Common pool resource management and risk perceptions

Can Askan Mavi, Nicolas Quérou

▶ **To cite this version:**

Can Askan Mavi, Nicolas Quérou. Common pool resource management and risk perceptions. 2022. hal-03052114v2

HAL Id: hal-03052114

https://hal.inrae.fr/hal-03052114v2

Preprint submitted on 19 Oct 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Common pool resource management and risk perceptions*

Can Askan Mavi[1] and Nicolas Quérou[2]

[1] *Université Paris-Saclay, INRAe, AgroParisTech, Paris-Saclay Applied Economics, 91120, Palaiseau*

[2] *CEE-M, Univ. Montpellier, CNRS, INRAe, Institut Agro, Montpellier, France*

October 5, 2022

Abstract

Motivated by recent discussions about the issue of risk perceptions for climate change related events, we introduce a non-cooperative game setting where agents manage a common pool resource under a potential risk, and agents exhibit different risk perceptions. We first highlight that risk and risk perceptions have qualitatively differing impacts on optimal decisions. Then, focusing on the effect of the polarization level and other population features, we show that the type of perception (overestimation, underestimation) and the pre- and post-shift resource quality levels have first-order importance on the qualitative nature of behavioral adjustments and on resource conservation. When there are non-uniform perceptions within the population, the intra-group structure qualitatively affects the degree of resource conservation. Moreover, science-based agents (using the probability estimate making consensus within the scientific community) may react in non-monotone ways to changes in the polarization level. The size of the science-based agents' sub-population does not qualitatively affect how an increase in the polarization level impacts behavioral adjustments, even though it affects the magnitude of this change. Finally, it is shown how risk perceptions affect the comparison between centralized and decentralized management, and several policies are discussed based on their likely effects on welfare.

Keywords: Conservation, Risk perception, Environmental risk, Renewable resources, Dynamic games

JEL Classification: Q20, Q54, D91, C72

1 Introduction

Environmental systems are likely to undergo drastic changes in response to exogenous shocks such as those driven by climate change. Several examples of these irreversible changes have been documented for ecosystems characterized by the collective management of natural resources such as (among other examples) fisheries, forests, groundwater (Costello and Ovando (2019); Oremus et al. (2020); Quaas et al. (2007)). Such irreversible events typically occur following regime shifts, that is, sudden changes in the dynamics of the

*We are grateful to participants in various conferences and seminars for helpful comments and suggestions. This research is supported by the ANR project GREEN-Econ (Grant ANR-16-CE03-0005) and the ANR-DFG project CRaMoRes (Grant ANR-19-FRAL-0010-01). Declarations of interest: none

natural resource which occurrence is uncertain.[1] There are multiple examples of such situations. A first one is related to the possibility driven by climate change of a shutdown of thermohaline circulation (Arnell et al. (2005)). The third IPCC assessment report stresses the possibility that fish stocks in the Northern Atlantic Ocean would be negatively affected in an irreversible manner by such an event (especially for thermally sensitive fish species). Other relevant examples relate to coral reef systems which may bleach due to warming ocean temperatures (see Treloar (2019)) and get permanent damages with considerable impacts on water quality and fish population (Tiller et al. (2019)), or to lakes which may shift from oligotrophic to eutrophic conditions due to an adverse uncertain event (excessive Phosphorus input, Carpenter et al. (1999), Scheffer (1998)).

In such situations agents usually exhibit heterogeneous views regarding the probability of occurrence of the regime shift (Lee et al. (2015)). Among other features, certain reinforcing (feedback) effects leading to irreversible changes in ecosystems may be overlooked by agents. We here consider three potential types of agents. *Science-based individuals* rely on the probability estimate emerging from consensus within the scientific community. *Overestimating (o-type) agents* rely on a higher estimate of occurrence probability. Finally, *underestimating (u-type) individuals* always rely on a lower estimate of occurrence probability compared to the estimate emerging from consensus within the scientific community. Such heterogeneity is often persistent and differences in type may be interpreted as different perceptions: For events related to climate change issues, few people tend to change their perception when new information is provided (about the features of the events, see Douenne and Fabre (2022)).[2]

Most of the related literature has abstracted from the issue of risk perception. The purpose of this paper is to analyze the effects of heterogeneous risk perception on individual behavioral adjustments and patterns of common-pool resource conservation by proposing a simple model. Indeed, a more complex model would be more real-world relevant but would not make a precise analysis of these effects possible. We thus remove the usual assumption of a unique regime shift probability estimate shared by every agent and show that assuming heterogeneous risk perceptions provides new insights about common pool resource management and the public policies to be implemented.

To that end, we introduce perception heterogeneity and environmental risk in a non-cooperative dynamic fish war game à la Levhari and Mirman (1980). We consider two cases: a first one where there is uniform perception, and a second one where science-based, o-type and u-type individuals co-exist within the population. We obtain several interesting results. First, risk and risk perception have qualitatively different effects on the individual extraction strategies. Second, the type of perception and the pre- and post-shift resource quality levels have first-order importance on the qualitative nature of behavioral adjustments and on the pattern of resource conservation. Third, when there are non-uniform perceptions within the population, the relative size of o-types and u-types' sub-populations qualitatively affects the degree of resource conservation. Moreover, science-based agents may react in non-monotone ways to changes in the polarization level. The size of the science-based agents' sub-population does not qualitatively impact the effect of changes in the polarization level on behavioral adjustments, even though it affects their magnitude.

We then characterize the socially efficient extraction policy in order to analyze the potential inefficiencies driven by decentralized management. The social planner may be populist and account for agents' potential perceptions, or she may be paternalist and does not take agents' perceptions into account. The comparison

[1] We follow the description made by Scheffer et al. (2001).

[2] A recent experiment by Abdellaoui et al. (2011) also provides evidence supporting the possibility that individuals exhibit heterogeneous risk perceptions.

between the pre- and post-shift resource quality levels qualitatively impacts both how the two social planner perspectives compare at the individual and aggregate levels, and how decentralized and centralized management approaches differ at the sub-population levels. While the comparison between the two centralized perspectives then mainly relates to the comparison of the sizes of o-types and u-types' sub-populations, the differences between centralized and decentralized management are more complex. For instance, while the overall size of the population does not qualitatively affect the comparison between decentralized and populist policies, there are cases where it does affect the comparison when the social planner is paternalist. While the tragedy of the commons[3] arises at the aggregate level, it does not always emerge within all sub-populations[4]. As such, a policy that would be designed on the basis of aggregate features only could face a serious acceptability problem. Indeed, if policy makers would only focus on aggregate scores, the use of a tax policy would be put forward: depending on the population structure, such a policy would likely face strong opposition that would not be based on social justice but on efficiency grounds. A sub-population could face a tax policy, while efficiency would have called for the use of a subsidy. Thus, while some current policy-related discussions focus on issues of social justice potentially raised by the existence of different perceptions, we highlight that such perception heterogeneity could actually raise serious efficiency problems. Finally, based on the comparison between the centralized and decentralized perspectives, several potential policies (some based on the environmental fundamentals, others on agents' perceptions) are discussed and contrasted depending on their likely effects on welfare and on their appropriate designs.

To shed light on the economic consequences of environmental shocks, a sizable literature focuses on the analysis of common-pool resource management under uncertainty (Bramoullé and Treich (2009); Costello et al. (2001); Fesselmeyer and Santugini (2013); Polasky et al. (2011); Ren and Polasky (2014); Sethi et al. (2005); Tsur and Zemel (1995); Mitra and Roy (2006)).[5] Ren and Polasky (2014) analyze the effect of exogenous/endogenous risk on the extraction decision of an infinitely lived agent, whereas Lucchetti and Santugini (2012) focus on the relationship between ownership risk and resource use and Fesselmeyer and Santugini (2013) study the strategic management of a common pool resource. Quaas et al. (2013) analyze the resilience of societies relying on natural resources when faced with exogenous shocks, while Quaas et al. (2019) focus on the insurance value of common-pool natural resources. Diekert (2017) introduces a dynamic game with a focus on a learning process related to the existence of a tipping point.[6] All these studies do not take risk perceptions into account and provide contributions differing from ours since risk and risk perception are shown to have qualitatively different effects on optimal extraction strategies. Another contribution (Agbo (2014)) studies the role of agents' heterogeneous beliefs about the future availability of a natural resources on extraction patterns. Yet, the risk of a regime shift is not considered in Agbo (2014), and perceptions are not accounted for either, since the regeneration of the natural resource depends on a stochastic variable whose distribution is unknown to all agents, and the focus is put on the effect of learning. In order to model perception, we follow some recent contributions analyzing the effect of such perceptions on economic activities (Farhi and Gabaix (2020); Gabaix (2019)).

One of the main contributions of this paper is to show how the polarization level of the population (as measured by the magnitude of the perception differences) results in different patterns of resource con-

[3]See Stavins (2001) for empirical evidence on tragedies of the commons.

[4]Dutta and Sundaram (1993) analyze cases where the tragedy of the commons may not occur and highlight that natural resource overuse may not always be a straightforward outcome.

[5]The issues of uncertainty and irreversibility also matter significantly for the timing of environmental policy adoption (Ulph and Ulph (1997)) or issues related to technological transfers (Elsayyad and Morath (2016)).

[6]See also Diekert and Nieminen (2015) for another potential effect of climate change, namely, a shift in the spatial distribution of the resource.

servation. Differently from the case of uniform perception, when several types of perception co-exist the polarization level has an interesting non-linear impact on science-based individuals' extractions. Our study thus complements an interesting literature focusing on the implications of risky events on natural resource conservation. Sakamoto (2014) provides a game setting where he shows how an endogenous regime shift probability alters the equilibrium structure and provides conditions under which a precautionary behavior emerges for the management of a common pool resource. A recent contribution by Costello et al. (2019) focuses on the spatial dimensions of the management of a common pool resource under the risk of a regime shift: the main question is how the regime shift probability affects the allocation of extraction levels in different patches. Miller and Nkuiya (2016) examines the relationship between the risk of a regime shift and the emergence of coalition formation between harvesters: conditions on the regime shift probabilities are provided under which harvesters have incentives to join or exit a coalition. Wagener (2003) analyzes the complex dynamics of resource extraction patterns in a shallow lake problem investigated by Mäler et al. (2003). Based on a body of literature exploring the issue of heterogeneous risk perceptions and considering the problem of decentralized extraction decisions within a common pool resource setting, we shed light on the link between the existence of risk perceptions and individual behavioral adjustments, their consequences for resource conservation and related implications for policy choice and design.

The remainder of the paper is organized as follows. In Section 2 we introduce the model. In section 3 we consider the generic case where agents have different types of perceptions, and we characterize the non-cooperative equilibrium outcomes. We then focus on the specific case where there is a uniform perception within the agents' population, as it allows to describe the main mechanisms at work and the main intuitions. We then come back to the generic case to highlight the resulting qualitative differences (for instance, the non-monotone effect of the polarization level on science-based agents' extraction levels). The socially efficient policies are characterized and compared with the decentralized outcome in Section 4, which in turn allows us to discuss several policies in Section 5 based on their likely effects on welfare. Section 6 concludes. All proofs are relegated to the end of the paper in an appendix.

2 Setting of the problem

We extend the model of Fesselmeyer and Santugini (2013) to allow for risk perceptions. Let us consider the Great Fish War dynamic game in which N agents derive utility from the extraction of a common and renewable resource within a discrete time infinite horizon setting. Formally, let y_t be the available stock of renewable resource at the beginning of period t. The stock evolves at the beginning of period $t+1$ according to the biological rule

$$y_{t+1} = y_t^{\alpha} \tag{1}$$

where $\alpha \in (0,1]$ models resource availability: An increase in α in a given time period reduces the availability of the exploited resource in all future periods. At $t=0$ the stock is below the carrying capacity ($y_0 < 1$). This functional form is extensively used in the literature on dynamic common-pool resource games (Fesselmeyer and Santugini (2013); Breton and Keoula (2014, 2011); Levhari and Mirman (1980); Kwon (2006)).

During period t, if agent j extracts a quantity $c_{j,t}$ of the natural resource, she derives utility $u_j(c_{j,t}) = \phi \ln c_{j,t}$ with $\phi > 0$. The parameter ϕ denotes the quality of the natural resource: A higher ϕ implies a higher utility and marginal utility of extraction. The present extraction decisions of the N agents affects the future stock level. Using (1), the stock of the natural resource evolves according to the following rule

$$y_{t+1} = \left(y_t - \sum_{j=1}^{N} c_{j,t} \right)^{\alpha} \tag{2}$$

where a total of $\sum_{j=1}^{N} c_{j,t}$ is extracted in period t by the agents and $y_t - \sum_{j=1}^{N} c_{j,t}$ denotes the remaining stock which is left to yield the stock y_{t+1} at the beginning of period $t+1$. In this setup, parameters α and ϕ are constant over time.

Let us introduce the environmental risk: resource characteristics ϕ and α now depend on the state of the environment s_t in the following manner. Given state of environment s_t, the resource available at the beginning of date t is

$$y_{t+1} = \left(y_t - \sum_{j=1}^{N} c_{j,t} \right)^{s_t} \tag{3}$$

and extracting $c_{i,t}$ yields agent i the utility level $u_i\left(c_{i,t}\right) = \phi_s \ln c_{i,t}$ at period t. We here consider a potential regime shift that, if it occurs, will affect the availability and quality levels of the resource. The process characterizing the regime shift can be described as follows. There are two possible states: $s_t \in \{\alpha_1, \alpha_2\}$ where $0 < \alpha_1 < 1$ and $0 < \alpha_2 < 1$. State α_1 represents the state of the environment prior to the regime shift. State α_2 denotes the state of the environment following the regime shift. After the occurrence of the shift, the natural resource is more scarce and available at the rate α_2.

The probability of the regime shift is $p \in (0, 1]$. In other words, if the state of the environment is α_1, then there is a constant probability p that there will be a permanent shift in the next periods. Probability p might not be the "exact" occurrence probability but rather a probability estimate that results from a consensus within the scientific community. Indeed, one cannot derive probability p from the frequency of event occurrence since the regime shift is irreversible once it occurs.[7] We make the following assumption:

Assumption 1. $\Pr\left[s_{t+1} = \alpha_2 \mid s_t = \alpha_1\right] = p$ and $\Pr\left[s_{t+1} = \alpha_2 \mid s_t = \alpha_2\right] = 1$

The first part of the assumption implies that if the economy is in state α_2, it will remain in this state forever. Moreover, we assume that the agents' extraction activity does not influence the regime shift probability. As explained in Fesselmeyer and Santugini (2013), one example is the abrupt shutdown of thermohaline circulation, which permanently affects stock of fishes in the North Atlantic Ocean. This assumption is also consistent with any situation where the risk is driven by the effects of climate change and the agents' extraction of the common-pool resource has a negligible effect on the risk probability. It could be interesting to analyze cases where the effects of the shift are reversible, while it may occur repeatedly in the future. Our paper is a first step in the analysis of the impact of perception for natural resource management problems, and as such we keep the model as simple as possible. Yet, we would not expect fundamental qualitative changes in such settings, as long as there is no learning: the reversibility of the shift might weaken some of

[7] One potential way to derive this estimate could be by somehow relying on the frequency of this shift occurring in a large sample of settings sharing similar characteristics than ours. There are other ways to construct the estimate, we abstract from discussing this point as it is not the focus of this paper.

the effects analyzed here, but the fact that the shift could repeat in the future might reinforce them on the other hand.

