Doctoral School
of Social Sciences

PhD programme in Developing Economics and Local Systems

Consequences of Environmental Degradation in Developing Countries:

Adaptation, Inequality, and International Trade Patterns

Giulio Galdi
Academic year 2017/18

Supervisor: Angelo Antoci

Co-supervisor: Nicolò Bellanca

April 19, 2019

This page intentionally left blank.

Acknowledgements

First of all, I acknowledge I have much to fill this section with. Nonetheless, I shall restrain myself to as few lines as I can muster, for half the thousand lines I wish to write would pay smaller tribute than a short but finite note to all.

I would deem myself lucky to have worked under the supervision of a person whose wisdom was not a cloud obscuring my own curiosity, but rather a light constantly feeding it. However, *lucky* would underestimate my condition, for I had two such supervisors, whose every effort was genuinely devoted to my development as a researcher. Embracing the mission of mentors while retaining the attitude of peers, they became the examples I still try to follow.

The academia leads many to distrust peers, to envy, to fight for the little sunshine while inevitably casting shadows upon others, to grow thick walls instead of dreaming what ideas we might breed together. I am proud to say that my colleagues were all my friends throughout this journey. Their support and solidarity made every success sweeter and every hardship not as bitter.

For her who nested in my heart and braved it through heavy mists and summer storms, I have no words ringing truer than the pledge I here renew. For naught I am thankful more than for meeting you.

This page intentionally left blank.

Contents

Introduction

The relevance of environmental degradation for the well-being of populations in less developed countries is manifold, making a complete analysis too extensive for this work, if not for any. I shall first clarify the reason why I deem a specific focus on less developed countries necessary. Although climate change is a global phenomenon affecting populations from countries at varying stages of development, a few of its consequences gain significantly more salience for less developed countries, specifically. Of course, many other relevant issues relate to less developed countries in particular. I focus on the adverse effects of environmental degradation[1] for two main reasons. On the one hand, it is my impression that the increased vulnerability of less developed countries to adverse environmental degradation needs more attention by scholars (Biermann and Möller, 2019). On the other hand, climate change is of the most pressing problems of our time, for which we have little time to act (IPCC, 2018) and whose consequences weight on the shoulders of many future generations. In other terms, the environment is deteriorating fast, and not only this is occurring faster in developing countries than in industrialised ones, the former are also less capable to protect against it for institutional, technological, and financial reasons (Barbier, 2010; Blaikie, 2016).

The asymmetry of environmental deterioration and of its effects is observable in many indicators. For instance, soil erosion is affecting ever larger areas, most of them in developing countries, whose institutions are not able or willing to intervene (Blaikie, 2016). Soil erosion mainly affects farmers and their productivity, but in developing countries they do not have the financial and technological means to defend themselves (Barbier, 2010; Blaikie, 2016), which induces some of them to resort to migration (when possible) (Blaikie, 2016). Farmers trying to adapt to the erosion of soil may adopt coping

[1] We note that climate change and environmental degradation are typically seen as two slightly different concepts, with the former considered to be one of the causes of the latter. In this work, we use the terms interchangeably in as much as they ultimately lead to the worsening of human well-being through the deterioration of some environmental indicator.

strategies[2] that endanger the sustainability of proximate water basins, fisheries (for instance, see Dejen et al., 2017) and of forestry (see Wondie, 2010). These practices ultimately undermine the sustainable extraction of natural resources and thus the long term productivity of the related activities. In these few introductory lines, I touched upon the issues I decided to investigate in this doctoral thesis: adaptation strategies and their possible negative side effects, inequality in the consequences of and in the means to face environmental degradation, and the long term effects of the latter on productivity. In my opinion, these are three very relevant pathways connecting environmental degradation to hindrances and obstacles to sustainable development of less developed countries. In the three chapters constituting the thesis, I investigate the economic dimensions of the problems above mentioned.

Firstly, I look deeper into the public good nature of the environment and what this implies for coping strategies of agents. Since the environment is a public good, its degradation can be viewed as a public bad. A public bad denotes a problem that is affecting a collective whose members cannot or do not have the incentives to solve such problem individually. Each member of the collective could tackle the problem in some measure, reducing its adverse effects also for everyone else. In other terms, individual effort towards reducing the public bad emits a positive externality for others. For this reason, the sum of the individual gains for each member, i.e. the collective gain, is higher than the individual benefit. Although the former is typically higher than the individual cost of tackling the problem, the latter is not, thus we have a stall in which no member puts in effort to reduce the public bad (environmental degradation, in this work), even if individual action is highly effective. In the first chapter of this doctoral thesis, I expand on this issue by considering the behaviour of populations that are granted access to an individual adaptation strategy that soften the problem for the adopter. Many adaptation strategies exist that allow individuals to defend against adverse effects of environmental degradation; some also contribute to tackle the problem for all, others actually fuel it. Typical examples of the first type are home insulation (Gupta and Gregg, 2012; IPCC, 2014), cutting on energy consumption and emissions (Tompkins et al., 2013; IPCC, 2014), sustainable agricultural practices (Smith et al., 2007), but many more strategies of this kind are already being employed (see Tompkins et al., 2013, for a review of such strategies in less developed countries). Examples of the second type of adaptation strategies are also abundant, the most famous ones being air conditioning and heating (Lundgren and Kjellstrom, 2013), unsustainable

[2]The main coping strategies are represented by internal migration, irrigation, overexploitation, and dams construction.

agricultural practices (Wilson and Tisdell, 2001), defensive expenditure in the tourism industry (Abegg et al., 2008). However, every coping strategy to remedy adverse effects of environmental degradation which includes the use of natural resources or energy, and thus poses an additional burden on the environment, could be ascribed to this type[3]. In order to identify the two different adaptation types, the literature defines as mitigation each strategy which also benefits other agents (IPCC, 2014), whereas the strategies damaging other agents are defined as maladaptation (Barnett and O'Neill, 2010). Building on this relevant differentiation, I show that the public good nature of the environment leads to sub-optimal effects whether the adaptation strategy represents a case of mitigation or one of adaptation. On the one hand, mitigation strategies are subject to the aforementioned discrepancy between individual and collective gain, leading agents to contribute less than the optimal amount to a public solution[4]. When agents have the choice to either adopt a mitigation technology or not, a lower than optimal share of the populations chooses to adopt it. We thus talk of under-adoption of mitigation strategies. On the other hand, if a private action is able to reduce personal damage from environmental degradation at the expense of others, i.e. it is maladaptive, then this strategy may gain a more than optimal diffusion. When the population as a whole pays an adoption cost higher than the individual benefit, a large diffusion of maladaptive strategies ultimately makes the whole population worse off with respect to a scenario in which the maladaptive strategy did not exist. I analyse the problem formally, stressing the issues it raises both at the level of relations between countries at different level of development and at the policy-making level. Indeed, there is the possibility that a more developed region is able to shift same-group externalities to a less developed region. This can be the case of some green policies, generating the renown Pollution Haven Effect (Copeland and Taylor, 2004). I also argue why such a policy can be counterproductive to adopt when a maladaptive technology is available in both regions. As concerns the considerations for the policy-maker, I identify the difficulties that can be encountered when adaptation strategy of either type enter the economy.

Secondly, when adaptation strategies as the ones analysed in the first chapter arise, the extant theoretical literature typically assumes that regions and countries at different stages of development are equally able and effec-

[3]On a note, the belief that environmental quality can be substituted with additional economic growth does not seem to hold, for the latter only seems to accelerate environmental degradation and its adverse effects (Luzzati et al., 2018)

[4]This is often referred to as the *free rider problem*, as agents rely on others to tackle the problem in their stead

tive in implementing them. This would imply that populations from more developed countries can protect themselves just as well as populations from less developed countries. This implicit assumption overlooks the financial and institutional limits restraining the set of viable alternatives for the former (Toya and Skidmore, 2007; Barbier, 2010). To connect with a notorious example above mentioned, not every person in a developing country might be able to purchase an air conditioning system, even if they would like to. In this scenario, a part of the global population would be able to soften the problem, whereas the other would have to endure the adverse effects of environmental degradation, worsened by the negative externalities generated by maladaptive strategies of the former. This constitutes an inequality issue which goes beyond the problems already discussed in the first chapter, relating coping strategies with externalities in general. Indeed, in this case maladaptation not only has the potential to worsen the well-being of all, but also disproportionately affects the most vulnerable (Barnett and O'Neill, 2010). In order to address this phenomenon, which is overlooked by the literature on maladaptation, I propose a theoretical model building on the evidences of many development economists. In particular, scholars showed that less developed countries are less effective in implementing adaptation strategies (Barbier, 2010) and also more exposed to the adverse effects of environmental degradation (Strömberg, 2007; Jodha, 1986). I prove that even when a maladaptation strategy can be adopted by the population residing in a less developed country, these two features lead to an unfair distribution of the additional burden caused by externalities of the maladaptation strategy. In other terms, initial differences in well-being due to development gaps are amplified by the presences of a maladaptation strategy, even when neglecting the impossibility of the most vulnerable agents to adopt it. Clean or potable water, soil quality, forestry and fishery resources, and sanitation are the main instances of ecological services for which populations of less developed countries cannot protect as effectively as more developed countries (e.g. water treatment plants, fertilisers, medication). The inadequacy of the coping strategies with respect to environmental degradation may have relevant indirect effect on the lives of resident populations (e.g. prices, productivity, conflicts over scarce resources), to the point of constituting an incentive to leave the region or country (Beine and Parsons, 2015; Hunter et al., 2015; Veronis and McLeman, 2014). We can see that if the first chapter presents the challenges for the policy maker in the presence of maladaptation (and mitigation) strategies, this chapter declines these challenges in a North-South perspective and warns that as the effects of Climate Change become direr, matters of international equality may gain salience. Considering that in developing countries most of the population is already living on non-favourable land (Barbier, 2010) and that the number of people living in the areas most exposed to environmental degradation is expected to increase (World Bank, 2003), combating this disproportionate distribution of the consequences is of the utmost importance (Hsiang et al., 2017).

Finally, I investigate deeper the international consequences of environmental degradation in the third chapter, focusing on its adverse effects on productivity. I have already hinted at the maladaptive nature of some productive activity in the agricultural sector, boosting productivity in the short term but ultimately curtailing it in the long term (Wilson and Tisdell, 2001). It is relevant to note that maladaptive practices affecting productivity are a concern also for more developed countries (Christian-Smith et al., 2015), whose technologies may be highly effective in coping with the deterioration of one environmental indicator, but negatively impact another one. The deterioration of several local environmental indicators (water, air, soil, etcetera) is undermining economic growth also in more developed countries (Hsiang et al., 2017), not to mention that global environmental degradation is found accountable of major GDP losses by several studies (Stern et al., 2006; Sterner and Persson, 2008; Ng and Zhao, 2011). All these (negative) variations do not alter only the absolute value of economic variables, but also the relative ones. In other terms, the productivity of countries is not symmetrically reduced by environmental degradation, partly due to the unequal efficacy of coping strategies studied in the previous chapter. As the relative productivity of countries varies, their comparative advantages change as well, altering the pattern of international trade. Even when laxer environmental regulation in less developed countries determines a starting comparative advantage in the most polluting goods due to international firms seeking to abate compliance costs(as maintained by works supporting the Pollution Haven Hypothesis, see e.g. Sapkota and Bastola, 2017; Candau and Dienesch, 2017), an asymmetrical vulnerability of the economic systems may change the industry specialisations of countries. I formally describe this connection between environmental degradation and international trade patterns in the last chapter of my thesis. I show that when the country currently detaining the comparative advantage in the production of one good is also the one whose productivity is most affected by environmental degradation, such country may lose its competitive advantage in that sector[5]. Although it is hard to predict the extent to which international trade patterns will be affected by environmental degradation, this work stresses that climate change not only affects the single economic systems, but the interactions among them.

[5]As in the previous chapter, I allow for different stages of development and exposure to environmental degradation for countries, as in the previous chapter.

Prompted by the growing importance of climate change for the life of all and by a particular attention on the well-being of populations from less developed countries, I decided to study some of the linkages between these two topics. As multiple such linkages exist, I chose the ones that in my opinion would contribute more to the extant literature, either by proposing a new perspective on a well-known issue or by introducing a novel concept. Furthermore, all chapters propose original theoretical models, mainly building on dynamical systems and comparative advantage analysis. Each work is followed by a dedicated bibliography section, for the sake of facilitating the inspection of the cited literature. Finally, in order to align with most scientific works, in what follows I employ the first-person plural.

Bibliography

Abegg, Bruno, R Bürki, and Hans Elsasser (2008) "Climate change and tourism in the Alps," in A. Borsdorf, J. Stötter, and E. Veulliet eds. *Managing Alpine Future*, pp. 73–80.

Barbier, Edward B (2010) "Poverty, development, and environment," *Environment and Development Economics*, Vol. 15, pp. 635–660.

Barnett, Jon and Saffron O'Neill (2010) "Maladaptation," *Global Environmental Change*, Vol. 2, pp. 211–213.

Beine, Michel and Christopher Parsons (2015) "Climatic factors as determinants of international migration," *The Scandinavian Journal of Economics*, Vol. 117, pp. 723–767.

Biermann, Frank and Ina Möller (2019) "Rich man's solution? Climate engineering discourses and the marginalization of the Global South," *International Environmental Agreements: Politics, Law and Economics*, pp. 1–17.

Blaikie, Piers (2016) *The political economy of soil erosion in developing countries*: Routledge, London, GB.

Candau, Fabien and Elisa Dienesch (2017) "Pollution haven and corruption paradise," *Journal of Environmental Economics and Management*, Vol. 85, pp. 171–192.

Christian-Smith, Juliet, Morgan C Levy, and Peter H Gleick (2015) "Maladaptation to drought: a case report from California, USA," *Sustainability Science*, Vol. 10, pp. 491–501.

Copeland, Brian R and M Scott Taylor (2004) "Trade, growth, and the environment," *Journal of Economic Literature*, Vol. 42, pp. 7–71.

Dejen, Eshete, Wassie Anteneh, and Jacobus Vijverberg (2017) "The decline of The Lake Tana (Ethiopia) fisheries: causes and possible solutions," *Land Degradation & Development*, Vol. 28, pp. 1842–1851.

Gupta, Rajat and Matthew Gregg (2012) "Using UK climate change projections to adapt existing English homes for a warming climate," *Building and Environment*, Vol. 55, pp. 20–42.

Hsiang, Solomon, Robert Kopp, Amir Jina, James Rising, Michael Delgado, Shashank Mohan, DJ Rasmussen, Robert Muir-Wood, Paul Wilson, Michael Oppenheimer et al. (2017) "Estimating economic damage from climate change in the United States," *Science*, Vol. 356, pp. 1362–1369.

Hunter, Lori M, Jessie K Luna, and Rachel M Norton (2015) "Environmental dimensions of migration," *Annual Review of Sociology*, Vol. 41, pp. 377–397.

IPCC (2014) *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*: Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

——— (2018) *Global Warming of 1.5° C: An IPCC Special Report on the Impacts of Global Warming of 1.5° C Above Pre-industrial Levels and Related Global Greenhouse Gas Emission Pathways, in the Context of Strengthening the Global Response to the Threat of Climate Change, Sustainable Development, and Efforts to Eradicate Poverty*: Intergovernmental Panel on Climate Change, Geneva, Switzerland.

Jodha, Narpat S (1986) "Common property resources and rural poor in dry regions of India," *Economic and Political Weekly*, pp. 1169–1181.

Lundgren, Karin and Tord Kjellstrom (2013) "Sustainability challenges from climate change and air conditioning use in urban areas," *Sustainability*, Vol. 5, pp. 3116–3128.

Luzzati, T, A Parenti, and T Rughi (2018) "Economic Growth and Cancer Incidence," *Ecological Economics*, Vol. 146, pp. 381–396.

Ng, Pin and Xiaobing Zhao (2011) "No matter how it is measured, income declines with global warming," *Ecological Economics*, Vol. 70, pp. 963–970.

Sapkota, Pratikshya and Umesh Bastola (2017) "Foreign direct investment, income, and environmental pollution in developing countries: Panel data analysis of Latin America," *Energy Economics*, Vol. 64, pp. 206–212.

Smith, Pete, Daniel Martino, Zucong Cai, Daniel Gwary, Henry Janzen, Pushpam Kumar, Bruce McCarl, Stephen Ogle, Frank O'Mara, Charles Rice et al. (2007) "Greenhouse gas mitigation in agriculture," *Philosophical Transactions of the Royal Society B: Biological Sciences*, Vol. 363, pp. 789–813.

Stern, Nicholas, Siobhan Peters, Vicki Bakhshi, Alex Bowen, Catherine Cameron, Sebastian Catovsky, Diane Crane, Sophie Cruickshank, Simon

Dietz, Nicola Edmonson et al. (2006) *Stern Review: The economics of climate change*, Vol. 30: HM treasury, London, GB.

Sterner, Thomas and U Martin Persson (2008) "An even sterner review: Introducing relative prices into the discounting debate," *Review of Environmental Economics and Policy*, Vol. 2, pp. 61–76.

Strömberg, David (2007) "Natural disasters, economic development, and humanitarian aid," *Journal of Economic Perspectives*, Vol. 21, pp. 199–222.

Tompkins, Emma L, Adelina Mensah, Lesley King, Tran Kim Long, Elaine T Lawson, Craig W Hutton, Viet Anh Hoang, Chris Gordon, Marianne Fish, Jen Dyer et al. (2013) "An investigation of the evidence of benefits from climate compatible development,"Technical report.

Toya, Hideki and Mark Skidmore (2007) "Economic development and the impacts of natural disasters," *Economics Letters*, Vol. 94, pp. 20–25.

Veronis, Luisa and Robert McLeman (2014) "Environmental influences on African migration to Canada: focus group findings from Ottawa-Gatineau," *Population and Environment*, Vol. 36, pp. 234–251.

Wilson, Clevo and Clem Tisdell (2001) "Why farmers continue to use pesticides despite environmental, health and sustainability costs," *Ecological Economics*, Vol. 39, pp. 449–462.

Wondie, Ayalew (2010) "Improving management of shoreline and riparian wetland ecosystems: the case of Lake Tana catchment," *Ecohydrology & Hydrobiology*, Vol. 10, pp. 123–131.

World Bank (2003) *World Development Report 2003: Sustainable Development in a Dynamic World: Transforming Institutions, Growth, and Quality of Life*: Oxford University Press, New York, DC, USA.

Chapter 1

Over and Under-adoption of Environmental Adaptation Technologies with Public Effects

Abstract

The global nature of the climatic challenge requires a high level of cooperation among agents, especially since most of the related coping strategies produce some kind of externalities toward others. Whether they are positive or negative, the presence of externalities may lead the system towards Pareto-dominated states. In this work, we study under and over-adoption of environmental adaptation technologies which enhance environmental quality for the individual while transferring externalities to other agents. We distinguish adaptation technologies between ones of maladaptation, when the externalities are negative, and ones of mitigation, when they are positive. In particular, we show that over-adoption may occur for maladaptive technologies, whereas under-adoption may occur in case of mitigation. We study a model with two regions at different stages of development, which allows us to draw considerations on well-being consequences of environmental dumping.

Keywords: Adaptation, Negative externalities, Evolutionary dynamics, Public good game, Environmental dumping

JEL classification: C70, D62, O13, O40, Q20

1.1 Introduction

Environmental hazards are requiring humans to change their behaviour and decision making criteria, as they gain relevance at an increasing pace. We had to forgo the comforting idea that "natural systems fluctuate within an unchanging envelope of variability"(Milly et al., 2008). As concerns environmental hazards, responses are typically characterised dichotomously: adaptation and/or mitigation (IPCC, 2014), with the implementation of one not excluding the other's. On the one hand, humans may (and do) adapt to a changing climate, reducing their exposure to the ensuing harm. This includes responding to abnormal hot or cold temperatures, adopting new agricultural techniques to cope with the impoverishment of soil, creating artificial snow in ski resorts, and much more (for a broad review on many other forms of adaptation, see Tompkins et al., 2010). On the other hand, humans may (and do) try to tackle the problem at its source and combat the causes of increased environmental risks. Efficient water management, restoration of soil, substitution of fossil fuels with agricultural by-products are some of the mitigation techniques currently under study for the agricultural sector (Smith et al., 2007). These strategies not only reduce the environmental hazards for the adopter, but for all agents, thus generating a positive externality to other agents. With respect to mitigation strategies, adaptation does not aim to reduce the problem, but rather to avoid at least part of its adverse affects. At times, this is done at the expense of other agents, i.e. adaptation strategies may generate negative externalities. For instance, a farmer suffering from reduced plot productivity due to soil degradation may decide to raze a forest area to expand her plot and compensate for her loss of income. However, in this way she is further contributing to the problem of soil erosion and to the general loss of regenerative capacity of the ecosystem. When a strategy is such that it worsen environmental hazards for everyone else or it disproportionally affects the most vulnerable, the literature defines it no more as adaptation, but rather as maladaptation (Barnett and O'Neill, 2010). We here study maladaptive and mitigation strategies and illustrate how global externalities lead to over-adoption of the former and under-adoption of the latter.

Since mitigation strategies actually reduce environmental hazards instead of (temporarily?) avoiding its effects, it is usually considered to be the most desirable strategy (IPCC, 2014). However, there are reasons why humans did not respond with enthusiasm to the emergence of mitigation solutions in the face of environmental hazards. Firstly, many mitigation strategies require long-term investments to pay off, with a time scale that may exceed the average life expectancy of a person before they become effective (Hallegatte, 2009). The incapacity of humans to make long-term investments and their preference for the present are additional threats to our capacity to make long-term commitments to stop environmental degradation (Warburton et al., 2018), leading to issues of intergenerational equity (Glotzbach and Baumgartner, 2012). We remark that the existence of mitigation solutions is not a sufficient condition for the abatement of environmental damage. The literature has uncovered several ways in which externalities of any type, either negative or positive (as is the case for mitigation strategies) may undermine the achievement of the social optimum. On the one hand, whenever an agent may transfer her cost to protect against environmental hazards onto others, in a way that is either anonymous or has no consequences on the self, she has little incentive to adopt a mitigation strategy. On the other hand, if a strategy actually reduces environmental risks not only for the adopter, but also for other agents, i.e. it has a positive externality, then it may happen that all agents wait for the others to tackle environmental degradation for everyone, but none is willing to pay the cost for the benefit of others[1]. Scholars studying these shiftable externalities highlighted that policy tools hindering maladaptive actions and promoting mitigation ones are desirable, e.g. a tax on negative externalities or a subsidy on positive ones (Bird, 1987; Shaw and Shaw, 1991; Shogren and Crocker, 1991; Geaun, 1993).

