geq \bar{t}_2 \geq \hat{t},\ r(\rho,t) = 1 - \lambda \Rightarrow \dot{r}(\rho,t) > 0)$ imply $\nexists t > \bar{t}_2$ s.t. $r(\rho,t) < 1 - \lambda$. Thus, $r(\rho,t) \geq 1 - \lambda > 1 - \alpha\ \forall t \geq \bar{t}_2$.

Combining the above yields for $\bar{t} = \min\{\bar{t}_1, \bar{t}_2\}$, $t > \bar{t} \Rightarrow r(\rho,t) \in [r_{min} - \lambda, r_{max} + \lambda] \subset [r_{min} - \alpha, r_{max} + \alpha]$. Since α was chosen arbitrarily, $r(\rho,t)$ converges to $[r_{min}, r_{max}]$ and, thus, $\xi(\rho,t)$ to $cl(\Omega_{BLC})$. Coupled with Lyapunov stability, this proves that $cl(\Omega_{BLC})$ is asymptotically stable. □

Lemma A.8. *For any specification of the model: If $\gamma > 0$, then $\exists\bar{p} > 0$ s.t. ($p < \bar{p}$ and $cl(\Omega_{BLC})$ of Proposition 6.5 exists) $\Rightarrow cl(\Omega_{BLC})$ is asymptotically stable.*

Proof. Consider any specification of the model s.t. $s(1,1) < c(1)$. Otherwise, $\nexists p > 0$ s.t. the boundary limit cycle exists. Suppose $s(1,0) > c(0)$. Otherwise, $\exists\bar{p}$ s.t. $p < \bar{p} \Rightarrow cl(\Omega_{BLC})$ of Proposition 6.5 does not exist, which renders the lemma never not true. Throughout, let $r_1 < r_2$ be s.t. $s(1,0) < c(r_1) < c(r_2) < s(1,1)$. Such r_1 and r_2 exist by continuity. Consider any $\bar{p}_1$ s.t. $s(1,0) + \bar{p}_1 < c(r_1) < s(1,1) + \bar{p}_1 < c(1)$, which exists since $s(1,0) <$

$c(r_1) < s(1,1) < c(1)$. Let $[c^{-1}](\cdot)$ be the inverse of $c(\cdot)$ and note that it must be (uniquely) defined on the domain $[c(0), c(1)]$ due to monotonicity of $c(\cdot)$. Consider any $\alpha > 0$ s.t. $0 < [c^{-1}](s(1,0)) - \alpha < [c^{-1}](s(1,1) + \bar{p}_1) + \alpha < 1$. Such $\alpha > 0$ exists since $c(0) < s(1,0) < s(1,1) + \bar{p}_1 < c(1)$. Let $\bar{R} := [[c^{-1}](s(1,0)) - \alpha, [c^{-1}](s(1,1) + \bar{p}_1) + \alpha] \subset (0,1)$. Let $\beta > 0$ be s.t. (1) $\min_{y \in \bar{R}} \delta(y) > e(1-\beta)$ and (2) $c(r_2) < s(1-\beta, 1-\beta)$, which must exist since (1) $\delta(r) > 0 \; \forall r \in (0,1)$ and (2) $c(r_2) < s(1,1)$. Let $\Delta t_{\phi\uparrow} > 0$ be s.t. $r_1 - r_2 = -\int_0^{\Delta t_{\phi\uparrow}} e(0) d\tau$. Given some $p > 0$, let $B_r := \max_{z \in [0,1]} \{[c^{-1}](s(z,z) + p) - [c^{-1}](s(z,z)) : c(0) \leq s(z,z) < s(z,z) + p \leq c(1)\} > 0$. Let $\Delta t_{\phi\downarrow} > 0$ be s.t. $B_r = \int_0^{\Delta t_{\phi\downarrow}} \min_{y \in \bar{R}} \delta(y) - e(1-\beta) d\tau$. Let $\bar{p} < \bar{p}_1$ be s.t. $p < \bar{p} \Rightarrow \Delta t_{\phi\downarrow} v c(1) < \Delta t_{\phi\uparrow} v \gamma$. Such $\bar{p}$ exists: As p approaches 0, so does B_r, and, since $\min_{y \in \bar{R}} \delta(y) - e(1-\beta) > 0$, so does $\Delta t_{\phi\downarrow}$.

Consider any $p < \bar{p}$ and suppose Ω_{BLC} of proposition 6.5 exists. We denote the minimum and maximum resource levels of Ω_{BLC} by r_{min} and r_{max}, respectively. Note that $p < \bar{p}_1 \Rightarrow [r_1, r_2] \subset (r_{min}, r_{max})$. We proceed to show that $cl(\Omega_{BLC})$ is asymptotically stable. For this purpose, consider an arbitrary neighborhood U of $cl(\Omega_{BLC})$ with distance ϵ. Consider any $\lambda_r \in (0, \min\{\epsilon, \alpha\})$ and $\lambda_\phi \in (0, \min\{\epsilon, \beta\})$ s.t. $s(1-\lambda_\phi, 0) > c(r_{min} - \lambda_r)$ and $s(1-\lambda_\phi, 1-\lambda_\phi) + p > s(1-\lambda_\phi, 1)$. Such λ_ϕ exists since $s(1,0) > c(r_{min} - \lambda_r)$ and $s(1,1) + p > s(1,1)$. Let $Z := [r_{min} - \lambda_r, r_{max} + \lambda_r] \times [1-\lambda_\phi] \times ([0, \lambda_\phi] \cup [1-\lambda_\phi, 1])$. Let $\tilde{\epsilon} \in (0, \min\{\lambda_r, \lambda_\phi\})$ be so small that $\frac{(1-\tilde{\epsilon})\exp(-vc(1)\Delta t_{\phi\downarrow})}{(1-\tilde{\epsilon})\exp(-vc(1)\Delta t_{\phi\downarrow}) + \tilde{\epsilon}} > 1 - \lambda_\phi$, which must exist since the term approaches 1 from below as $\tilde{\epsilon}$ approaches 0. Let $\tilde{U}$ be the neighborhood of $cl(\Omega_{BLC})$ with distance $\tilde{\epsilon}$.