We now introduce a second assumption:

Assumption 2. $\phi_1 > \phi_2$ and $\alpha_1 < \alpha_2$

Thus, the effect of an environmental shift is twofold. The quality of the natural resource decreases: $\phi_1 > \phi_2$. Moreover, its availability decreases following the shift: $\alpha_2 > \alpha_1$.

A remark is that the functional form we use is quite conventional in the existing literature (Fesselmeyer and Santugini (2013), Levhari and Mirman (1980), Breton and Keoula (2014)). The literature on regime shifts considered the change in the carrying capacity of the resource stock instead of the changes in its regenerative capacity (Polasky et al. (2011)). Indeed, for the analytical tractability of our model, we assume the carrying capacity equal to one and a concave regeneration function in line with the existing literature. It would be interesting to extend the analysis by considering a shift in the carrying capacity and comparing both cases. This is left for future research.

We next introduce the final feature: risk perception. The existence of risk perception is documented in different situations, one is related to climate change related risks (see Lee et al. (2015) for instance). There have been several recent contributions in the economics literature on how to model heterogeneous risk perceptions, we follow Gabaix (2019) or Farhi and Gabaix (2020) in terms of the modeling assumptions. Later on in the analysis, we will assume that agents may overestimate or underestimate the probability of occurrence of the regime shift compared to its true value p. Following Gabaix (2019), an overestimating/underestimating agent perceives the value of the regime shift probability as follows

$$p^S = (1-m)\,p + mx \tag{4}$$

where x denotes the default value (prior mean) and parameter m the magnitude of risk perception, so that $m = 0$ corresponds to science-based individuals while $m = 1$ corresponds to agents exhibiting full perception. When the prior mean x is higher (lower) than the true value p, an agent overestimates (underestimates) the occurrence of the regime shift. We consider risk perceptions: agents do not revise their estimate of the regime shift occurrence as time goes by. We will denote the sub-population of overestimating agents by N^o, the sub-population of underestimating agents by N^u, and by N^j the sub-population of agents who have a science-based perception of the regime shift probability. The following figure provides an illustration[8]:

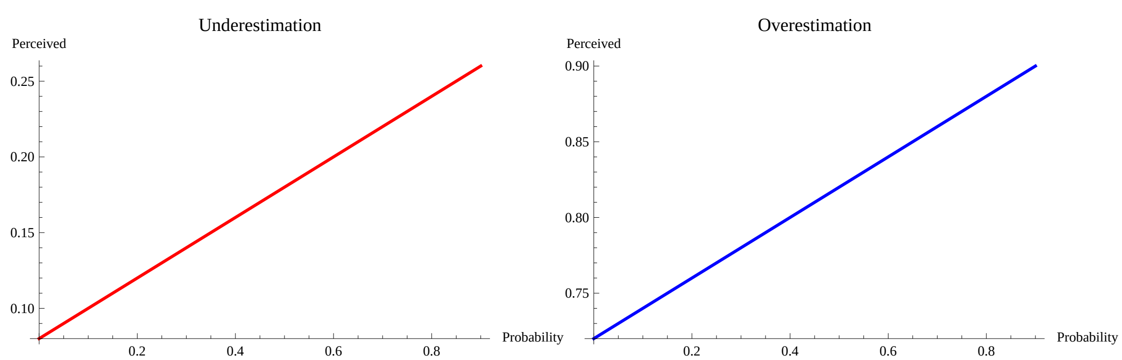

Figure 1: Underestimation ($x < p$)) and overestimation ($x > p$) of the regime shift probability

[8]The parameter values for overestimation is $x = 0.9$, $m = 0.8$ and for underestimation $x = 0.1$, $m = 0.8$.

The case of science-based individuals might be understood as one where there is a reasonable scientific consensus on an estimate of the probability, and science-based agents follow it. Heterogeneous perceptions might emerge for different reasons: manipulations of the agents' beliefs about environmental issues by environmental or political lobbies (McCright and Dunlap (2003)); individual or collective denial of environmental problems (Opotow and Weiss (2000)). The main feature here is that we analyze persistent differences in perception, that is, cases where individuals do not change their perception when new information is provided. We assume that agents know the distribution of types in the population.[9]

3 The analysis

Since we consider a decentralized and dynamic setting, we focus on Markov Perfect Nash equilibria, and more specifically on linear strategies Markov Perfect equilibria (although non-linear equilibria might exist). We will show that there exists a unique linear Markov Perfect Nash equilibrium: since it is stationary we hereafter drop script t when characterizing the equilibrium.

3.1 Non-cooperative equilibria

We first characterize interior MPNE in the general case. The number of science-based agents is N^j, and there are N^o o-type and N^u u-type individuals. The total population size is thus $N = N^j + N^u + N^o$. The pre-shift value function of a type-l agent is:

$$V_1^l\left(y\right) = \max_{0\leq c_l \leq y-\sum_{k\neq l}^{N} c_k} \phi_1 \ln c_l + \left(1-p^l\right)\delta V_1^l\left(\left(y - c_l - \sum_{k\neq l}^{N} c_k\right)^{\alpha_1}\right)$$
$$+ p^l \delta V_2^l\left(\left(y - c_l - \sum_{k\neq l}^{N} c_k\right)^{\alpha_2}\right) \tag{5}$$

In the subsequent period, this agent anticipates the occurrence of the shift with probability p^l. Her value function in state 2 (once the regime shift has occurred) is then:

$$V_2^l\left(y\right) = \max_{0\leq c_l \leq y-\sum_{k\neq l}^{N} c_k} \phi_2 \ln c_j + \delta V_2^l\left(\left(y - c_l - \sum_{k\neq l}^{N} c_k\right)^{\alpha_2}\right) \tag{6}$$

Regarding the expressions of pre- and post-shift value functions, the difference is that a type-l agent takes into account the extraction of natural resources of the other types of agents. We obtain:

Proposition 1. *There exists a unique linear MPNE, which is characterized as follows: o-type, u-type and science-based agents extract the following amounts of the natural resource:*

$$g_o\left(y\right) = y\frac{\phi_1 z_j z_u}{z_j z_o z_u + \phi_1\left[N^j z_o z_u + N^o z_j z_u + N^u z_j z_o\right]} = g_o^{dec} y \tag{7}$$

[9] For certain situations it could well be that agents "agree to disagree" on their risk perceptions. It could be interesting in future research to consider situations where involved agents' biases could be influenced by those of other agents, for instance, through decision-making processes aimed at collectively addressing the potential occurrence of the shift.

$$g_u(y) = y\frac{\phi_1 z_j z_o}{z_j z_o z_u + \phi_1 \left[N^j z_o z_u + N^o z_j z_u + N^u z_j z_o\right]} = g_u^{dec} y \tag{8}$$

$$g_j(y) = y\frac{\phi_1 z_o z_u}{z_j z_o z_u + \phi_1 \left[N^j z_o z_u + N^o z_j z_u + N^u z_j z_o\right]} = g_j^{dec} y \tag{9}$$

where $z_j = \left[(1-p)\delta\alpha_1 a_1^j + p\delta\alpha_2 a_2^j\right]$, $z_o = [(1-p^o)\delta\alpha_1 a_1^o + p\delta\alpha_2 a_2^o]$ and $z_u = [(1-p^u)\delta\alpha_1 a_1^u + p\delta\alpha_2 a_2^u]$ with $a_2^j = a_2^o = a_2^u = \frac{\Phi_2}{1-\delta\alpha_2}$ and $a_1^k = \frac{\Phi_1 + p^k\delta\alpha_2 a_2^k}{1-(1-p^k)\delta\alpha_1}$ for $k = j, o, u$.

Proof. See Appendix (A)

We will first restrict the analysis to the "uniform" perception case where science-based agents co-exist with agents exhibiting only one type of perception.

3.2 Uniform perception

The number of science-based agents' sub-population is N^j and there are N^i o-type (or u-type) individuals with $i = u$ or $i = o$. The size of the whole population is thus $N = N^j + N^i$. We obtain:

$$g_j(y) = c_j = \frac{\phi_1 y}{\phi_1 N^j + (1-p)\, a_1^j \alpha_1 \delta + p a_2^j \alpha_2 \delta} - \underbrace{\frac{\phi_1 N^i c_i}{\phi_1 N^j + (1-p)\, a_1^j \alpha_1 \delta + p a_2^j \alpha_2 \delta}}_{\text{Indirect effect}} \tag{10}$$

$$g_i(y) = c_i = \frac{\phi_1 y}{\phi_1 N^i + (1-p^S)\, a_1^i \alpha_1 \delta + p^S a_2^i \alpha_2 \delta} - \underbrace{\frac{\phi_1 N^j c_j}{\phi_1 N^i + (1-p^S)\, a_1^i \alpha_1 \delta + p^S a_2^i \alpha_2 \delta}}_{\text{Indirect effect}} \tag{11}$$

In expression (10) one notices that the second term relates to the effect of o-type (u-type) individuals' decisions on the science-based agents' extraction decisions. Using Proposition 1 we deduce that there exists a unique linear MPNE. At this equilibrium, any science-based agent j and o-type (u-type) individual i extract the following amount, respectively:

$$g_j(y) = \phi_1\left(\frac{1 - N^i\gamma}{\phi_1 N^j + z_j}\right) y \tag{12}$$

$$g_i(y) = \gamma y \tag{13}$$

where $\gamma = \frac{\phi_1 z_j}{(\phi_1 N^j + z_j)(\phi_1 N^i + z_i) - (\phi_1)^2 N^i N^j}$, $z_i = \left(1 - p^i\right) a_1^i \delta\alpha_1 + p^i a_2^i \delta\alpha_2$ and $z_j = (1-p)\, a_1^j \delta\alpha_1 + p a_2^j \delta\alpha_2$.

The only term that depends on the o-type (u-type) individuals' risk perception p^i is the term γ. The expressions of optimal extraction strategies highlight that o-types (u-types) and science-based agents react to a change in the magnitude of the polarization level m in an opposite way, the formal proof is provided in Appendix B. A first important point to be highlighted is that risk perception and risk have qualitatively different effects. Specifically, we now highlight that higher risk perception and higher risk do not have the same effect on individual resource extraction decisions and hence on the aggregate extraction level. In this sense as well, our paper thus differs from Fesselmeyer and Santugini (2013). We do so by analyzing the effect of risk on the extraction decisions of science-based agents and o-type individuals, since in this case larger values of m will result in a larger difference between p and the o-types' estimate.

Proposition 2. *Higher risk perception and higher risk do not have the same qualitative impact on individual resource extraction decisions.*

Proof. See Appendix C

To understand this result, let us focus on the effects of a marginal increase in the polarization level m and of a marginal increase in the regime probability on the extraction rate of an o-type individual. We have:

$$\frac{\partial g_o(y)}{\partial m} = -\frac{\phi_1 \frac{\partial z_o}{\partial m}\left[\phi_1 N^j + z_j\right] y}{\left[(\phi_1 N^j + z_j)(\phi_1 N^i + z_i) - (\phi_1)^2 N^i N^j\right]^2} \leq 0 \Longleftrightarrow \phi_1 \leq \phi_2 \frac{\alpha_2(1-\delta\alpha_1)}{\alpha_1(1-\delta\alpha_2)} \tag{14}$$

Equation (14) shows that the effect of the polarization level m on the extraction level of o-type individuals is negative when the quality of the natural resource after the regime shift is sufficiently high. The threshold is independent of the polarization level m. By contrast, the effect of a higher risk is:

$$\frac{\partial g_o(y)}{\partial p} = \phi_1 y \frac{\frac{\partial z_j}{\partial p}\phi_1 N^j z_o - z_j \frac{\partial z_o}{\partial p}\left[\phi_1 N^j + z_j\right]}{\left[(\phi_1 N^j + z_j)(\phi_1 N^i + z_i) - (\phi_1)^2 N^i N^j\right]^2} \tag{15}$$

It can be checked that the impact of the probability is not monotone and thus qualitatively different from the effect of higher risk perceptions. Indeed, when m gets close to zero (individual perceptions get close to science-based estimates) it can be checked that $\frac{\partial g_o(y)}{\partial p} \leq 0$ when $\phi_1 \leq \phi_2 \frac{\alpha_2(1-\delta\alpha_1)}{\alpha_1(1-\delta\alpha_2)}$ is satisfied. When m increases and gets close to one, o-type individuals perceive a lower marginal increase (discounted by $(1-m)$) of the probability, which results in a lower value of $\frac{\partial z_o}{\partial p}$ together with a higher value of m. This in turn results in $\frac{\partial z_o}{\partial p}$ getting close to zero and $\frac{\partial g_o(y)}{\partial p} \geq 0$ when $\phi_1 \leq \phi_2 \frac{\alpha_2(1-\delta\alpha_1)}{\alpha_1(1-\delta\alpha_2)}$ is satisfied.

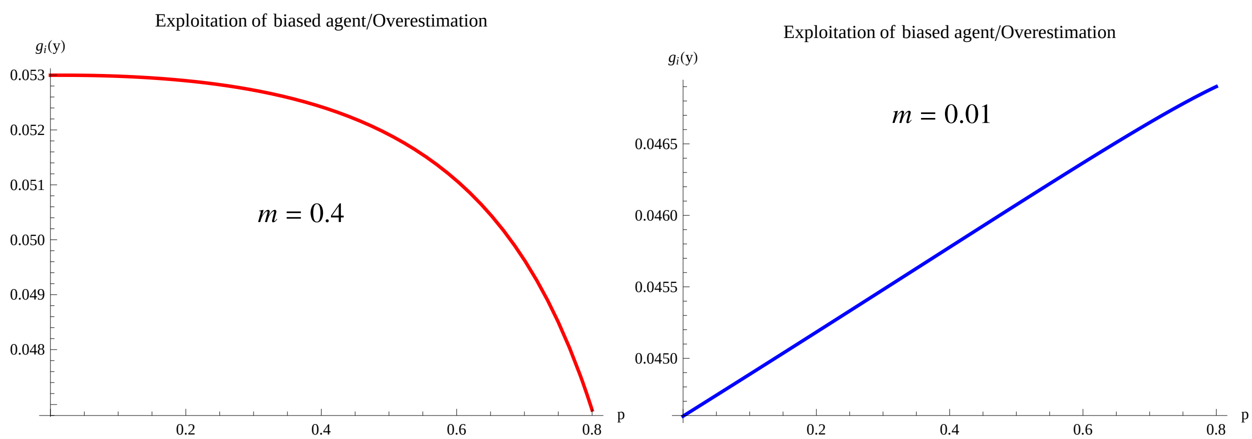

Figure 2: Effect of polarization m through the effect of the probability p on natural resource extraction

When the value of parameter m is low, the elasticity of term $z_o^{'}$ is higher, which is likely to make $\frac{\partial g_o(y)}{\partial p}$ positive. This result is interesting as it shows that a higher risk perception and a higher risk do not yield the same qualitative conclusions. This is one reason for why we will later come back to the generic case to understand the interactions between different groups (in the next subsection). We now use Proposition 1 to assess the effect of an increase in the polarization level of the population, as measured by an increase in the magnitude of polarization level m. We also assess how the number of o-type (u-type) individuals and science-based agent affects the aggregate extraction of the resource (thus keeping N constant, denoting the size of the science-based agents' sub-population by $N - N^i$). We obtain:

Proposition 3. *(i) Before the shift occurs, an o-type individual's (science-based agent's) extraction level*

increases (decreases) as m increases if and only if $\phi_1 \geq \frac{\alpha_2(1-\delta\alpha_1)\phi_2}{\alpha_1(1-\delta\alpha_2)}$ holds. Aggregate extraction increases as either m or the number of o-type individuals increases if and only if $\phi_1 \geq \frac{\alpha_2(1-\delta\alpha_1)\phi_2}{\alpha_1(1-\delta\alpha_2)}$ holds.
ii) The same conclusions than in point (i) apply to the case of u-type individuals if and only if $\phi_2 < \phi_1 \leq \frac{\alpha_2(1-\delta\alpha_1)\phi_2}{\alpha_1(1-\delta\alpha_2)}$ holds.