In this work, we study the adoption dynamics when agents are faced with either a maladaptation strategy or a mitigation one. We assume that individuals from a more developed region and a less developed one have the possibility to adopt a technology which enhances environmental quality for the adopters, but also generates externalities on other agents. In particular, each region has a local environmental indicator which is affected by the adoption dynamics of both regions, so that the externalities have a global effect. These externalities may be either negative, in case of a maladaptation technology, or positive, in case of a mitigation technology. We highlight what is the underlying mechanism which leads to over-adoption of maladaptation strategies and under-adoption of mitigation ones. In addition, we show what are the effects on the less developed region if the maladaptation technology is such that it disproportionally burdens its population with respect to the agents from the more developed region. The adoption dynamics is modelled

[1]In an experimental setting, Hasson et al. (2010) show that agents rarely contribute to the mitigation solution and that their contributions to a common mitigation policy are not sensitive to the likelihood of extremely adverse events. In a somewhat similar experiment, Milinski et al. (2006) show that reputation effects may nudge agents to contribute in an environmental framing.

by a two population evolutionary game which employs replicator equations, so that all agents may imitate their peers in the region, if the well-being of the latter is greater. Our analysis leads to three major conclusions: 1) when only a maladaptive technology is available, either all agents adopt it or none, depending on the initial distribution of strategies; 2) when only a mitigation technology is available, the system *typically* reaches a state in which a part of the population adopts the technology while the rest does not; no path dependency arises 3) if the more developed region dumps negative externalities onto the less developed one, it might happen the the well-being of all agents decreases.

1.2 The model

We here formally analyse the dynamics in two regions, which we call N and S for illustrative purposes, when agents from both regions can either adopt a strategy **A** enhancing personal environmental quality at a cost C^j or choose not to do so (strategy **NA**). We can think of strategy **A** as to a technology which improves the personal quality of a local environmental indicator, but which also affects the well-being of others. When an agent from region j chooses the environment enhancing strategy **A**, she is improving the benefits she derives from a local environmental variable E^j by an amount $p^j \geq 0$, while also possibly altering environmental quality for other agents from both regions. Indeed, the overall quality of the environmental variable for agent i at any given time t is determined by the sum of the private effects of her actions and the public effects of all agents:

$$E_t^j = \begin{cases} \overline{E}^j + P_t^j & \text{if } i \text{ chooses strategy } \mathbf{NA} \\ \overline{E}^j + p^j + P_t^j & \text{if } i \text{ chooses strategy } \mathbf{A} \end{cases}$$

where j indexes for regions N, S; $\overline{E}^j > 0$ is the autonomous level of environmental quality when there is no agent adopting strategy **A**; P_t^j measures the public effect (i.e. the externalities) of the actions of all agents on E^j. We shall later specify how the former is derived; for now, it suffices to say that it can take either sign. The well-being of an agent i from region j depends on E^j and on whether she incurred in the adoption cost:

$$\Pi_i^j := \begin{cases} \ln(\overline{E}^j + P_t^j) & \text{if } i \text{ chooses strategy } \mathbf{NA} \\ \ln(\overline{E}^j + P_t^j + p^j) - C^D & \text{if } i \text{ chooses strategy } \mathbf{A} \end{cases} \tag{1.1}$$

where the adoption cost C^j is strictly positive. We now define the public effect P_t^j, which depends on the shares of agents $x_t, z_t \in [0,1]$ adopting strategy **A** at time t in regions N and S, respectively. More in detail, we differentiate between domestic and foreign effects of the adaptation technology. The former describes the impact on a local environmental indicator of same-region adopters, whereas the latter describes the impact of cross-region adopters. For the sake of simplicity, we assume that the public effects are determined by linear functions:

$$P_t^N := -d^N \cdot x_t - f^N \cdot z_t \tag{1.2a}$$

$$P_t^S := -f^S \cdot x_t - d^S \cdot z_t \tag{1.2b}$$

where d^j and f^j are the domestic and the foreign public effects, respectively, for country $j = N, S$. They represent the public impact of adoption of all agents on the local environmental indicator of region j, distinguished according to the source of such impact. Domestic effects d^j are caused by agents in region j and worsen the quality of their own local environmental indicator, whereas foreign effects f^j affect the local environmental indicator of region j but are caused by agents in the other region. We do not apply any sign restriction on the public effects, so that externalities of adoption of the environmental adaptation technology may take either sign. When a public effect P^j is positive, adoption of strategy **A** by an agent carries part of its benefits over to other agents. This case qualifies as a mitigation case, in which an agent is working for the cooperative improvement of environmental quality, or equivalently towards the abatement of pollution. By contrast, when the public effect is negative, an agent adopting strategy **A** is actually benefiting herself by worsening environmental quality for others. From the concavity of (1.1), we may add that a negative public effect affects relatively more (reduces well-being by a higher amount) the agents who are not adopting the environment enhancing strategy **A**. By the definition provided by Barnett and O'Neill (2010), this is a case of maladaptation.

In order to study the dynamics of this system, we now describe the way in which the share of agents adopting strategy **A** in either country varies. We assume that if the difference in well-being $\Delta\Pi^j = \Pi_A^j - \Pi_{NA}^j$ between strategy **A** and strategy **NA** is positive for region j, then the share of agents adopting the adaptation technology in that region will increase, since it provides higher payoffs. The opposite holds if the payoff difference is negative. Finally, if the payoff difference equals zero, economic agents are indifferent between adopting or not adopting the technology, so that the population shares of agents adopting the technology keeps constant over time. Therefore, we have that:

$$\Delta\Pi^N(x_t, z_t) \gtreqless 0 \Rightarrow \dot{x} \gtreqless 0 \quad \Delta\Pi^S(x_t, z_t) \gtreqless 0 \Rightarrow \dot{z} \gtreqless 0 \tag{1.3}$$

where $\dot{x}$ and $\dot{z}$ are the time derivatives of x_t and z_t, respectively. Hence, in each region the payoff difference $\Delta\Pi^j(x_t, z_t)$ in N and $\Delta\Pi^S(x_t, z_t)$ in S) has the same sign as the time derivative of the population share that adopts the environmental adaptation technology in that region. Referring to the well-being definition (1.1), we may explicit the payoff difference $\Delta\Pi^j$:

$$\Delta\Pi^j(x_t, z_t) = \Pi^j_A(x_t, z_t) - \Pi^j_{NA}(x_t, z_t) = \ln \frac{\overline{E}^j + p^j + P^j_t}{\overline{E}^j + P^j_t} - C^j \tag{1.4}$$

For the sake of simplicity, we assume that the dynamics of x_t and z_t is given by the so-called "replicator dynamics"(see e.g. Weibull, 1995):

$$\begin{cases} \dot{x} = x(1-x)\Delta\Pi^N(x, z) \\ \dot{z} = z(1-z)\Delta\Pi^S(x, z) \end{cases} \tag{1.5}$$

where we omitted the temporal subscript t to improve readability. Dynamics (1.5) describes an adaptive process based on an imitation mechanism: every period t, a (very) small fraction of the population changes its strategy adopting the more remunerative one. Differently from the "classical" contexts where replicator dynamics are introduced (in which economic agents are pairwise randomly matched), here the well-being of each agent depends on the technological choice by *all* agents, in both regions, and at the same instant; that is, we analyse a *population game*. Replicator dynamics may be generated by several learning mechanisms in a random matching context (see e.g. Börgers and Sarin, 1997; Schlag, 1998); however, rationales for such dynamics can be found also in our context (see e.g. Sacco, 1994). Sethi and Somanathan (1996) propose an application of replicator equations in a context similar to ours.

1.1.1 Basic mathematical results

As the shares of agents adopting strategy **A** are defined in the interval $[0, 1]$, the dynamic system (1.5) is defined in the square Q:

$$Q = \{(x, z) : 0 \leq x \leq 1, \quad 0 \leq z \leq 1\}.$$

We will henceforth denote with $Q_{x=0}$ the side of Q along which $x = 0$, and with $Q_{x=1}$ the side along which $x = 1$. Similar interpretations apply to $Q_{z=0}$ and $Q_{z=1}$. All sides of this square are invariant; in other terms, if the pair (x, z) initially lies on one of the sides, then the whole correspondent trajectory also lies on that side.

Note that the states $\{(x, z) = (0, 0), (0, 1), (1, 0), (1, 1)\}$ are always stationary states of the dynamic system (1.5). In such states, only one strategy (either **A** or **NA**) is played in each region. Other stationary states are the points of intersection between the interior of the sides $Q_{x=0}$, $Q_{x=1}$ (where $\dot{x} = 0$) and the locus $\Delta\Pi^S(x, z) = 0$ (where $\dot{z} = 0$) and the points of intersection between the interior of sides $Q_{z=0}$, $Q_{z=1}$ (where $\dot{z} = 0$) and the locus $\Delta\Pi^N(x, z) = 0$ (where $\dot{x} = 0$). In such stationary states, there is a region in which both available strategies are played by a positive share of agents, while in the other region all agents choose the same strategy. Finally, other possible stationary states are those in the interior of Q where the loci $\Delta\Pi^N(x, z) = 0$ and $\Delta\Pi^S(x, z) = 0$ meet; in such states, both strategies are adopted by a positive share of agents in both regions.

Besides the vertices of **Q**, we find that the loci $\dot{x} = 0$ and $\dot{z} = 0$ are respectively represented by the lines:

$$z = \frac{\overline{E}^N}{f^N} - \frac{p^N}{f^N\left(e^{C^N} - 1\right)} - \frac{d^N}{f^N}x \tag{1.6a}$$

$$z = \frac{\overline{E}^S}{d^S} - \frac{p^S}{d^S\left(e^{C^S} - 1\right)} - \frac{f^S}{d^S}x \tag{1.6b}$$

where we recall that $e^{Cj} - 1 > 0$. This is obtained by substituting the public effects (1.2) into the well-being differential (3.15). Note that the slope of (1.6a) is negative if the domestic effect d^N and the foreign effect f^N in N have the same sign, whereas the slope of (1.6b) is negative if the domestic effect d^S and the foreign effect f^S in S have the same sign. Furthermore, the slope of (1.6a) is greater than the slope of (1.6b) if $\frac{d^N}{f^N} < \frac{f^S}{d^S}$. Finally, we note that $\Delta\Pi^N(x, z)$ is positive (i.e. $\dot{x} > 0$) above (1.6a) if $f^N > 0$ (vice versa if $f^N < 0$) and that $\Delta\Pi^S(x, z)$ is positive (i.e. $\dot{z} > 0$) above (1.6b) if $d^S > 0$ (vice versa if $d^S < 0$). Since both (1.6a) and (1.6b) are straight lines, there generally[2] exists at most one stationary state in the interior of each side of Q and at most one in the interior of Q. Consequently, by recalling that all vertices are stationary states, as well, the highest number of stationary states that can be generally observed is nine (four vertices, four points on the sides, and an internal point).

[2]In the unlikely circumstance that lines (1.6a) and (1.6b) have the same slope and the same intercept, the two lines completely overlap and all their points in the interior of **Q** are stationary states.

1.2 Technologies with negative public effects

Let us now outline the possible scenarios the system may reach when the adaptation technology is characterised by: 1) negative public effects towards all agents (maladaptation); 2) positive public effects towards all agents (mitigation). Other relevant cases could be investigated, yet we restrain the analysis to these two cases for the sake of parsimony. In this section we study the first case, in which the adaptation technology is maladaptive, i.e. it is such that it lowers the environmental quality for all. Formally, this maladaptation technology has both a domestic and a foreign negative public effect. One common example of such technology in the literature is air conditioning: it provides the person with an improvement of her environmental quality at the cost of a small deterioration of the environmental quality (and energy security) for all other people (Deschênes and Greenstone, 2011; Lundgren and Kjellstrom, 2013). From an analytical perspective, this translates into all public effect parameters being strictly positive: d^N, d^S, f^S, $f^N > 0$.

1.2.1 Dynamic regimes

First of all, we note that if d^N, d^S, f^S, $f^N > 0$, then both lines (1.6a) and (1.6b), along which $\dot{x} = 0$ and $\dot{z} = 0$, respectively, have negative slope. Above these lines, we have that the share of agents adopting strategy **A** increases. In particular, $\dot{x} > 0$ above line (1.6a) and $\dot{z} > 0$ above line (1.6b), whereas the reverse occurs below these lines. This is very informative with respect to the behaviour of agents: for a higher value of x, z must be lower in order for agents in either region to be indifferent to the maladaptation technology, or else they would all adopt strategy **A**. From another perspective, for a given point (x, z) which lies on either line (1.6a) or (1.6b), a translation to the right would destabilise the system towards full adoption by agents in region N or S (or both), respectively. The adoption process of the maladaptation technology is thus self-reinforcing: the higher is the proportion of agents adopting it in either group, the higher is the incentive for others to do the same. Moreover, we note that lines (1.6a) and (1.6b) move downwards if the autonomous environmental quality for region N or S is lower. For sufficiently low values of $\overline{E}^N$ and $\overline{E}^S$, we have that $\dot{x} > 0$ and $\dot{z} > 0$, respectively, for all points in **Q**. The reverse applies when $\overline{E}^N$ and $\overline{E}^S$ are sufficiently high.

The following proposition characterises the dynamics of the system when d^N, d^S, f^S, $f^N > 0$.

Proposition 1 *Under the assumption that d^N, d^S, f^S, $f^N > 0$, the system* (1.5) *has the following features:*

(a) Every trajectory of the system approaches a stationary state.

(b) Only the vertices of $\boldsymbol{Q}$*, i.e. the stationary states* $(0,0),(0,1),(1,0),(1,1)$*, can be attractive.*

1.2.2 Stability properties of the vertices

In order to assess the stability properties of the vertices of $\mathbf{Q}$, we derive the Jacobian matrix of the system (1.5 evaluated at the stationary state $(x,z)=(i,k)$, $i=0,1$ and $k=0,1$:

$$\begin{pmatrix} (1-2i)\Delta\Pi^N(i,k) & 0 \\ 0 & (1-2k)\Delta\Pi^S(i,k) \end{pmatrix} \tag{1.7}$$

which has the eigenvalues: $(1-2i)\Delta\Pi^N(i,k)$ and $(1-2k)\Delta\Pi^S(i,k)$.

The analysis of the sign of the eigenvalues allows us to illustrate the stability properties of the stationary states $(0,0)$, $(0,1)$, $(1,0)$, $(1,1)$.

Stability of the stationary state $(0,0)$ In this scenario no agent adopts the technology. In order for this non-adoption scenario to be attractive, it must be individually convenient to adopt strategy **NA** in both regions. In order for this to hold, both the eigenvalues in the direction of $Q_{z=0}$ and $Q_{x=0}$ must be negative. This is verified when it holds that:

$$\overline{E}^j > \frac{p^j}{e^{C^j}-1} \qquad \textit{with } j=N,S \tag{1.8}$$

whereas the eigenvalues are strictly positive iff the opposite of (1.8) holds. To the right hand side of this inequality we have the ratio of the positive private effect of the technology over its cost of adoption, which we may interpret as its efficiency in region j. We note that the denominator is strictly positive since $C^j>0$. To the left hand side we have the autonomous environmental quality in j, which also coincides with the overall environmental quality since no agent is adopting strategy **A** $(x=0,z=0)$. Condition (1.8) thus requires that in both regions the efficiency of the technology is lower than the environmental quality.

Stability of the stationary state $(0,1)$ In this case, only agents in S adopt the technology, while no agent does so in N. We now have that the eigenvalue in direction of $Q_{z=1}$ of the Jacobian matrix (1.7), evaluated at $(0,1)$, is strictly negative iff:

$$\overline{E}^N - f^N > \frac{p^N}{e^{C^N}-1} \tag{1.9}$$

whereas it is strictly positive iff the opposite of (1.9) holds. We note that this condition is similar to condition (1.8), but we now have that the autonomous environmental quality is adjusted by the public effect of the agents in S adopting strategy $\mathbf{A}$ (since $z=1$). In other terms, in order for the agents in N to be more convenient not to adopt the technology, its efficiency needs to be lower than the overall environmental quality, which includes the public effects of agents in S. We remark that environmental quality in this case can be either lower or higher than in the non-adoption scenario, since the public effect can be either positive or negative; condition (1.9) can thus be either less or more restrictive than (1.8), respectively. As concerns the eigenvalue in direction of $Q_{x=0}$, it is strictly negative iff:

$$\overline{E}^S - d^S < \frac{p^S}{e^{C^S}-1} \tag{1.10}$$

whereas it is strictly positive iff the opposite of (1.10) holds. In order for agents in S to adopt the technology, its efficiency needs to be greater than the environmental quality, which includes the domestic public effect d^S.

Stability of the stationary state $(1,0)$ This case is specular to the previous one, with all agents in N adopting the technology and no agent adopting it in S. We find that the eigenvalue in direction of $Q_{z=0}$ of the Jacobian matrix (1.7), evaluated at $(1,0)$, is strictly negative iff:

$$\overline{E}^N - d^N < \frac{p^N}{e^{C^N}-1} \tag{1.11}$$

whereas it is strictly positive iff the opposite of (1.11) holds. This condition states that all agents in N adopt strategy $\mathbf{A}$ only if its efficiency is greater than the environmental quality adjusted by the domestic public effect d^N. The eigenvalue in direction of $Q_{x=1}$ is strictly negative iff:

$$\overline{E}^S - f^S > \frac{p^S}{e^{C^S}-1} \tag{1.12}$$

whereas it is strictly positive iff the opposite of (1.12) holds. Agents in S do not adopt strategy $\mathbf{A}$ only if its efficiency is lower than their environmental quality, adjusted by the foreign public effect f^S.

Stability of the stationary state $(1,1)$ Finally, this case represents a full adoption scenario, in which all agents from both regions adopt the technology. We have that the eigenvalues in direction of $Q_{z=1}$ and $Q_{x=1}$ of the Jacobian matrix (1.7), evaluated at $(1,1)$, are strictly negative iff:

$$\overline{E}^j - (d^j + f^j) < \frac{p^j}{e^{C^j} - 1} \qquad \text{with } j = N, S \tag{1.13}$$

whereas they are strictly positive iff the opposite of (1.13) holds. On the left hand side of condition (1.13) we see that now the environmental quality is affected by both domestic and foreign public effects, since all agents are adopting **A**. The condition requires the efficiency of the technology for both regions to be greater than the environmental quality.

Finally, we remark that the vertices of **Q** can be simultaneously attractive, which occurs when the following condition holds:

$$\frac{p^j}{e^{C^j} - 1} + f^j < \overline{E}^j < \frac{p^j}{e^{C^j} - 1} + d^j \qquad \text{with } j = N, S \tag{1.14}$$

We note that in order for condition (1.14) to hold, it is necessary that $f^j < d^j$ for $j = N, S$. By checking their definitions in (1.2), we can see that this implies that foreign public effects must be lower than domestic public effects, in both N and S. If foreign public effects were stronger than domestic ones in at least one region, then the stationary states $(0,1)$ and $(1,0)$ could not be simultaneously attractive. Indeed, it would not be otherwise convenient for an agent not to adopt strategy **A** when all agents in the other region are doing so unless foreign public effects were neglectable with respect to domestic ones.

Some examples of multistability are shown in Figures 1.1–1.5, where attractive stationary states are represented by full dots $\bullet$, repulsive ones by open dots $\circ$, and saddles by squares $\square$. In all cases graphically represented, agents in each region coordinate on one of the two strategies. The most interesting dynamics of this kind is the one represented in Figure 1.1, where condition (1.14) is satisfied. In this case all vertices of **Q** are attractive, whereas all other stationary states along the sides of **Q** are saddle points and the stationary state inside **Q** is a source. As Figure 1.1 shows, almost every trajectory[3] will lead to a vertex of **Q**, where each region ends up choosing a single strategy (either adopting the environmental maladaptation technology

[3] The stable branches of the saddles are exceptions, as they lead the system toward the saddle point, which is stationary.

or not). The basins of attraction of the vertices are delimited by the stable manifolds of the saddle point in the interior of the sides of **Q**.

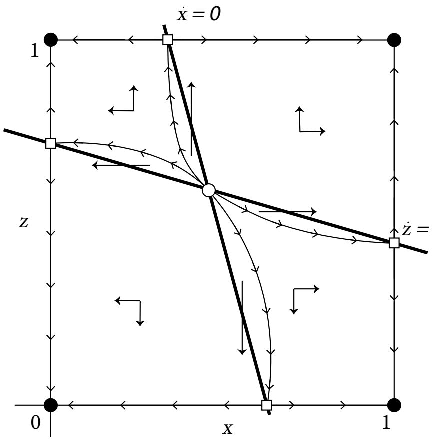


Figure 1.1: All nine stationary states exist: the vertices are attractors, the ones on the sides are saddles and the internal one is a repulsor.

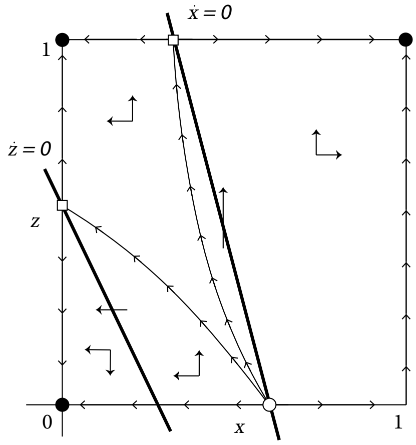


Figure 1.2: In this case, there are three attractors: $(0,0)$, $(0,1)$, $(1,1)$, whereas the other stationary states are either repulsors or saddles.

1.2.3 Well-being analysis

We will now examine the average level of well-being in the two regions when all public effects are negative, i.e. the coefficients are positive: d^N, d^S, f^S, $f^N > 0$. The average level of well-being in N and in S is equal to the weighted average of the well-being of agents adopting strategy **A** and the well-being of agents adopting **NA**, where the weights are given by share of adopters in the region. Formally, we have that:

$$\widetilde{\Pi}^N(x,z) \coloneqq x \cdot \Pi^N_A(x,z) + (1-x) \cdot \Pi^N_{NA}(x,z) \tag{1.15}$$

$$\widetilde{\Pi}^S(x,z) \coloneqq z \cdot \Pi^S_A(x,z) + (1-z) \cdot \Pi^S_{NA}(x,z) \tag{1.16}$$

so that $\widetilde{\Pi}^N(0,z) = \Pi^N_{NA}(0,z)$ represents the average well-being in N when no agent is adopting **A** in this region, whereas $\widetilde{\Pi}^N(1,z) = \Pi^N_A(1,z)$ represents the opposite case. The interpretation is analogous for region S. The following proposition applies:

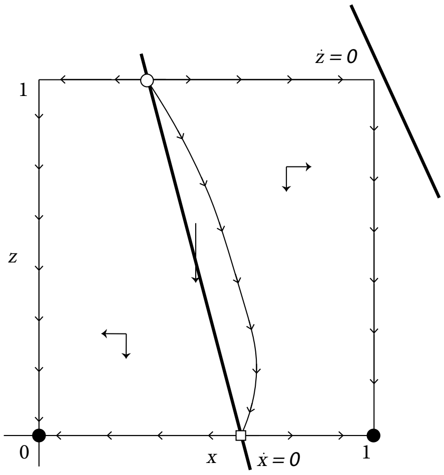


Figure 1.3: In this case only the stationary states $(0,0)$ and $(1,0)$ are attractors. The stationary state in the interior of the top side of $\mathbf{Q}$ is a repulsor whereas the one lying in the interior of the botom side is a repulsor.