Note that $\lambda_\phi < \beta$ and $\lambda_r < \alpha$ imply: (a) $s(x,1) < s(x,x) + p \; \forall x \in [1-\lambda_\phi, 1]$, (b) $c(r_2) < s(x,x) + p \; \forall x \in [1-\lambda_\phi, 1]$, (c) $[r_{min} - \lambda_r, r_{max} + \lambda_r] \subset \bar{R} \subset [0,1]$, (d) $\delta(z) - e(x) > \min_{y \in \bar{R}} \delta(y) - e(1-\beta) \; \forall z \in [r_{min} - \lambda_r, r_{max} + \lambda_r], x \in [1-\lambda_\phi, 1]$, and (e) $s(1-\lambda_\phi, 0) > c(r_{min} - \lambda_r)$.

Throughout, consider any $\rho \in \tilde{U} \subset Z$. Below, we show $\xi(\rho, t)$ always remains in $Z \subset U$ and converges to $cl(\Omega_{BLC})$. For any $t > 0$ s.t. $\xi(\rho, t) \in Z$, (a) $r(\rho, t) = r_{min} - \lambda_r \Rightarrow s(\phi(\rho, t), 0) + p > c(r(\rho, t)) \Rightarrow \psi^*(\rho, t) \neq 0 \Rightarrow \dot{r}(\rho, t) \geq \min_{y \in \bar{R}} \delta(y) - e(1-\lambda_\phi) > 0$ and (b) $r(\rho, t) = r_{max} + \lambda_r \Rightarrow s(\phi(\rho, t), 0) < s(\phi(\rho, t), \phi(\rho, t)) + p < c(r(\rho, t)) \Rightarrow \psi^*(\rho, t) = 0 \Rightarrow \dot{r}(\rho, t) \leq \min_{y \in \bar{R}} \delta(y) - e(0) < 0$. Analogously to the proof of Proposition 6.6, the above allows us to infer that for any $\tilde{t}$ s.t. $\phi(\rho, t) > 1 - \lambda_\phi \; \forall t \leq \tilde{t}$, $r(\rho, t) \in [r_{min} - \lambda_r, r_{max} + \lambda_r] \; \forall t \leq \tilde{t}$. Moreover, if $\phi(\rho, t)$ converges to 1, then $\psi^*(\rho, t)$ converges to $\{0, 1\}$ and, analogously to the proof of Proposition 6.6, we can show $r(\rho, t)$ converges to $[r_{min}, r_{max}]$. Hence, to prove asymptotic stability, it suffices to show that $\phi(\rho, t) > 1 - \lambda_\phi > 1 - \epsilon\ \forall t \geq 0$ (ensuring Lyapunov stability) and $\lim_{t\to\infty} \phi(\rho, t) = 1$ (ensuring convergence).

$\phi(\rho, 0) = 1 \Rightarrow \dot{\phi}(\rho, t) = 0\ \forall t \geq 0 \Rightarrow \phi(\rho, t) = 1\ \forall t \geq 0$. Hence, $\xi(\rho, t)$ stays close and converges to $cl(\Omega_{BLC})$. Below, we study the other case: $\phi(\rho, 0) < 1$. Due to the asymptotic nature of $\dot{\phi}(\rho, t)$ when $\phi(\rho, t)$ approaches 1, $\nexists t > 0$ s.t. $\phi(\rho, t) = 1$. In the following, we present Lemmas A.9, A.10, A.11, and A.12 and their corresponding proofs, which we then use to complete the proof of this proposition.

Lemma A.9. *For any $t_0 \geq 0$ s.t. $\xi(\rho, t_0) \in \tilde{U}$ and $\phi(\rho, t_0) < 1$, there is $\bar{t} \geq t_0$ s.t. $\psi^*(\rho, \bar{t}) = \phi(\rho, \bar{t})$.*