Proof. See Appendix (D)

In order to understand this result, one should compare the marginal cost of extraction for each type of agent. We derive the ratio of marginal utility by using the corresponding optimality conditions:

$$\frac{\frac{\phi_1}{c_i}}{\frac{\phi_1}{c_j}} = \frac{a_1^i\delta\alpha_1 - p^S\left(a_1^i\delta\alpha_1 - \delta\alpha_2 a_2^i\right)}{a_1^j\delta\alpha_1 - p\left(a_1^j\delta\alpha_1 - \delta\alpha_2 a_2^j\right)} \tag{16}$$

The right hand side of equality (16) corresponds to the ratio of marginal cost of natural resource exploitation for o-type (or u-type) individuals and science-based agents. The left hand side corresponds to the ratio of marginal utility of resource extraction. The term $p^S a_1^i\delta\alpha_1$ may be interpreted as a discounted weight awarded by an o-type (or u-type) individual to the state before regime shift, accounting for the regeneration rate of the environment α, and the risk perception parameter p^S. A higher perceived probability p^S results in a higher weight related to the situation before the regime shift $V_1^i(y)$. This in turn implies that the weight $(1-p^S)a_1^i\delta\alpha_1$ related to the option to stay in the same state 1 $V_1^i(y)$ decreases. The same applies to the terms related to a science-based agent.

We first discuss the case when $p^S > p$ holds. Let us come back to the right-hand side of (16). A higher polarization level m decreases the ratio through the effect on term $p^S\left(\delta a_1^i\alpha_1 - \delta\alpha_2 a_2^i\right)$ if $a_1^i\delta\alpha_1 > \delta\alpha_2 a_2^i$ holds, which is the case if condition $\phi_1 > \frac{\phi_2\alpha_2(1-\delta\alpha_1)}{\alpha_1(1-\delta\alpha_2)}$ is satisfied. This implies that the ratio of the marginal costs of extraction related to o-type individuals and science-based agents decreases. As a result, an o-type individual increases her extraction level as the situation becomes more polarized (m increases). The same reasoning applies to a u-type individual. If condition $a_1^i\delta\alpha_1 > \delta\alpha_2 a_2^i$ is satisfied, a u-type individual decreases her natural extraction rate before shift as the situation becomes more polarized. This is so because the ratio of marginal costs of extraction between o-type individuals and science-based agents (see right-hand side of equation (16)) increases when parameter m increases.

Proposition 3 also highlights that the effect of parameter m on the aggregate extraction level does not depend on the population structure: the effect resulting from o-type (or u-type) individuals' optimal decisions drives the effect on aggregate extraction. In the case of o-type individuals, their marginal cost of extraction is lower, and this implies that their extraction level is higher. Now to understand the effect of N^i, we use

$$\frac{\partial\left(N^i g_i + \left(N-N^i\right)g_j\right)}{\partial N^i} = g_i - g_j + \frac{\partial g_i}{\partial N^i}N^i\left(1 + \frac{N-N^i}{N^i}\frac{g_j}{g_i}\right) \tag{17}$$

where $\frac{\partial g_i}{\partial N^i} = -\frac{(\phi_1)^2 z_1}{(k_1)^2}\left(z_1 - z_2\right)y$. An increase in N^i has two effects. First, the distribution of o-type (u-type) individuals and science-based agents changes with respect to N^i (an increase in the number of o-type individuals) . Second, it affects the optimal extraction levels g_i and g_j. The term $g_i - g_j$ can be interpreted as the direct effect (the composition effect). The term $\frac{\partial g_i}{\partial N^i}N^i\left(1+\frac{N-N^i}{N^i}\frac{g_j}{g_i}\right)$ is the indirect effect through the effect on equilibrium extractions. If $g_i > g_j$ then the composition effect is positive. This implies that

$z_1 > z_2$ holds (see equation (16)). Therefore, the indirect effect is negative. The conclusion is that the two effects work in opposite directions. In the present setting the indirect effect offsets the direct effect.

3.3 Coexistence of several types of perception

We come back to the general case so that we now analyze the qualitative differences when one introduces non-uniform perception. We first consider the effect of an increase in the polarization level:

Proposition 4. *i) When $\phi_1 \leq \frac{\phi_2\alpha_2(1-\delta\alpha_1)}{\alpha_1(1-\delta\alpha_2)}$ a science-based agent's extraction together with the welfare of this sub-population increase, and the aggregate extraction level decreases, as m increases if and only if $\frac{N^o}{N^u} \geq -\frac{z_u^{'}(z_o)^2}{z_o^{'}(z_u)^2}$ hold, with $z_r^{\prime} = \frac{\partial z_r}{\partial m}$ for $r = o, u$. When $\phi_1 > \frac{\phi_2\alpha_2(1-\delta\alpha_1)}{\alpha_1(1-\delta\alpha_2)}$ the same conclusions apply if and only if $\frac{N^o}{N^u} \leq -\frac{z_u^{'}(z_o)^2}{z_o^{'}(z_u)^2}$ hold.*
ii) o-type and u-type individuals react in opposite ways to an increase in the magnitude of the polarization.

Proof. See Appendix (E)

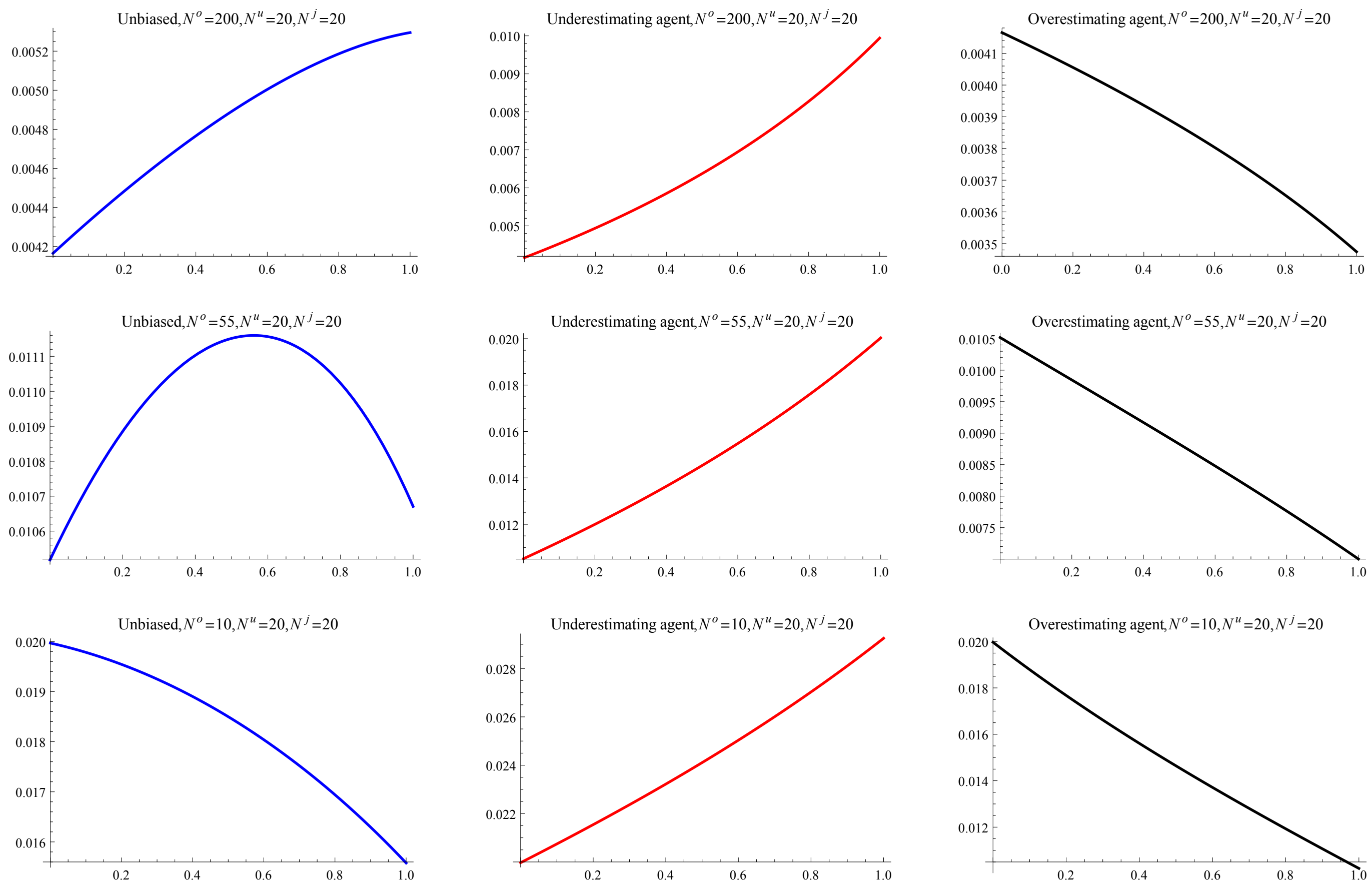

Figure 3: Resource extractions as functions of m

The main qualitative difference when allowing for non-uniform perceptions is as follows. First, when there is a uniform perception within the population, the intra-group structure does not qualitatively affect individual behavioral adjustments. However, under non-uniform perceptions, this internal structure has an effect. Indeed, the extraction rate of a science-based agent is $g_j(y) = \frac{\phi_1(y-N^o g_o(y)-N^u g_u(y))}{N^j\phi_1+(1-p)a_1^j\alpha_1\delta+pa_2^j\alpha_2\delta}$ and thus the marginal effect of an increase in m is as follows:

$$\frac{\partial g_j(y)}{\partial m} = -\frac{1}{N^j\phi_1+(1-p)\,a_1^j\alpha_1\delta+pa_2^j\alpha_2\delta}\left[N^o\frac{\partial g_o(y)}{\partial m}+N^u\frac{\partial g_u(y)}{\partial m}\right] \tag{18}$$

Expression (18) highlights that the effect of m on the science-based agents' extraction rate is driven by the number of o-type and u-type individuals and by the direct effect of m on these agents' optimal decision. Another important implication of Proposition 4 is that science-based agents may exhibit non-monotone behaviors (see Figure (3)) as the polarization level changes: Specifically, g_j might initially *increase* for low polarization levels and *decrease* for high enough values of m.[10] The term $N^o\frac{\partial g_o(y)}{\partial m}+N^u\frac{\partial g_u(y)}{\partial m}$ in expression (18) is either positive or negative since both marginal effects have opposite signs. The marginal effect on the aggregate extraction is

$$\frac{\partial\left(N^o g_o+N^u g_u+N^j g_j\right)}{\partial m} = -\left(\phi_1\right)^2 z_j\frac{\left(z_o'\left(z_u\right)^2N^o+z_u'\left(z_o\right)^2N^u\right)}{\left[z_jz_oz_u+\phi_1\left(N^jz_oz_u+N^oz_jz_u+N^uz_jz_u\right)\right]^2} \tag{19}$$

The effect of m on the aggregate extraction is thus driven by the value of ratio $\frac{N^o}{N^u}$ as stated in Proposition 4. The intuition is as follows. For sufficiently low ratio of pre- to post-shift resource quality levels, o-type (u-type) individuals' dynamic marginal costs get higher (lower) compared to that of science-based individuals as the polarization level increases, and their extraction level decreases (increases). The effect on aggregate extraction is then driven by the magnitude of the spillover resulting from each type of individuals. When the size of the o-type individuals' population is high enough compared to that of u-type individuals, the effect on aggregate extraction is then negative. Finally, the marginal effect is:

$$\frac{\partial g_j}{\partial m} = \frac{\left(\phi_1\right)^2 z_j\left(z_o'\left(z_u\right)^2N^o+z_u'\left(z_o\right)^2N^u\right)}{\left[z_jz_oz_u+\phi_1\left(N^jz_oz_u+N^oz_jz_u+N^uz_jz_u\right)\right]^2} \tag{20}$$

As such, a larger sub-population of science-based agents only affects the magnitude of the change in the extraction level, but it does not affect whether the change is positive or negative.

4 Social optimum versus decentralized management

We will characterize the socially efficient extraction path and then contrast it with decentralized extraction policies. We provide details on the notion of optimality we will use since there exist different perceptions in the society. The social planner's problem is to maximize the sum of the agents' individual value functions $\sum_{i=1}^N V_{i,1}$ and here two scenarios are considered. In the first one, the social planner is populist and accounts for the agents' perceptions. In the second one, the social planner is paternalistic: the agents' perceptions are unaccounted for.