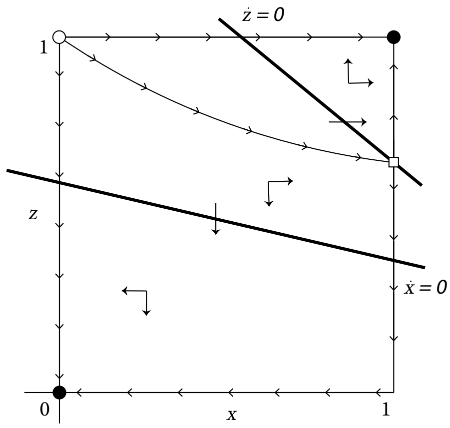


Figure 1.4: There are two attractors, corresponding to the full adoption $(1,1)$ and the non-adoption $(0,0)$ scenarios. There are also a saddle point on the right hand side and a repulsor in the asymmetric state $(0,1)$.

Proposition 2 *If d^N, d^S, f^S, $f^N > 0$, then:*

(a) for agents in N, the non-adoption stationary state $(0,0)$ Pareto-dominates all other stationary states in $\mathbf{Q}$, when they exist, with $0 \leq x < 1$ and $0 \leq z \leq 1$. Equivalently, $\widetilde{\Pi}^N(0,0) > \widetilde{\Pi}^N(x,z)$ for every $(x,z) \neq (0,0)$ with x and z such that $0 \leq x < 1$ and $0 \leq z \leq 1$.

(b) for agents in S, the non-adoption stationary state $(0,0)$ Pareto-dominates all other stationary states in $\mathbf{Q}$, when they exist, with $0 \leq x \leq 1$ and $0 \leq z < 1$. Equivalently, $\widetilde{\Pi}^S(0,0) > \widetilde{\Pi}^S(x,z)$ for every $(x,z) \neq (0,0)$ with x and z such that $0 \leq x \leq 1$ and $0 \leq z < 1$.

(c) for agents in both regions N and S, the non-adaptation stationary state $(0,0)$ Pareto-dominates also the full adoption stationary state $(1,1)$ when the efficiency of strategy $\boldsymbol{A}$, net of domestic public effects, is lower than local autonomous environmental quality. Equivalently, $\widetilde{\Pi}^j(0,0) > \widetilde{\Pi}^j(x,z)$ for every $(x,z) \neq (0,0)$ when $\overline{E}^j > \frac{p^j - d^j}{e^{C^j}-1}$, with $j = N, S$.

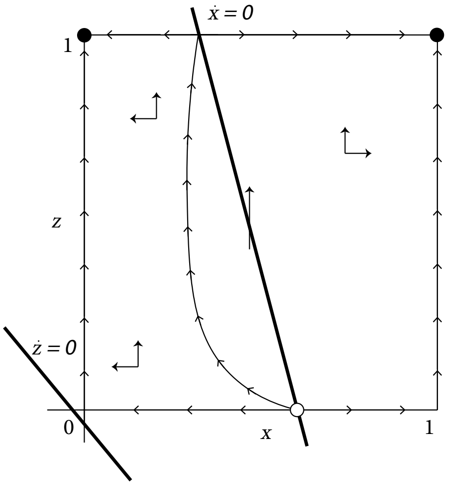


Figure 1.5: In this case, the vertices $(0,1)$ and $(1,1)$ are attractors, whereas a repulsor lies on the interior of the bottom side of **Q**.

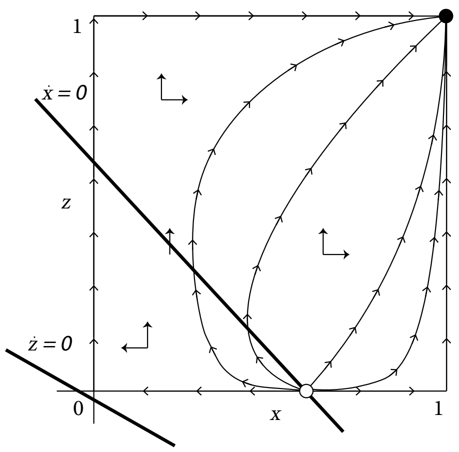


Figure 1.6: There exist a global attractor, corresponding to the full adoption scenario $(1,1)$. There is alsoa repulsor in the interior of the bottom side of **Q**.

By the above proposition, and by virtue of section 1.2.2, it is easy to check that when the stationary state $(0,0)$ is locally attractive, then it Pareto-dominates all others. Furthermore, the non-adoption stationary state $(0,0)$ may Pareto-dominate the stationary state $(1,1)$ (in both regions) even if $(1,1)$ is the only attractive stationary state (see Figure 1.6), provided that $\overline{E}^N$ and $\overline{E}^S$ are sufficiently high. In such case, the adoption of maladaptation technologies in both regions reduces the well-being of agents as system moves from the repulsive non-adoption state $(0,0)$ to the attractive full adoption state $(1,1)$. One could also check that if $(0,0)$ does not Pareto-dominate all other stationary states (in both N and S), then the dynamics (1.5) is trivial, i.e. $\dot{x}$ and $\dot{z}$ are always positive in Q. In such case, the stationary state $(1,1)$ is globally attractive and Pareto-dominates any other possible state (x,z) in N and S.

Remark *From the well-being analysis above, in the context represented in Figure 1.1, every agent, from each region, achieves its highest level of well-being in* $(0,0)$*. Therefore, only one of the four attractive vertices yields the maximum level of well-being. Furthermore, the lowest level of well-being is achieved in* $(1,1)$*, whereas intermediate levels are reached in* $(0,1)$ *and* $(1,0)$.

1.2.4 Environmental dumping

At the centre of debates of both environmental and development economists, environmental dumping is the phenomenon for which an economic activity in an industrialised country results in the disproportionate degradation of the environment of a developing country. Some scholars even argued that policies targeted to improve environmental quality in industrialised countries lead to increased pollution in developing ones. For instance, scholars investigating the Pollution Haven Effect[4] maintain that carbon taxes and stricter environmental regulation are a push factor for firms, which offshore to developing countries with laxer environmental institutions. Opponents of this theory argue that international trade and offshoring incentivise developing countries to raise their environmental standards and thus help tackling the problem of environmental degradation. The analysis of the North American Free Trade Agreement performed by Gallagher (2000) seems to partly support both claims: Mexican firms reduced their emission intensity following the agreement, yet overall emissions increased due to the relatively lower Mexican standards with respect to the US ones. Since CO_2 emissions are a public bad (their negative effects affect the whole world), this increased pollution might have damaged industrialised countries, as well.

We here investigate this hypothesis, for which shifting environmental burden from one country to the other might worsen the well-being of all agents. More precisely, our model allows to study the adoption dynamics of an environmental maladaptation technology with negative public effects and which asymmetrically degrades the environmental quality indicator of one of the two regions. We here discuss what happens when an exogenous factor (e.g. a new policy) raises the foreign effect f^S in S, whereas it decreases domestic effect d^N in N. This is the case of a green tax or policy in the industrialised region which decreases the domestic effect on the local resource but increases the foreign effect on the resource of the other region, further degrading it. By means of a simple comparative dynamics analysis, we note that a smaller value of d^N improves the environmental quality in N and decreases the well-being differential of adopters of the maladaptation technology. Since the foreign effect f^S on the local environmental indicator of S is greater, the environmental degradation of agents in S is greater, and the well-being differential of the adopters increases and leads more agents to adopt the maladaptation technology. The overall well-being effects for agents in N cannot be assessed a priori. If the reduction in the domestic effect d^N

[4] See Copeland and Taylor (2004) for a definition of the concept and its differences with the slightly similar Pollution Haven Hypothesis.

is sufficiently large, it might counterbalance the additional degradation deriving from more adopters in the S region, who emit the foreign effect f^N affecting the environmental quality in region N. Vice versa, if the domestic effect is weaker with respect to the increased adoption induced in the foreign region, then the well-being of N decreases as a consequence of the exogenous change.

A graphical illustration is provided by Figures 1.1 and 1.5. In the former figure, we recall that the non adoption state $(0,0)$ is Pareto-dominant. However, a change in the value of f^S may cause the stationary states $(0,0)$ and $(1,0)$ to become unstable (see section 1.2.2), giving rise to the dynamic regime represented in Figure 1.5. In this case, the Pareto-dominant state $(0,0)$ would no more be an attractor, and the system would lose its social optimum. By contrast, the Pareto-dominated state $(1,1)$ would still be attractive. This analysis highlighted that environmental policies in an industrialised region may have either a positive or a negative effect for its agents, depending on the feedback effects of agents in the developing region.

1.3 Technologies with positive public effects

We now study the case in which all public effects of the environmental adaptation technology are positive, that is: d^N, d^S, f^S, $f^N < 0$. This case describes the adoption dynamics of a mitigation technology, which thus improves environmental quality for all agents (see Gupta and Gregg, 2012; Hallegatte, 2009, for instances of adaptation technologies with mitigation features). If we think of the agents as firms, instances of such technologies might be the installation of a water treatment plant on a common water basin or, equivalently, of a technology which reduces emissions or waste water usage. Other examples might draw from businesses dealing with the management of common environmental resources, such as fisheries or forestries (Olson, 1965).

1.3.1 Dynamic regimes

We first note that if d^N, d^S, f^S, $f^N < 0$, both the straight lines (1.6a) (where $\dot{x} = 0$) and (1.6b) (where $\dot{z} = 0$) have negative slope. Differently from the case with negative public effects, in this case $\dot{x} > 0$ below line (1.6a), whereas $\dot{x} < 0$ above it. Analogously, $\dot{z} > 0$ below line (1.6b), whereas $\dot{z} < 0$ above it. In contrast to the previous case, now the adoption dynamics is not self-reinforcing: more specifically, the incentive to adopt the environmental mitigation technology decreases if the share of agents adopting the technology in either group increases. This is the well known free riding problem, for which agents are not willing to contribute to a public good and would rather benefit from the contributions of others without paying the cost of their own contribution. In addition, the concavity of the well-being function with respect to the environmental quality accentuates the effect, as it makes any further improvement of the environment less desirable. Since the returns from the mitigation technology decrease with the share of adopters while the cost is constant, we may see why this context favours coexistence between strategies **A** and **NA**. Indeed, as more and more agents adopt the mitigation technology, the well-being differential of such strategy falls to zero, allowing for a stationary state in which in the same region there are agents adopting strategy **A** and agents adopting **NA**. Moreover, we remark that if the autonomous environmental quality $\overline{E}$ is sufficiently high in N and S, then the well-being differential is always negative, i.e. $\dot{x} < 0$ and $\dot{z} < 0$, leading agents to drop the mitigation technology and shift from **A** to **NA**. In this case, the autonomous level of environmental quality is so high that no agent finds it convenient to increase it further by an amount equal to the private effect p^j, with $j = 0, 1$. This might also be due to the inefficiency of the mitigation technology (a low value of $\frac{p^j}{e^{C^j}-1}$). In formal terms, we may say that lines (1.6a) and (1.6b) move downwards if the autonomous environmental quality for region N or S is higher. For sufficiently high values of $\overline{E}^N$ and $\overline{E}^S$ or for sufficiently low values of p^N and p^S, we have that $\dot{x} < 0$ and $\dot{z} < 0$, respectively. The reverse applies when $\overline{E}^N$ and $\overline{E}^S$ are sufficiently low or p^N and p^S sufficiently high.

We find that the following proposition characterises the adoption dynamics when: d^N, d^S, f^S, $f^N < 0$.

Proposition 3 *Under the assumption that d^N, d^S, f^S, $f^N < 0$, the system 1.5 has the following features:*

(a) Every trajectory of the system approaches a stationary state.

(b) When the stationary state $(0, 0)$ is attractive (see section 1.2.2), then it is globally attractive, i.e. there is no other attractive stationary state (see Figure 1.7).

(c) When the stationary state $(1, 1)$ is attractive (see section 1.2.2), then it is globally attractive (see Figure 1.8).

(d) If there is no stationary state in the interior of $\boldsymbol{Q}$, then there exists only one attractive stationary state in the boundary of $\boldsymbol{Q}$; it may either be one of the vertices or lie on the interior of the edges of **Q**.

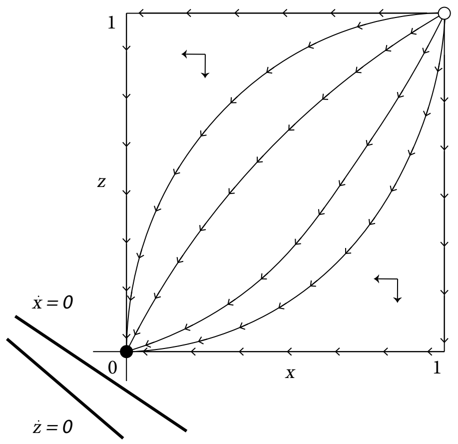


Figure 1.7: The non-adoption stationary state $(0,0)$ is globally attractive, whereas the full adoption one $(1,1)$ is repulsive.

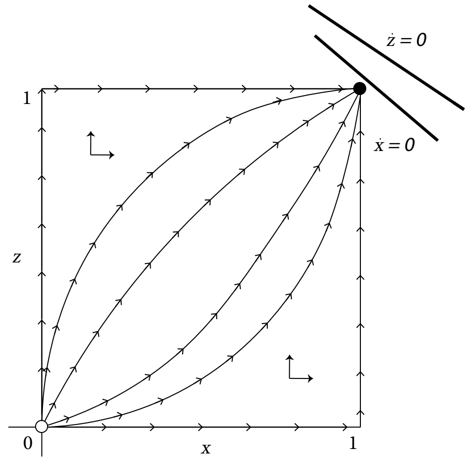


Figure 1.8: The full adoption stationary state $(1,1)$ is globally attractive, whereas the non-adoption one $(0,0)$ is repulsive.

(e) *If $d^N d^S - f^N f^S > 0$, i.e. the domestic effects are larger than the foreign ones, the stationary state in the interior of $\boldsymbol{Q}$ (in which both strategies are played in both regions) is globally attractive, when it exists (see Figure 1.9).*

(f) *If $d^N d^S - f^N f^S < 0$, i.e. the domestic effects are smaller than the foreign ones, the stationary state in the interior of $\boldsymbol{Q}$ is a saddle point, when it exists. In addition, there exist two attractive stationary states lying in the edges of $\boldsymbol{Q}$: they may be the vertices $(0,1)$ and $(1,0)$ or lie in the interior of the edges $\boldsymbol{Q}_{h,k}$ (see Figures 1.10–1.13).*

(g) *If $p^N = p^S = 0$ (i.e. the private effect of strategy $\boldsymbol{A}$ is 0 in both regions), then non-adoption is individually convenient for all agents: $\Pi^N_{NA} > \Pi^N_A$ and $\Pi^S_{NA} > \Pi^S_A$, whatever the values of x and z are. This implies that $\dot{x} < 0$ and $\dot{z} < 0$ always hold and consequently $(0,0)$ is globally attractive (the classical free-riding problem arises for public goods provision).*

Remark *The coordinates of the internal stationary state are:*

$$\overline{x} = \frac{d^S\left(\overline{E}^N - \frac{p^N}{e^{C^N}-1}\right) - f^N\left(\overline{E}^S - \frac{p^S}{e^{C^S}-1}\right)}{d^N d^S - f^S f^N} \tag{1.17a}$$

$$\overline{z} = \frac{d^N\left(\overline{E}^S - \frac{p^S}{e^{C^S}-1}\right) - f^S\left(\overline{E}^N - \frac{p^N}{e^{C^N}-1}\right)}{d^N d^S - f^S f^N} \tag{1.17b}$$

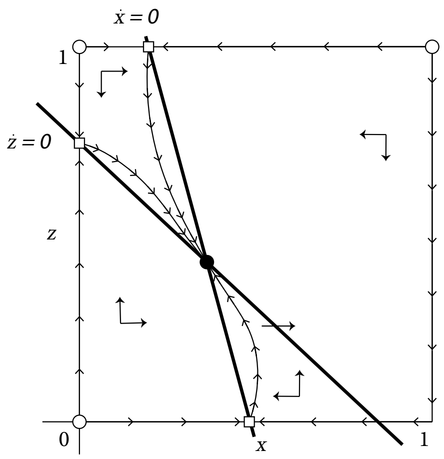


Figure 1.9: The internal steady state is an attractor. There are also three saddles on the sides and three repulsors on the vertices $(0,0)$, $(1,0)$, $(1,1)$.

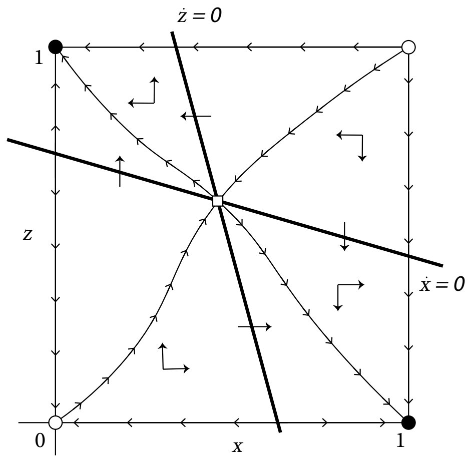


Figure 1.10: The internal point is a saddle, whereas the asymmetric states $(0,1)$ and $(1,0)$ are attractors. The non-adoption state $(0,0)$ and the full adoption one $(1,1)$ are repulsors.

Thus, if $d^N/f^N > f^S/d^S$, i.e. the stationary state is attractive, the internal stationary state exists if and only if:

$$0 < d^S\left(\overline{E}^N - \frac{p^N}{e^{C^N}-1}\right) - f^N\left(\overline{E}^S - \frac{p^S}{e^{C^S}-1}\right) < d^N d^S - f^S f^N$$

$$0 < d^N\left(\overline{E}^S - \frac{p^S}{e^{C^S}-1}\right) - f^S\left(\overline{E}^N - \frac{p^N}{e^{C^N}-1}\right) < d^N d^S - f^S f^N$$

which can be rewritten as:

$$f^N\left(\overline{E}^S - \frac{p^S}{e^{C^S}-1}\right) < d^S\left(\overline{E}^N - \frac{p^N}{e^{C^N}-1}\right) < d^N d^S - f^S f^N$$

$$f^S\left(\overline{E}^N - \frac{p^N}{e^{C^N}-1}\right) < d^N\left(\overline{E}^S - \frac{p^S}{e^{C^S}-1}\right) < d^N d^S - f^S f^N$$

These conditions require both numerator and denominator of coordinates (1.17) to be positive, with the former being greater than the latter. The condition that the numerators of (1.17) be lower than the related denominators restricts $\bar{x}$ and $\bar{z}$ to be lower than 1, thus making the point $(\bar{x}, \bar{z})$ belong to the interior of $\mathbf{Q}$.

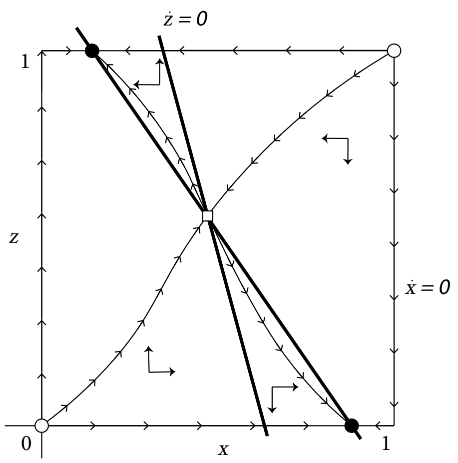


Figure 1.11: The internal fixed point is a saddle and both the non-adoption $(0, 0)$ and the full adoption $(1, 1)$ states are repulsors. Two attractors lie on the interiors of the bottom side and on the top side of **Q**.

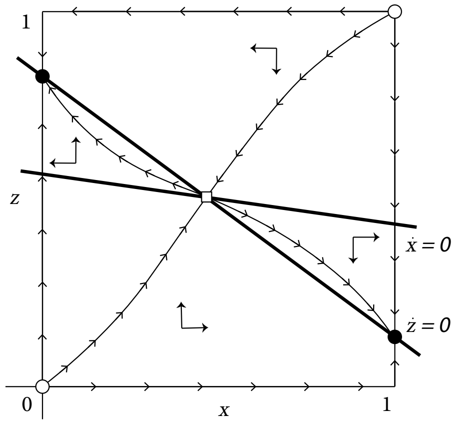


Figure 1.12: The internal fixed point is a saddle and both the non-adoption $(0, 0)$ and the full adoption $(1, 1)$ states are repulsors. Two attractors lie on the interiors of side to the left and on the side to the right of **Q**.

1.3.2 Well-being in the context with positive externalities

We now examine the average level of well-being in the two regions when all public effects are positive: d^N, d^S, f^S, $f^N < 0$ (see (2.24) and (2.25) in the previous section for a comparison). The following proposition applies.

Proposition 4 *Assume d^N, d^S, f^S, $f^N < 0$. In such context, it holds:*

(a) The stationary state $(0, 0)$ is Pareto-dominated (in both regions) by any attractive stationary state with $x > 0$ and/or $z > 0$. When $(0, 0)$ is attractive[5]*, it may be Pareto-dominated by other stationary states* [6]*.*

*(b) The stationary state $(1, 1)$ Pareto dominates (in both regions) any other stationary state when it is attractive (remember that, in such case, no other stationary state can be attractive). Furthermore, even if it is unstable, it Pareto dominates the stationary state in the interior of **Q**, when it exists.*

[5]As stated in **Proposition 3**, point (b), in this case $(0, 0)$ is globally attractive.

[6]This occurs, for instance, when $\frac{p^N}{e^{C^N}-1} < \overline{E}^N < \frac{p^N-d^N-f^N}{e^{C^N}-1}$ and $\frac{p^S}{e^{C^S}-1} < \overline{E}^S < \frac{p^S-d^S-f^S}{e^{C^S}-1}$ hold. Indeed, in this case $(0, 0)$ is attractive but is Pareto-dominated by $(1, 1)$.