Proof. Recall that $\phi(\rho, t_0) < 1 \Rightarrow \phi(\rho, t) < 1\ \forall t > t_0$. Let $\hat{t} := \min\{t \geq t_0 : \psi^*(\rho, \hat{t}) \neq 0\}$. Such $\hat{t}$ exists since otherwise $\dot{r}(\rho, t) < \max_{y\in[0,1]} \delta(y) - e(0) < 0\ \forall t \geq 0$ and $\dot{\phi}(\rho, t) \geq 0\ \forall t \geq 0$, which imply $\exists \tilde{t}$ s.t. $r(\rho, \tilde{t}) < r_{min} - \lambda_r$ and $\phi(\rho, t) \geq \phi(\rho, t_0) > 1 - \lambda_\phi\ \forall t \leq \tilde{t}$, which contradicts the results on $r(\rho, t)$. Let $\check{t} := \min\{t \geq \hat{t} : \psi^*(\rho, \hat{t}) \neq 1\}$. Such $\check{t}$ exists since otherwise $\dot{r}(\rho, t) \geq \min_{y\in[r_{min}-\lambda_r, r_{max}+\lambda_r]} \delta(y) > 0\ \forall t \geq \hat{t}$ s.t. $r(\rho, t) \leq r_{max} + \lambda_r$ and $\dot{\phi}(\rho, t) \geq 0\ \forall t \geq \hat{t}$, which imply $\exists \tilde{t}$ s.t. $r(\rho, \tilde{t}) > r_{max} + \lambda_r$ and $\phi(\rho, t) \geq \phi(\rho, t_0) > 1 - \lambda_\phi\ \forall t \leq \tilde{t}$, which is also a contradiction. If $\check{t} = \hat{t}$, then $\psi^*(\rho, \hat{t}) = \phi(\rho, \hat{t})$. Alternatively, suppose $\check{t} > \hat{t}$. Hence, $\psi^*(\rho, t) = 1\ \forall t \in [\hat{t}, \check{t}) \Rightarrow s(\phi(\rho, t), 1) > c(r(\rho, t))\ \forall t \in [\hat{t}, \check{t})$. $\psi^*(\rho, \check{t}) \neq 1 \Rightarrow s(\phi(\rho, \check{t}), 1) \leq c(r(\rho, \check{t}))$, and, by continuity, $s(\phi(\rho, \check{t}), 1) = c(r(\rho, \check{t}))$. $\psi^*(\rho, t) \in \{0, 1\}\ \forall t \in [t_0, \check{t}) \Rightarrow \dot{\phi}(\rho, t) = 0\ \forall t \in [t_0, \check{t}) \Rightarrow \phi(\rho, \check{t}) \geq \phi(\rho, t_0) > 1 - \lambda_\phi \Rightarrow s(\phi(\rho, \check{t}), \phi(\rho, \check{t})) + p > s(\phi(\rho, \check{t}), 1) \Rightarrow s(\phi(\rho, \check{t}), \phi(\rho, \check{t})) < c(r(\rho, \check{t})) < s(\phi(\rho, \check{t}), \phi(\rho, \check{t})) + p$. Hence, $\phi(\rho, \check{t}) \in \Psi^*(r(\rho, \check{t}), \phi(\rho, \check{t}))$. Assumption 5.2 implies $\psi^*(\rho, \check{t}) = \phi(\rho, \check{t})$, which proves the lemma. □

Lemma A.10. *Consider some t_1 s.t. $\xi(\rho, t_1) \in \tilde{U}$ and $\psi^*(\rho, t_1) = \phi(\rho, t_1) < 1$, let $t_2 := \min\{t > t_1 : \psi^*(\rho, t) \neq \phi(\rho, t)\}$. It holds that*

1. *such t_2 exists,*
2. *$\psi^*(\rho, t_2) = 0$,*
3. *$r(\rho, t_2) > r_2$, and*

4. $\phi(\rho,t) \geq \frac{\phi(\rho,t_1)\exp(-vc(1)\Delta t_{\phi\downarrow})}{\phi(\rho,t_1)\exp(-vc(1)\Delta t_{\phi\downarrow})+1-\phi(\rho,t_1)} > 1-\lambda_\phi \; \forall t \in [t_1,t_2]$.

Proof. Consider any situation as described in the lemma. Before proving the four conditions, we establish that for any $\check{t} > t_1$ s.t. $\psi^*(\rho,t) = \phi(\rho,t) \; \forall t \in [t_1,\check{t})$, $\phi(\rho,t) > 1-\lambda_\phi \; \forall t \in [t_1,\check{t}]$. Note that such $\check{t}$ exists due to $\psi^*(\rho,t_1) = \phi(\rho,t_1)$ and the piecewise nature of $\xi(\rho,t)$. Assume by contradiction that $\exists t \in [t_1,\check{t}]$ s.t. $\phi(\rho,t) \leq 1-\lambda_\phi$. Let $\bar{t} := \min\{t \in (t_1,\check{t}] : \phi(\rho,t) = 1-\lambda_\phi\}$. Such $\bar{t}$ exists by continuity of $\phi(\rho,t)$. For all $t \in [t_1,\bar{t}]$, $\dot{\phi}(\rho,t)$ is bounded below by $-\phi(\rho,t)(1-\phi(\rho,t))vc(1) < 0$, implying that $f(t)$ s.t. (a) $f(t_1) = \phi(\rho,t_1)$ and (b) $\dot{f}(t) = -f(t)(1-f(t))vc(1)$ is a lower bound of $\phi(\rho,t)$ for all $t \in [t_1,\bar{t}]$. Solving the initial value problem for $f(t)$ and substituting $f(t_1) = \phi(\rho,t_1)$ yields $f(t) = \frac{\phi(\rho,t_1)\exp(-vc(1)(\bar{t}-t_1))}{\phi(\rho,t_1)\exp(-vc(1)(\bar{t}-t_1))+1-\phi(\rho,t_1)}$. $\phi(\rho,\bar{t}) = 1-\lambda_\phi = \phi(\rho,t_1) + \int_{t_1}^{\bar{t}} \dot{\phi}(\rho,\tau)d\tau \geq f(t_1) - vc(1)\int_{t_1}^{\bar{t}} f(\tau)(1-f(\tau))d\tau = \frac{\phi(\rho,t_1)\exp(-vc(1)(\bar{t}-t_1))}{\phi(\rho,t_1)\exp(-vc(1)(\bar{t}-t_1))+1-\phi(\rho,t_1)}$. Recall, $1-\tilde{\epsilon} < \phi(\rho,t_1) \Rightarrow 1-\lambda_\phi < \frac{\phi(\rho,t_1)\exp(-vc(1)\Delta t_{\phi\downarrow})}{\phi(\rho,t_1)\exp(-vc(1)\Delta t_{\phi\downarrow})+1-\phi(\rho,t_1)}$. Hence, $(\frac{\phi(\rho,t_1)\exp(-vc(1)(\bar{t}-t_1))}{\phi(\rho,t_1)\exp(-vc(1)(\bar{t}-t_1))+1-\phi(\rho,t_1)} \leq 1-\lambda_\phi < \frac{\phi(\rho,t_1)\exp(-vc(1)\Delta t_{\phi\downarrow})}{\phi(\rho,t_1)\exp(-vc(1)\Delta t_{\phi\downarrow})+1-\phi(\rho,t_1)}) \Rightarrow \Delta t_{\phi\downarrow} < \bar{t}-t_1$. $(t_1+\Delta t_{\phi\downarrow} < \bar{t} \wedge \psi^*(\rho,t) = \phi(\rho,t) \geq 1-\lambda_\phi \; \forall t \in [t_1,\bar{t}]) \Rightarrow r(\rho,\bar{t}) = r(\rho,t_1) + \int_{t_1}^{\bar{t}} \dot{r}(\rho,\tau)d\tau \geq r(\rho,t_1) + \int_{t_1}^{\bar{t}} \min_{y\in\bar{R}}[\delta(y)-e(1-\beta)]d\tau \geq r(\rho,t_1) + \int_{t_1}^{t_1+\Delta t_{\phi\downarrow}} \min_{y\in\bar{R}}[\delta(y)-e(1-\beta)]d\tau = r(\rho,t_1)+B_r \Rightarrow r(\rho,\bar{t})-r(\rho,t_1) \geq B_r$. However, (a) $[c^{-1}](s(\phi(\rho,t_1),\phi(\rho,t_1))+p) - [c^{-1}](s(\phi(\rho,t_1),\phi(\rho,t_1))) \leq B_r$ by definition of B_r, (b) $\phi(\rho,t_1) = \psi^*(\rho,t_1) \Rightarrow r(\rho,t_1) > [c^{-1}](s(\phi(\rho,t_1),\phi(\rho,t_1)))$, (c) $\phi(\rho,\bar{t}) \leq \phi(\rho,t_1) \Rightarrow [c^{-1}](s(\phi(\rho,\bar{t}),\phi(\rho,\bar{t}))+p) \leq [c^{-1}](s(\phi(\rho,t_1),\phi(\rho,t_1))+p)$, and (d) $\phi(\rho,\bar{t}) = \psi^*(\rho,\bar{t}) \Rightarrow r(\rho,\bar{t}) < [c^{-1}](s(\phi(\rho,\bar{t}),\phi(\rho,\bar{t}))+p)$ imply: $\phi(\rho,t_1) = \psi^*(\rho,t_1) > \phi(\rho,\bar{t}) = \psi^*(\rho,\bar{t}) \Rightarrow r(\rho,\bar{t})-r(\rho,t_1) < B_r$. We reached a contradiction. We continue to prove the conditions of the lemma.