4.1 The socially efficient policy: two perspectives

For a populist social planner, the conjecture is that the pre-shift and post-shift value functions satisfy $V_i^1(y) = a_i^1 ln(y) + b_i^1$ and $V_i^2(y) = a_i^2 ln(y) + b_i^2$ for any agent i. We thus have:

[10] The parameter values are $x_o = 0.9$, $x_u = 0.15$, $\delta = 0.25$, $\alpha_2 = 0.5$, $\alpha_1 = 0.45$, $p = 0.5$, $\phi_1 = 1$, $\phi_2 = 0.9$ The proof of the general claim about non-monotonicity is available upon request.

$$\sum_{i=1}^{N}\left[a_1^j \ln y + b_1^j\right] = \max_{c_i}\sum_{i=1}^{N}\left[\phi_1 \ln c_i + \left(1-p^i\right)\delta V_i^1\left(\left(y-\sum_{k=1}^{N}c_k\right)^{\alpha_1}\right) + p^i \delta V_i^2\left(\left(y-\sum_{k=1}^{N}c_k\right)^{\alpha_2}\right)\right] \quad (21)$$

subject to $0 \leq \sum_{j=1}^{N} c_j \leq y$ and

$$\sum_{i=1}^{N}\left[a_2^i \ln y + b_2^i\right] = \max_{c_i}\sum_{i=1}^{N}\left[\phi_2 \ln c_i + \delta V_i^2\left(\left(y-\sum_{k=1}^{N}c_k\right)^{\alpha_2}\right)\right] \quad (22)$$

By constrast, a paternalistic social planner only uses the occurrence probability based on the scientific consensus. We obtain the following result:

Proposition 5. *i) The solution to the populist social planner's program is characterized by a uniform level of extraction within the entire population:*

$$g_j^{pop}(y) = g_o^{pop}(y) = g_u^{pop}(y) = l_1^{pop} y \quad (23)$$

where $l_1^{pop} = \frac{\phi_1}{N\phi_1 + \sum_{k=1}^{N}\left[(1-p^k)\delta\alpha_1 \frac{\phi_1 + \delta p^k \alpha_2 \frac{\phi_2}{1-\delta\alpha_2}}{1-\delta\alpha_1(1-p^k)} + p^k \delta\alpha_2 \frac{\phi_2}{1-\delta\alpha_2}\right]} = \frac{\phi_1}{N\phi_1 + \sum_{k=1}^{N} z_k} = \frac{\phi_1}{N\phi_1 + N^o z_o + N^j z_j + N^u z_u}.$

ii) The solution to the paternalistic social planner's program is characterized by a uniform level of extraction within the entire population:

$$g_j^{pat}(y) = l_1^{pat} y \quad (24)$$

where $l_1^{pat} = \frac{\phi_1}{N}\frac{[1-\delta\alpha_1(1-p)](1-\delta\alpha_2)}{\phi_1(1-\delta\alpha_2)+\delta p \alpha_2 \phi_2}$.

Proof. See Appendix (F).

Even though agents exhibit heterogeneous perceptions and the social planner accounts for these perceptions, all agents follow the same extraction policy. This is due to an externality effect: each agent has to account for those with different perceptions than his own, and follows the same extraction pattern. Let us briefly contrast the two perspectives. First, when $\phi_1 \leq \frac{\alpha_2(1-\delta\alpha_1)}{\alpha_1(1-\delta\alpha_2)}\phi_2$ holds then $g^{pop}(y) \geq g^{pat}(y)$ if and only if $\frac{N^u}{N^o} \geq \frac{z_o - z_j}{z_j - z_u}$ holds. Indeed, under this condition $z_j < z_o$ and $z_j > z_u$ hold, and the expressions of $g^{pop}(y) \geq g^{pat}(y)$ imply that a populist planner would implement a higher extraction policy if and only if $N^u(z_j - z_u) + N^o(z_j - z_o) \geq 0$ holds. A similar argument implies that, when the ratio $\frac{\phi_1}{\phi_2}$ is large enough then the same conclusion follows if and only if $\frac{N^u}{N^o}$ is low enough. Intuitively, when the pre-shift resource quality level is low enough, the effect of a given agent gets larger as the polarization level increases. In the populist policy this implies that, compared to u-type individuals, a science-based agent results in a decrease in aggregate extraction. The same property goes for the effect of an o-type individual compared to a science-based agent. Then, for a populist policy to yield higher extraction levels, the ratio $\frac{N^u}{N^o}$ must be large enough.

4.2 Comparison with the decentralized outcome

In this subsection, we compare the socially efficient and decentralized policies. We will then rely on this comparison in order to derive policy implications. We first have:

Proposition 6. *Consider the case where several types of perception coexist, and assume that* $\phi_1 \leq \frac{\alpha_2(1-\delta\alpha_1)}{\alpha_1(1-\delta\alpha_2)}\phi_2$ *is satisfied. We have:*

- *u-type individuals extract more under decentralization than under a social planner policy.*
- *Under a populist planner o-type individuals extract more under decentralization if and only if* N^o *is large enough. Under a paternalistic planner they extract less under decentralization when* $N \leq \frac{z_o}{z_j}$ *holds. When* $N > \frac{z_o}{z_j}$ *they extract more under decentralization if and only if* N^o *is large enough.*
- *When* $\frac{N^o}{N^u}$ *is large enough science-based agents extract more under decentralization. Otherwise they extract more under decentralization provided* N^j *is large enough.*
- *Decentralized management always results in a sub-optimally high aggregate extraction level.*

Proof: See Appendix (G).

The case where $\phi_1 > \frac{\alpha_2(1-\delta\alpha_1)}{\alpha_1(1-\delta\alpha_2)}\phi_2$ is similar with the roles of u-type and o-type agents being reversed. Since the interpretation remains consistent, we provide some discussion on the first case under a populist social planner. When pre-shift resource quality level is low enough, a u-type individual puts more weight on the pre-shift state than the other categories, and the population is characterized by the highest extraction level under decentralization. The effect of a centralized policy is to internalize part of the externalities driven by decentralization: it only decreases the level of extraction corresponding to this sub-population. By contrast, o-type individuals are characterized by the lowest extraction levels under decentralization. A populist policy, even though it accounts for agents' perceptions, still induces a uniform extraction level within the population. Whether this level lies above or below the corresponding level under decentralization depends on the magnitude of the externality driven by this sub-population. As such, when the size of this sub-group lies above a threshold value, decentralized management results in suboptimally high extraction levels. It is interesting to notice that the magnitude of the polarization affects the level of the threshold value.

The case of science-based agents is more complex. The comparison then depends on the relative size of the o-type and u-type individuals' sub-populations. When the ratio of the number of o-type to that of u-type individuals is large enough, the dominant effect is still driven by the negative externality resulting from decentralized management, and science-based agents extract more than under a populist policy. When this ratio is small enough, the spillover effect imposed by u-type individuals on the science-based population is dominant: their higher extraction levels induce science-based agents to potentially decrease their decentralized extraction levels. It then comes back to the science-based agents' population size: When it is large enough the classical conclusion prevails. When it is small enough, the spillover effect prevails and science-based agents extract less under decentralization than under a populist planner policy.

Finally, at the aggregate level, the dominant effect is mainly driven by the negative externalities resulting from decentralized management, and the tragedy of the commons emerges at the overall population level. We now discuss some differences induced by the two approaches adopted for centralized management.

First, when the qualitative conclusions are similar for both approaches, the magnitude of the threshold values related to the size of sub-populations differs in both cases. A more qualitative difference is that the size of the overall population might have a direct effect on the comparison when the planner is a paternalist. Specifically, it might affect the comparison for the sub-population characterized by the lowest extraction levels under decentralization. When the size of the overall population lies below a threshold value, the spillover effect induced by the other sub-populations (which tends to decrease extraction levels) outweighs the direct externality effect driven by decentralization (which results in higher extraction levels). Decentralized management then always results in lower extraction levels compared to a centralized policy.

5 Policy discussions and implications

5.1 Some direct implications of the results

The results are useful to discuss an important issue related to the form a public policy should take to solve the efficiency problem. There are several insights resulting from the previous comparisons. First, the comparison between the pre- and post-shift resource quality levels has first-order importance as it qualitatively impacts both how the two social planner perspectives compare and how decentralized and centralized management approaches differ. Policy discussions that would not account for these fundamentals would miss an important part of the problem. Second, while the overall size of the population does not qualitatively affect the comparison between decentralized and populist management policies, there are cases where it does affect the comparison when the social planner is paternalistic. Third, while the tragedy of the commons still arises at the aggregate level for the two centralized perspectives, it does not arise at all sub-population levels. A policy that would be designed relying on aggregate properties only could face a serious acceptability problem. If policy makers focus on aggregate scores they would propose the use of a tax policy: depending on the population composition, such a policy would face strong opposition that would not be based on social justice but on efficiency grounds. Specifically, a sub-population could face a tax policy while efficiency would have called for the use of a subsidy. While many policy-related discussions tend to focus on issues of social justice potentially raised by the existence of different perceptions, we highlight this could also raise serious efficiency problems at the sub-population levels.

Now we discuss the effects of some potentially available policies. There are different instruments that might be designed to tackle the externalities caused by risk perceptions. Here we discuss interventions aimed at decreasing aggregate extraction: indeed, we know from Proposition 11 that aggregate extraction is always suboptimally high. As such, policy interventions that would result in an overall decrease in aggregate extraction would have a positive effect on social welfare (even though they might not result in a Pareto improvement). We do not introduce the costs of implementing a policy, but rather focus on discussing whether the qualitative effect of certain policies is likely to be positive.[11]

5.2 Adaptation and mitigation policies

In our setting adaptation policies would consist in diminishing the negative effects of the shift on the resource. They could consist in either changing the post-shift quality level of the resource from Φ_2 to Φ_2^ε such that $\Phi_1 - \Phi_2 > \Phi_1 - \Phi_2^\varepsilon$ or affecting its availability α_2 such that the post-shift availability becomes α_2^ε such that $\alpha_1 < \alpha_2^\varepsilon < \alpha_2$. We thus need to understand the comparative statics effects of either increasing Φ_2 or decreasing α_2 on the decentralized aggregate extraction level. We obtain the following result:

Proposition 7. *An adaptation policy aimed at increasing the post-shift quality level of the resource will positively affect social welfare (that is, decreases aggregate extraction). By contrast, an adaptation policy aimed at decreasing α_2 will have a negative effect on welfare.*

[11] We thus only focus on cases where the change would partially internalize the existing externality.

See Appendix (H).

An overall implication is that adaptation policies might be effective when focused on the resource quality level, but might be counterproductive when focused on the post-shift availability of the resource. This result follows in a direct manner from the characterizations provided in Proposition 1. For instance, the effect of Φ_2 on the aggregate extraction is negative, implying that increasing the post-shift quality level results in decreasing the equilibrium aggregate extraction level. As we know that the socially optimal aggregate extraction level is always lower than the decentralized one, the effect of such an adaptation policy is positive. Moving on to mitigation policies, they would consist in decreasing the occurrence probability of the shift. A question arises: does this policy affect science-based agents only, or does it affect all agents?

Proposition 8. *The qualitative effect of a mitigation policy is the same whether it affects all agents or science-based agents only. Its effectiveness is ambiguous and depends on the relative values of growth and quality parameters: specifically, it will positively affect social welfare when* $\Phi_1 \geq \frac{\alpha_2(1-\delta\alpha_1)}{\alpha_1(1-\delta\alpha_2)}\Phi_2$ *is satisfied.*

Proof: See Appendix (I).

Again, relying on the characterizations provided in Proposition 1, focusing on the case where the policy affects all agents, the change in aggregate extraction is driven by the opposite effect on $(1-p^l)\delta\alpha_1 a_1^l + p^l \delta a_2^l$ (interpreted in condition (16) as the marginal cost of natural resource harvesting by a l-type agent, $l = j, o, u$). This change is - $[\frac{\partial p^l}{\partial p}(-\delta\alpha_1 a_1^l + \delta a_2^l) + (1-p^l)\delta\alpha_1 \frac{\partial a_1^l}{\partial p}]$, where the first term denotes the direct effect of the policy on the agent's perception, and the second term denotes the indirect effect on the value related to staying in the no-shift state. Aggregate extraction will thus decrease as p decreases when $\phi_1\alpha_1 - \alpha_2\phi_2\frac{1-\delta\alpha_1}{1-\delta\alpha_2} > 0$ holds: a mitigation policy will be effective at increasing social welfare when the ratio of pre- to post-shift resource quality levels is large enough. This highlights one advantage of adaptation policies over mitigation policies when the focus is put on the wedge between resource quality levels.

5.3 Convincing u-type individuals, or making the population more cohesive?

A public authority may consider a policy aimed at targeting the u-type individuals to affect their risk perception. We focus on a stylized form of the instrument which consists in shifting the distribution of types within the agents' population: the authority may, at some cost, affect the initial distribution of agents (N^j, N^o, N^u) and change it as $(N^j + N^c, N^o, N^u - N^c)$. What would be the effect of such a policy?

Proposition 9. *The effectiveness of a policy aimed at convincing u-type individuals is ambiguous and depends on the relative values of growth and quality parameters: it positively affect social welfare when* $\Phi_1 < \frac{\alpha_2(1-\delta\alpha_1)}{\alpha_1(1-\delta\alpha_2)}\Phi_2$ *holds.*

Proof: See Appendix (J).

The implication of this result is that policies aimed at influencing u-type individuals might be counterproductive in certain cases, namely when the ratio of pre- to post-shift resource quality levels is large enough. Moreover, it suggests that such a type of policy and mitigation policies are unlikely to be complementary, as they might be effective only for different cases (with respect to the relative quality levels of the resource). Another approach would aim at making the population more cohesive, which would correspond to a policy aimed at decreasing the polarization level as modeled by parameter m.

Proposition 10. *The effectiveness of a policy aimed at making the population more cohesive is ambiguous as it both depends on the relative values of growth and quality parameters and is group-specific (that is, it depends on the type distribution). Specifically, when* $\Phi_1 < \Phi_2\frac{\alpha_2(1-\delta\alpha_1)}{\alpha_1(1-\delta\alpha_2)}$ *is satisfied, the policy will be* *effective when* $\frac{N^o}{N^u} < \left(\frac{z_o}{z_u}\right)^2 \frac{-\frac{\partial z_u}{\partial m}}{\frac{\partial z_o}{\partial m}}$ *(the ratio of o-type to u-type individuals is low enough). By contrast, when* $\Phi_1 \geq \Phi_2 \frac{\alpha_2(1-\delta\alpha_1)}{\alpha_1(1-\delta\alpha_2)}$ *is satisfied, the policy will be effective when* $\frac{N^u}{N^o} < \left(\frac{z_u}{z_o}\right)^2 \frac{-\frac{\partial z_o}{\partial m}}{\frac{\partial z_u}{\partial m}}$ *holds.*

Proof: See Appendix (K).

The effectiveness of a policy aimed at making the population more cohesive is ambiguous as it both depends on the environmental fundamentals and on the population structure. In particular, the effectiveness does depend on the relative share of o-type and u-type individuals. This is in stark contrast with respect to the other policies discussed here.

6 Conclusion

We analyze the effects of introducing risk perceptions in a group of agents managing a renewable resource. First, the type of perception and the pre- and post-shift resource quality levels have first-order importance. Second, when there are non-uniform perceptions, the group structure qualitatively affects the degree of resource conservation. The science-based agents may also react in non-monotone ways to changes in the polarization level. Moreover, the size of the science-based agents' sub-population does not qualitatively affect how an increase in the polarization level impacts behavioral adjustments, even though it affects the magnitude of this change. We then characterize the socially efficient extraction policies. The comparison between the pre- and post-shift resource quality levels qualitatively impacts both the comparison between the two social planner perspectives, and the comparison between decentralized and centralized management approaches. While the comparison between the two centralized perspectives mainly relates the relative sizes of the o-type and u-type individuals' sub-populations, the differences between centralized and decentralized management are more complex. For instance, the overall size of the population may affect it. Thirdly, while the tragedy of the commons arises at the aggregate level, it does not always emerge at the sub-population level. As such, a policy designed on the basis of aggregate features only could face serious acceptability problems. Indeed, the use of a tax policy could be put forward in such situations. Yet, depending on the population distribution, some sub-populations should face a subsidy for efficiency reasons. As such, perceptions could actually raise serious efficiency problems and not only issues of social justice. Several other potential policies (based on the environmental fundamentals or on agents' perceptions) are discussed and contrasted depending on their likely effects on welfare and on their appropriate designs.