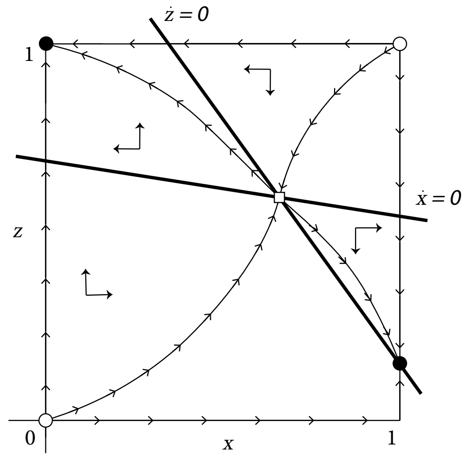


Figure 1.13: In this case, one attractor lies on the asymmetric state $(0, 1)$ and another lies on the interior of the right-hand of square $\mathbf{Q}$. The internal point is a saddle and both the non-adoption $(0, 0)$ and the full adoption $(1, 1)$ states are repulsors.

Remark *From the well-being analysis above, in the context in which the stationary state* $(\overline{x}, \overline{z})$ *in the interior of* $\mathbf{Q}$ *is attractive, we have that* $(\overline{x}, \overline{z})$ *Pareto-dominates* $(0, 0)$ *but is Pareto-dominated by* $(1, 1)$*; however, the latter stationary state cannot be reached because it is not attractive.*

These results are reversed with respect to the case with negative public effects. Indeed, in the previous case $(0, 0)$ Pareto-dominates all stationary states in most cases, although it is not attractive. The selfish nature of the maladaptation technology leads agents towards adoption, although it results in a lower level of well-being for all. The technology is thus over-adopted with respect to the Pareto-optimum. With positive public effects, we have that $(0, 0)$ is Pareto-dominated by all other stationary states whereas $(1, 1)$ Pareto-dominates them when it is attractive. All agents benefit from the mitigation technology adopted by others, but they are less willing to pay its cost as they do not internalise the well-being of other. In this case, the technology is under-adopted, as the full adoption scenario would be the Pareto optimum. This last result is in line with the results by Shogren and Crocker (1991).

1.4 Discussion and conclusions

In this work, we excluded altruistic consideration on the part of agents towards either same-region and cross-region agents. In other terms, we assumed that the actions of agents are only driven by self-interest considerations. We then studied the case of two regions whose agents may adopt an environmental adaptation technology which yields a private benefit to the adopter, while also transferring a negative or positive externality both to agents in the same region and to agents in the other one. We defined same-region externalities as domestic public effects and cross-region externalities as foreign public effects. The model here proposed is very broad, so that a complete analysis of all possible specifications is beyond the scope of this chapter. Instead, we focused on two salient characterisations. On the one hand, we analysed the case of a maladaptation technology, whose domestic and foreign public effects are both negative. In this case, an adopter shifts the environmental load to agents from both regions. A common example of this kind of technologies is air conditioning (Lundgren and Kjellstrom, 2013). On the other hand, we analysed the case of a mitigation technology, whose domestic and foreign public effects are both positive. In this case, each adopter is improving the well-being of agents from both regions. In analogy with the previous example, we may think of home insulation, as it allows each household to reduce both heating and air conditioning, benefiting the environment on a global scale. Our results show that for the maladaptation technology the social optimum is represented by the non-adoption scenario, unless the efficiency of the technology is extremely high (greater than the autonomous level of environmental quality). However, Pareto-dominated states may be reached, because agents do not internalise the externalities of the technology. In this case, we talk of over-adoption of the maladaptation technology. The reverse occurs with a mitigation technology, which would have a full adoption scenario as its Pareto-optimum. However, an intermediate state (in which only some agents are adopters) is *typically* reached, since the returns on adoption decrease for each additional adopter. Also in this case, agents do not take into account the (positive) externalities of adoption on other agents, this time leading to under-adoption of the technology with respect to the Pareto-dominant state. Finally, under the hypothesis of a maladaptation technology, with negative public effects, we analysed the effects of an environmental dumping strategy. This represents a stronger characterisation of maladaptation, as it requires that in the more developed region both the domestic and the foreign public effect are relatively low (null, in the extreme case) with respect to the public effects in the less developed region. Although it is intuitive that the agents from the less developed region would be worse off in this case, the implications for agents from the more developed region are not straightforward. Indeed, according to relative magnitude of foreign effects in the two regions, the well-being of agents from the more developed region could either increase or decrease.

This last result is particularly interesting, although its plausibility should be verified by further research. Indeed, instances of such negative feedbacks could provide greater insight on the cost-benefit analysis of many maladaptation strategies available to the more developed regions. In addition, further research should try to map the specifications which are not illustrated in this work. Interesting dynamics could arise, for example, if the public effects had different signs according to whether they are domestic or foreign. In particular, a case in which all domestic public effects are null or positive, while all foreign effects are negative would depict a situation in which all adopters shift the environmental burden to foreign agents, although they increase the well-being of same-region individuals. In this case, it is not intuitive which state the system would reach. Another relevant case would be represented by technological differences between the two regions allowing the agents from the more developed region to adopt a mitigation technology, whereas agents in the less developed region could only adopt a maladaptation technology. Well-being analysis could highlight which region is relatively more affected by the negative externalities and which state is more likely to be reached. All similar research directions, focusing on translating real phenomena and dynamics into the model, would provide a fine extension to this work and a contribution to the understanding of the relationship between regions and countries at different stages of development and their environmental quality.

Appendix A

Proofs of the propositions in text

Proof of Proposition 1 The proof of point (b) is straightforward and follows immediately from the local stability analysis (which can be found in the mathematical appendix). To prove point (a) we have to show that limit cycles cannot exist (see e.g. Lefschetz, 1963, pp. 230 ff). This is obviously the case when the internal stationary state $(\overline{x}, \overline{z})$, with $0 < \overline{x}, \overline{z} < 1$, does not exist or, if it does, is a saddle point. If $(\overline{x}, \overline{z})$ is a source, then $d^N d^S - f^S f^N > 0$ (see (B.5) in the appendix), that is the straight line (1.6a) (where $\dot{x} = 0$) crosses from above the straight line (1.6b) (where $\dot{z} = 0$). In such case, it is easy to see that the regions in Q where $\dot{x}$ and $\dot{z}$ have the same sign are positively invariant, so that no oscillatory behaviour of trajectories can occur. This implies, by the Poincaré-Bendixson Theorem, that any trajectory starting in Q approaches a stationary state. This concludes the proof of the proposition.

Proof of Proposition 2 To prove point (1) of proposition 2, we have to show that the average payoff in N, evaluated at $(0,0)$, is higher than at any point $(\bar{x}, \bar{z})$ along the line $\Delta\Pi^N(x,z) = 0$ (where $\dot{x} = 0$) and along the side $Q_{x=0}$. The average level of well-being in $(0,0)$ is:

$$\widetilde{\Pi}^N(0,0) = \Pi^N_{NA}(0,0) = \ln \overline{E}^N$$

Let us now take a point $(\bar{x}, \bar{z}) \in \left\{\Delta\Pi^N(x,z) = 0\right\}$. We have that both strategies yield the same level of well-being: $\Pi^N_A(\bar{x}, \bar{z}) = \Pi^N_{NA}(\bar{x}, \bar{z})$, which implies:

$$\widetilde{\Pi}^N(\bar{x},\bar{z}) = \Pi^N_{NA}(\bar{x},\bar{z}) = \ln\left(\overline{E}^N - d^N \cdot \bar{x} - f^N \cdot \bar{z}\right)$$

Therefore, if $\bar{x}$ and/or $\bar{z} > 0$, it follows that: $\widetilde{\Pi}^N(0,0) > \widetilde{\Pi}^N(\bar{x},\bar{z})$. This means that the average well-being in the non-adoption state $(0,0)$ is higher than in any stationary state in the interior of Q and in any stationary state in the interior of the sides $Q_{z=h}$ $(h = 1,2)$. Furthermore, it is easy to check that $(0,0)$ always Pareto-dominates any stationary state with $z > 0$ in the side $Q_{x=0}$. In order to prove point (c), we now show that $(0,0)$ Pareto-dominates any stationary state in the side $Q_{x=1}$ if $\overline{E}^N > \frac{p^N - d^N}{e^{C^N}-1}$. It can be easily verified that $(1,0)$ always Pareto-dominates any other stationary state in the side $Q_{x=1}$. Therefore, we simply have to compare well-being in $(0,0)$ with the one in $(1,0)$. By very simple computations, we obtain that, if $\overline{E}^N > \frac{p^N - d^N}{e^{C^N}-1}$, then $(0,0)$ Pareto-dominates $(1,0)$. With similar arguments, it is easy to check that $(1,1)$ is Pareto-dominated by all the other stationary states when $\overline{E}^N > \frac{p^N - d^N}{e^{C^N}-1}$. To prove that analogous results hold for the well-being of region S, it suffices to apply the same arguments.

Proof of Proposition 3 The proof of point (b) is straightforward and follows immediately from graphical analysis: if $(0,0)$ is attractive, then it must lie above the straight lines (1.6a) and (1.6b). Consequently, in the interior of Q, it holds $\dot{x} < 0$ and $\dot{z} < 0$, which implies the global attractiveness of $(0,0)$. With similar arguments, point (c) can be proved. In order to prove point (e), it suffices to check that when $d^N/f^N > f^S/d^S$, the internal stationary state is locally attractive (see **Proposition 6**). Graphical analysis then allows to see that no other attractive stationary state can exist. It remains to show that limit cycles cannot exist. To do so, we note that the straight line (1.6a), along which $\dot{x} = 0$, crosses the straight line (1.6b), along which $\dot{z} = 0$, from above. In such case, the regions of **Q** where $\dot{x}$ and $\dot{z}$ have opposite signs are positively invariant; this implies that no oscillatory behaviour of trajectories may occur and consequently that the internal stationary state is globally attractive by the Poincaré-Bendixson Theorem. We now prove point (f): if $d^N/f^N < f^S/d^S$, the internal stationary state is a saddle point (see section 1.2.2); consequently, no limit cycle may exist. Furthermore, we note that the straight line (1.6a) crosses the straight line (1.6b) from below. In such case, the regions of Q where $\dot{x}$ and $\dot{z}$ have opposite sign are positively invariant and, in each of these regions, the trajectories approach a stationary state lying on the boundary of Q. Finally, the proof of points (a), (d) and (g) is straightforward.

Proof of Proposition 4 To prove point (a) of the proposition, we first consider the average well-being in N, which in $(0,0)$ is equal to:

$$\widetilde{\Pi}^N(0,0) = \Pi^N_{NA}(0,0) = \ln \overline{E}^N$$

Let us now consider a point $(\bar{x}, \bar{z}) \in \mathbf{Q}$. If $(\bar{x}, \bar{z})$ is a stationary state belonging to the curve $\Delta\Pi^N(x,z) = 0$, then it holds that $\Pi^N_A(\bar{x}, \bar{z}) = \Pi^N_{NA}(\bar{x}, \bar{z})$, and consequently we have:

$$\widetilde{\Pi}^N(\bar{x}, \bar{z}) = \Pi^N_{NA}(\bar{x}, \bar{z}) = \ln\left(\overline{E}^N - d^N \cdot \bar{x} - f^N \cdot \bar{z}\right)$$

Therefore, since d^N, d^S, f^S, $f^N < 0$, if either $\bar{x}$ or $\bar{z} > 0$, we have that: $\widetilde{\Pi}^N(0,0) < \widetilde{\Pi}^N(\bar{x}, \bar{z})$. Thus, average payoff in $(0,0)$ is lower than in any stationary state in the interior of Q and in any stationary state in the interior of the sides $Q_{z=k}$ $(k = 1,2)$. Furthermore, it is easy to check that $(0,0)$ is always Pareto-dominated by any stationary state in the side $Q_{x=0}$. It remains to prove that $(0,0)$ is Pareto-dominated by any attractive stationary state in the side $Q_{x=1}$. Easy algebraic manipulations show that $\widetilde{\Pi}^N(0,0) < \widetilde{\Pi}^N(1,1)$ if and only if $\overline{E}^N < \frac{p^N - d^N - f^N}{e^{C^N} - 1}$. The latter condition is always satisfied if $(1,1)$ is attractive (see section 1.2.2). In the same way, it can be checked that $\widetilde{\Pi}^N(0,0) < \widetilde{\Pi}^N(1,0)$ when $(1,0)$ is attractive. Finally, it is left to prove that $(0,0)$ is Pareto-dominated by any attractive stationary state $(1, \bar{z})$ lying in the interior of $Q_{x=1}$. As already seen above, the well-being in $(0,0)$ is lower than in any stationary state, so that $\widetilde{\Pi}^N(0,0) = \Pi^N_{NA}(0,0) < \Pi^N_{NA}(1, \bar{z})$. Furthermore, we note that if $(1, \bar{z})$ is attractive, then the curve $\Delta\Pi^N(x,z) = 0$ must lie on the right of it (see **Proposition 5**); consequently, on the left of $\Delta\Pi^N(x,z) = 0$, it holds that $\Delta\Pi^N(x,z) > 0$. This implies that $\Pi^N_{NA}(1, \bar{z}) < \Pi^N_A(1, \bar{z})$. Therefore, $\widetilde{\Pi}^N(0,0) < \widetilde{\Pi}^N(1, \bar{z})$, being $\widetilde{\Pi}^N(1, \bar{z}) = \Pi^N_A(1, \bar{z})$. The corresponding results for S can be proved following the same steps. To check the remaining part of point (a), we simply have to solve the inequality $\widetilde{\Pi}^N(0,0) < \widetilde{\Pi}^N(1,1)$ and draw from the stability results in section 1.2.2 about the stationary state $(0,0)$. The proof of point (b) follows very similar steps.

Appendix B

Stability properties of the stationary states

We here study the stability of the stationary states beyond the vertices of the region **Q**, in order to understand toward which the system may converge. Indeed, the attractive states are of particular interest, as they are the only states that can actually be reached by the system. We recall that the condition for a stationary state to be attractive is that both the eigenvalues of the Jacobian matrix evaluated on it are negative [1].

B.1 Stability properties of the stationary states in the interior of the edges of Q

The following proposition concerns the stability properties of the stationary states belonging to the interior of the edges of the square Q, i.e. those where both adoption choices coexist in N while all agents in S play the same strategy or vice versa.

Proposition 5 *The Jacobian matrix of the system* (1.5) *evaluated at the stationary states in the interior of the edges* $Q_{x=h}$ *(*$h = 0,\, 1$*) is:*

$$\begin{pmatrix} (1-2h)\Delta\Pi^N(h,\overline{z}) & 0 \\ \overline{z}(1-\overline{z})\frac{\partial\Delta\Pi^S(h,\overline{z})}{\partial x} & \overline{z}(1-\overline{z})\frac{\partial\Delta\Pi^S(h,\overline{z})}{\partial z} \end{pmatrix} \tag{B.1}$$

[1] If the eigenvalues are both positive, then the state is repulsive and cannot be reached by system (unless it coincides with its initial condition). If they have opposite signs, instead, the state is a saddle and can only be reached if the initial condition of the system lies on its stable branch.

where $\overline{z}$ *is the value of* z *at the stationary state, and has the eigenvalues:* $\overline{z}(1-\overline{z})\frac{\partial \Delta\Pi^S(h,\overline{z})}{\partial z}$ *(in direction of* $Q_{x=h}$*) and* $(1-2h)\Delta\Pi^N(h,\overline{z})$ *(in direction of the interior of* Q*). The Jacobian matrix of the system (1.5) evaluated at the stationary states in the interior of the edges* $Q_{z=h}$ *(*$h = 0,\, 1$*) is:*

$$\begin{pmatrix} \overline{x}(1-\overline{x})\frac{\partial \Delta\Pi^N(\overline{x},h)}{\partial x} & \overline{x}(1-\overline{x})\frac{\partial \Delta\Pi^N(\overline{x},h)}{\partial z} \\ 0 & (1-2h)\Delta\Pi^S(\overline{x},h) \end{pmatrix} \tag{B.2}$$

where $\overline{x}$ *is the value of* x *at the stationary state, and has the eigenvalues:* $\overline{x}(1-\overline{x})\frac{\partial \Delta\Pi^N(\overline{x},h)}{\partial x}$ *(in direction of* $Q_{z=h}$*) and* $(1-2h)\Delta\Pi^S(\overline{x},h)$ *(in direction of the interior of* Q*).*

Proof. Straightforward. ■

We remark that, given a stationary state in an edge $Q_{h=k}$, $h = x, z$ and $k = 0, 1$, the sign of its eigenvalue in direction of $Q_{h=k}$ is negative if and only if the stationary states at the extrema of $Q_{h=k}$ which are the vertices of $\mathbf{Q}$, have positive eigenvalues in direction of $Q_{h=k}$.

The conditions for the attractiveness of the steady states within the edges of $\mathbf{Q}$ deserve further comment. Indeed, the attractiveness conditions (B.1) and (B.2) require that:

$$\frac{\partial \Delta\Pi^N(\overline{x},i)}{\partial x} < 0 \tag{B.3a}$$

$$\frac{\partial \Delta\Pi^S(i,\overline{z})}{\partial z} < 0 \tag{B.3b}$$

Inequalities (B.3) describe a nonlinear dynamics of strategies **A** and **NA** in N and S, respectively, similar to the "elitist"narratives in Antoci et al. (2018). Since the well-being differential of adopting strategy **A** decreases with the share of adopters, strategy **A** yields the highest payoffs when only a minority of agents adopts it. As strategy **A** diffuses, so the incentive to adopt it decreases, to the point that agents become indifferent toward the technology. Intuitively, the presence of this dynamics in (only) one of the two regions is necessary in order to have coexistence of strategies in such region and a pure population strategy in the other region.

B.2 Stability properties of stationary states in the interior of Q

The following proposition deals with the stability of stationary states in the interior of the square Q, in which a positive share of agents adopts each strategy in both regions.

Proposition 6 *The Jacobian matrix of the system (1.5) evaluated at a stationary state* $(\overline{x}, \overline{z})$ *in the interior of* Q *(i.e.* $0 < \overline{x}$, $\overline{z} < 1$*) is:*

$$\begin{pmatrix} \overline{x}(1-\overline{x})\frac{\partial \Delta\Pi^N(\overline{x},\overline{z})}{\partial x} & \overline{x}(1-\overline{x})\frac{\partial \Delta\Pi^N(\overline{x},\overline{z})}{\partial z} \\ \overline{z}(1-\overline{z})\frac{\partial \Delta\Pi^S(\overline{x},\overline{z})}{\partial x} & \overline{z}(1-\overline{z})\frac{\partial \Delta\Pi^S(\overline{x},\overline{z})}{\partial z} \end{pmatrix} \tag{B.4}$$

where the sign of the determinant of (B.4) is equal to the sign of the expression:

$$d^N d^S - f^S f^N \tag{B.5}$$

and the trace of (B.4) is equal to:

$$d^N(e^{C^N} - 1)\overline{x}(1-\overline{x}) + d^S(e^{C^S} - 1)\overline{z}(1-\overline{z}) \tag{B.6}$$

Proof. Straightforward. ∎

According to the above proposition, we have that if expression (B.5) is strictly negative, then the internal stationary state is a saddle (i.e. it is unstable). If it is positive, then the stationary state may be a source (i.e. a repulsor) or a sink (i.e. an attractor). In the context in which expression (B.5) is strictly positive, the condition d^N, $d^S > 0$ (< 0) is a sufficient condition for the repulsiveness (attractiveness) of the internal stationary state. Moreover, if the determinant is negative, then the stationary state is a saddle, whereas it is attractive when the determinant is positive and the trace is negative. The trace is given by the sum:

$$\overline{x}(1-\overline{x})\frac{\partial \Delta\Pi^N(\overline{x},\overline{z})}{\partial x} + \overline{z}(1-\overline{z})\frac{\partial \Delta\Pi^S(\overline{x},\overline{z})}{\partial z}$$

This means that in order for the trace to be negative, at least one of the partial derivatives above must be negative, meaning that in the corresponding region strategy **A** has an elitist dynamics as previously defined.

Bibliography

Antoci, Angelo, Nicolò Bellanca, and Giulio Galdi (2018) "At the relational crossroads: Narrative Selection, Contamination, Biodiversity in Trans-Local Contexts," *Journal of Economic Behavior & Organization*, Vol. 150, pp. 98–113.

Barnett, Jon and Saffron O'Neill (2010) "Maladaptation," *Global Environmental Change*, Vol. 2, pp. 211–213.

Bird, Peter JWN (1987) "The transferability and depletability of externalities," *Journal of Environmental Economics and Management*, Vol. 14, pp. 54–57.

Börgers, Tilman and Rajiv Sarin (1997) "Learning through reinforcement and replicator dynamics," *Journal of Economic Theory*, Vol. 77, pp. 1–14.

Copeland, Brian R and M Scott Taylor (2004) "Trade, growth, and the environment," *Journal of Economic Literature*, Vol. 42, pp. 7–71.

Deschênes, Olivier and Michael Greenstone (2011) "Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the US," *American Economic Journal: Applied Economics*, Vol. 3, pp. 152–85.

Gallagher, Kevin (2000) *Trade Liberalization and Industrial Pollution in Mexico: Lessons of the FTAA*, Vol. 00-07: Tufts University, Boston, Massachussetts, USA.

Geaun, Jerome C (1993) "On the shiftable externalities," *Journal of Environmental Economics and Management*, Vol. 25, pp. 30–44.

Glotzbach, Stefanie and Stefan Baumgartner (2012) "The relationship between intragenerational and intergenerational ecological justice," *Environmental Values*, Vol. 21, pp. 331–355.

Gupta, Rajat and Matthew Gregg (2012) "Using UK climate change projections to adapt existing English homes for a warming climate," *Building and Environment*, Vol. 55, pp. 20–42.

Hallegatte, Stéphane (2009) "Strategies to adapt to an uncertain climate change," *Global Environmental Change*, Vol. 19, pp. 240–247.

Hasson, Reviva, Åsa Löfgren, and Martine Visser (2010) "Climate change in a public goods game: investment decision in mitigation versus adaptation," *Ecological Economics*, Vol. 70, pp. 331–338.

IPCC (2014) *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*: Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Lefschetz, Solomon (1963) *Differential equations: Geometric theory*, Vol. 6: Interscience publishers, New York, NY, USA.

Lundgren, Karin and Tord Kjellstrom (2013) "Sustainability challenges from climate change and air conditioning use in urban areas," *Sustainability*, Vol. 5, pp. 3116–3128.

Milinski, Manfred, Dirk Semmann, Hans-Jürgen Krambeck, and Jochem Marotzke (2006) "Stabilizing the Earth's climate is not a losing game: Supporting evidence from public goods experiments," *Proceedings of the National Academy of Sciences*, Vol. 103, pp. 3994–3998.

Milly, Paul Christopher D, Julio Betancourt, Malin Falkenmark, Robert M Hirsch, Zbigniew W Kundzewicz, Dennis P Lettenmaier, and Ronald J Stouffer (2008) "Stationarity is dead: Whither water management?" *Science*, Vol. 319, pp. 573–574.