(1) $\exists t_2 > t_1$ s.t. $\psi^*(\rho,t_2) \neq \phi(\rho,t_2)$, since otherwise $\phi(\rho,t) \geq 1-\lambda_\phi \; \forall t \geq t_1$ and, hence, $\dot{r}(\rho,t) > \min_{y\in[r_{min}-\lambda_r,r_{max}+\lambda_r]} \delta(y)-e(1-\lambda_\phi) > 0 \; \forall t \geq t_1$ s.t. $r(\rho,t) \leq r_{max}+\lambda_r$, which implies $\exists\tilde{t}$ s.t. $r(\rho,\tilde{t}) > r_{max}+\lambda_r$ and $\phi(\rho,t) > 1-\lambda_\phi \; \forall t \in [t_1,\tilde{t}]$, which is a contradiction to our results on r. Throughout the following, we consider such t_2. (2) $\psi^*(\rho,t) = \phi(\rho,t) < 1 \; \forall t \in [t_1,t_2) \Rightarrow s(\phi(\rho,t),\phi(\rho,t)) < c(r(\rho,t)) < s(\phi(\rho,t),\phi(\rho,t))+p \; \forall t \in [t_1,t_2)$. $(\psi^*(\rho,t) = \phi(\rho,t) \; \forall t \in [t_1,t_2) \wedge \psi^*(\rho,t_2) \neq \phi(\rho,t_2)) \Rightarrow \phi(\rho,t_2) \notin \Psi^*(r(\rho,t_2),\phi(\rho,t_2))$. $(\phi(\rho,t_2) \notin \Psi^*(r(\rho,t_2),\phi(\rho,t_2)) \wedge \phi(\rho,t_1) \in \Psi^*(r(\rho,t_1),\phi(\rho,t_1)) \wedge \phi(\rho,t_2) < \phi(\rho,t_1) \wedge r(\rho,t_2) > r(\rho,t_1)) \Rightarrow s(\phi(\rho,t_2),\phi(\rho,t_2)) < s(\phi(\rho,t_2),\phi(\rho,t_2))+p \leq c(r(\rho,t_2))$. $\phi(\rho,t_2) > 1-\lambda_\phi \Rightarrow s(\phi(\rho,t_2),1) < s(\phi(\rho,t_2),\phi(\rho,t_2))+p \leq c(r(\rho,t_2)) \Rightarrow$