This paper is a first step in the analysis of natural resource management problems driven by risk perceptions. We focus on the case of a one-shot irreversible regime shift: it could be interesting to analyze cases where the effects of the shift are reversible, while it may occur repeatedly in the future. We would not expect fundamentally different findings, as long as there is no learning: the reversibility of the shift might weaken some of the effects analyzed here, but the fact that the shift could repeat in the future might reinforce them on the other hand. The in-depth analysis of different types of policy instruments (combining economic and psychological interventions, as suggested by Stern (2011)) could also constitute a next step. Finally, it could also be worth studying different types of perception and their implications for resource management.

Appendix

A Proof of Proposition 1.

A.1 Post-shift problem

Plugging $V_2^{r=o,u,j}(y) = a_2^r \ln y + b_2^r$ into (6) yields

$$V_2^l(y) = \max_{0 \le c_l \le y - \sum_{k \ne l} c_k} \phi_2 \ln c_l + \delta \alpha_2 a_2^l \ln \left(y - c_l - \sum_{k \ne l} c_k \right) + \delta b_2^l \tag{25}$$

with $l = o, u, j$, the first-order conditions are $\frac{\phi_2}{c_l} = \frac{a_2^l \alpha_2 \delta}{y - N^j c_j - N^u c_u - N^o c_o}$ and solving for c_j, c_o and c_u

$$c_u = y \frac{\phi_2 a_2^j a_2^o}{N^j a_2^u a_2^o + N^o a_2^u a_2^j + N^u a_2^o a_2^j + a_2^u a_2^o a_2^j \delta \alpha_2}; c_j = c_u \frac{a_2^u}{a_2^j}; c_o = c_u \frac{a_2^u}{a_2^o} \tag{26}$$

Using (26) and (25) yields $V_2^{r=j,o,u}(y) = \phi_2 \ln c_r(y) + \delta \alpha_2 a_2^r \ln \left(y - N^j c_j(y) - N^o c_o(y) - N^u c_u(y) \right) + \delta b_2^r$ Then, we deduce quickly that

$$a_2^{r=j,o,u} = \frac{\phi_2}{1 - \delta \alpha_2}; b_2^{r=j,o,u} = \frac{\phi_2 ln \bar{c}_r + \delta \alpha_2 a_2^r \ln \left(1 - N^j c_j - N^o c_o - N^u c_u \right)}{1 - \delta}$$

A.2 Pre-shift problem

Plugging the conjecture $V_1^{r=o,u,j}(y) = a_1^r \ln y + b_1^r$ we obtain the pre-shift value function of a type-l agent ($l = j, o, u$)

$$V_1^l(y) = \max_{0 \le c_l \le y - \sum_{k \ne l} c_k} \phi_1 \ln c_l + \left(1 - p^l \right) \delta \alpha_1 a_1^l \ln \left(y - c_l - \sum_{k \ne l} c_k \right)$$

$$+ p^l \delta \alpha_2 a_2^l \ln \left(y - c_l - \sum_{k \ne l} c_k \right) + \delta \left(1 - p^l \right) b_1^l + \delta p^l b_2^l \tag{27}$$

The first-order conditions are $\frac{\phi_1}{c_r} = \frac{(1-p^r)\delta \alpha_1 a_1^r + p^r \delta \alpha_2 a_2^r}{y_t - N^j c_j - N^o c_o - N^u c_u}$ and solving for g_j, g_o and g_u we obtain:

$$g_o(y) = y \frac{\phi_1 z_j z_u}{z_j z_o z_u + \phi_1 \left[N^j z_o z_u + N^o z_j z_u + N^u z_j z_o \right]} = g_o^{dec} y; g_u(y) = y \frac{\phi_1 z_j z_o}{z_j z_o z_u + \phi_1 \left[N^j z_o z_u + N^o z_j z_u + N^u z_j z_o \right]} = g_u^{dec} y$$

$$g_j(y) = y \frac{\phi_1 z_o z_u}{z_j z_o z_u + \phi_1 \left[N^j z_o z_u + N^o z_j z_u + N^u z_j z_o \right]} = g_j^{dec} y$$

where

$$z_j = \left[(1-p) \delta \alpha_1 a_1^j + p \delta \alpha_2 a_2^j \right]; z_o = \left[(1-p^o) \delta \alpha_1 a_1^o + p^o \delta \alpha_2 a_2^o \right]; z_u = \left[(1-p^u) \delta \alpha_1 a_1^u + p^u \delta \alpha_2 a_2^u \right]$$

Using the conjecture $V_1^{r=o,u,j}(y) = a_1^r \ln y + b_1^r$, we write the value of the problem before shift

$$V_1^{r=o,u,j}(y) = \phi_1 \ln g_r(y) + ((1-p^r)\,\delta\alpha_1 a_1^r + p^r\delta\alpha_2 a_2^r)\ln\left(y\left(1 - N^o g_o^{dec} - N^u g_u^{dec} - N^j g_j^{dec}\right)\right) + \delta(1-p^r)\,b_1^r + \delta p^r b_2^r$$

Then, we can write

$$a_1^r = \frac{\phi_1 + p^r\delta\alpha_2 a_2^r}{1-(1-p^r)\,\delta\alpha_1};\, b_1^r = \frac{\phi_1 ln g_r^{dec} + ((1-p^r)\,\delta\alpha_1 a_1^r + p^r\delta\alpha_2 a_2^r)\ln\left(1 - N^o g_o^{dec} - N^u g_u^{dec} - N^j g_j^{dec}\right) + \delta p^r b_2^r}{1-\delta(1-p^r)}$$

To prove uniqueness, agent r's post-shift first-order condition is so:

$$\frac{\Phi_2}{h_r^2} - \delta(V_2^r)'\left((Y - h_r^2 - \sum_{j\neq r} h_j^2)^{\alpha_2}\right)\alpha_2(Y - h_r^2 - \sum_{j\neq r} h_j^2)^{\alpha_2-1} = 0$$

Thus the post-shift value function of agent r now satisfies $V_2^r(Y) = \Phi_2 ln h_r^2(Y) + \delta V_2^r\left((Y - h_r^2(Y) - \sum_{j\neq r} h_j^2(Y))^{\alpha_2}\right)$ Differentiating both sides with respect to Y we obtain:

$$(V_2^r)'(Y) = \Phi_2\frac{(h_r^2)'(Y)}{h_r^2(Y)} + \delta(V_2^r)'\left((Y - h_r^2(Y) - \sum_{j\neq r} h_j^2(Y))^{\alpha_2}\right)\alpha_2(1-(h_r^2)'(Y) - \sum_{j\neq r}(h_j^2)'(Y))(Y - h_r^2(Y) - \sum_{j\neq r} h_j^2(Y))^{\alpha_2-1}$$

Using the first order condition, this simplifies to:

$$(V_2^r)'(Y) = \delta(V_2^r)'\left((Y - h_r^2(Y) - \sum_{j\neq r} h_j^2(Y))^{\alpha_2}\right)\alpha_2(1 - \sum_{j\neq r}(h_j^2)'(Y))(Y - h_r^2(Y) - \sum_{j\neq r} h_j^2(Y))^{\alpha_2-1}$$

Again, using the first order condition, we have:

$$(V_2^r)'(Y) = \delta(V_2^r)'\left((Y - h_r^2(Y) - \sum_{j\neq r} h_j^2(Y))^{\alpha_2}\right)\alpha_2(1 - \sum_{j\neq r}(h_j^2)'(Y))(Y - h_r^2(Y) - \sum_{j\neq r} h_j^2(Y))^{\alpha_2-1} = \frac{\Phi_2}{h_r^2(Y)}(1 - \sum_{j\neq r}(h_j^2)'(Y))$$

This condition must hold for any state Y, so evaluating it for $Y = \left(Y_t - h_r^2(Y_t) - \sum_{j\neq r} h_j^2(Y_t)\right)^{\alpha_2}$ we obtain:

$$(V_2^r)'((Y_t - h_r^2(Y_t) - \sum_{j\neq r} h_j^2(Y_t))^{\alpha_2}) = \frac{\Phi_2}{h_r^2((Y_t - h_r^2(Y_t) - \sum_{j\neq r} h_j^2(Y_t))^{\alpha_2})}(1 - \sum_{j\neq r}(h_j^2)'((Y_t - h_r^2(Y_t) - \sum_{j\neq r} h_j^2(Y_t))^{\alpha_2}))$$

And the first order condition finally writes as:

$$\frac{\Phi_2}{h_r^2(Y_t)} = \delta\alpha_2\frac{\Phi_2}{h_r^2((Y_t - h_r^2(Y_t) - \sum_{j\neq r} h_j^2(Y_t))^{\alpha_2})}(1 - \sum_{j\neq r}(h_j^2)'((Y_t - h_r^2(Y_t) - \sum_{j\neq r} h_j^2(Y_t))^{\alpha_2}))(Y_t - h_r^2(Y_t) - \sum_{j\neq r} h_j^2(Y_t))^{\alpha_2-1}$$

Now, within the class of linear strategies MPNE we know that $h_r^2(Y_t) = g_r^2 Y_t$ for any agent r, and so:

$$1 = \delta\alpha_2\frac{1 - \sum_{j\neq r} g_j^2}{1 - \sum_j g_j^2}$$

This necessarily implies that $g_r^2 = g^2$ for any agent r and thus g^2 is characterized by $1 = \delta\alpha_2 \frac{1-(N-1)g^2}{1-Ng^2}$. We deduce that $\frac{1-(N-1)g^2}{g^2} = \frac{1}{1-\delta\alpha_2}$. Coming back at the pre-shift stage, a type-r agent's function is:

$$V_1^r(Y) = \max_{0 \le h_r^1 \le Y - \sum_{k \neq r} h_j^1} \Phi_1 ln h_r^1 + (1-p^r)\delta V_1^r \left((Y - h_r^1 - \sum_{j \neq r} h_r^1)^{\alpha_1} \right)$$

$$+ p^r \delta V_2^r \left((Y - h_r^1 - \sum_{j \neq r} h_j^1)^{\alpha_2} \right)$$

Using the same reasoning than in the post-shift case, the first-order condition can be rewritten as:

$$\frac{\Phi_1}{h_r^1(Y_t)} =$$

$$(1-p^r)\delta\alpha_1(Y_t - h_r^1(Y_t) - \sum_{j \neq r} h_j^1(Y_t))^{\alpha_1 - 1}(1 - \sum_{j \neq r}(h_r^1)'((Y_t - h_r^1(Y_t) - \sum_{j \neq r} h_j^1(Y_t))^{\alpha_1}) \frac{\Phi_1}{h_r^1((Y_t - h_r^1(Y_t) - \sum_{j \neq r} h_j^1(Y_t))^{\alpha_1})}$$

$$+ p^r \delta\alpha_2(Y_t - h_r^1(Y_t) - \sum_{j \neq r} h_j^1(Y_t))^{\alpha_2 - 1}(1 - \sum_{j \neq r}(h_r^2)'((Y_t - h_r^1(Y_t) - \sum_{j \neq r} h_j^1(Y_t))^{\alpha_2}) \frac{\Phi_2}{h_r^2((Y_t - h_r^1(Y_t) - \sum_{j \neq r} h_j^1(Y_t))^{\alpha_2})}$$

For each type of agents it can be checked that the first-order condition is identical, and thus we are looking for (within the class of linear MPNE) $h_j^1(Y) = g_j^{dec} Y$, $h_o^1(Y) = g_o^{dec} Y$ and $h_u^1(Y) = g_u^{dec} Y$, respectively. The corresponding first-order conditions can then be rewritten as

$$\Phi_1 = \frac{(1-p)\delta\alpha_1\Phi_1[1 - (N^j - 1)g_j^{dec} - N^u g_u^{dec} - N^o g_o^{dec}] + p\delta\alpha_2\Phi_2 g_j^{dec} \frac{1-(N-1)g^2}{g^2}}{1 - N^j g_j^{dec} - N^o g_o^{dec} - N^u g_u^{dec}}$$

for any science-based agent,

$$\Phi_1 = \frac{(1-p^u)\delta\alpha_1\Phi_1[1 - (N^u - 1)g_u^{dec} - N^j g_j^{dec} - N^o g_o^{dec}] + p\delta\alpha_2\Phi_2 g_u^{dec} \frac{1-(N-1)g^2}{g^2}}{1 - N^j g_j^{dec} - N^o g_o^{dec} - N^u g_u^{dec}}$$

for any u-type individual, and

$$\Phi_1 = \frac{(1-p^o)\delta\alpha_1\Phi_1[1 - (N^o - 1)g_o^{dec} - N^j g_j^{dec} - N^u g_u^{dec}] + p\delta\alpha_2\Phi_2 g_o^1 \frac{1-(N-1)g^2}{g^2}}{1 - N^j g_j^{dec} - N^o g_o^{dec} - N^u g_u^{dec}}$$

for any o-type individual. Now plugging $\frac{1-(N-1)g^2}{g^2} = \frac{1}{1-\delta\alpha_2}$ and equating the right hand side terms of all three conditions, we deduce that $g_o^{dec} = \frac{z_j}{z_o} g_j^{dec}$ and $g_u^{dec} = \frac{z_j}{z_u} g_j^{dec}$ hold. Plugging these two equalities into the first condition we obtain the desired expressions of g_j^{dec}, g_u^{dec} and g_o^{dec}.

B Proof of the statement

We differentiate equations (12) and (13) with respect to m

$$\frac{\partial g_j(y)}{\partial m} = -\phi_1 \frac{\frac{\partial \gamma}{\partial m}}{\phi_1 N^j + z_j} y; \frac{\partial g_i(y)}{\partial m} = \frac{\partial \gamma}{\partial m} y \tag{28}$$

Thus m affects the science-based agents and o-type individuals' extraction levels in opposite way.

C Proof of Proposition 2.

The only thing that remains to show is to notice that:

$$\frac{\partial z_j}{\partial m} = 0; \frac{\partial z_o}{\partial m} = \delta(x^o - p)\frac{\alpha_2 \frac{\phi_2}{1-\delta\alpha_2}(1-\delta\alpha_1) - \alpha_1\phi_1}{\left[1-(1-p^o)\delta\alpha_1\right]^2} \tag{29}$$

and

$$\frac{\partial z_j}{\partial p} = \delta\frac{\alpha_2 \frac{\phi_2}{1-\delta\alpha_2}(1-\delta\alpha_1) - \alpha_1\phi_1}{\left[1-(1-p)\delta\alpha_1\right]^2}; \frac{\partial z_o}{\partial p} = \delta(1-m)\frac{\alpha_2 \frac{\phi_2}{1-\delta\alpha_2}(1-\delta\alpha_1) - \alpha_1\phi_1}{\left[1-(1-p^o)\delta\alpha_1\right]^2} \tag{30}$$

We thus conclude (as $(x^o - p) > 0$) that $\frac{\partial z_o}{\partial m} \geq 0$ if and only if $\alpha_2 \frac{\phi_2}{1-\delta\alpha_2}(1-\delta\alpha_1) - \alpha_1\phi_1 \geq 0$ or $\phi_1 \leq \phi_2 \frac{\alpha_2(1-\delta\alpha_1)}{\alpha_1(1-\delta\alpha_2)}$ is satisfied. Moreover, we conclude that $\frac{\partial z_j}{\partial p} \geq 0$ and $\frac{\partial z_o}{\partial p} \geq 0$ if and only if $\phi_1 \leq \phi_2 \frac{\alpha_2(1-\delta\alpha_1)}{\alpha_1(1-\delta\alpha_2)}$ is satisfied, with $\frac{\partial z_o}{\partial p}$ getting arbitrarily low as m increases. The rest of the proof then follows.