Olson, Mancur (1965) *Logic of collective action: Public goods and the theory of groups*, Vol. 124 of Harvard economic studies: Harvard University Press, Cambridge, Massachussetts, USA.

Sacco, Pier Luigi (1994) "Selection mechanisms in economics," *Behavioral Science*, Vol. 39, pp. 311–325.

Schlag, Karl H (1998) "Why imitate, and if so, how?: A boundedly rational approach to multi-armed bandits," *Journal of Economic Theory*, Vol. 78, pp. 130–156.

Sethi, Rajiv and Eswaran Somanathan (1996) "The evolution of social norms in common property resource use," *The American Economic Review*, pp. 766–788.

Shaw, Daigee and Rong-Dean Shaw (1991) "The resistibility and shiftability of depletable externalities," *Journal of Environmental Economics and Management*, Vol. 20, pp. 224–233.

Shogren, Jason F and Thomas D Crocker (1991) "Cooperative and noncooperative protection against transferable and filterable externalities," *Environmental and Resource Economics*, Vol. 1, pp. 195–214.

Smith, Pete, Daniel Martino, Zucong Cai, Daniel Gwary, Henry Janzen, Pushpam Kumar, Bruce McCarl, Stephen Ogle, Frank O'Mara, Charles Rice et al. (2007) "Greenhouse gas mitigation in agriculture," *Philosophical Transactions of the Royal Society B: Biological Sciences*, Vol. 363, pp. 789–813.

Tompkins, Emma L, W Neil Adger, Emily Boyd, Sophie Nicholson-Cole, Keith Weatherhead, and Nigel Arnell (2010) "Observed adaptation to climate change: UK evidence of transition to a well-adapting society," *Global Environmental Change*, Vol. 20, pp. 627–635.

Warburton, Christopher ES et al. (2018) "Positive Time Preference And Environmental Degradation: The Effects Of World Population Growth And Economic Activity On Intergenerational Equity, 1970-2015," *Applied Econometrics and International Development*, Vol. 18, pp. 5–24.

Weibull, Jörgen W (1995) *Evolutionary game theory*: MIT press, Cambridge, Massachussetts, USA.

Chapter 2

Maladaptation and Inequality in an Evolutionary Model with Heterogeneous Agents

Abstract

Developing countries are disproportionally affected by phenomena of environmental degradation. Several scholars approached this issue and researched its causes. Each stream of literature framed the problem differently: environmental dumping, pollution haven, pollution terms of trade. In this work, we look at the consequences in terms of inequality in well-being of a pre-existing inequality in the capacity to cope with environmental adversities. We study the case in which agents may protect themselves from environmental degradation by adopting a maladaptive strategy which emits negative externalities to others. We show that Pareto-dominated states may be reached and that inequality in well-being is increased when all agents adopt the maladaptive strategy.

Keywords: Self-protective strategies, Inequality, Maladaptation, Negative externalities, Developing countries

JEL classification: C70, D62, O13, O40, Q20

2.1 Introduction

Climate Change (and environmental degradation in general) has long become a phenomenon of the highest relevance, whose alarming effects on the economies have been noticed and signalled by many scholars. Perhaps most notably, Meadows et al. (1972) warned that economic growth could be slowed down or even halted by negative feedbacks coming from an increasing degradation of the environment and the related services. More recently, the Stern Review (Stern et al., 2006) delivered a cost-benefits analysis showing that the costs of countering environmental degradation (air pollutants, in particular) would be much lower that the costs in which economies are likely to incur in a Business-As-Usual scenario. The response from economic scholars to the Stern Review and the suggested policy implementations has been diverse. Few scholars maintained that the review built on strong or even radical assumptions on the deterioration of environmental quality or on its feedback on economic activity (Lea, 2006). By contrast, many authors argued that environmental risks were severely underestimated, as only the losses that could be monetised were taken into account, whereas ecosystemic damages are hardly quantifiable and doubtfully reversible (Neumayer, 2007; Spash, 2007; Sterner and Persson, 2008). However controversial its reception, the Stern Review stimulated the debate and led other scholars to conduct economic investigations on the matter, with results confirming the negative relation between environmental degradation and economic activity (Ng and Zhao, 2011).

It would be misleading to consider only the economic effects of Climate Change when assessing its negative impacts on human life, though. The lives of all people are affected also on other crucial aspects. Local environmental degradation is acknowledged as a relevant push factor of many migration phenomena for developing countries (Reuveny, 2007), even if scholars point to a more indirect role (Veronis and McLeman, 2014; Beine and Parsons, 2015). In the words of (Reuveny, 2007), people living in developing countries don't have many possibilities to mitigate the adverse effect (voice) and are left with the decision to either resign themselves to accept the additional burden (loyalty) or to leave their home country (exit). Besides the migration it induces, environmental degradation compromises the health of many people, also in more developed countries (Luzzati et al., 2018): heat-waves, increased frequency of catastrophic events, crop failures, deterioration of air quality. Many are the channels through which Climate Change has been found accountable for the loss of over 150,000 lives per year (Patz et al., 2005; McMichael et al., 2006). Humans are not suffering through the negative effects of environmental degradation without developing new strategies to cope with it. Indeed, people adapt their behaviour in order to protect themselves. In formal terms, agents are increasingly adopting costly strategies to soften the damage caused by environmental degradation. These self-protective choices increase individual well-being of the adopter, while sometimes worsening the conditions of others. It is often the case that the environmental harm is shifted upon the poorest or otherwise most vulnerable individuals by the majority. According to Barnett and O'Neill (2010) and Juhola et al. (2016), such disproportionate burdening of the most vulnerable suffices to characterise these instances as maladaptation. Several instances of maladaptive strategies relating to environmental matters have been put forward by the literature. A common example is the use of air conditioning (Deschênes and Greenstone, 2011), which dampens the effects of climate change for the individual while it contributes to raise energy consumption and ultimately leads to carbon leakage, while also increasing the risk of energy shortages (Lundgren and Kjellstrom, 2013). Other examples include snow-making (Abegg et al., 2008), water desalination (Barnett and O'Neill, 2010), and urban planning of coastal areas (Macintosh, 2013). In this paper, we focus on the effects of maladaptive strategies on the most vulnerable and show that the existence of a maladaptive strategy may increase inequality in well-being across regions or countries. In other terms, we highlight the mechanism through which a pre-existing divergence in the well-being of agents may be amplified by heterogeneity in their capacity to respond to environmental hazards effectively.

The effects of heterogeneity of agents on the capacity to defend against environmental hazards have been debated in the literature since the work by Olson (1965), whose argument was later rephrased with an effective example by Baland and Platteau (1999). These authors claim that if the rights to extract from the common (here, environmental) goods were distributed among many, and the free entry condition held, then the individual agents would be incentivised to supply the resources extracted from the common until the expected return is null (perfect competition scenario). If we assume decreasing productivity of the common, each new agent extracting from it damages all other agents at the same time. This would lead to selfish and inefficient overexploitation of the common, with its depletion not even being compensated by the profits from extraction (because of the perfect competition assumption implying null profits). By contrast, the authors argue that if extraction rights are reserved to few, then the agents would be more affected by the fall in productivity caused by each additional unit extracted. In the end, the common would suffer from less overexploitation, i.e. the amount extracted would be closer to the social optimum. For this reason, Baland and Platteau (1999) state that heterogeneity fosters successful management of environmental resources. Baland and Platteau (1999) also add that high initial costs need to be withstood by the agents engaging in consensus building within the group and such concertation costs might be too high for most agents to be undertaken. In this perspective, inequality allows for unequal distribution of returns upon a successful arrangement, thus constituting an incentive for the wealthiest ones to engage in the costly consensus building.

However, as Agrawal et al. (2002) argue in their seminal paper, heterogeneity of agents (e.g. wealth inequality) likely comes with heterogeneity in interests, leading to unsuccessful environmental preservation. Asymmetry in power and conflicts of interests between different groups (or subgroups) is an obstacle in the design of successful institutions to prevent environmental degradation (Agrawal and Gibson, 1999). We add that insofar as an unequal distribution allows some individuals to accumulate enough resources to adopt maladaptive strategies, it also deteriorates inequality of well-being. Moreover, the ones that hold the least power, i.e. the poors, are also the ones who depend on environmental resources the most (Jodha, 1986), and may thus be the most affected by the depletion of the common (in this case, the environment). People living in developing countries and in rural areas in particular are especially vulnerable to environmental degradation and extreme natural catastrophes (Strömberg, 2007), as they often lack the financial, institutional, and technological means to defend against environmental degradation (Toya and Skidmore, 2007; Barbier, 2010). Here we characterise two different regions and the related populations as heterogeneous in their efficacy to employ defensive strategies against environmental hazards. In particular, we describe the population living in the less developed region as being both less effective in combating environmental degradation and more exposed to its adverse effects, as the above mentioned literature suggests. Furthermore, in this paper we argue that higher levels of environmental degradation (lower environmental quality) exacerbate the connection between inequality in the capacity to cope with environmental hazards and inequality in well-being. Indeed, the unequal capacity to cope with an even more degraded environment magnifies the payoff divergence of more developed regions and less developed ones.

This work is an attempt to formally describe the interconnection among maladaptive strategies, asymmetrical exposure to environmental degradation, and inequality. In particular, we show that if there exists a costly self-protective maladaptive strategy which has reduced efficacy for the poorest, then the divergence in well-being with the wealthiest increases. With respect to the theoretical framework of the previous chapter, we here assume that the degradation concerns a common environmental resource, instead of a dedicated one for each group. To this respect, we remark that any division of a population into two distinct groups characterised by an asymmetry in capabilities is valid. In this work, we use a two-country exemplification, although the same conclusions hold for two different social classes differing in their capacity to cope with any adversity. A further difference with the previous study is that we here deal with a more specifically contextualised system, which is characterised by a common (and negative) public effect of strategies, which better fits the cases of environmental maladaptation. This allows for the addition of a meaningful analysis of the effects of inequality on exposure to environmental degradation. From the methodological point of view, we also draw from the Climate Change literature by employing a damage function which explicitly models the relation between the environment and economic activity. Our analysis points to two major results. First, under given conditions, we find that non-adoption of the self-protective strategy by all agents may be the Pareto optimal equilibrium. In particular, when non-adoption is individually preferable, i.e. when the cost of adoption is higher than its benefits even for the wealthiest agents, the non-adoption equilibrium is always the Pareto optimal one. Second, if all agents (even the poorest ones) adopt the maladaptive self-protective strategy, then inequality increases with respect to the non-adoption scenario. With this work, we aim to contribute to the debate on the future of environmental adaptation, arguing that inequality in the capacity to adopt coping strategies is a crucial element of the debate which cannot be overlooked.

2.2 Modelisation

Let us assume that there are two social groups, which differ from each other in terms of three elements: gross output, exposure to environmental hazards, efficacy in reducing such exposure. All three elements are exogenous and shall be presented in this section. As concerns the two groups, for illustrative purposes we here characterise them as two countries at different stages of development, although many other characterisations of this phenomenon are plausible. Moreover, this description of the problem allows us to relate to the asymmetrical capability of less developed countries and more developed ones to cope with environmental degradation (Barbier, 2010). We name the wealthier country N and poorer country S. Agents from both countries benefit from a gross output represented by $\overline{Y}_i$, with i indexing for N, S. We impose that the gross output of agents from more developed countries is higher than in less developed ones:

$$\overline{Y}_N > \overline{Y}_S > 0$$

The net output Y_i of every agent is equal to their gross output minus an economic damage $\Omega_i(P)$ suffered from the environmental degradation P:

$$Y_i = \overline{Y}_i - \Omega_i(P), \qquad\qquad i = N, S$$

In this perspective, we see that $\overline{Y}_i$ is the output of country i when there is no environmental degradation or the damage function is otherwise equal to 0. In order to reduce the damage from environmental degradation (and thus increase net output), agents are allowed to choose whether to defend against it or not. As in Antoci and Borghesi (2010b), this translates into two strategies available to agents from both countries:

1. [**D**] Protecting from environmental degradation;

2. [**ND**] Not protecting from environmental degradation.

In order to adopt strategy **D**, agents have to incur into an additional cost equal to C^D. The strategy chosen by each agent from country i determines the final level of the net output:

$$Y_i^{ND} = \overline{Y}_i - \Omega_i^{ND}(P), \qquad \text{if } i \text{ chooses } \mathbf{ND} \tag{2.1}$$

$$Y_i^{D} = \overline{Y}_i - \Omega_i^{D}(P), \qquad \text{if } i \text{ chooses } \mathbf{D} \tag{2.2}$$

Hence, the difference in net output yielded by the two strategies depends exclusively on the different form of the damage function. In particular, for an agent from country i it holds that:

$$\Omega_i^{ND}(P) = \alpha_i P, \qquad \text{if } i \text{ chooses } \mathbf{ND}$$

$$\Omega_i^{D}(P) = \frac{\alpha_i P}{1 + d_i}, \qquad \text{if } i \text{ chooses } \mathbf{D}$$

where α_i measures the negative impact of P on the economic activities of agents from country i and d_i measures the efficacy of protecting against environmental degradation. In line with Strömberg (2007) and Barbier (2010), we assume that $\alpha_S > \alpha_N > 0$ and that $d_N > d_S > 0$, i.e. environmental degradation is not only more damaging to the economy of S, but the latter is also less effective in contrasting its adverse effects. Under the stated conditions on the parameters, we have that:

$$\Omega_i^{D}(P) < \Omega_i^{ND}(P), \qquad i = N, S \tag{2.3}$$

$$\Omega_N^{j}(P) < \Omega_S^{j}(P), \qquad j = \mathbf{ND}, \mathbf{D} \tag{2.4}$$

On the one hand, (2.3) states that the adverse effects of environmental degradation are stronger if an agent does not protecting herself. On the other hand, (2.4) states that if both countries adopt the same strategy, the damage from environmental degradation in N would still be lower than in S. We recall that this describes a scenario in which industrialised countries are more effective to tackle negative environmental feedbacks on economic productivity. Moreover, it holds that:

$$\frac{d\Omega_i^{D}(P)}{dP} < \frac{d\Omega_i^{ND}(P)}{dP}, \qquad i = N, S \tag{2.5}$$

$$\frac{d\Omega_N^{j}(P)}{dP} < \frac{d\Omega_S^{j}(P)}{dP}, \qquad j = \mathbf{ND}, \mathbf{D} \tag{2.6}$$

Condition (2.5) specifies that as the environment degrades, i.e. P increases, self-protecting choices become increasingly convenient, as the damage function increases less rapidly for agents adopting **D**. Condition (2.6) adds that for any chosen common strategy, the damage from environmental degradation increases faster in the developing country, i.e. S. We note that conditions (2.5) and (2.6) make conditions (2.3) and (2.4) more stringent as P increases.

Similarly to other works, we assume that the level of environmental degradation is only affected by the strategy distribution in the two countries (Antoci and Borghesi, 2010a). In this case, the dynamics of P depends on the shares of agents x_t and z_t adopting strategy **D** at time t in country N and S, respectively:

$$P = \overline{P} + \beta_N \cdot x_t + \beta_S \cdot z_t \tag{2.7}$$

where $\overline{P} > 0$ is the environmental degradation when all agents in both countries adopt strategy **ND**, i.e. $x_t = z_t = 0$, whereas $\beta_N, \beta_S > 0$ measure the impact of self-protecting choices of agents from N and S, respectively. Given the positive sign of these parameters, if the share of agents adopting strategy **D** increases, then the environmental degradation becomes higher. We remark that if all agents successfully coordinated on strategy **ND**, they could be better off, as they would enjoy a lower value of P while not incurring into the increased cost of adopting **D**. Ultimately, if β_N and β_S are sufficiently high, then the well-being might be higher in case no agent adopts strategy **D**, although it is individually convenient to do so. This mechanism describes a particular case of maladaptation, in which even the adopters of the maladaptive strategy are worse off, in the end.

In order to estimate the well-being of agents, we need to define their payoff functions. Recalling that agents adopting strategy **D** pay a cost $C^D > 0$[1], we outline the payoff of an agent from country i:

$$U_i = \begin{cases} U_i^{ND}(x,z) = \left[\overline{Y}_i - \Omega_i^{ND}(P)\right], & \text{for strategy ND} \\ U_i^{D}(x,z) = -C^D + \left[\overline{Y}_i - \Omega_i^{D}(P)\right], & \text{for strategy D} \end{cases} \tag{2.8}$$

whereas x and z follow a replicator dynamics (see, for example, Weibull, 1995):

$$\dot{x} = x(1-x)\ \Pi_N \tag{2.9a}$$

$$\dot{z} = z(1-z)\ \Pi_S \tag{2.9b}$$

where $\dot{x}$ and $\dot{z}$ are the time derivatives of x and z, respectively and Π_i is the payoff differential of country i:

$$\Pi_i = U_i^D(x,z) - U_i^{ND}(x,z) \qquad \text{with } i = N, S \tag{2.10}$$

We remark that the share of agents in country i adopting strategy **D** increases ($\dot{x} > 0$) when its payoff differential is positive ($U_i^D(x,z) - U_i^{ND}(x,z) > 0$), whereas it decreases ($\dot{x} < 0$) if the payoff differential is negative ($U_i^D(x,z) - U_i^{ND}(x,z) < 0$).

2.3 Stationary states

Given the payoff functions (2.8), we may rewrite the payoff difference (2.10) as follows:

$$\begin{aligned} \Pi_i &= -C^D + \overline{Y}_i - \Omega_i^D(P) - \left[\overline{Y}_i - \Omega_i^{ND}(P)\right] \\ &= -C^D + \frac{\alpha_i d_i}{1+d_i} P, \end{aligned} \tag{2.11}$$

[1] We could also assume additional positive costs $C > 0$ common to all agents, whether they choose **D** or **ND**. However, since we are concerned with differential costs (as can be seen later in this section), this would have no impact on the choices of agents.

with $i = N, S$ and P given by (2.7). Intuitively, the payoff differential is positive ($\Pi_i > 0$) if the costs C^D of adopting strategy $\mathbf{D}$ are lower than its benefits, represented by the damage functions differential:

$$\Omega_i^{ND}(P) - \Omega_i^{D}(P) = \frac{\alpha_i d_i}{1+d_i} P, \tag{2.12}$$

where we omitted the subscripts t for the sake of simplicity. We note that this damage differential is always positive as required by condition (2.3). By substituting equation (2.7) in (2.11) we obtain:

$$\Pi_i = -C^D + \frac{\alpha_i d_i}{1+d_i} \left(\overline{P} + \beta_N \cdot x + \beta_S \cdot z\right), \ i = N, S \tag{2.13}$$

The dynamic system (2.9a)–(2.9b) is defined in the square Q:

$$Q = \{(x, z) : 0 \leq x \leq 1, \ 0 \leq z \leq 1\}$$

In order to find the stationary states of system (2.9a)–(2.9b), we study the points in which $\dot{x} = 0$ and $\dot{z} = 0$. We find that $\dot{x} = 0$ holds when all agents in N adopt the same strategy, i.e. for $x = 0$, $x = 1$. In these cases, even if the payoff difference is different from zero and would otherwise induce a change in strategies, the agents could not imitate the better performing strategy from anyone in their country. More interestingly, the agents from N have no incentive in changing strategy if the two yield the same payoff, which occurs on all couplets (x, z) belonging to the straight line:

$$z = f_N(x) := \frac{1+d_N}{\alpha_N \beta_S d_N} C^D - \frac{\overline{P}}{\beta_S} - \frac{\beta_N}{\beta_S} \cdot x \tag{2.14}$$

along which the two strategies yield the same payoff for agents in N, i.e. $U_N^D(x, z) = U_N^{ND}(x, z)$. It is easy to check that the share of agents in N adopting strategy $\mathbf{D}$ increases ($\dot{x} > 0$) above the line (2.14), whereas it decreases ($\dot{x} < 0$) below it.

Analogously, $\dot{z} = 0$ holds when all agents in S adopt the same strategy, i.e. for $z = 0$, $z = 1$, and in all couplets (x, z) belonging to the straight line:

$$z = f_S(x) := \frac{1+d_S}{\alpha_S \beta_S d_S} C^D - \frac{\overline{P}}{\beta_S} - \frac{\beta_N}{\beta_S} \cdot x \tag{2.15}$$

along which $U_S^D(x, z) - U_S^{ND}(x, z) = 0$. In analogy with country N, the share of agents adopting $\mathbf{D}$ in S increases above the line (2.15) and decreases below it. We note that the lines (2.14) and (2.15) have the same slope: $f'_N(x) = f'_S(x) = -\frac{\beta_N}{\beta_S} < 0$. Furthermore, $f_N(0) = f_S(0)$ holds for:

$$\frac{\alpha_S d_S}{1+d_S} = \frac{\alpha_N d_N}{1+d_N} \tag{2.16}$$

We shall see that (2.16) proves to be a fundamental condition for the existence of stationary states internal to **Q**. Indeed, we now summarise the stationary points, i.e. the points of system (2.9a)–(2.9b) in which $\dot{x} = \dot{z} = 0$:

1. All vertices $(0,0)$, $(1,1)$, $(0,1)$, $(1,0)$ of Q; each of them represents a scenario in which a single strategy is adopted in both populations. These are "pure population stationary states" in that they describe scenarios in which each population "specialises" in a single strategy.

2. The intersection points (when they exist) between the line (2.14) and the sides of Q with either $z = 0$ or $z = 1$.

3. The intersection points (when they exist) between the curve (2.15) and the sides of Q with either $x = 0$ or $x = 1$.

4. The points, when they exist, internal to the square region Q and belonging to both lines (2.14) and (2.15). According to the above analysis, no internal stationary exists if (2.16) is not satisfied. In the non robust case in which (2.16) holds, the lines (2.14) and (2.15) coincide and all the points belonging to the intersection between them and the interior of the square Q are stationary states. For simplicity, in the following analysis we shall not consider such a non robust case.

2.4 Dynamic regimes

According to the results illustrated in the Mathematical appendix, we have that almost all[2] the trajectories starting in the interior of the square Q converge to one of the vertices $(0,0), (0,1), (1,0), (1,1)$. **Proposition 3**, in the appendix, gives the conditions allowing for the stability of such stationary states. According to such a proposition, the stationary state $(0,0)$ (in which all agents adopt the non-defensive strategy **ND** in both countries) is (locally) attractive if the cost C^D is high enough; that is, if both the following

[2]The unique trajectories not converging to a vertex of Q are those belonging to the one-dimensional stable manifolds of the saddle points in the interior of the sides of Q (when existing).

conditions hold:

$$C^D > \frac{\alpha_N d_N}{1+d_N}\overline{P} \tag{2.17}$$

$$C^D > \frac{\alpha_S d_S}{1+d_S}\overline{P} \tag{2.18}$$

We recall that parameter C^D measures the cost difference between the defensive strategy **D** and the non defensive one **ND** whereas parameter $\overline{P}$ is the environmental degradation when all agents choose strategy **ND**. Finally, we also recall that α_i measures the economic damage for each unit of environmental degradation, whereas d_i discounts the damage function and measures the efficacy of strategy **D**.