$1 \notin \Psi^*(r(\rho,t_2),\phi(\rho,t_2))$. $(1 \notin \Psi^*(r(\rho,t_2),\phi(\rho,t_2)) \wedge \phi(\rho,t_2) \notin \Psi^*(r(\rho,t_2),\phi(\rho,t_2))) \Rightarrow \psi^*(\rho,t_2) = 0$. (3) $r_2 < r(\rho,t_2)$ follows from $s(1-\lambda_\phi, 1-\lambda_\phi)+p > c(r_2)$, $\phi(\rho,t_2) > 1-\lambda_\phi$, $s(\phi(\rho,t_2),\phi(\rho,t_2))+p \leq c(r(\rho,t_2))$, and $s(\cdot,\cdot)$ and $c(\cdot)$ being increasing. (4) If $\phi(\rho,t_1) \leq \phi(\rho,t_2)$, Condition 4 of the lemma is true since $\phi(\rho,t_1) \geq \frac{\phi(\rho,t_1)\exp(-vc(1)\Delta t_{\phi\downarrow})}{\phi(\rho,t_1)\exp(-vc(1)\Delta t_{\phi\downarrow})+1-\phi(\rho,t_1)} > 1-\lambda_\phi$. In the following, we suppose $\phi(\rho,t_1) > \phi(\rho,t_2)$. To prove the condition, we first show that $t_2 < t_1 + \Delta t_{\phi\downarrow}$. $(\psi^*(\rho,t) = \phi(\rho,t) > 1-\lambda_\phi \wedge r(\rho,t) \leq r_{max}+\lambda_r \; \forall t \in [t_1,t_2)) \Rightarrow \dot{r}(\rho,t) > \min_{y\in[r_{min}-\lambda_r, r_{max}+\lambda_r]} \delta(y) > 0 \; \forall t \in [t_1,t_2) \Rightarrow r(\rho,t_1) < r(\rho,t_2)$. By analogous reasoning to the last implication in the first paragraph of this lemma's proof, $\psi^*(\rho,t_2) = \phi(\rho,t_2) < \psi^*(\rho,t_1) = \phi(\rho,t_1) \Rightarrow r(\rho,t_2) - r(\rho,t_1) < B_r$. t_2 is s.t. $r(\rho,t_2) - r(\rho,t_1) = \int_{t_1}^{t_2} \dot{r}(\rho,\tau)d\tau$. Hence, t_2 increases in $r(\rho,t_2) - r(\rho,t_1)$ and decreases in $\dot{r}(\rho,t)$ (for any $t \in [t_1,t_2)$). Consequently, $B_r > r(\rho,t_2) - r(\rho,t_1)$ and $\min_{y\in\bar{R}}[\delta(y) - e(1-\beta)] < \dot{r}(\rho,t) \; \forall t \in [t_1,t_2]$ imply $t_2 - t_1 < \Delta t_{\phi\downarrow}$ (recall $\Delta t_{\phi\downarrow}$ solves $B_r = \int_0^{\Delta t_{\phi\downarrow}} \min_{y\in\bar{R}} \delta(y) - e(1-\beta)d\tau$) and, thus, $t < t_1 + \Delta t_{\phi\downarrow} \; \forall t \in [t_1,t_2]$. Recall that for all $t \in [t_1,t_2]$, $\phi(\rho,t)$ is bounded below by $f(t) = \frac{\phi(\rho,t_1)\exp(-vc(1)(t-t_1))}{\phi(\rho,t_1)\exp(-vc(1)(t-t_1))+1-\phi(\rho,t_1)} \; \forall t \in [t_1,t_2]$. For all $t \in [t_1,t_2]$, $\phi(\rho,t) = \phi(\rho,t_1) + \int_{t_1}^{t} \dot{\phi}(\rho,\tau)d\tau \geq f(t) = f(t_1) - vc(1)\int_{t_1}^{t} f(\tau)(1-f(\tau))d\tau = \frac{\phi(\rho,t_1)\exp(-vc(1)(t-t_1))}{\phi(\rho,t_1)\exp(-vc(1)(t-t_1))+1-\phi(\rho,t_1)} \geq \frac{\phi(\rho,t_1)\exp(-vc(1)\Delta t_{\phi\downarrow})}{\phi(\rho,t_1)\exp(-vc(1)\Delta t_{\phi\downarrow})+1-\phi(\rho,t_1)}$. Lastly, $\phi(\rho,t_1) > 1-\tilde{\epsilon} \Rightarrow \frac{\phi(\rho,t_1)\exp(-vc(1)\Delta t_{\phi\downarrow})}{\phi(\rho,t_1)\exp(-vc(1)\Delta t_{\phi\downarrow})+1-\phi(\rho,t_1)} > 1-\lambda_\phi$. □

Lemma A.11. *Consider $t_2 > 0$ s.t. $\psi^*(\rho,t_2) \neq \phi(\rho,t_2) \in (1-\lambda_\phi,1)$, $\lim_{t\to t_2^-} \psi^*(\rho,t) = \phi(\rho,t_2)$, and $r(\rho,t_2) \in (r_2, r_{max}+\lambda_r]$. Let $t_3 := \min\{t > t_2 : \psi^*(\rho,t_2) = \phi(\rho,t_2)\}$. It holds that*

1. *such t_3 exists*
2. *$t_3 - t_2 > \Delta t_{\phi\uparrow}$, and*
3. *$\phi(\rho,t_3) \geq \frac{\phi(\rho,t_2)\exp(v\gamma\Delta t_{\phi\uparrow})}{\phi(\rho,t_2)\exp(v\gamma\Delta t_{\phi\uparrow})+1-\phi(\rho,t_2)} > \phi(\rho,t_2)$.*

Proof. By analogous reasoning as in Lemma A.10, $\psi^*(\rho,t_2) = 0$. Lemma A.9 implies t_3 exists. Let $\hat{t} := \min\{t > t_2 : \psi^*(\rho,t) \neq 0\}$. Such $\hat{t}$ exists since t_3 exists. $\hat{t} = \min\{t > t_2 : \psi^*(\rho,t) \neq 0\} \Rightarrow \psi^*(\rho,t) = 0 \; \forall t \in [t_2,\hat{t}) \Rightarrow s(\phi(\rho,0)) + p < c(r(\rho,t)) \; \forall t \in [t_2,\hat{t})$ and $\psi^*(\rho,\hat{t}) \neq 0 \Rightarrow s(\phi(\rho,\hat{t}),0) + p \geq c(r(\rho,\hat{t}))$. By continuity, $s(\phi(\rho,\hat{t}),0) + p = c(r(\rho,\hat{t}))$. $(r_{min} = s(1,0) + p \wedge \phi(\rho,\hat{t}) < 1) \Rightarrow r(\rho,\hat{t}) < r_{min}$. Continuity of $r(\rho,t)$, $\dot{r}(\rho,t) < 0 \; \forall t \in$