D Proof of Proposition 3.

In order to simplify the exposition, we write g_i instead of $g_i(y)$. The reader is reminded that we have

$$g_j(y) = \phi_1\left(\frac{1-N^i\gamma_1(m)}{\gamma_2}\right)y; g_i(y) = \gamma_1(m)\,y \tag{31}$$

where $\gamma_1(m) = \frac{\phi_1 z_1}{(\phi_1 N^j + z_1)(\phi_1 N^i + z_2(m)) - \phi_1\phi_1 N^i N^j}$, $z_2(m) = \left(1-p^S\right)a_1^i\delta\alpha_1 + p^S a_2^i\delta\alpha_2$, $\gamma_2 = \phi_1 N^j + z_1$ and $z_1 = (1-p)\,a_1^j\delta\alpha_1 + pa_2^j\delta\alpha_2$. We obtain $\frac{\partial g_i(y)}{\partial m} = \gamma_1'(m)\,y$; $\frac{\partial g_j(y)}{\partial m} = -\frac{N^i\phi_1}{\gamma_2}\gamma_1'(m)\,y$.

We then obtain $\gamma_1'(m) = -\frac{\phi_1 z_1\left(\phi_1 N^j + z_1\right)z_2'(m)}{\left((\phi_1 N^j + z_1)(\phi_1 N^i + z_2(m)) - (\phi_1)^2 N^j N^i\right)^2}$ and the expression of $z_2(m)$ is

$$z_2(m) = (1-(m(x-p)+p))\,\delta\alpha_1\left(\frac{\phi_1 + (m(x-p)+p)\,\delta\alpha_2 a_2}{1-(1-(m(x-p)+p))\,\delta\alpha_1}\right) + (m(x-p)+p)\,\delta\alpha_2 a_2^i \tag{32}$$

Differentiating equation (32) with respect to m yields

$$z_2'(m) = \frac{a+b+c+d}{\left[1-(1-(m(x-p)+p))\,\delta\alpha_1\right]^2}$$

where

$$a = -(x-p)\,\delta\alpha_1\left(\phi_1 + (m(x-p)+p)\,\delta\alpha_2 a_2^i\right)(1-(1-(m(x-p)+p))\,\delta\alpha_1)$$

$$b = +(x-p)\,\delta\alpha_1(1-(m(x-p)+p))\,\delta\alpha_2 a_2^i(1-(1-(m(x-p)+p))\,\delta\alpha_1)$$

$$c = -(x-p)(1-(m(x-p)+p))(\delta\alpha_1)^2\left(\phi_1 + (m(x-p)+p)\,\delta\alpha_2 a_2^i\right)$$

$$d = +\left(x-p\right)\delta\alpha_2 a_2^i \left(1-\left(1-\left(m\left(x-p\right)+p\right)\right)\delta\alpha_1\right)^2$$

So for o-type individuals ($x-p>0$ holds) $\phi_1 \leq \phi_2 \frac{\alpha_2(1-\delta\alpha_1)}{\alpha_1(1-\delta\alpha_2)} \iff z_2'(m) \geq 0$ *and* $\gamma_1'(m) \leq 0$. Then this implies $\frac{\partial g_i(y)}{\partial m} \leq 0$ and $\frac{\partial g_j(y)}{\partial m} \geq 0$. The other case follows from similar arguments.

To proceed with the case of the aggregate extraction, we first differentiate g_i and g_j with respect to m. We can write[12] $\frac{\partial\psi(y)}{\partial m} = N^i \frac{\partial g_i(y)}{\partial m} + \left(N-N^i\right)\frac{\partial g_j(y)}{\partial m}$ or $\frac{\partial\psi(y)}{\partial m} = \gamma_1^{'}\left(m\right)N^i\left(1-\left(N-N^i\right)\frac{\phi_1}{\gamma_2}\right)y \lesseqgtr 0$ where $\gamma_2 = \phi_1 N^j + z_1$. From the proof of Proposition 2, when $x-p>0$ is satisfied we know that

$$\gamma_1'(m) \leq 0 \iff \phi_1 \leq \phi_2 \frac{\alpha_2\left(1-\delta\alpha_1\right)}{\alpha_1\left(1-\delta\alpha_2\right)}$$

We deduce that $\frac{\partial\Psi}{\partial m} \geq 0$ if and only if either $\gamma_1' \geq 0$ and $\frac{\gamma_2}{\phi_1} \geq \left(N-N^i\right)$ or $\gamma_1' \leq 0$ and $\frac{\gamma_2}{\phi_1} \leq \left(N-N^i\right)$ hold. Yet, from the definition of γ_2 we deduce that $\left(N-N^i\right) \leq \frac{\gamma_2}{\phi_1}$ always holds. All together $\frac{\partial\Psi}{\partial m} \geq 0$ if and only if $\gamma_1' \geq 0$ holds or $\phi_1 \geq \phi_2\frac{\alpha_2(1-\delta\alpha_1)}{\alpha_1(1-\delta\alpha_2)}$. The other case follows from similar arguments.

Regarding the effect of N^i, we obtain $\frac{\partial g_i(y)}{\partial N^i} = \left(\frac{\partial\gamma_1}{\partial N^i}\right)y$; $\frac{\partial g_j(y)}{\partial N^i} = -\phi_1\left(\frac{\gamma_1}{\gamma_2} + \left(\frac{N^i}{\gamma_2}\frac{\partial\gamma_1}{\partial N^i}\right)\right)y$ where $\frac{\partial\gamma_1}{\partial N^i} = -\frac{(\phi_1)^2 z_1}{(k_1)^2}\left(z_1-z_2\right)$ and $k_1 = \phi_1 z_2\left(N-N^i\right) + \phi_1 z_1 N^i + z_1 z_2$. We now rewrite the total extraction rate as $N^i g_i + \left(N-N^i\right)g_j = N^i g_i\left(1+\frac{N-N^i}{N^i}\frac{g_j}{g_i}\right)$. Since the ratio $\frac{g_j}{g_i}$ does not depend on N^i, we obtain:

$$\frac{\partial\left(N^i g_i + \left(N-N^i\right)g_j\right)}{\partial N^i} = \left(g_i + N^i\frac{\partial g_i}{\partial N^i}\right)\underbrace{\left(1+\frac{N-N^i}{N^i}\frac{g_j}{g_i}\right)}_{>0} - \frac{N}{N^i}g_j \tag{33}$$

We deduce that $\frac{\partial g_i}{\partial N^i} = -g_i\frac{\phi_1}{k_1}\left(z_1-z_2\right)$ and thus $\frac{\partial\left(N^i g_i+\left(N-N^i\right)g_j\right)}{\partial N^i} = g_i\frac{z_2(z_1-z_2)}{k_1}$. We obtain $z_1 - z_2 = \frac{p-p^S}{1-\delta\alpha_2}\frac{-\phi_1\delta\alpha_1(1-\delta\alpha_2)+\phi_2\delta\alpha_2(1-\delta\alpha_1)}{[1-(1-p)\delta\alpha_1][1-(1-p^S)\delta\alpha_1]}$. For u-type individuals $p-p^S>0$ and thus $z_1 \geq z_2$ (also $\frac{\partial\left(N^i g_i+\left(N-N^i\right)g_j\right)}{\partial N^i} \geq 0$) if and only if $-\phi_1\delta\alpha_1\left(1-\delta\alpha_2\right) + \phi_2\delta\alpha_2\left(1-\delta\alpha_1\right) \geq 0$ holds or $\phi_1 \leq \phi_2\frac{\alpha_2(1-\delta\alpha_1)}{\alpha_1(1-\delta\alpha_2)}$. The other case follows from similar arguments.

E Proof of Proposition 4.

i) We differentiate $g_j\left(y\right) = y\frac{\phi_1 z_o z_u}{z_j z_o z_u + \phi_1[N^j z_o z_u + N^o z_j z_u + N^u z_j z_o]} = g_j^{dec} y$ with respect to m

$$\frac{\partial g_j}{\partial m} = y\frac{\phi_1\left(z_o^{'}z_u + z_o z_u^{'}\right)}{z_j z_o z_u + \phi_1\left(N^j z_o z_u + N^o z_j z_u + N^u z_j z_u\right)}$$
$$- y\frac{\phi_1 z_o z_u\left[z_j z_o^{'} z_u + z_j z_o z_u^{'} + \phi_1\left(N^j z_o^{'} z_u + N^j z_o z_u^{'} + N^o z_u^{'} z_j + N^u z_j z_o^{'}\right)\right]}{\left[z_j z_o z_u + \phi_1\left(N^j z_o z_u + N^o z_j z_u + N^u z_j z_u\right)\right]^2} \tag{34}$$

After arranging all terms, we obtain

$$\frac{\partial g_j}{\partial m} = y\frac{\left(\phi_1\right)^2 z_j\left(z_o^{'}\left(z_u\right)^2 N^o + z_u^{'}\left(z_o\right)^2 N^u\right)}{\left[z_j z_o z_u + \phi_1\left(N^j z_o z_u + N^o z_j z_u + N^u z_j z_u\right)\right]^2} \tag{35}$$

If $\phi_1 \leq \frac{\phi_2\alpha_2(1-\delta\alpha_1)}{\alpha_1(1-\delta\alpha_2)}$ we have $z_o^{'} \geq 0$. Thus $\frac{\partial g_j}{\partial m} \geq 0$ if and only if $N^o \geq -\frac{z_u^{'}(z_o)^2 N^u}{z_o^{'}(z_u)^2}$ holds. We know from

[12] We use either ∂f or $f^{'}$ to denote the (partial) derivative of function f.

the proof of Proposition 1 that a science-based agent's value function satisfies $V_1^j(Y) = a_1^j ln(Y) + b_1^j$ where $a_1^j = \frac{\Phi_1 + p\delta\alpha_2 \frac{\Phi_2}{1-\delta\alpha_2}}{1-(1-p)\delta\alpha_1}$ and

$$b_1^j = \frac{\Phi_1 ln \frac{\Phi_1 z_u z_o}{z_j z_o z_u + \Phi_1[N^o z_j z_u + N^u z_j z_o + N^j z_o z_u]} + \delta\left[(1-p)\alpha_1 a_1^j + p\alpha_2 \frac{\Phi_2}{1-\delta\alpha_2}\right] ln \frac{z_j z_o z_u}{z_j z_o z_u + \Phi_1[N^o z_j z_u + N^u z_j z_o + N^j z_o z_u]} + \delta p b_2^j}{1-\delta(1-p)}$$

with $b_2^j = \frac{\Phi_2 ln g^2 + \delta\alpha_2 \frac{\Phi_2}{1-\delta\alpha_2} ln[1-Ng^2]}{1-\delta}$. It is easily checked that $\frac{\partial b_2^j}{\partial m} = 0$ due to the characterization of g^2. We also deduce that $\frac{\partial a_1^j}{\partial m} = 0$ holds, and thus $\frac{\partial V_1^j}{\partial m} = \frac{\partial b_1^j}{\partial m}$ with

$$\frac{\partial b_1^j}{\partial m} = \frac{\Phi_1 \frac{\frac{\partial g_j^1}{\partial m}}{g_j^1} + \delta[(1-p)\alpha_1 a_1^j + p\alpha_2 \frac{\Phi_2}{1-\delta\alpha_2}] \frac{-\frac{\partial G^{dec}}{\partial m}}{G_1}}{1-\delta(1-p)}$$

We obtain that $\frac{\frac{\partial g_j^{dec}}{\partial m}}{g_j^{dec}} = \frac{\Phi_1 z_j [N^o (z_u)^2 \frac{\partial z_o}{\partial m} + N^u (z_o)^2 \frac{\partial z_u}{\partial m}]}{z_j z_o z_u + \Phi_1[N^o z_j z_u + N^u z_j z_o + N^j z_o z_u]}$ while $-\frac{\frac{\partial G^{dec}}{\partial m}}{G^{dec}} = \frac{\Phi_1 z_j [N^o (z_u)^2 \frac{\partial z_o}{\partial m} + N^u (z_o)^2 \frac{\partial z_u}{\partial m}]}{z_o z_u (z_j z_o z_u + \Phi_1[N^o z_j z_u + N^u z_j z_o + N^j z_o z_u])}$ and the sign of $\frac{\partial b_1^j}{\partial m}$ is thus given by that of

$$\Phi_1 \frac{\frac{\partial g_j^{dec}}{\partial m}}{g_j^{dec}} + \delta[(1-p)\alpha_1 a_1^j + p\alpha_2 \frac{\Phi_2}{1-\delta\alpha_2}] \frac{-\frac{\partial G^{dec}}{\partial m}}{G^{dec}} =$$

$$\left[\Phi_1 + \delta \frac{(1-p)\alpha_1\Phi_1 + p\alpha_2 \frac{\Phi_2}{1-\delta\alpha_2}}{z_o z_u [1-(1-p)\delta\alpha_1]}\right] \frac{\Phi_1 z_j}{(z_j z_o z_u + \Phi_1[N^o z_j z_u + N^u z_j z_o + N^j z_o z_u]} \left[N^o (z_u)^2 \frac{\partial z_o}{\partial m} + N^u (z_o)^2 \frac{\partial z_u}{\partial m}\right]$$

We conclude that the sign of $\frac{\partial b_1^j}{\partial m}$ is given by that of $N^o (z_u)^2 \frac{\partial z_o}{\partial m} + N^u (z_o)^2 \frac{\partial z_u}{\partial m}$. Since

$$\frac{\partial z_r}{\partial m} = \frac{\delta(x^r - p)}{[1-(1-p)\delta\alpha_1]^2} \left[\alpha_2 \frac{\Phi_2}{1-\delta\alpha_2}(1-\delta\alpha_1) - \alpha_1\Phi_1\right]$$

we conclude that, when $\Phi_1 \leq \frac{\alpha_2}{\alpha_1} \frac{1-\delta\alpha_1}{1-\delta\alpha_2} \Phi_2$ is satisfied then $\frac{\partial z_u}{\partial m} \leq 0$ and $\frac{\partial z_o}{\partial m} \geq 0$ are satisfied. Then $\frac{\partial b_1^j}{\partial m} \geq 0$ if and only if $\frac{N^0}{N^u} \geq \left(\frac{z_o}{z_u}\right)^2 \frac{-\frac{\partial z_u}{\partial m}}{\frac{\partial z_o}{\partial m}}$ when $\Phi_1 \leq \frac{\alpha_2}{\alpha_1} \frac{1-\delta\alpha_1}{1-\delta\alpha_2} \Phi_2$ is satisfied. By contrast, when $\Phi_1 > \frac{\alpha_2}{\alpha_1} \frac{1-\delta\alpha_1}{1-\delta\alpha_2} \Phi_2$ is satisfied then $\frac{\partial z_u}{\partial m} \geq 0$ and $\frac{\partial z_o}{\partial m} \leq 0$ are satisfied, then $\frac{\partial b_1^j}{\partial m} \geq 0$ if and only if $\frac{N^u}{N^o} \geq \left(\frac{z_u}{z_o}\right)^2 \frac{-\frac{\partial z_o}{\partial m}}{\frac{\partial z_u}{\partial m}}$ is satisfied.