The stationary state $(1,1)$ (in which all agents adopt the defensive strategy **D** in both countries) is attractive if C^D is sufficiently low, that is, if both of the following conditions hold:

$$C^D < \frac{\alpha_N d_N}{1+d_N}\left(\overline{P}+\beta_N+\beta_S\right) \tag{2.19}$$

$$C^D < \frac{\alpha_S d_S}{1+d_S}\left(\overline{P}+\beta_N+\beta_S\right) \tag{2.20}$$

We note that the stationary states $(0,0)$ and $(1,1)$ can be simultaneously (locally) attractive. In this case, which stationary state the system converges to depends on the initial distribution of strategies $x(0)$ and $z(0)$. If the latter is sufficiently close to a stationary state, then the trajectory starting from $(x(0), z(0))$ will approach such stationary state. This bistability scenario occurs if both of the following conditions hold:

$$\frac{\alpha_N d_N}{1+d_N}\left(\overline{P}+\beta_N+\beta_S\right) > C^D > \frac{\alpha_N d_N}{1+d_N}\overline{P}$$

$$\frac{\alpha_S d_S}{1+d_S}\left(\overline{P}+\beta_N+\beta_S\right) > C^D > \frac{\alpha_S d_S}{1+d_S}\overline{P}$$

Also the asymmetric stationary states $(0,1)$ and $(1,0)$, in which the populations of the two countries specialise in different strategies (either **D** or **ND**), can be attractive. In particular, $(0,1)$ is attractive if the following conditions hold:

$$\frac{\alpha_N d_N}{1+d_N}\left(\overline{P}+\beta_S\right) < C^D < \frac{\alpha_S d_S}{1+d_S}\left(\overline{P}+\beta_S\right) \tag{2.21}$$

The condition becomes intuitive recalling that $(0,1)$ is an asymmetric stationary state in which all agents in N adopt strategy **ND**, while the agents from S adopt strategy **D**[3]. In this case, the environmental degradation P is equal to $\overline{P}+\beta_S$, so that the terms to the left and to the right of inequality (2.21) are the damage differential for N and S, respectively. The condition for the stability of $(0,1)$ thus requires the costs of adopting strategy **D** to be lower than the benefits only for country S, leading to a share $x=0$ adopting the self-protective strategy. Analogously $(1,0)$ is attractive if the following conditions hold:

$$\frac{\alpha_S d_S}{1+d_S}\left(\overline{P}+\beta_N\right) < C^D < \frac{\alpha_N d_N}{1+d_N}\left(\overline{P}+\beta_N\right) \tag{2.22}$$

This condition is the reverse of the previous one and its interpretation is indeed specular. As the costs of adopting strategy **D** are higher than its benefits for S, whereas they are lower than its benefits for N, only the population of the latter ends up adopting the maladaptive technology. We note that the stationary states $(0,1)$ and $(1,0)$ cannot be simultaneously attractive. On the one hand, $(0,1)$ is attractive only if the following inequality holds:

$$\frac{\alpha_N d_N}{1+d_N} < \frac{\alpha_S d_S}{1+d_S} \tag{2.23}$$

or equivalently, if $f_N(0) > f_S(0)$. On the other hand, $(1,0)$ is attractive only if the opposite condition holds. We remark that the terms to the left and to right of condition (2.23) describe the damage differentials per unit of environmental degradation for N and S, respectively. Moreover, it is relevant to observe that the damage differential of a country i deriving from adopting strategy **D** depends positively on both the impact coefficient α_i and on self-protection efficacy d_i. It is intuitive that a higher efficacy would increase the damage differential. It is less intuitive, perhaps, that a higher impact coefficient has the same effect. This is due to the fact that a country that suffers more from environmental degradation has more to gain from adopting a strategy which dampens this adverse effect. Figure 1 illustrates a bistable dynamic regime in which only the stationary states $(0,0)$ and $(1,1)$ are attractive. Figures 2 and 3 illustrate the dynamic regimes in which, respectively, $(0,0),(1,1),(0,1)$ and $(0,0),(1,1),(1,0)$ are attractive.

[3]The reverse case is represented by stationary state $(1,0)$, in which all agents from N adopt the self-protective strategy while the agents from S do not.

2.5 Well-being and inequality

We are now interested in studying what is the well-being of the two populations in the stationary states of the system. To do so, we define the average well-being of country i as the average of the utility of agents adopting **D** and the ones adopting **ND**, weighted by their shares in the population. The average well-being in N and S is thus given respectively by:

$$\widetilde{U}_N(x,z) := x \cdot U_N^D(x,z) + (1-x) \cdot U_N^{ND}(x,z) \tag{2.24}$$

$$\widetilde{U}_S(x,z) := z \cdot U_S^D(x,z) + (1-z) \cdot U_S^{ND}(x,z) \tag{2.25}$$

We note that $\widetilde{U}_N(0,z) = U_N^D(0,z)$ and $\widetilde{U}_N(1,z) = U_N^{ND}(1,z)$ represent the average well-being in N when the whole population adopts strategy **D** or strategy **ND**, respectively. The interpretation of $\widetilde{U}_S(x,0) = U_S^{ND}(x,0)$ and $\widetilde{U}_S(x,1) = U_S^D(x,1)$ for country S is analogous. We shall limit our analysis to the comparison of well-being levels evaluated at the stationary states $(0,0),(0,1),(1,0),(1,1)$, which are the only ones that can be attractive. The following proposition holds.

Proposition 1 *When the stationary state* $(0,0)$ *is attractive, then it always Pareto-dominates the stationary states* $(0,1)$, $(1,0)$, *and* $(1,1)$, *that is:* $\widetilde{U}_i(0,0) > \widetilde{U}_i(0,1), \widetilde{U}_i(1,0), \widetilde{U}_i(1,1)$, $i = N, S$. *Furthermore,* $(0,0)$ *may Pareto-dominate* $(0,1)$, $(1,0)$, *and* $(1,1)$ *even if it is not attractive (more details in the proof below).*

Proof. We have:

$$(0,0): \begin{cases} \widetilde{U}_N(0,0) = U_N^{ND}(0,0) = \overline{Y}_N - \alpha_N \overline{P} \\ \widetilde{U}_S(0,0) = U_S^{ND}(0,0) = \overline{Y}_S - \alpha_S \overline{P} \end{cases} \tag{2.26}$$

$$(1,1): \begin{cases} \widetilde{U}_N(1,1) = U_N^D(1,1) = -C^D + \overline{Y}_N - \dfrac{\alpha_N}{1+d_N}\left(\overline{P} + \beta_N + \beta_S\right) \\ \widetilde{U}_S(1,1) = U_S^D(1,1) = -C^D + \overline{Y}_S - \dfrac{\alpha_S}{1+d_S}\left(\overline{P} + \beta_N + \beta_S\right) \end{cases} \tag{2.27}$$

$$(0,1): \begin{cases} \widetilde{U}_N(0,1) = U_N^{ND}(0,1) = \overline{Y}_N - \alpha_N\left(\overline{P} + \beta_S\right) \\ \widetilde{U}_S(0,1) = U_S^D(0,1) = -C^D + \overline{Y}_S - \dfrac{\alpha_S}{1+d_S}\left(\overline{P} + \beta_S\right) \end{cases} \tag{2.28}$$

$$(1,0): \begin{cases} \widetilde{U}_N(1,0) = U_N^D(1,0) = -C^D + \overline{Y}_N - \dfrac{\alpha_N}{1+d_N}\left(\overline{P} + \beta_N\right) \\ \widetilde{U}_S(1,0) = U_S^{ND}(1,0) = \overline{Y}_S - \alpha_S\left(\overline{P} + \beta_N\right) \end{cases} \tag{2.29}$$

Note that:

$$\widetilde{U}_i(0,0) > \widetilde{U}_i(1,1), \qquad \text{with } i = N, S$$

if and only if:

$$C^D > \frac{\alpha_i d_i}{1+d_i}\overline{P} - \frac{\alpha_i}{1+d_i}\left(\beta_N + \beta_S\right), \qquad \text{with } i = N, S \tag{2.30}$$

In other terms, stationary state $(0,0)$ yields a higher utility than $(1,1)$ if the cost of adopting strategy **D** is greater than its benefits in terms of damage reduction, balanced for the increased level of environmental degradation $\frac{\alpha_i}{1+d_i}\left(\beta_N + \beta_S\right)$. In other words, condition (2.30) requires the costs of coping with additional degradation to offset the benefits from a reduced environmental damage. This can occur when the value of C^D, or any of the impact coefficients β_N, β_S are sufficiently high. It is easy to check that the local attractiveness of the stationary state $(0,0)$ (see conditions (2.17)) – (2.18)) implies the Pareto-dominance of $(0,0)$ on the stationary state $(1,1)$. Furthermore, $(0,0)$ may Pareto-dominate $(1,1)$ even if the former is not attractive while the latter is (see conditions (2.19)–(2.20)). In this last case, agents from both countries would be better off if nobody adopted the maladaptive strategy **D**, defending against environmental degradation but ultimately fuelling it. However, as long as strategy **D** is individually convenient, agents will end up in the Pareto-dominated equilibrium $(1,1)$. This poses a dilemma to the policy maker, since enforcing a norm to abandon an attractive state of the system can be very hard, making the norm ultimately ineffective.

We can also compare the well-being in $(0,0)$ against the well-being in the asymmetric stationary states $(0,1)$, $(1,0)$, in which the population of one country adopts unanimously the self-protecting strategy **D** whereas in the other country no one does. Let us consider the stationary state $(0,1)$. Non-adoption of the self-protective strategy **D** in both countries is Pareto-dominant if:

$$\widetilde{U}_N(0,0) > \widetilde{U}_N(0,1)$$
$$\widetilde{U}_S(0,0) > \widetilde{U}_S(0,1)$$

which holds true when:

$$\alpha_N\left(\overline{P}+\beta_S\right) \;>\; \alpha_N\overline{P} \tag{2.31}$$

$$C^D \;>\; \frac{\alpha_S d_S}{1+d_S}\overline{P}-\frac{\alpha_S}{1+d_S}\beta_S \tag{2.32}$$

Condition (2.31) is always satisfied, while condition (2.32) holds if the stationary state $(0,0)$ is attractive (see condition (2.18)). Very similar steps allow to show that the stationary state $(0,0)$ Pareto-dominates the stationary state $(1,0)$, when the former is attractive. ■

Let us now look at inequality in well-being. First of all, we recall that the payoff of an agent is constituted by the gross output $\overline{Y}_i$, the damage term $\Omega_i(P)$, and the cost C^D (in case of adoption of strategy **D**). We stress that the only variable part of the payoff is thus identified in the damage term, while the gross output and the adoption cost are fixed. This leads us to the consideration that any change in the inequality in well-being between agents of N and S is accountable to a variation in the differential of environmental damage. In particular, we find that the following proposition holds.

Proposition 2 *Inequality in $(0,0)$ is lower than in $(1,1)$ when:*

$$\widetilde{U}_N(0,0)-\widetilde{U}_S(0,0)<\widetilde{U}_N(1,1)-\widetilde{U}_S(1,1) \tag{2.33}$$

which occurs if and only if:

$$\frac{\beta_N+\beta_S}{\overline{P}}>\frac{\alpha_S-\alpha_N}{\frac{\alpha_S}{1+d_S}-\frac{\alpha_N}{1+d_N}}-1 \tag{2.34}$$

where $\alpha_S-\alpha_N>0$ and $\alpha_S/(1+d_S)-\alpha_N/(1+d_N)>0$ always.[4]

Proof.

The above proposition can be checked by substituting:

$$\widetilde{U}_N(0,0)-\widetilde{U}_S(0,0)=\overline{Y}_N-\overline{Y}_S+(\alpha_S-\alpha_N)\overline{P} \tag{2.35}$$

and

$$\widetilde{U}_N(1,1)-\widetilde{U}_S(1,1)=\overline{Y}_N-\overline{Y}_S+\left(\frac{\alpha_S}{1+d_S}-\frac{\alpha_N}{1+d_N}\right)\left(\overline{P}+\beta_N+\beta_S\right) \tag{2.36}$$

[4]By assumption, we have that $\alpha_S>\alpha_N$ and $d_N>d_S$.

into inequality (2.33) and solving it. Note that equation (2.34) always holds when the term on its right hand side is non-positive. If it is negative, we have:

$$\frac{\alpha_S - \alpha_N}{\frac{\alpha_S}{1+d_S} - \frac{\alpha_N}{1+d_N}} > 1 \tag{2.37}$$

which holds if and only if:

$$\frac{\alpha_N d_N}{1+d_N} < \frac{\alpha_S d_S}{1+d_S}$$

that is, if condition (2.23) holds (and, consequently, $f_N(0) > f_S(0)$). We recall that the above inequality requires that the damage differential of N due to the adoption of strategy **D** must be lower than the one of S. If this condition holds, then inequality (2.34) holds only for sufficiently high values of the ratio $(\beta_N + \beta_S)/\overline{P}$. This occurs when negative externalities of agents adopting self-protective strategy **D** are much higher than autonomous degradation $\overline{P}$. In other terms, environmental degradation must be driven mainly by human action, rather than by exogenous forces. Under these circumstances, inequality is bound to increase moving from $(0,0)$ to $(1,1)$, as human action makes environmental degradation increase more sharply and thus causes the gap in well-being between agents from N and S to widen, by virtue of equation (2.6). ■

2.6 Discussion and conclusions

In this work we posed the problem of the interconnection between inequality in the capacity to adopt self-protective strategies and inequality in well-being, highlighting the role that environmental degradation plays in this relation. In addition, we explicitly modelled the heterogeneity of agents, making the environmental damage they suffer and the efficacy of their responses dependant on the groups they belong to. We recall that we chose a country-wise exemplification to illustrate our model only for the relevance of the issue at hand and for its descriptive efficacy. Other instances in which two groups are differentiated according to their capacity to cope with adversities would be equally valid, and so our results. We also provided evidence that the model specification chosen did not affect the results, which are robust to a more general formulation of the problem (check section C.2 in the appendix to this chapter). We derived two major conclusions from our analysis.

Firstly, the non-adoption case, in which no agent in either country adopts the maladaptive self-protective strategy, is Pareto-optimal whenever it is attractive, i.e. it is individually beneficial. This represents the case in which the cost of adoption of the maladaptive strategy is too high to make it convenient for any agent, thus disincentivising agents from the emission of negative externalities related to such strategy. Intuitively, no agent would thus be better off adopting the Pareto-dominated maladaptive strategy. However, our analysis shows that non-adoption could be Pareto-optimal even if it is not attractive. This happens when the adoption cost of the maladaptive strategy is sufficiently low for at least the wealthier agents and its negative externalities are sufficiently high to make every agent worse off. In this case maladaptive strategies make the system reach a Pareto-dominated stationary state. This undesirable outcome could be overcome if agents successfully coordinated on non-adoption or if an institution were established or a policy enforced to prevent agents from adopting maladaptive strategies.

Secondly, we also found that the inequality in well-being between the different groups of agents is higher in the full adoption scenario, in which all agents from both countries adopt the maladaptive strategy. Indeed, even if agents from the relatively less developed country adopt the maladaptive strategy, the fact that they are also more vulnerable to environmental degradation makes inequality increase. In the full adoption scenario, the environment is further degraded by the negative externalities coming from all agents, which goes to disproportionally increase the burden on the most vulnerable.

We remark that this model describes the dynamics and well-being consequences of maladaptive strategies, but it does not deal with the innovation process in any group. In other terms, in this work we do not model the appearance of adopters in a group which does not have any already. An expansion of this model might indeed implement a cross-group imitation, discounting the payoff difference by a factor relating to the belonging of the imitating and the imitated agents to different groups. Moreover, we did not include within-group heterogeneity since it did not seem to directly affect our results. However, relevant insights might be drawn from an analysis that studies both cross-group and stratified within-group externalities. Finally, in this work we were concerned with the degradation of an environmental variable which affected both groups, in order to reflect its interpretation in terms of Climate Change. Maladaptive strategies could also affect environmental variables which are only relevant to a homogeneous group, e.g. a lake. An analysis of the interaction of *local* degradation with respect to a *global* one would make for a compelling extension of this work, with a focus on the dynamics of maladaptive strategies affecting the former or the latter.

Appendix C

Mathematical appendix

C.1 Stability properties of the stationary states

C.1.1 Stationary states on the corners of Q

The following proposition concerns the stability of the stationary states $(0,0),(0,1),(1,0),(1,1)$.

Proposition 3 *The Jacobian matrix of the system* (2.9) *evaluated at the stationary state* $(x,z)=(i,k)$, $i=0,1$ *and* $k=0,1$, *is:*

$$\begin{pmatrix} (1-2i)\left[U_N^D(i,k)-U_N^{ND}(i,k)\right] & 0 \\ 0 & (1-2k)\left[U_S^D(i,k)-U_S^{ND}(i,k)\right] \end{pmatrix} \tag{C.1}$$

and has eigenvalues:

$$(1-2i)\left[U_N^D(i,k)-U_N^{ND}(i,k)\right]$$

and

$$(1-2k)\left[U_S^D(i,k)-U_S^{ND}(i,k)\right]$$

Proof. Straightforward. ■

The analysis of the sign of the eigenvalues given in **Proposition 3** allows us to illustrate the stability properties of the stationary states (0,0),(0,1),(1,0),(1,1). In what follows we will denote with $Q_{x=0}$ the side of Q where $x=0$, and with $Q_{x=1}$ the side where $x=1$. Similar interpretations apply to $Q_{z=0}$ and $Q_{z=1}$. All sides of this square are invariant; namely, if the pair (x,z) initially lies on one of the sides, then the whole correspondent trajectory also lies on that side.

Stability of the stationary state $(0, 0)$

The eigenvalue in direction of $Q_{z=0}$ of the Jacobian matrix (C.1), evaluated at $(0, 0)$, is strictly negative if and only if (iff, hereafter):

$$C^D > \frac{\alpha_N d_N}{1 + d_N} \overline{P} \tag{C.2}$$

whereas it is strictly positive iff the opposite of (C.2) holds. The eigenvalue in direction of $Q_{x=0}$ is strictly negative iff:

$$C^D > \frac{\alpha_S d_S}{1 + d_S} \overline{P} \tag{C.3}$$

whereas it is strictly positive iff the opposite of (C.3) holds.

Stability of the stationary state $(0, 1)$

The eigenvalue in direction of $Q_{z=1}$ of the Jacobian matrix (C.1), evaluated at $(0, 1)$, is strictly negative iff:

$$C^D > \frac{\alpha_N d_N}{1 + d_N} \left(\overline{P} + \beta_S\right) \tag{C.4}$$

whereas it is strictly positive iff the opposite of (C.4) holds. The eigenvalue in direction of $Q_{x=0}$ is strictly negative iff:

$$C^D < \frac{\alpha_S d_S}{1 + d_S} \left(\overline{P} + \beta_S\right) \tag{C.5}$$

whereas it is strictly positive iff the opposite of (C.5) holds.

Stability of the stationary state $(1, 0)$

The eigenvalue in direction of $Q_{z=0}$ of the Jacobian matrix (C.1), evaluated at $(1, 0)$, is strictly negative iff:

$$C^D < \frac{\alpha_N d_N}{1 + d_N} \left(\overline{P} + \beta_N\right) \tag{C.6}$$

whereas it is strictly positive iff the opposite of (C.6) holds. The eigenvalue in direction of $Q_{x=1}$ is strictly negative iff:

$$C^D > \frac{\alpha_S d_S}{1 + d_S} \left(\overline{P} + \beta_N\right) \tag{C.7}$$

whereas it is strictly positive iff the opposite of (C.7) holds.

Stability of the stationary state $(1,1)$

The eigenvalue in direction of $Q_{z=1}$ of the Jacobian matrix (C.1), evaluated at $(1,1)$, is strictly negative iff:

$$C^D < \frac{\alpha_N d_N}{1+d_N}\left(\overline{P}+\beta_N+\beta_S\right) \tag{C.8}$$

whereas it is strictly positive iff the opposite of (C.8) holds. The eigenvalue in direction of $Q_{x=1}$ is strictly negative iff:

$$C^D < \frac{\alpha_S d_S}{1+d_S}\left(\overline{P}+\beta_N+\beta_S\right) \tag{C.9}$$

whereas it is strictly positive iff the opposite of (C.9) holds.

Remarks Proposition 3 associates to each stationary state threshold values for the parameter C^D, which measures the cost difference between the defensive strategy **D** and the non defensive one **ND**. When C^D takes a value beyond these thresholds, then the signs of the eigenvalues of the stationary states change and consequently their stability properties change. The stationary state $(0,0)$ (where no agent adopts the defensive strategy in both regions) is (locally) attractive if the cost C^D is sufficiently high (conditions (C.2)–(C.3)). The stationary state $(1,1)$ (where all agents adopt the defensive strategy **D** in both regions) is attractive if C^D is low enough (conditions (C.8)–(C.9)). We note that both the vertices $(0,0)$ and $(1,1)$ can be simultaneously (locally) attractive. The stationary state $(0,1)$ (in which no agent in N adopts the defensive strategy and all agents in S adopt it) can be attractive only if:

$$\frac{\alpha_N d_N}{1+d_N} < \frac{\alpha_S d_S}{1+d_S} \tag{C.10}$$

holds (see conditions (C.4)–((C.5)), while the stationary state $(1,0)$ (where all agents in N adopt the defensive strategy **D** and no agent in S does so) can be attractive only if the opposite of condition (C.10) holds (see conditions (C.6)–(C.7)). This implies that $(0,1)$ and $(1,0)$ cannot be simultaneously attractive. Finally, we note that condition (C.10) is satisfied iff $f_N(0) > f_S(0)$ holds (see (2.16)).

C.1.2 Stability properties of the stationary states in the interior of the edges of Q

The following proposition concerns the stability properties of the stationary states belonging to the interior of the edges of the square Q, i.e. those where both adoption choices coexist in N while all agents in S play the same strategy or vice versa.