$[t_2, \hat{t})$, and $r(\rho, \hat{t}) < r_{min} < r_1 < r_2 < r(\rho, t_2))$ imply $\exists \tilde{t}_1, \tilde{t}_2 \in (t_2, \hat{t})$ s.t. $r(\rho, \tilde{t}_1) = r_2$, $r(\rho, \tilde{t}_2) = r_1$, and $\tilde{t}_2 > \tilde{t}_1$. $0 > r(\rho, \tilde{t}_2) - r(\rho, \tilde{t}_1) = \int_{\tilde{t}_1}^{\tilde{t}_2} \dot{r}(\rho, \tau) d\tau$, which indicates that $\tilde{t}_2$ decreases in $\dot{r}(\rho, t)$ (for any $t \in [r_1, r_2]$). Since $-e(0) \leq \dot{r}(\rho, t)\ \forall t \geq 0$, $\Delta t_{\phi\uparrow}$ provides a lower bound on $\tilde{t}_2 - \tilde{t}_1 < \hat{t} - t_2 < t_3 - t_2$, which proves the second condition. Moreover, $\phi(\rho, t_3) = \phi(\rho, t_2) + \int_{t_2}^{t_3} \dot{\phi}(\rho, \tau) d\tau \geq \phi(\rho, t_2) + \int_{t_2}^{\hat{t}} \dot{\phi}(\rho, \tau) d\tau = \phi(\rho, t_2) + \int_{t_2}^{\hat{t}} v\gamma\phi(\rho, \tau)(1 - \phi(\rho, \tau)) d\tau = \frac{\phi(\rho,t_2)\exp(v\gamma(\hat{t}-t_2))}{\phi(\rho,t_2)\exp(v\gamma(\hat{t}-t_2))+1-\phi(\rho,t_2)} \geq \frac{\phi(\rho,t_2)\exp(v\gamma\Delta t_{\phi\uparrow})}{\phi(\rho,t_2)\exp(v\gamma\Delta t_{\phi\uparrow})+1-\phi(\rho,t_2)} > \phi(\rho, t_2)$, where we obtain $\frac{\phi(\rho,t_2)\exp(v\gamma(\hat{t}-t_2))}{\phi(\rho,t_2)\exp(v\gamma(\hat{t}-t_2))+1-\phi(\rho,t_2)}$ by solving the differential equation. □

Lemma A.12. *Consider some t_1 s.t. $\xi(\rho, t_1) \in \tilde{U}$ and $\psi^*(\rho, t_1) = \phi(\rho, t_1) < 1$. Let*

- $t_2 := \min\{t > t_1 : \psi^*(\rho, t) \neq \phi(\rho, t)\}$,
- $t_3 := \min\{t > t_2 : \psi^*(\rho, t) = \phi(\rho, t)\}$, *and*
- $t_4 := \min\{t > t_3 : \psi^*(\rho, t) \neq \phi(\rho, t)\}$.

The following is true:

1. $\phi(\rho, t_4) \geq \frac{\phi(\rho,t_2)\exp(v\gamma\Delta t_{\phi\uparrow} - vc(1)\Delta t_{\phi\downarrow})}{\phi(\rho,t_2)\exp(v\gamma\Delta t_{\phi\uparrow} - vc(1)\Delta t_{\phi\downarrow})+1-\phi(\rho,t_2)} > \phi(\rho, t_2) > 1 - \lambda_\phi$,
2. $\phi(\rho, t) > \phi(\rho, t_2)\ \forall t \in [t_2, t_4]$, *and*
3. $t_4 - t_2 > \Delta t_{\phi\uparrow}$.

Proof. Consider any t_1, t_2, t_3, t_4 as described. The previous lemmas imply their existence. Moreover, we know that $\phi(\rho, t_2) \geq 1 - \lambda_\phi$ (from Lemma A.10), $\phi(\rho, t_3) \geq \frac{\phi(\rho,t_2)\exp(v\gamma\Delta t_{\phi\uparrow})}{\phi(\rho,t_2)\exp(v\gamma\Delta t_{\phi\uparrow})+1-\phi(\rho,t_2)}$ (from Lemma A.11), and $\phi(\rho, t) \geq \frac{\phi(\rho,t_3)\exp(-vc(1)\Delta t_{\phi\downarrow})}{\phi(\rho,t_3)\exp(-vc(1)\Delta t_{\phi\downarrow})+1-\phi(\rho,t_3)}\ \forall t \in (t_3, t_4]$ (from Lemma A.10). Since $\frac{\phi(\rho,t_3)\exp(-vc(1)\Delta t_{\phi\downarrow})}{\phi(\rho,t_3)\exp(-vc(1)\Delta t_{\phi\downarrow})+1-\phi(\rho,t_3)}$ increases in $\phi(\rho, t_3)$ and we know $\phi(\rho, t_3) \geq \frac{\phi(\rho,t_2)\exp(v\gamma\Delta t_{\phi\uparrow})}{\phi(\rho,t_2)\exp(v\gamma\Delta t_{\phi\uparrow})+1-\phi(\rho,t_2)}$, we can substitute $\frac{\phi(\rho,t_2)\exp(v\gamma\Delta t_{\phi\uparrow})}{\phi(\rho,t_2)\exp(v\gamma\Delta t_{\phi\uparrow})+1-\phi(\rho,t_2)}$ for $\phi(\rho, t_3)$ into $\frac{\phi(\rho,t_3)\exp(-vc(1)\Delta t_{\phi\downarrow})}{\phi(\rho,t_3)\exp(-vc(1)\Delta t_{\phi\downarrow})+1-\phi(\rho,t_3)}$ to obtain a lower bound for $\phi(\rho, t)\ \forall t \in (t_3, t_4]$. Doing so and simplifying yields $\frac{\phi(\rho,t_2)\exp(v\gamma\Delta t_{\phi\uparrow} - vc(1)\Delta t_{\phi\downarrow})}{\phi(\rho,t_2)\exp(v\gamma\Delta t_{\phi\uparrow} - vc(1)\Delta t_{\phi\downarrow})+1-\phi(\rho,t_2)}$, which is strictly larger than $\phi(\rho, t_2)$ since $v\gamma\Delta t_{\phi\uparrow} - vc(1)\Delta t_{\phi\downarrow} > 0$. Hence, the first statement of the Lemma is true. $\phi(\rho, t) > \phi(\rho, t_2)\ \forall t \in [t_2, t_4]$ follows since (a) $\psi^*(\rho, t) \neq \phi(\rho, t)\ \forall t \in [t_2, t_3) \Rightarrow \dot{\phi}(\rho, t) > 0\ \forall t \in [t_2, t_3) \Rightarrow \phi(\rho, t) \geq \phi(\rho, t_2) \geq \forall t \in [t_2, t_3)$, (b) $\phi(\rho, t_3) \geq \frac{\phi(\rho,t_2)\exp(v\gamma\Delta t_{\phi\uparrow})}{\phi(\rho,t_2)\exp(v\gamma\Delta t_{\phi\uparrow})+1-\phi(\rho,t_2)} > \phi(\rho, t_2)$, and (c) $\phi(\rho, t) \geq$