ii) We obtain $\frac{\partial g_o}{\partial m} = \frac{\partial g_j}{\partial m} \frac{z_j}{z_o} - g_j \frac{z_j}{(z_o)^2} z_o^{'}$; $\frac{\partial g_u}{\partial m} = \frac{\partial g_j}{\partial m} \frac{z_j}{z_u} - g_j \frac{z_j}{(z_u)^2} z_u^{'}$ then using equation (35) and arranging terms yields

$$\frac{\partial g_o}{\partial m} = y \frac{\phi_1 z_j \left[\phi_1 z_j N^u \left(z_o z_u^{'} - z_u z_o^{'}\right) - (z_u)^2 z_o^{'} \left(z_j + \phi_1 N^j\right)\right]}{\left(z_j z_o z_u + \phi_1 \left(N^j z_o z_u + N^o z_j z_u + N^u z_j z_u\right)\right)^2}; \frac{\partial g_u}{\partial m} = y \frac{\phi_1 z_j \left[\phi_1 z_j N^o \left(z_o^{'} z_u - z_u^{'} z_o\right) - (z_o)^2 z_u^{'} \left(z_j + \phi_1 N^j\right)\right]}{\left(z_j z_o z_u + \phi_1 \left(N^j z_o z_u + N^o z_j z_u + N^u z_j z_u\right)\right)^2} \quad (36)$$

If $\phi_1 \leq \frac{\phi_2\alpha_2(1-\delta\alpha_1)}{\alpha_1(1-\delta\alpha_2)}$, then $z_o^{'} \geq 0$ and $z_u^{'} \leq 0$. It follows that $\frac{\partial g_o}{\partial m} \leq 0$ and $\frac{\partial g_u}{\partial m} \geq 0$. The same argument applies also when $\phi_1 \geq \frac{\phi_2\alpha_2(1-\delta\alpha_1)}{\alpha_1(1-\delta\alpha_2)}$ and we obtain $z_o^{'} \leq 0$ and $z_u^{'} \geq 0$.

We now rewrite the total extraction as $N^o g_o + N^u g_u + N^j g_j = g_j \left(N^o \frac{g_o}{g_j} + N^u \frac{g_u}{g_j} + N^j\right)$ then differentiating with respect to m and using $\frac{g_u}{g_j} = \frac{z_j}{z_u}$ and $\frac{g_o}{g_j} = \frac{z_j}{z_o}$ yield

$$\frac{\partial\left(N^o g_o+N^u g_u+N^j g_j\right)}{\partial m}=\frac{\partial g_j}{\partial m}\left(N^o\frac{g_o}{g_j}+N^u\frac{g_u}{g_j}+N^j\right)-g_j z_j\left(\frac{N^u z_u^{'}}{\left(z_u\right)^2}+\frac{N^o z_o^{'}}{\left(z_o\right)^2}\right) \tag{37}$$

We finally obtain $\frac{\partial\left(N^o g_o+N^u g_u+N^j g_j\right)}{\partial m}=-\left(\phi_1\right)^2 z_j\frac{\left(z_o^{'}(z_u)^2N^o+z_u^{'}(z_o)^2N^j\right)}{[z_j z_o z_u+\phi_1(N^j z_o z_u+N^o z_j z_u+N^u z_j z_u)]^2}>0$ if and only if $\frac{N^o}{N^u}$ satisfies $\frac{N^o}{N^u}<-\frac{z_u^{'}}{z_o^{'}}\left(\frac{z_o}{z_u}\right)^2$ and ϕ_1 satisfies $\phi_1>\frac{\phi_2\alpha_2(1-\delta\alpha_1)}{\alpha_1(1-\delta\alpha_2)}$.

F Proof of Proposition 5.

i) Plugging $\sum_{j=1}^N\left[a_1^j\text{ln}y+b_1^j\right]$ into (21) we obtain:

$$\sum_{j=1}^N\left[a_2^j\text{ln}y+b_2^j\right]=\max_{c_{j,2}}\sum_{j=1}^N\left[\phi_2\text{ln}c_{j,2}+\delta\alpha_2 a_2^j\text{ln}\left(y-\sum_{k=1}^N c_{k,2}\right)+\delta b_2^j\right] \tag{38}$$

The first order conditions are $\left(\sum_{k=1}^N\delta\alpha_2 a_2^k\right)c_{j,2}+\phi_2\sum_{k=1}^N c_{k,2}=\phi_2 y$. We deduce that $c_{j,2}=c_{i,2}=c_2$ for any $i\neq j$ and we conclude that $c_2=l_2^{pop}y; y-Nc_2=q_2^{pop}y$ where $l_2^{pop}=\frac{\phi_2}{N\phi_2+\left(\sum_{k=1}^N\delta\alpha_2 a_2^k\right)}$ and $q_2^{pop}=\frac{\left(\sum_{k=1}^N\delta\alpha_2 a_2^k\right)}{N\phi_2+\left(\sum_{k=1}^N\delta\alpha_2 a_2^k\right)}$. Arranging terms in (38) yields

$$\sum_{j=1}^N\left[a_2^j\text{ln}y+b_2^j\right]=\sum_{j=1}^N\left[\left(\phi_2+\delta\alpha_2 a_2^j\right)\text{ln}y+\delta\alpha_2 a_2^j\text{ln}q_2^{pop}+\delta b_2^j+\phi_2\text{ln}l_2^{pop}\right] \tag{39}$$

From (39), consistency yields $a_2^j=\frac{\phi_2}{1-\delta\alpha_2}$ and $b_2^j=\frac{\delta\alpha_2 a_2^j\text{ln}q_2^{pop}+\phi_2\text{ln}l_2^{pop}}{1-\delta}$. Coming back to the expressions of l_2^{pop} and q_2^{pop} we now obtain

$$l_2^{pop}=\frac{\phi_2}{N\phi_2+\left(\sum_{k=1}^N\delta\alpha_2 a_2^k\right)}=\frac{1-\delta\alpha_2}{N};q_2^{pop}=\frac{\left(\sum_{k=1}^N\delta\alpha_2 a_2^k\right)}{N\phi_2+\left(\sum_{k=1}^N\delta\alpha_2 a_2^k\right)}=\delta\alpha_2 \tag{40}$$

In the pre-shift problem we plug $\sum_{j=1}^N\left[a_1^j\text{ln}y+b_1^j\right]$ and we obtain

$$\sum_{j=1}^N\left[a_1^j\text{ln}y+b_1^j\right]=\max_{c_{i,1}}\sum_{j=1}^N\left[\phi_1\text{ln}c_{j,1}+\left(1-p^j\right)\delta\alpha_1 a_1^j\text{ln}\left(y-\sum_{k=1}^N c_{k,1}\right)+p^j\delta\alpha_2 a_2^j\text{ln}\left(y-\sum_{k=1}^N c_{k,1}\right)+\left(1-p^j\right)\delta b_1^j+p^j\delta b_2^j\right] \tag{41}$$

The first order conditions are $\frac{\phi_1}{c_{j,1}}=\frac{\sum_{k=1}^N\left[\left(1-p^k\right)\delta\alpha_1 a_1^k+p^k\delta a_2^k\right]}{y-\sum_{k=1}^N c_{k,1}}$. We deduce that $c_{j,1}=c_{i,1}=c_1$ for any $j\neq i$ and the optimality conditions can be rewritten as

$$\left[\sum_{k=1}^N\left(\left(1-p^k\right)\delta\alpha_1 a_1^k+p^k\delta a_2^k\right)+N\phi_1\right]c_1=\phi_1 y \tag{42}$$

Using (42) we conclude that $c_1=l_1^{pop}y$ and $y-Nc_1=q_1^{pop}y$ where $l_1^{pop}=\frac{\phi_1}{N\phi_1+\sum_{k=1}^N\left(\left(1-p^k\right)\delta\alpha_1 a_1^k+p^k\delta a_2^k\right)}$ and $q_1^{pop}=\frac{\sum_{k=1}^N\left(\left(1-p^k\right)\delta\alpha_1 a_1^k+p^k\delta a_2^k\right)}{N\phi_1+\sum_{k=1}^N\left(\left(1-p^k\right)\delta\alpha_1 a_1^k+p^k\delta a_2^k\right)}$, thus

$$\sum_{j=1}^{N}[a_1^j \ln y + b_1^j] = \sum_{j=1}^{N}[(\phi_1 + (1-p^j)\delta\alpha_1 a_1^j + p^j\delta\alpha_2 a_2^j)\ln y + ((1-p^j)\delta\alpha_1 a_1^j + p^j\delta\alpha_2 a_2^j)\ln q_1^{pop} + \phi_1 \ln l_1^{pop} + p^j\delta b_2^j + (1-p^j)\delta b_1^j] \tag{43}$$

From (43) consistency yields $a_1^j = \frac{\phi_1 + p^j\delta\alpha_2 a_2^j}{1-(1-p^j)\delta\alpha_1} = \frac{\phi_1 + p^j\delta\alpha_2 \frac{\phi_2}{1-\delta\alpha_2}}{1-(1-p^j)\delta\alpha_1}$; $b_1^j = \frac{\left((1-p^j)\delta\alpha_1 a_1^j + p^j\delta\alpha_2 a_2^j\right)\ln q_1^{pop} + \phi_1 \ln l_1^{pop} + p^j\delta b_2^j}{1-(1-p^j)\delta}$.

Coming back to the expressions of l_1^{pop} and q_1^{pop}, we obtain:

$$l_1^{pop} = \frac{\phi_1}{N\phi_1 + \sum_{k=1}^{N}[(1-p^k)\delta\alpha_1 \frac{\phi_1 + \delta p^k \alpha_2 \frac{\phi_2}{1-\delta\alpha_2}}{1-\delta\alpha_1(1-p^k)} + p^k\delta\alpha_2 \frac{\phi_2}{1-\delta\alpha_2}]}; q_1^{pop} = \frac{\sum_{k=1}^{N}[(1-p^k)\delta\alpha_1 \frac{\phi_1 + \delta p^k \alpha_2 \frac{\phi_2}{1-\delta\alpha_2}}{1-\delta\alpha_1(1-p^k)} + p^k\delta\alpha_2 \frac{\phi_2}{1-\delta\alpha_2}]}{N\phi_1 + \sum_{k=1}^{N}[(1-p^k)\delta\alpha_1 \frac{\phi_1 + \delta p^k \alpha_2 \frac{\phi_2}{1-\delta\alpha_2}}{1-\delta\alpha_1(1-p^k)} + p^k\delta\alpha_2 \frac{\phi_2}{1-\delta\alpha_2}}$$

The proof of point ii) follows directly from point i).

G Proof of Proposition 6.

We first consider the case of a populist social planner. We have:

$$g_o^{dec}(y) \geq g^{pop}(y) \Longleftrightarrow \phi_1 \left[N^j z_u (z_j - z_o) + N^u z_j (z_u - z_o)\right] + z_j z_u \left[N_j z_j + (N^o - 1) z_o + N^u z_u\right] \geq 0 \tag{44}$$

Condition $\phi_1 \leq \frac{\alpha_2(1-\delta\alpha_1)}{\alpha_1(1-\delta\alpha_2)}\phi_2$ is necessary and sufficient to ensure that $z_j < z_o$ and $z_u < z_o$ hold. The first term in the above sum is thus negative, while the second one is positive. When $\phi_1 > \frac{\alpha_2(1-\delta\alpha_1)}{\alpha_1(1-\delta\alpha_2)}\phi_2$ holds $z_j > z_o$ and $z_u > z_o$ hold. Both terms in the above sum are positive, and the second conclusion follows:

$$g_o^{dec}(y) \geq g^{pop}(y) \Longleftrightarrow N^o - 1 \geq \frac{N^j z_u \left[\phi_1 (z_o - z_j) - (z_j)^2\right] + N^u z_j \left[\phi_1 (z_o - z_u) - (z_u)^2\right]}{z_o z_j z_u}$$

We now have:

$$g_u^{dec}(y) \geq g^{pop}(y) \Longleftrightarrow \phi_1 \left[N^j z_o (z_j - z_u) + N^o z_j (z_o - z_u)\right] + z_j z_o \left[N_j z_j + (N^u - 1) z_u + N^o z_o\right] \geq 0$$

The conclusions follow from similar arguments. We also have:

$$g_j^{dec}(y) \geq g^{pop}(y) \Longleftrightarrow \phi_1 \left[N^o z_u (z_o - z_j) + N^u z_o (z_u - z_j)\right] + z_u z_o \left[N^u z_u + \left(N^j - 1\right) z_j + N^o z_o\right] \geq 0$$

We conclude by noticing (when $\Phi_1 \leq \frac{\alpha_2(1-\delta\alpha_1)}{\alpha_1(1-\delta\alpha_2)}\Phi_2$) that $N^o \geq N^u \frac{z_o}{z_u}\frac{z_j - z_u}{z_o - z_j}$ implies that $g_j^{dec}(y) \leq g^{pop}(y)$ holds, and that otherwise $g_j^{dec}(y) \geq g^{pop}(y)$ holds if and only if N^j is large enough. The aggregate extraction will be sub-optimally high if and only if:

$$\frac{N^j z_o z_u + N^o z_j z_u + N^u z_j z_o}{z_j z_o z_u + \phi_1 \left[N^j z_o z_u + N^o z_j z_u + N^u z_j z_o\right]} \geq \frac{N}{N\phi_1 + N^j z_j + N^o z_o + N^u z_u}$$

which always holds. For a paternalistic social planner the proof follows from similar arguments, with:

$$g_o^{dec}(y) \geq g^{pat}(y) \Longleftrightarrow \phi_1 \left[N^j z_u (z_j - z_o) + N^u z_j (z_u - z_o)\right] + z_j z_u \left[N z_j - z_o\right] \geq 0$$

$$g_u^{dec}(y) \geq g^{pat}(y) \Longleftrightarrow \phi_1\left[N^j z_o\left(z_j - z_u\right) + N^o z_j\left(z_o - z_u\right)\right] + z_j z_o\left[N z_j - z_u\right] \geq 0$$

$$g_j^{dec}(y) \geq g^{pat}(y) \Longleftrightarrow \phi_1\left[N^o z_u\left(z_o - z_j\right) + N^u z_o\left(z_u - z_j\right)\right] + (N-1)\, z_o z_u z_j \geq 0$$

Finally, the aggregate extraction will be sub-optimally high if and only if $\frac{N^j z_o z_u + N^o z_j z_u + N^u z_j z_o}{z_j z_o z_u + \phi_1[N^j z_o z_u + N^o z_j z_u + N^u z_j z_o]} \geq \frac{N}{N\phi_1 + N z_j}$ which always holds.