Proposition 4 *The Jacobian matrix of the system ((2.9)) evaluated at the stationary states $(i, \overline{z})$ in the interior of the edges $Q_{x=i}$ ($i = 0, 1$) is:*

$$\begin{pmatrix} (1-2i)\left[U_N^D(i,\overline{z}) - U_N^{ND}(i,\overline{z})\right] & 0 \\ \overline{z}(1-\overline{z})\frac{\partial\left[U_S^D(i,\overline{z}) - U_S^{ND}(i,\overline{z})\right]}{\partial x} & \overline{z}(1-\overline{z})\frac{\partial\left[U_S^D(i,\overline{z}) - U_S^{ND}(i,\overline{z})\right]}{\partial z} \end{pmatrix} \tag{C.11}$$

where $\overline{z}$ is the value of z at the stationary state, and has the eigenvalues:

$$\overline{z}(1-\overline{z})\frac{\partial\left[U_S^D(i,\overline{z}) - U_S^{ND}(i,\overline{z})\right]}{\partial z} = \overline{z}(1-\overline{z})\frac{\alpha_S d_S}{1+d_S}\beta_S > 0 \tag{C.12}$$

in direction of $\mathbf{Q_{x=i}}$, and

$$(1-2i)\left[U_N^D(i,\overline{z}) - U_N^{ND}(i,\overline{z})\right] \tag{C.13}$$

in direction of the interior of $\mathbf{Q}$*. The Jacobian matrix of the system* (2.9) *evaluated at the stationary states $(\overline{x}, i)$ in the interior of the edges $Q_{z=i}$ ($i = 0$, 1) is:*

$$\begin{pmatrix} \overline{x}(1-\overline{x})\frac{\partial\left[U_N^D(\overline{x},i) - U_N^{ND}(\overline{x},i)\right]}{\partial x} & \overline{x}(1-\overline{x})\frac{\partial\left[U_N^D(\overline{x},i) - U_N^{ND}(\overline{x},i)\right]}{\partial z} \\ 0 & (1-2i)\left[U_S^D(\overline{x},i) - U_S^{ND}(\overline{x},i)\right] \end{pmatrix} \tag{C.14}$$

where $\overline{x}$ is the value of x at the stationary state, and has the eigenvalues:

$$\overline{x}(1-\overline{x})\frac{\partial\left[U_N^D(\overline{x},i) - U_N^{ND}(\overline{x},i)\right]}{\partial x} = \overline{x}(1-\overline{x})\frac{\alpha_N d_N}{1+d_N}\beta_N > 0 \tag{C.15}$$

in direction of $\mathbf{Q_{z=i}}$, and

$$(1-2i)\left[U_S^D(\overline{x},i) - U_S^{ND}(\overline{x},i)\right] \quad (\textit{in direction of the interior of } Q) \tag{C.16}$$

Proof. Straightforward. ■

We note that the Jacobian matrix (C.14) has at least one positive eigenvalue (the one in direction of one of the sides of Q); therefore, the stationary states in the interior of the sides of Q cannot be attractive. In particular, they are either saddle points (if the eigenvalue in direction of the interior of Q is negative) or repulsors.

C.2 The general case

In this work we employed a specification of the model whose simplicity allows for a simpler illustration of the phenomenon. However, we remark that a more general modelisation does not alter our results, which are robust to other formal specifications. Although a complete study of the general case is beyond the scope of this work, we here show that the results hold also if only the following assumptions are made on the payoff functions $U_i^D(P)$ e $U_i^{ND}(P)$:

$$\frac{dU_i^{ND}(P)}{dP} < \frac{dU_i^D(P)}{dP} < 0, \quad i = N, S \tag{C.17}$$

$$\frac{dU_S^j(P)}{dP} < \frac{dU_N^j(P)}{dP} < 0, \quad j = D, ND$$

We recall that such conditions require that the payoff is always decreasing in the level of environmental degradation. In addition, they require that: 1) for all agents from any country, the payoff is lower if the self-protective strategy **D** is not adopted and that 2) for any given strategy, agents from S obtain a lower payoff. Assuming that P is a function of x and z, with partial derivatives $\frac{\partial P(x,z)}{\partial x} > 0$ and $\frac{\partial P(x,z)}{\partial z} > 0$ (which are thus increasing in x and z), we have that $\dot{x} = 0$ for $x = 0$, $x = 1$, and along the line:

$$U_N^D(P) - U_N^{ND}(P) = 0 \tag{C.18}$$

Analogously, it holds that $\dot{z} = 0$ for $z = 0$, $z = 1$, and along the line:

$$U_S^D(P) - U_S^{ND}(P) = 0 \tag{C.19}$$

The two equations (C.18) and (C.19) implicitly define two functions: $z = f_N(x)$ and $z = f_S(x)$, with equal slope:

$$f_i'(x) = -\frac{\frac{d\left[U_i^D(P) - U_i^{ND}(P)\right]}{dP} \cdot \frac{\partial P(x,z)}{\partial x}}{\frac{d\left[U_i^D(P) - U_i^{ND}(P)\right]}{dP} \cdot \frac{\partial P(x,z)}{\partial z}} = -\frac{\frac{\partial P(x,z)}{\partial x}}{\frac{\partial P(x,z)}{\partial z}} < 0, \quad j = D, ND$$

Therefore, the graph of function $z = f_N(x)$ is a translation (either upward or downward) of the graph of $z = f_S(x)$. This implies that, just as in the model analysed in this work, there are *generally* no stationary states internal to square **Q**, in which strategies **D** and **ND** coexist in both countries.

As concerns the stationary states internal to the sides of **Q** (where the two strategies coexist in only one of the countries), we have that the Jacobian matrix evaluated at $(i, \overline{z})$, with $i = 0, 1$ and $1 > \overline{z} > 0$, is given by:

$$\begin{pmatrix} (1-2i)\left\{U_N^D\left[P\left(i,\overline{z}\right)\right]-U_N^{ND}\left[P\left(i,\overline{z}\right)\right]\right\} & 0 \\ \overline{z}(1-\overline{z})\frac{\partial\left\{U_S^D[P(i,\overline{z})]-U_S^{ND}[P(i,\overline{z})]\right\}}{\partial x} & \overline{z}(1-\overline{z})\frac{\partial\left\{U_S^D[P(i,\overline{z})]-U_S^{ND}[P(i,\overline{z})]\right\}}{\partial z} \end{pmatrix}$$

and has a positive eigenvalue in the direction of the side of $\mathbf{Q}$ with $x = i$:

$$\overline{z}(1-\overline{z})\frac{\partial\left\{U_S^D\left[P\left(i,\overline{z}\right)\right]-U_S^{ND}\left[P\left(i,\overline{z}\right)\right]\right\}}{\partial z} > 0 \quad \text{(by virtue of (C.17))}$$

Analogously, the Jacobian matrix evaluated at the stationary states $(\overline{x}, i)$, with $i = 0, 1$ and $1 > \overline{x} > 0$, is given by:

$$\begin{pmatrix} \overline{x}(1-\overline{x})\frac{\partial\left\{U_N^D[P(\overline{x},i)]-U_N^{ND}[P(\overline{x},i)]\right\}}{\partial x} & \overline{x}(1-\overline{x})\frac{\partial\left\{U_N^D[P(\overline{x},i)]-U_N^{ND}[P(\overline{x},i)]\right\}}{\partial z} \\ 0 & (1-2i)\left\{U_S^D\left[P\left(\overline{x},i\right)\right]-U_S^{ND}\left[P\left(\overline{x},i\right)\right]\right\} \end{pmatrix}$$

and has a positive eigenvalue in direction of the side of $\mathbf{Q}$ on the side with $z = i$:

$$\overline{x}(1-\overline{x})\frac{\partial\left\{U_N^D\left[P\left(\overline{x},i\right)\right]-U_N^{ND}\left[P\left(\overline{x},i\right)\right]\right\}}{\partial x} > 0 \quad \text{(by virtue of (C.17))}$$

Therefore, also in this more general context, we find that only the vertices $(0,0)$, $(0,1)$, $(1,0)$, and $(1,1)$ can be attractive and that all trajectories starting from the interior of $\mathbf{Q}$ converge to one of such vertices, excluding the ones belonging to the stable manifolds of the stationary states $(i, \overline{z})$ e $(\overline{x}, i)$ above mentioned, when they exist and are saddle points. The less general specification adopted in this work is thus meant to provide an easier interpretative framework.

Bibliography

Abegg, Bruno, R Bürki, and Hans Elsasser (2008) "Climate change and tourism in the Alps," in A. Borsdorf, J. Stötter, and E. Veulliet eds. *Managing Alpine Future*, pp. 73–80.

Agrawal, Arun et al. (2002) "Common resources and institutional sustainability," in Elinot Ostrom et al. eds. *The drama of the commons*: National Academy Press, Chap. 2, pp. 41–85.

Agrawal, Arun and Clark C Gibson (1999) "Enchantment and disenchantment: the role of community in natural resource conservation," *World Development*, Vol. 27, pp. 629–649.

Antoci, Angelo and Simone Borghesi (2010a) "Environmental degradation, self-protection choices and coordination failures in a North–South evolutionary model," *Journal of Economic Interaction and Coordination*, Vol. 5, pp. 89–107.

——— (2010b) "Transferring negative externalities: feedback effects of self-protection choices in a two hemispheres model," in *Nonlinear Dynamics in Economics, Finance and Social Sciences*: Springer, pp. 1–11.

Baland, Jean-Marie and Jean-Philippe Platteau (1999) "The ambiguous impact of inequality on local resource management," *World Development*, Vol. 27, pp. 773–788.

Barbier, Edward B (2010) "Poverty, development, and environment," *Environment and Development Economics*, Vol. 15, pp. 635–660.

Barnett, Jon and Saffron O'Neill (2010) "Maladaptation," *Global Environmental Change*, Vol. 2, pp. 211–213.

Beine, Michel and Christopher Parsons (2015) "Climatic factors as determinants of international migration," *The Scandinavian Journal of Economics*, Vol. 117, pp. 723–767.

Deschênes, Olivier and Michael Greenstone (2011) "Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the US," *American Economic Journal: Applied Economics*, Vol. 3, pp. 152–85.

Jodha, Narpat S (1986) "Common property resources and rural poor in dry regions of India," *Economic and Political Weekly*, pp. 1169–1181.

Juhola, Sirkku, Erik Glaas, Björn-Ola Linnér, and Tina-Simone Neset (2016) "Redefining maladaptation," *Environmental Science & Policy*, Vol. 55, pp. 135–140.

Lea, Ruth (2006) "Just another excuse for higher taxes."

Lundgren, Karin and Tord Kjellstrom (2013) "Sustainability challenges from climate change and air conditioning use in urban areas," *Sustainability*, Vol. 5, pp. 3116–3128.

Luzzati, T, A Parenti, and T Rughi (2018) "Economic Growth and Cancer Incidence," *Ecological Economics*, Vol. 146, pp. 381–396.

Macintosh, Andrew (2013) "Coastal climate hazards and urban planning: how planning responses can lead to maladaptation," *Mitigation and Adaptation Strategies for Global Change*, Vol. 18, pp. 1035–1055.

McMichael, Anthony J, Rosalie E Woodruff, and Simon Hales (2006) "Climate change and human health: present and future risks," *The Lancet*, Vol. 367, pp. 859–869.

Meadows, Donella H, Dennis L Meadows, Jørgen Randers, and William W Behrens III (1972) *The limits to growth*: Universe Books, New York, NY, USA.

Neumayer, Eric (2007) "A missed opportunity: The Stern Review on climate change fails to tackle the issue of non-substitutable loss of natural capital," *Global Environmental Change*, Vol. 17, pp. 297–301.

Ng, Pin and Xiaobing Zhao (2011) "No matter how it is measured, income declines with global warming," *Ecological Economics*, Vol. 70, pp. 963–970.

Olson, Mancur (1965) *Logic of collective action: Public goods and the theory of groups*, Vol. 124 of Harvard economic studies: Harvard University Press, Cambridge, Massachussetts, USA.

Patz, Jonathan A, Diarmid Campbell-Lendrum, Tracey Holloway, and Jonathan A Foley (2005) "Impact of regional climate change on human health," *Nature*, Vol. 438, p. 310.

Reuveny, Rafael (2007) "Climate change-induced migration and violent conflict," *Political Geography*, Vol. 26, pp. 656–673.

Spash, CL (2007) “Fallacies of economic growth in addressing environmental losses: Human induced climatic change,” *Newsletter of the Australia New Zealand Society for Ecological Economics (ANZSEE)*, pp. 2–4.

Stern, Nicholas, Siobhan Peters, Vicki Bakhshi, Alex Bowen, Catherine Cameron, Sebastian Catovsky, Diane Crane, Sophie Cruickshank, Simon Dietz, Nicola Edmonson et al. (2006) *Stern Review: The economics of climate change*, Vol. 30: HM treasury, London, GB.

Sterner, Thomas and U Martin Persson (2008) “An even sterner review: Introducing relative prices into the discounting debate,” *Review of Environmental Economics and Policy*, Vol. 2, pp. 61–76.

Strömberg, David (2007) “Natural disasters, economic development, and humanitarian aid,” *Journal of Economic Perspectives*, Vol. 21, pp. 199–222.

Toya, Hideki and Mark Skidmore (2007) “Economic development and the impacts of natural disasters,” *Economics Letters*, Vol. 94, pp. 20–25.

Veronis, Luisa and Robert McLeman (2014) “Environmental influences on African migration to Canada: focus group findings from Ottawa-Gatineau,” *Population and Environment*, Vol. 36, pp. 234–251.

Weibull, Jörgen W (1995) *Evolutionary game theory*: MIT press, Cambridge, Massachussetts, USA.

Chapter 3

Environmental Degradation and Comparative Advantage Reversals

Abstract

Many scholars argue that one of the main drivers of international trade patterns is the varying costs, either monetary or not, firms incur into to pollute the environment. This claim has been studied by different streams of research, some of which obtained particular attention: Pollution Terms of Trade, Pollution Haven Hypothesis, Environmental Dumping. We here briefly review some works from these research streams and build a Comparative Advantage model which provides a theoretical foundation to the claim that lower environmental standards shape the patterns of international trade. In this work, productivity of both more developed and less developed countries is assumed to be negatively affected by environmental degradation, with larger environmental impacts on the productivity of less developed countries. We show that, under these conditions, Comparative Advantages may invert if the environment is sufficiently degraded, leading to a reversal in product specialisation of countries. Finally, we also investigate the conditions under which international trade yields a lower well-being with respect to the autarky level for at least one of the countries.

Keywords: Comparative advantages, Negative externalities, Evolutionary dynamics, Environmental degradation, International trade

JEL classification: C70, D62, O13, O40, Q20

3.1 Introduction

Intensive agricultural practices, involving continuous usage of fertilisers, pesticides, and herbicides, while also increasing the demand for irrigation, might undermine the sustainability of the agricultural sector in less developed countries. Although this type of farming provides farmers with a short-term increase in their outcome, the negative externalities associated with it may endanger ecological services and reduce the returns of farmers in the long-term. Wilson and Tisdell (2001) provide a sketch of the argument, providing an explanation of the reasons why this is more likely to happen in less developed countries. Higher real discount rates make short-term gains relatively larger with respect to long-term losses. Since real discount rates are commonly assumed to be higher in less developed countries, the authors claim that the unsustainable production processes are more likely to take place in these countries. This resonates with the literature on the Pollution Haven Hypothesis (Scoppola and Raimondi, 2018; Gallagher, 1999), which maintains that the most polluting processes are offshored to less developed countries due to laxer regulation and cheaper waste disposal. However, this ends up shifting the environmental load onto the already fragile ecologies of less developed countries (Peng et al., 2016; World Bank, 2003). In order to boost output, farmers in less developed countries make excessive use of fertilisers, which threatens farm productivity due to soil degradation, salinisation and waterlogging (Hazell and Wood, 2008; Scherr, 1999; Tisdell, 1999). The impoverishment of soil due to fertilisers and other chemical inputs may push farmers to increase their demand for such products, locking them into a poverty-environment trap (Phélinas and Choumert, 2017; Barbier, 2010). A similar argument holds for pesticides: their employment may eliminate natural predators of pests from the local ecology, depriving the farms of natural viable alternatives to pesticides (Pelletier and Tyedmers, 2010) and binding the farmers into economic dependence toward suppliers of these products (Wilson and Tisdell, 2001).

We briefly sketched the reasons why long-term productivity of less developed countries might be undermined, in exchange of short-term benefits in the form of reduced costs and increased exports in the related sectors. Less developed countries might indeed be able to build a competitive advantage by either employing products which are forbidden in more developed countries for their environmental hazard or by adopting short-sighted practices (excessive use of fertilisers, pesticides, or herbicides). For instance, Argentina was able to create a comparative advantage in soybean production, but this has taken a toll in terms of deforestation and soil erosion (Phélinas and

Choumert, 2017). The expansion of GM soybean production induced crucial changes in the Argentinian agrarian structure. In particular, the land under tenancy increased and the rotational cropping pattern was discarded in favour of permanent soybean cultivation, to the detriment of long-term soil quality and long term productivity (Phélinas and Choumert, 2017). By recalling the aforementioned threats to the sustainability of agricultural practices, can current comparative advantages be taken for granted?

The motivation of this study lies in the evidence provided by the economic literature maintaining that less developed countries might already be using laxer environmental standards as a leverage to build a comparative advantage in pollution-intensive sectors (Broner et al., 2012; Kozluk and Timiliotis, 2016). This phenomenon became well-known to scholars as the Pollution Haven Hypothesis (PHH) and received much empirical confirmation (see e.g. Millimet and Roy, 2016; Candau and Dienesch, 2017; Sapkota and Bastola, 2017), but also some criticism over its actual magnitude (Zugravu-Soilita, 2017; Garsous and Kozluk, 2017; Kozluk and Timiliotis, 2016). The Pollution Haven Effect (PHE) received analogous feedback from literature (see e.g. Kellenberg, 2009; Cole, 2004; Tang, 2015). Although very similarly defined, the PHH and the PHE describe slightly different concepts. Indeed, the latter posits that stricter environmental regulation reduces inbound FDIs or even pushes firms to relocate elsewhere, while the former specifies that firms move to countries with laxer environmental regulation, usually represented by less developed countries (Copeland and Taylor, 2004). Both literature streams study the investment and location decisions of firms based on stringency and enforcement of extant environmental regulation. A stream of literature focusing more on the pollution content of traded good is the one investigating the Pollution Terms of Trade. This is an index proposed by Antweiler (1996) and empirically employed by scholars (Grether and Mathys, 2013) showing that more developed economies displace environmental load to less developed ones and that trade imbalances magnify this phenomenon. Similar approaches exist which investigate the embodied pollution of trade flows, reaching the same results (Muradian et al., 2002; Bruckner et al., 2012). In the perspective of the three streams of literature cited, increasing trade and investment openness may damage the environmental quality of less developed countries to the extent that it results in their specialisation on polluting industries. Critics of such streams argued that increasing openness (as with Free Trade Areas) actually decreases the pollution intensity in less developed countries, thanks to the technological transfer performed by firms from more developed countries[1]. At the very least, this effect does not seem to outbal-

[1] This concept is commonly defined as Pollution Halo Hypothesis.

ance the regulatory gap effect (PHH): Duan and Jiang (2017) find that the positive technological innovation effect in China is offset by a widening of the trade imbalance; Zugravu-Soilita (2017) underlines that the PHH effect is stronger than the technological effect in countries with lax environmental standards; Gallagher (2000) stresses that total emissions increased as a result of the North American Free Trade Agreement in the interested area. The linkages presented between environmental degradation and trade urge scholars to further inquiry into the long term consequences of this interaction, which is what we propose to do in this work.

We provide an additional piece to the trade and environment puzzle, investigating not only the consequences of regulatory gap-driven comparative advantages on the environment, but also the long run effects of environmental degradation on international trade patterns. We illustrate why comparative advantages defining the patterns of international trade are not time invariant, but are actually shaped by environmental degradation which may vary in time. We drew examples from the agricultural sector for illustrative purposes, but we do not restrain our analysis to a specific industry. We are interested to study how environmental degradation can have an asymmetrical affect on the productivity of sectors, possibly altering the competitive advantages of countries in time. To the extent of our knowledge, the formal analysis of such phenomenon is new to the literature and could contribute in making more accurate assessments of future trajectories of international trade. Finally, we also look at the well-being consequences for the countries involved and find that, under given conditions on production technology, international trade may be Pareto dominated with respect to autarky.

3.2 The model

We now present our analytical model, which investigates the possible adverse effects of international trade on the environment. With respect to Asako (1979), we consider global environmental degradation instead of local pollution. All countries thus contribute to a global fall in environmental quality, which feedbacks on the productivity of all countries. In this work, we employ a comparative advantage model of international trade with two goods x and y traded between a more developed country N and a less developed country S. We assume that the two countries are constituted by homogeneous agents with identical production rates x_i, y_i. The size of the population is normalised country-wise so that each country has a population size of 1. Under this specification, the aggregate production of country i

for each good is given by the individual production rate times 1. We can thus interchangeably refer to the aggregate production of the country or to the individual production rate, which coincides with the production rate of the representative agent and whose utility is described by a Cobb-Douglas function:

$$(x_i^C)^\theta \; (y_i^C)^{1-\theta} \qquad\qquad \textit{with } i = N,\, S \tag{3.1}$$

where x_i^C and y_i^C define the quantities consumed of goods x and y in country i. The consumption decision of the representative agent from country i is subject to the budget constraint:

$$p_x x_i + p_y y_i = p_x x_i^C + p_y y_i^C. \tag{3.2}$$

with x_i and y_i representing the quantities produced of each good by the representative agent of country i, who acts as a price-taker with respect to both goods, their prices identified by p_x and p_y. In addition, Market Clearing Conditions (MCC) require that:

$$x_N + x_S = x_N^C + x_S^C \tag{3.3a}$$

$$y_N + y_S = y_N^C + y_S^C \tag{3.3b}$$

These conditions require that all goods produced by both countries are consumed. Solving system (3.3) together with the budget constraint (3.2) yields the equilibrium prices, which are expressed in terms of the production rate of the representative agent i. We describe the production function of each representative agent as:

$$x_i = \frac{\rho_i^x}{1 + \eta_i^x P} L_i^x \tag{3.4}$$

$$y_i = \frac{\rho_i^y}{1 + \eta_i^y P} (1 - L_i^x) \tag{3.5}$$

where ρ_i^x and ρ_i^y are (positive) productivity coefficients for countries, with i indexing for N, S. The representative agents devote a share L_x of their time to the production of good x and its complement $1 - L_x$ to the production of good y. Finally, the parameters η_i^x and η_i^y are positive and account for the negative impact of pollution on the production of goods x and y, respectively. We can thus see that the production of both goods is negatively affected by environmental degradation P, which in turn is the by-product of production in both countries:

$$\dot{P} = \gamma_N^x x_N + \gamma_N^y y_N + \gamma_S^x x_S + \gamma_S^y y_S - \varepsilon P e^{-\varphi P} \tag{3.6}$$

where $\gamma_i^x, \gamma_i^y > 0$ are the environmental impact coefficients for country i in sectors x and y, respectively. The parameter ε is the rate of environmental regeneration. We note that the environmental impact coefficients differ across both countries and sectors. This reflects the possibility that countries adopt diverging environmental standards and the production processes of goods are not equally polluting. Furthermore, we note that the rate of environmental degradation depends negatively on its current level, so that a share ε of P decays continuously. However, the fact that this term is multiplied by $e^{-\varphi P}$ makes the decay rate dampen as the value of P increases. For extremely high values of P, the decay is close to 0, i.e. the ecosystem is no longer able to regenerate. Most of the literature does not include the exponential term as we do in this work, thus assuming that there exists a constant decay rate ε of the environment. We remark that the same result may be achieved in our model by setting $\varphi = 0$. The proposed modelisation is a generalisation of the more specific case featuring a constant decay rate.