$\frac{\phi(\rho,t_3)\exp(-vc(1)\Delta t_{\phi\downarrow})}{\phi(\rho,t_3)\exp(-vc(1)\Delta t_{\phi\downarrow})+1-\phi(\rho,t_3)} > \frac{\phi(\rho,t_2)\exp(v\gamma\Delta t_{\phi\uparrow}-vc(1)\Delta t_{\phi\downarrow})}{\phi(\rho,t_2)\exp(v\gamma\Delta t_{\phi\uparrow}-vc(1)\Delta t_{\phi\downarrow})+1-\phi(\rho,t_2)} > \phi(\rho,t_2)\ \forall t \in (t_3, t_4]$. The last condition follows from $t_4 - t_2 > t_3 - t_2$ and $t_3 - t_2 > \Delta t_{\phi\uparrow}$ (see Lemma A.11). □

Recall $\rho \in \tilde{U}$. Lemma A.9 implies there is $t_1 := \min\{t \geq 0 : \psi^*(\rho, \bar{t}) = \phi(\rho, \bar{t})\}$. If $t_1 = 0$, then $\psi^*(\rho, t_1) = \phi(\rho, 0) > 1 - \lambda_\phi$. If $t_1 > 0$, then for all $t \in [0, t_1)$, $\psi^*(\rho, t) \neq \phi(\rho, t) \Rightarrow \dot{\phi}(\rho, t) \geq 0 \Rightarrow \phi(\rho, t_1) \geq \phi(\rho, 0) > 1 - \lambda_\phi$.

Let T be the set of all discontinuities t_i of $\psi^*(\rho, t)$ s.t. $\lim_{x \to t_i} \psi^*(\rho, x) = \phi(\rho, t_i) \neq \psi^*(\rho, t_1)$ (such as t_2 and t_4 in Lemma A.12), where the elements of T are indexed in ascending order with elements from the set of even integers $2\mathbb{Z}$ (e.g., $t_2 < t_4 < t_6 < ...$). Note that T is infinite, since Lemma A.10 and $\phi(\rho, t_1) > 1 - \lambda_\phi$ imply $\exists t_2 \in T$ and Lemmas A.10 and A.11 imply that for each $t_i \in T$ there is $t_{i+2} \in T$. Condition 3 of Lemma A.12 implies that t_i tends to infinity as i does. Lemma A.12 states that for all $t_i, t_{i+2} \in T$, $\phi(\rho, t) \geq \phi(\rho, t_i)\ \forall t \in [t_i, t_{i+2}]$. It follows that $\phi(\rho, t) \geq \phi(\rho, t_i)\ \forall t \geq t_i$. Moreover, since $\phi(\rho, t) > 1 - \lambda_\phi\ \forall t \in [0, t_2]$, it must hold that $\phi(\rho, t) > 1 - \lambda_\phi\ \forall t \geq 0$. Hence, $cl(\Omega_{BLC})$ is Lyapunov stable. Lemma A.12 also states that for all $t_i, t_{i+2} \in T$, $\phi(\rho, t_i) - \phi(\rho, t_{i+2}) \geq \frac{\phi(\rho,t_i)\exp(v\gamma\Delta t_{\phi\uparrow}-vc(1)\Delta t_{\phi\downarrow})}{\phi(\rho,t_i)\exp(v\gamma\Delta t_{\phi\uparrow}-vc(1)\Delta t_{\phi\downarrow})+1-\phi(\rho,t_i)} - \phi(\rho, t_i) > 0$. $\frac{\phi(\rho,t_i)\exp(v\gamma\Delta t_{\phi\uparrow}-vc(1)\Delta t_{\phi\downarrow})}{\phi(\rho,t_i)\exp(v\gamma\Delta t_{\phi\uparrow}-vc(1)\Delta t_{\phi\downarrow})+1-\phi(\rho,t_i)} - \phi(\rho, t_i)$ approaches 0 only as $\phi(\rho, t_i)$ converges to 1. Hence, as $i \in 2\mathbb{Z}$ approaches infinity, $\phi(\rho, t_i)$ converges to 1. Since $\phi(\rho, t) \geq \phi(\rho, t_i)\ \forall t \geq t_i$, $\phi(\rho, t)$ converges to 1. Coupled with Lyapunov stability, we can infer that $cl(\Omega_{BLC})$ is asymptotically stable. □