H Proof of Proposition 7.

The decentralized aggregate extraction level is given by $G^{dec}(Y) = \left[1 - \frac{z_j z_o z_u}{z_j z_o z_u + \Phi_1[N^o z_j z_u + N^u z_j z_o + N^j z_o z_u]}\right] Y$. We obtain $\frac{\partial G^{dec}(Y)}{\partial \Phi_2} = -\frac{\Phi_1[N^j (z_o z_u)^2 \frac{\partial z_j}{\partial \Phi_2} + N^u (z_j z_o)^2 \frac{\partial z_u}{\partial \Phi_2} + N^o (z_j z_u)^2 \frac{\partial z_o}{\partial \Phi_2}]}{(z_j z_o z_u + \Phi_1[N^o z_j z_u + N^u z_j z_o + N^j z_o z_u])^2} Y$. Since $\frac{\partial z_r}{\partial \Phi_2} = \frac{p^r \delta \alpha_2}{(1-\delta\alpha_2)[1-(1-p^r)\delta\alpha_1]} > 0$ for any $r = j, o, u$ we conclude that $\frac{\partial G^{dec}(Y)}{\partial \Phi_2} < 0$ so a policy aimed at increasing Φ_2 would result in lower aggregate extraction. Differentiating with respect to α_2:

$$\frac{\partial G^{dec}(Y)}{\partial \alpha_2} = -\frac{\Phi_1[N^j (z_o z_u)^2 \frac{\partial z_j}{\partial \alpha_2} + N^u (z_j z_o)^2 \frac{\partial z_u}{\partial \alpha_2} + N^o (z_j z_u)^2 \frac{\partial z_o}{\partial \alpha_2}]}{\left(z_j z_o z_u + \Phi_1[N^o z_j z_u + N^u z_j z_o + N^j z_o z_u]\right)^2} Y$$

Since the sign of $\frac{\partial z_r}{\partial \alpha_2}$ for any $r = j, o, u$ is given by that of $\frac{p^r \delta^2 \alpha_2 \Phi_2}{(1-\delta\alpha_2)^2} > 0$ we conclude that $\frac{\partial G^{dec}(Y)}{\partial \alpha_2} < 0$ so a policy aimed at decreasing α_2 would result in a higher decentralized aggregate extraction level.

I Proof of Proposition 8.

Let us first consider the case where the mitigation policy affects all agents. Differentiating the decentralized aggregate extraction level with respect to p:

$$\frac{\partial G^{dec}(Y)}{\partial p} = -\frac{\Phi_1[N^j (z_o z_u)^2 \frac{\partial z_j}{\partial p} + N^u (z_j z_o)^2 \frac{\partial z_u}{\partial p} + N^o (z_j z_u)^2 \frac{\partial z_o}{\partial p}]}{\left(z_j z_o z_u + \Phi_1[N^o z_j z_u + N^u z_j z_o + N^j z_o z_u]\right)^2} Y$$

Since the sign of $\frac{\partial z_r}{\partial p}$ for any $r = j, o, u$ is given by that of $(1-m)\delta\left(\alpha_2 \frac{\Phi_2}{1-\delta\alpha_2}(1-\delta\alpha_1) - \alpha_1\Phi_1\right)$ we conclude that $\frac{\partial z_r}{\partial p} > 0$ if and only if $\Phi_1 < \Phi_2 \frac{\alpha_2(1-\delta\alpha_1)}{\alpha_1(1-\delta\alpha_2)}$ is satisfied. We would then obtain $\frac{\partial G^{dec}(Y)}{\partial p} < 0$ so a policy aimed at decreasing p would result in a higher aggregate extraction level. By contrast, such a policy would result in a lower decentralized aggregate extraction level when $\Phi_1 \geq \Phi_2 \frac{\alpha_2(1-\delta\alpha_1)}{\alpha_1(1-\delta\alpha_2)}$ is satisfied. Let us now assume that the mitigation policy affects science-based agents only: This would correspond to a case where $\frac{\partial z_r}{\partial p} = 0$ for $r = o, u$ and thus

$$\frac{\partial G^{dec}(Y)}{\partial p} = -\frac{\Phi_1 N^j (z_o z_u)^2 \frac{\partial z_j}{\partial p}}{\left(z_j z_o z_u + \Phi_1[N^o z_j z_u + N^u z_j z_o + N^j z_o z_u]\right)^2} Y$$

The remainder of the proof is entirely similar.

J Proof of Proposition 9.

The aggregate extraction level corresponding to the initial population structure is $G^{dec}(Y)$. If the distribution is affected and becomes $(N^j + N^c, N^o, N^u - N^c)$ then the aggregate extraction level becomes

$$G^c(Y) = \left[1 - \frac{z_j z_o z_u}{z_j z_o z_u + \Phi_1[N^o z_j z_u + (N^u - N^c) z_j z_o + (N^j + N^c) z_o z_u]}\right] Y$$

To be effective the policy must result in $G^c(Y) < G^{dec}(Y)$ or $z_o z_u < z_o z_j$ or $z_u < z_j$, which holds (as $p^u < p$) if and only if $\frac{\partial z^r}{\partial p} > 0$ or $\delta\alpha_2 \frac{\Phi_2}{1-\delta\alpha_2}(1 - \delta\alpha_1) - \delta\alpha_1\Phi_1 > 0$ so finally $\Phi_1 < \frac{\alpha_2(1-\delta\alpha_1)}{\alpha_1(1-\delta\alpha_2)}\Phi_2$.

K Proof of Proposition 10.

We have $\frac{\partial G^{dec}(Y)}{\partial m} = -\frac{\Phi_1[N^u(z_j z_o)^2 \frac{\partial z_u}{\partial m} + N^o(z_j z_u)^2 \frac{\partial z_o}{\partial m}]}{(z_j z_o z_u + \Phi_1[N^o z_j z_u + N^u z_j z_o + N^j z_o z_u])^2} Y$. Since the sign of $\frac{\partial z_r}{\partial p}$ for any $r = o, u$ is given by that of $(x^r - p)\left(\delta\alpha_2 \frac{\Phi_2}{1-\delta\alpha_2}(1 - \delta\alpha_1) - \delta\alpha_1\Phi_1\right)$ we conclude that, when $\Phi_1 < \Phi_2 \frac{\alpha_2(1-\delta\alpha_1)}{\alpha_1(1-\delta\alpha_2)}$ holds then $\frac{\partial z_o}{\partial m} > 0$ while $\frac{\partial z_u}{\partial m} < 0$ holds. In order for the policy to be effective $\frac{\partial G^{dec}(Y)}{\partial m} > 0$ has to hold (decreasing m results in lower aggregate extraction) or $\frac{N^o}{N^u} < \left(\frac{z_o}{z_u}\right)^2 \frac{-\frac{\partial z_u}{\partial m}}{\frac{\partial z_o}{\partial m}}$, which can be rewritten as $\frac{N^o}{N^u} < \left(\frac{z_o}{z_u}\right)^2 \frac{p-x^u}{x^o-p} \left(\frac{1-\delta\alpha_1(1-p^o)}{1-\delta\alpha_1(1-p^u)}\right)^2$. The first and third ratios on the right hand side term of the inequality are greater than one, while the second one might be greater or smaller than one (depending on the relative values of x^o and x^u). By contrast, when $\Phi_1 > \Phi_2 \frac{\alpha_2(1-\delta\alpha_1)}{\alpha_1(1-\delta\alpha_2)}$ holds then $\frac{\partial z_o}{\partial m} < 0$ while $\frac{\partial z_u}{\partial m} > 0$ holds. In order for the policy to be effective $\frac{\partial G^{dec}(Y)}{\partial m} > 0$ has to hold or $\frac{N^u}{N^o} < \left(\frac{z_u}{z_o}\right)^2 \frac{-\frac{\partial z_o}{\partial m}}{\frac{\partial z_u}{\partial m}}$.

References

Abdellaoui, M., O. L'Haridon, and C. Paraschiv (2011). Experienced vs. described uncertainty: Do we need two prospect theory specifications? *Management Science 57*(10), 1879–1895.

Agbo, M. (2014). Strategic exploitation with learning and heterogeneous beliefs. *Journal of Environmental Economics and Management 67*(2), 126–140.

Arnell, N. W., E. L. Tompkins, and W. N. Adger (2005, nov). Eliciting information from experts on the likelihood of rapid climate change. *Risk Analysis 25*(6), 1419–1431.

Bramoullé, Y. and N. Treich (2009). Can uncertainty alleviate the commons problem? *Journal of the European Economic Association 7*, 1042–1067.

Breton, M. and M. Y. Keoula (2011, jul). Farsightedness in a coalitional great fish war. *Environmental and Resource Economics 51*(2), 297–315.

Breton, M. and M. Y. Keoula (2014). A great fish war model with asymmetric players. *Ecological Economics 97*, 209–223.

Carpenter, S. R., D. Ludwig, and W. A. Brock (1999, aug). Management of eutrophication for lakes subject to potentially irreversible change. *Ecological Applications 9*(3), 751–771.

Costello, C., B. Nkuiya, and N. Quérou (2019). Spatial renewable resource extraction under possible regime shift. *American Journal of Agricultural Economics 101*(2), 507–527.

Costello, C. and D. Ovando (2019). Status, institutions, and prospects for global capture fisheries. *Annual Review of Environment and Resources 44*, 177–200.

Costello, C., S. Polasky, and A. Solow (2001). Renewable resource management with environmental prediction. *Canadian Journal of Economics 34*, 196–211.

Diekert, F. and E. Nieminen (2015). International fisheries agreements with a shifting stock. *Dynamic Games and Applications 7*, 185–211.

Diekert, F. K. (2017). Threatening thresholds? The effect of disastrous regime shifts on the non-cooperative use of environmental goods and services. *Journal of Public Economics 147*, 30–49.

Douenne, T. and A. Fabre (2022, feb). Yellow vests, pessimistic beliefs, and carbon tax aversion. *American Economic Journal: Economic Policy 14*(1), 81–110.

Dutta, P. and R. Sundaram (1993). The tragedy of the commons ? *Economic Theory 3*(3), 413–426.

Elsayyad, M. and F. Morath (2016). Technology transfers for climate change. *International Economic Review 57*, 1057–1084.

Farhi, E. and X. Gabaix (2020). Optimal taxation with behavioral agents. *American Economic Review 110*(1), 298–336.

Fesselmeyer, E. and M. Santugini (2013). Strategic exploitation of a common resource under environmental risk. *Journal of Economic Dynamics and Control 37*(1), 125–136.

Gabaix, X. (2019). Behavioral Inattention. *Handbook of Behavioral Economics 2*, 261–343.

Kwon, O. S. (2006, apr). Partial international coordination in the great fish war. *Environmental & Resource Economics 33*(4), 463–483.

Lee, T. M., E. M. Markowitz, P. D. Howe, C.-Y. Ko, and A. Leiserowitz (2015). Predictors of public climate change awareness and risk perception around the world. *Nature Climate Change 5*, 1014–1020.

Levhari, D. and L. J. Mirman (1980). The Great Fish War: An Example Using a Dynamic Cournot-Nash Solution. *The Bell Journal of Economics 11*(1), 322.

Lucchetti, L. and M. Santugini (2012). Ownership risk and the use of common-pool natural resources. *Journal of Environmental Economics and Management 63*, 242–259.

Mäler, K. G., A. Xepapadeas, and A. De Zeeuw (2003). The Economics of Shallow Lakes. *Environmental and Resource Economics 26*(4), 603–624.

McCright, R. and A. Dunlap (2003). Defeating Kyoto: The conservative movement's impact on U.S. climate change policy. *Social Problems 50*, 348–373.

Miller, S. and B. Nkuiya (2016). Coalition formation in fisheries with potential regime shift. *Journal of Environmental Economics and Management 79*, 189–207.

Mitra, T. and S. Roy (2006). Optimal exploitation of renewable resources under uncertainty and the extinction of species. *Economic Theory 28*(1), 1–23.

Opotow, S. and L. Weiss (2000). Denial and the process of moral exclusion in environmental conflict. *Journal of Social Issues 56*, 475–490.

Oremus, K., J. Bone, C. Costello, J. Garcia Molinos, A. Lee, T. Mangin, and J. Salzman (2020). Governance challenges for tropical nations losing fish species due to climate change. *Nature Sustainability 3*, 277–280.

Polasky, S., A. D. Zeeuw, and F. Wagener (2011). Optimal management with potential regime shifts. *Journal of Environmental Economics and Management 62*, 229–240.

Quaas, M., S. Baumgartner, S. Becker, K. Frank, and B. Muller (2007, apr). Uncertainty and sustainability in the management of rangelands. *Ecological Economics 62*(2), 251–266.

Quaas, M., S. Baumgartner, and M. De Lara (2019). Insurance value of natural capital. *Ecological Economics 165*, 106388.

Quaas, M., D. van Soest, and S. Baumgartner (2013). Complementarity, impatience, and the resilience of natural-resource-dependent economies. *Journal of Environmental Economics and Management 66*(1), 15–32.

Ren, B. and S. Polasky (2014). The optimal management of renewable resources under the risk of potential regime shift. *Journal of Economic Dynamics and Control 40*, 195–212.

Sakamoto, H. (2014). Dynamic resource management under the risk of regime shifts. *Journal of Environmental Economics and Management 68*(1), 1–19.

Scheffer, M. (1998). The ecology of shallow lakes. *Chapman and Hall London*.

Scheffer, M., S. Carpenter, J. Foley, C. Folke, and B. Walker (2001). Catastrophic shifts in ecosystems. *Nature 413*, 591–596.

Sethi, G., C. Costello, A. Fisher, M. Hanemann, and L. Karp (2005). Fishery management under multiple uncertainty. *Journal of Environmental Economics and Management 50*, 300–318.

Stavins, R. (2001). The problem of the commons: Still unsettled after 100 years. *American Economic Review 101*, 81–108.

Stern, P. (2011). Contributions of psychology to limiting climate change. *American Psychologist 66*, 303–314.

Tiller, R., E. Nyman, D. Dankel, and Y. Liu (2019, jan). Resilience to exogenous shocks in environmental management regimes in the arctic – lessons learned from survivors. *The Polar Journal 9*(1), 133–153.

Treloar, N. (2019). A proposed exogenous cause of the global temperature hiatus. *Climate 7*(2), 31.

Tsur, Y. and A. Zemel (1995). Uncertainty and irreversibility in groundwater resource management. *Journal of environmental Economics and Management 29*(2), 149–161.

Ulph, A. and D. Ulph (1997). Global warming, irreversibility and learning. *Economic Journal 107*, 636–650.

Wagener, F. (2003). Skiba points and heteroclinic bifurcations, with applications to the shallow lake system. *Journal of Economic Dynamics and Control 27*(9), 1533–1561.