Lemma A.13. *Consider two specifications of the dynamic model that differ only in social sanctions and material costs: $\hat{s}(\cdot,\cdot), \hat{c}(\cdot)$ and $\check{s}(\cdot,\cdot), \check{c}(\cdot)$. Let $\hat{\xi}(\rho, t)$ and $\check{\xi}(\rho, t)$ be the solutions to the dynamic systems with $\hat{s}(\cdot,\cdot), \hat{c}(\cdot)$ and $\check{s}(\cdot,\cdot), \check{c}(\cdot)$ respectively. Suppose (1) $\hat{\Omega}_{BLC}$ of Proposition 6.5 exists at $\hat{s}(\cdot,\cdot), \hat{c}(\cdot)$, (2) $\check{\Omega}_{BLC}$ of Proposition 6.5 exists at $\check{s}(\cdot,\cdot), \check{c}(\cdot)$, and (3) $\check{s}(\phi, \psi) \geq \hat{s}(\phi, \psi)\ \forall \phi, \psi \in [0,1]$ and $\check{c}(r) \leq \hat{c}(r)\ \forall r \in (0,1]$.*

1. *Suppose that for some $N < 1$ and all $\phi > N$, (a) $[\hat{c}^{-1}](\kappa\hat{s}(\phi,\phi) + \gamma) = [\check{c}^{-1}](\kappa\check{s}(\phi,\phi) + \gamma) \in (0,1)$ and (b) $x < \kappa x + \gamma < x + p\ \forall x \in \{\hat{s}(\phi,\phi), \check{s}(\phi,\phi)\}$. There is U of $\hat{\Omega}_{BLC} \cup \check{\Omega}_{BLC}$ s.t.*

   (a) $\{\rho \in U : \check{\xi}(\rho, 0) = (r, \phi, \phi) \wedge \dot{\phi} > 0$ at $\check{s}(\cdot,\cdot), \check{c}(\cdot)\} \subseteq \{\rho \in U : \hat{\xi}(\rho, 0) = (r, \phi, \phi) \wedge \dot{\phi} > 0$ at $\hat{s}(\cdot,\cdot), \hat{c}(\cdot)\}$ and

(b) $\{\rho \in U : \hat{\xi}(\rho, 0) = (r, \phi, \phi) \wedge \dot{\phi} < 0 \textit{ at } \hat{s}(\cdot,\cdot), \hat{c}(\cdot)\} \subseteq \{\rho \in U : \check{\xi}(\rho, 0) = (r, \phi, \phi) \wedge \dot{\phi} < 0 \textit{ at } \check{s}(\cdot,\cdot), \check{c}(\cdot)\}$.

2. *For all ρ s.t. $\hat{\xi}(\rho, 0) = \check{\xi}(\rho, 0) = (r, \phi, \phi)$, $\dot{\phi}$ is larger at $\hat{s}(\phi, \psi), \hat{c}(r)$ than at $\check{s}(\phi, \psi), \check{c}(r)$.*

Proof. First, consider $N < 1$ and suppose that for all $\phi > N$, (a) $[\hat{c}^{-1}](\kappa \hat{s}(\phi, \phi) + \gamma) = [\check{c}^{-1}](\kappa \check{s}(\phi, \phi) + \gamma)$ and (b) $x < \kappa x + \gamma < x + p \ \forall x \in \{\hat{s}(\phi, \phi), \check{s}(\phi, \phi)\}$. Let us look at Condition 1a. Consider some U of $\check{\Omega}_{BLC} \cup \hat{\Omega}_{BLC}$ s.t. $\phi > N \ \forall \rho \in U$. Consider some $\rho \in U$ s.t. $\check{\xi}(\rho, 0) = (r, \phi, \phi)$ and $\dot{\phi} > 0$ at $\check{s}(\cdot,\cdot), \check{c}(\cdot)$. Note that $\dot{\phi} > 0 \Rightarrow \phi < 1$. $(\psi^* = \phi \wedge \dot{\phi} > 0$ at $\check{s}(\cdot,\cdot), \check{c}(\cdot)) \Rightarrow \kappa \check{s}(\phi, \phi) + \gamma > \check{c}(r) \Rightarrow \kappa \hat{s}(\phi, \phi) + \gamma > \hat{c}(r)$. To see why this holds, consider $\bar{r}$ s.t. $\kappa \check{s}(\phi, \phi) + \gamma = \check{c}(\bar{r})$. Since $[\hat{c}^{-1}](\kappa \hat{s}(\phi, \phi) + \gamma) = [\check{c}^{-1}](\kappa \check{s}(\phi, \phi) + \gamma)$, it follows $\kappa \hat{s}(\phi, \phi) + \gamma = \hat{c}(\bar{r})$. $\kappa \check{s}(\phi, \phi) + \gamma > \check{c}(r) \Rightarrow r < \bar{r} \Rightarrow \kappa \hat{s}(\phi, \phi) + \gamma > \hat{c}(r)$. $\kappa \hat{s}(\phi, \phi) + \gamma > \hat{c}(r) \Rightarrow \hat{s}(\phi, \phi) + p > \hat{c}(r)$. $\check{\xi}(\rho, 0) = (r, \phi, \phi) \Rightarrow \check{c}(r) > \check{s}(\phi, \phi) \Rightarrow \hat{c}(r) > \hat{s}(\phi, \phi)$. Combining these insights, yields $\hat{s}(1, 1) < \hat{c}(r) < \hat{s}(1, 1) + p$ and $\kappa \hat{s}(1, 1) + \gamma > \hat{c}(r)$, which imply $\hat{\xi}(\rho, 0) = (r, \phi, \phi)$ and $\dot{\phi} > 0$ at $\hat{s}(\cdot,\cdot), \hat{c}(\cdot)$. The proof for Condition 1b is analogous. Lastly, Condition 2 follows straightaway from how $\dot{\phi}$ depends on $s(\phi, \psi)$ and $c(r)$ for equilibrium behavior $\psi^* = \phi$. □