For the sake of readability, we rewrite equations (3.4) and (3.5) in a more compact form throughout the rest of this section:

$$x_i = \alpha_i^x L_i^x \tag{3.7}$$

$$y_i = \alpha_i^y (1 - L_i^x) \tag{3.8}$$

where α_i^x and α_i^y are a synthetic measure of productivity of country i in sectors x and y, respectively, and are defined as:

$$\alpha_i^x = \frac{\rho_i^x}{1 + \eta_i^x P} \tag{3.9a}$$

$$\alpha_i^y = \frac{\rho_i^y}{1 + \eta_i^y P} \tag{3.9b}$$

3.2.1 Autarky solution

Before moving to the analysis of the two-country model with international trade, we here study the case of isolated markets. The results from this part will serve as a benchmark against the free trade case. The representative agent of country i maximises her utility choosing the optimal level of production, consumption, and labour share to allocate in the two sectors, subject to their budget constraints. By inserting equations (3.7) and (3.8) into the budget constraint (3.2), we derive the Lagrangian function for the agent of country i:

$$\mathscr{L}_i = (x_i^C)^\theta \, (y_i^C)^{1-\theta} - \lambda \left(p_x x_i^C + p_y y_i^C - p_x \alpha_i^x L_i^x - p_y \alpha_i^y (1 - L_i^x) \right) \tag{3.10}$$

We thus derive the following First Order Conditions (FOCs):

$$\frac{\partial \mathscr{L}_i}{\partial x_i^c} = 0 \implies \theta \left(\frac{y_i^c}{x_i^c} \right)^{1-\theta} = \lambda p_x \tag{3.11}$$

$$\frac{\partial \mathscr{L}_i}{\partial y_i^c} = 0 \implies (1-\theta) \left(\frac{y_i^c}{x_i^c} \right)^{-\theta} = \lambda p_y \tag{3.12}$$

$$\frac{\partial \mathscr{L}_i}{\partial L_i^x} = 0 \implies p_x \alpha_i^x = p_y \alpha_i^y \tag{3.13}$$

$$\frac{\partial \mathscr{L}_i}{\partial \lambda} = 0 \implies p_x x_i^c + p_y y_i^c = p_x \alpha_i^x L_i^x + p_y \alpha_i^y (1 - L_i^x) \tag{3.14}$$

Solving this system provides the following solutions:

$$\begin{aligned} x_i^c &= x_i = \alpha_i^x \theta \\ y_i^c &= y_i = \alpha_i^y (1-\theta) \\ L_i^x &= \theta \end{aligned}$$

It should be remarked that the amount produced and consumed by country i is not constant over time. Indeed, we recall that α_i^x and α_i^y represent the productivity of sectors x and y, respectively, and that both depend negatively on P. By contrast, the allocation of labour between the two sectors keeps constant, as L_i^x is always equal to θ, which is the preference toward good x and is not affected by the productivity coefficients α_i^x,α_i^y. As the environmental degradation increases, the amounts produced and consumed of the two goods decreases, until the system reaches a point in which $\dot{P} = 0$.

3.2.2 Free trade solution

We consider the case in which the technology of the two countries is such that they specialise on the production of different goods (see section D in the appendix, where we also conduct an analysis of other scenarios). In this case, the goods are traded at the international prices p_x, p_y. The representative agents of the two countries solve the same maximisation problem of the autarky case, but the optimal quantities of x_i^C, y_i^C, L_i^x for countries $i = N, S$ are interdependent. In Tables 3.1–3.3 we show the results of the optimisation, which is presented in detail in the appendix to this work. In particular, in Table 3.2 we see the value of all variables of interest when country N has a comparative advantage in the production of good x, as represented by the term $\delta \coloneqq \alpha_S^x \alpha_N^y - \alpha_N^x \alpha_S^y < 0$. Indeed, by rearranging the terms we see that:

$$\frac{\alpha_N^y}{\alpha_N^x} < \frac{\alpha_S^y}{\alpha_S^x} \tag{3.15}$$

Table 3.1: R_j, j=1,2,3, identify the different regimes for $\delta < 0$.

	$\theta < \frac{\alpha_N^y}{\alpha_N^y + \alpha_S^y} \quad (R_1)$	$\frac{\alpha_N^y}{\alpha_N^y + \alpha_S^y} \leq \theta \leq \frac{\alpha_N^x}{\alpha_N^x + \alpha_S^x} \quad (R_2)$	$\theta > \frac{\alpha_N^x}{\alpha_N^x + \alpha_S^x} \quad (R_3)$
x_N	$\theta \frac{\alpha_N^x}{\alpha_N^y}(\alpha_N^y + \alpha_S^y)$	α_N^x	α_N^x
y_N	$(1-\theta)\alpha_N^y - \theta\alpha_S^y$	0	0
x_S	0	0	$\theta\alpha_S^x - (1-\theta)\alpha_N^x$
y_S	α_S^y	α_S^y	$(1-\theta)\frac{\alpha_S^y}{\alpha_S^x}(\alpha_N^x + \alpha_S^x)$
x_N^c	$\theta\alpha_N^x$	$\theta\alpha_N^x$	$\theta\alpha_N^x$
y_N^c	$(1-\theta)\alpha_N^y$	$\theta\alpha_S^y$	$(1-\theta)\alpha_S^y \frac{\alpha_N}{\alpha_S^x}$
x_S^c	$\theta \frac{\alpha_N^x}{\alpha_N^y}\alpha_S^y$	$(1-\theta)\alpha_N^x$	$\theta\alpha_S^x$
y_S^c	$(1-\theta)\alpha_S^y$	$(1-\theta)\alpha_S^y$	$(1-\theta)\alpha_S^y$
L_N^x	$\theta \frac{\alpha_N^y + \alpha_S^y}{\alpha_N^y}$	1	1
L_S^x	0	0	$1-(1-\theta)\frac{\alpha_N^x + \alpha_S^x}{\alpha_S^x}$
$\dot{P}$	$\gamma_N^x L_N^x + \gamma_N^y(1-L_N^x) + \gamma_S^y - \varepsilon P e^{-\varphi P}$	$\gamma_N^x + \gamma_S^y - \varepsilon P e^{-\varphi P}$	$\gamma_N^x + \gamma_S^x L_S^x + \gamma_S^y(1-L_S^x) - \varepsilon P e^{-\varphi P}$

where the term to the left is the opportunity cost of producing good x in country N and the term to the right is the opportunity cost of producing x in S. This implies that if countries N and S start to trade, then the former would specialise (partly or completely) in the production of x, whereas S would specialise (partly or completely) in the production of y. Analogously, Table 3.1 reports the value of the main variables when the specialisations and the international trade pattern are reversed. This condition is characterised by $\delta > 0$ which can be rearranged so that the opposite of condition (3.15) holds. Finally, for the sake of completeness, we show in Table 3.3 the value of the variables when the two countries have the same opportunity cost and $\delta = 0$ and no specialisation occurs. We anticipate that this may represent a transitory regime between one specialisation regime and the other.

We note that three possible regimes are depicted in Tables 3.1 and 3.2, according to the productivity coefficients of countries. We recall that such coefficients depend (negatively) on environmental degradation P, whereas the preference parameter θ is fixed. As the environment degrades, productivity coefficients change value and the system may move from one regime to the

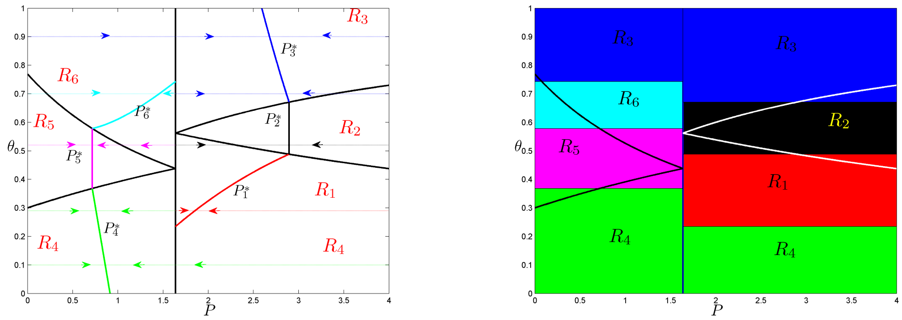
Table 3.2: R_j, j=4,5,6 identifies the different regimes for $\delta > 0$.

	$\theta < \dfrac{\alpha_S^y}{\alpha_N^y + \alpha_S^y} \quad (R_4)$	$\dfrac{\alpha_S^y}{\alpha_N^y + \alpha_S^y} \leq \theta \leq \dfrac{\alpha_S^x}{\alpha_N^x + \alpha_S^x} \quad (R_5)$	$\theta > \dfrac{\alpha_S^x}{\alpha_N^x + \alpha_S^x} \quad (R_6)$
x_N	0	0	$\theta(\alpha_N^x + \alpha_S^x) - \alpha_S^x$
y_N	α_N^y	α_N^y	$(1-\theta)\dfrac{\alpha_N^y}{\alpha_N^x}(\alpha_N^x + \alpha_S^x)$
x_S	$\theta\dfrac{\alpha_S^x}{\alpha_S^y}(\alpha_N^y + \alpha_N^y)$	α_S^x	α_S^x
y_S	$(1-\theta)\alpha_S^y - \theta\alpha_N^y$	0	0
x_N^c	$\theta\dfrac{\alpha_S^x}{\alpha_S^y}\alpha_N^y$	$(1-\theta)\alpha_S^x$	$\theta\alpha_N^x$
y_N^c	$(1-\theta)\alpha_N^y$	$(1-\theta)\alpha_N^y$	$(1-\theta)\alpha_N^y$
x_S^c	$\theta\alpha_S^x$	$\theta\alpha_S^x$	$\theta\alpha_S^x$
y_S^c	$(1-\theta)\alpha_S^y$	$\theta\alpha_N^y$	$(1-\theta)\dfrac{\alpha_S^x}{\alpha_N^x}\alpha_N^y$
L_N^x	0	0	$1-(1-\theta)\dfrac{\alpha_N^x + \alpha_S^x}{\alpha_N^x}$
L_S^x	$\theta\dfrac{\alpha_N^y + \alpha_S^y}{\alpha_S^y}$	1	1
$\dot{P}$	$\gamma_N^y + \gamma_S^x L_S^x + \gamma_S^y(1-L_S^x) - \varepsilon P e^{-\varphi P}$	$b + c - \varepsilon P e^{-\varphi P}$	$\gamma_N^x L_N^x + \gamma_N^y(1-L_N^x) + \gamma_S^x - \varepsilon P e^{-\varphi P}$

other. The central column represents the case of complete specialisation in both tables (as can be seen from the labour share allocation), whereas the side columns represent the cases of incomplete specialisation. As the results shown in Table 3.3 relate to the value of the variables of interest for a specific value of P (the one for which $\delta = 0$), there are no multiple cases represented.

Table 3.3: The regime under $\delta = 0$.

x_N	y_N	x_S	y_S	x_N^c	y_N^c	x_S^c	y_S^c
$\alpha_N^x L_N^x$	$\alpha_N^y L_N^x$	$\alpha_S^y(1-L_S^x)$	$\alpha_S^y(1-L_S^x)$	$\theta\alpha_N^x$	$(1-\theta)\alpha_N^y$	$\theta\alpha_S^x$	$(1-\theta)\alpha_S^y$
$L_N^{x\,*} = \theta\dfrac{\alpha_N^x + \alpha_S^x}{\alpha_N^x} - \dfrac{\alpha_S^x}{\alpha_N^x}L_S^{x\,*}, \quad \max\left(0, \theta\dfrac{\alpha_N^x + \alpha_S^x}{\alpha_S^x} - \dfrac{\alpha_S^x}{\alpha_S^x}\right) < L_S^{x\,*} < \min\left(1, \theta\dfrac{\alpha_N^x + \alpha_S^x}{\alpha_S^x}\right)$							



(a) (b)

Figure 3.1: Both figures inform on the regime in which the system lies according to the values of θ and P. The blue vertical lines identify the value of P for which a specialisation reversal occurs. Parameter set: $\rho_N^x = 0.3, \rho_N^y = 0.7, \rho_S^x = 1, \rho_S^y = 0.3, \eta_N^x = 0, \eta_N^y = 0.5, \eta_S^x = 2, \eta_S^y = 0, \gamma_N^x = 0.8, \gamma_N^y = 0, \gamma_S^x = 0.4, \gamma_S^y = 0.5, \varepsilon = 0.6, \varphi = 0.1$.

3.3 Regime change

We now study the linkage between trade patterns and the degradation of the environment when N is better off selling x and importing y whereas the reverse applies for S. We note that the different regimes that the system may reach are characterised by the conditions reported at the top of Tables (3.1) and (3.2). By recalling that the productivities α_i^x and α_i^y of country i depend on the current level of environmental degradation P (see equations (3.9a) and (3.9b)), we may see how different specialisation regimes hold as P varies. For instance, suppose $\eta_S^y > \eta_N^y$ and the system is currently in regime R_4, which describes an incomplete specialisation of S, whereas N fully specialises on sector y. If P increases, then $\frac{\alpha_S^y}{\alpha_S^y+\alpha_N^y}$ eventually falls below θ and the system reaches regime R_5, i.e. full specialisation for both countries. Subsequent changes in specialisation regimes would depend on the relative magnitude of η_S^x and η_N^x and on the dynamics of P.

In Figures 3.1a and 3.1b we illustrate how the initial level of environmental degradation and the level of preferences θ determine the stationary state the system reaches, i.e. a state in which $\dot{P} = 0$. In particular, Figure 3.1a shows the attraction basins of each stationary point and the ones the system converges to, each denoted by P_k^*, where k is the regime in which it is contained. For instance, we note that for a value of $\theta = 0.7$, when the environment is not yet much degraded (P is to the left of the vertical threshold) the system lies in regime R_6 which represents the case of complete specialisation of S in the production of x. In this case, the system converges

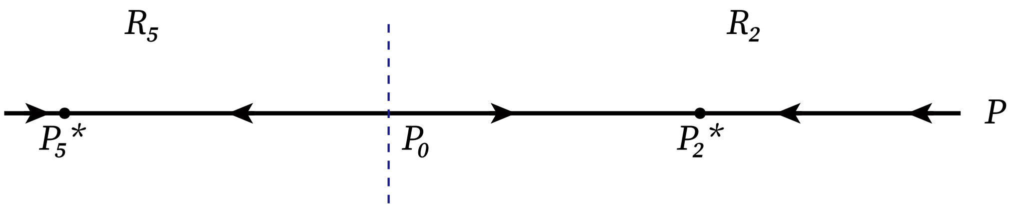


Figure 3.2: We here show an occurrence of bistability: according to the initial condition of P, the system may reach the stationary state P_5^* in regime R_5 (where N specialises in the production of y and S is in the production of x) or the stationary state P_2^* in regime R_2 (where the specialisations are inverted). If the system starts to the right of a threshold P_0, it will pass regimes R_6, R_3, and R_2 before reaching the stationary state P_2^*.

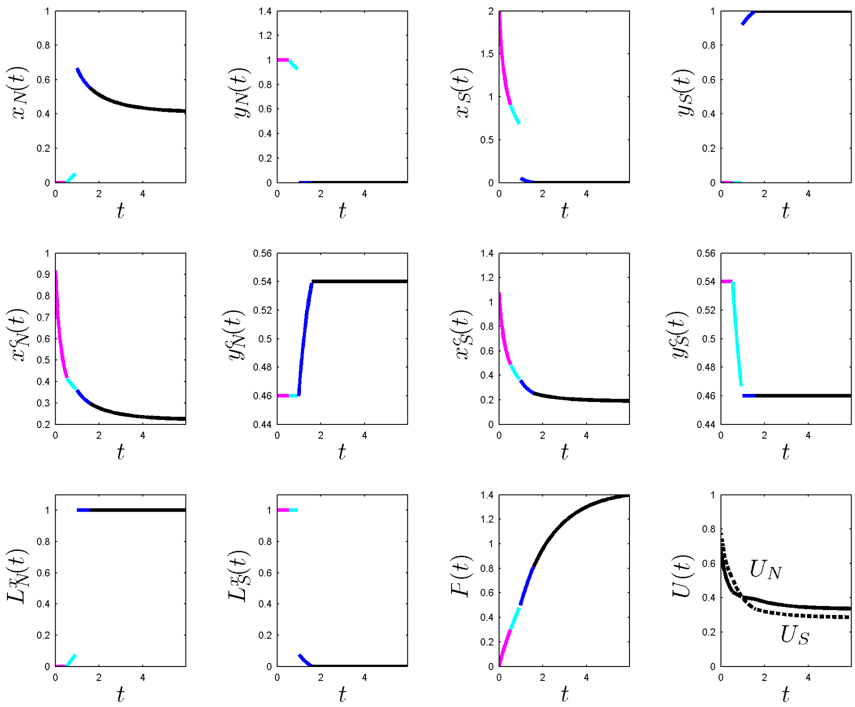
to attractor P_6^* and no comparative advantage reversal occurs. For the same value of θ, we see that if the initial environmental degradation is beyond the vertical threshold, the system converges to another stationary state, i.e. P_3^*, which lies in regime R_3 identifying the opposite case in which country N specialises in the production of x. To which of the two states the system converges thus depends on the initial level of environmental degradation. This is a case of bistability, similar to the one represented in Figure 3.2, in which the case for $\theta = 0.58$ is analysed. As can be seen from this figure, two stationary states exist: P_5^* and P_2^*, whose basins of attraction are separated by a threshold value of environmental degradation equal to P_0. In this case, the dynamics are path-dependent, and the system may thus reach either of the two stationary states, according to the initial value of P. It might also happen that a single attractive stationary state exists, for instance when $\theta = 0.9$. In this case, even when environmental degradation is initially below the threshold level (under regime R_6), the system eventually converges to P_3^*, in regime R_3. If the threshold is crossed in the process, a comparative advantage reversal occurs. We provide a graphical illustration of which specialisation regime the system reaches, for each combination of the preference parameter and of the initial environmental degradation, in Figure 3.1b. In both figures, each colour identifies a specialisation regime.

3.3.1 Dynamics of the main economic variables

The main contribution of this work is to show that comparative advantages may not be time invariant, due to environmental degradation asymmetrically affecting the productivity of less developed countries with respect to more developed ones. Let us assume that a less developed country S has an advantage in the agricultural sector, for instance x, whose production is also relatively more vulnerable to environmental degradation with respect to the one of the more developed country, N. This part of the story is coherent with the documented higher exposure of less developed countries to natural disasters (Strömberg, 2007) and their reduced capability to cope with environmental distress due to the lack of credit access, weaker institutions, and lower average educational attainment (Toya and Skidmore, 2007; Barbier, 2010). In a scenario in which production of agricultural goods is concentrated in the less developed countries, the productivity of the latter is doomed to decrease in time due to the ensuing environmental degradation. This leads to a reduction in the productivity of less developed countries in this sector, or equivalently an increase in costs. As the relative cost to produce agricultural goods increases for S (and thus decreases for N), the comparative advantage of the two countries might reverse, with the production of the non-agricultural goods now being relatively cheaper for the less developed country.

In Fig. 3.3 we illustrate a simulation of the above scenario, in which each colour represents a different specialisation regime, which can be retrieved from Tables 3.2 and 3.1. In particular, under R_5, which is depicted in magenta, there is a complete specialisation of the two countries, with N and S producing exclusively good y and x, respectively. As environmental degradation accumulates, the relative cost of producing x in the less developed country increases leading the countries to regime R_6, which is shown in cyan in the figure and describes a scenario of incomplete specialisation for N. Indeed, N keeps being the exclusive producer of good y, although it also produces some quantity of good x, which is also the only good produced by S. This is apparent also by noting that in Fig. 3.3 we can observe that under R_6 there is a continuous increase in L_N^x, i.e. the labour share of country N devoted to the x-sector, whereas no change occurs in L_S^x. As the environmental degradation further accumulates, the relative cost of producing y in N becomes so high that a reversal in comparative advantages takes place, and the two countries go under regime R_3, shown in blue. In this situation, S is now the only producer of good y, although it still devotes part of its labour share to the production of x. By contrast, N specialises in the production of x and assigns no share of its labour force to the production of y. In fact, we could identify a specific moment in time in which the two countries have the same relative costs for the goods and hence no comparative advantage exists. However, asymmetrical vulnerability to environmental degradation soon breaks the balance and the relative costs of the two countries diverge again, leading to the formation of comparative advantages. Finally, for high lev-

Figure 3.3: Dynamics of the main variables. Parameter set: $\theta = 0.58, \rho_N^x = 1, \rho_N^y = 1, \rho_S^x = 2, \rho_S^y = 1, \eta_N^x = 1, \eta_N^y = 0, \eta_S^x = 4, \eta_S^y = 0, \gamma_N^x = 1, \gamma_N^y = 0, \gamma_S^x = 0.1, \gamma_S^y = 0, \varepsilon = 0.8, \varphi = 0.1$. Initial condition $P(0) = 0$. Colours magenta, cyan, blue, black identify regimes R_5, R_6, R_3, R_2.



els of environmental degradation the two countries come to be under regime R_2, shown in black, in which they both fully specialise their production. In particular, country N produces only good x whereas country S produces only good y. No more regime changes occur and eventually the system sets on an equilibrium level in which the quantities produced and consumed depend on the environmental degradation P^*. Finally, we should note that as time passes and P increases, the well-being of agents from both countries decreases.

3.3.2 Adverse effects of international trade with comparative advantage reversal

We already argued that when the country that is the least pollution-intensive is also the one whose productivity is most exposed to environmental degradation there is the possibility that comparative advantages reverse. Intuitively, this means that if at an initial state the production of the two countries is externality-wise optimal (minimises emission), when the comparative advantage reversal occurs the opposite applies. In this case, the countries specialise in a way that the production is inefficient in terms of environmental externalities. We now expand on this consideration, highlighting that when this inefficiency is a result of a comparative advantage reversal, it is possible that the utility of all agents falls below the autarky level. The increased amount of externalities due to the reversal decreases the productivity of both countries and the overall quantities produced (and consumed) of the affected goods. In Figure 3.3 we may see how a reduction in the quantity consumed of x by agents of both countries leads to a decrease in the utility. If this reduction is sufficiently large, the utility might fall below the autarky level. In Table 3.4 we show an instance of this occurrence, using the same parameter set of Figure 3.1a and setting the preference parameter θ to 0.9.

	x_N	y_N	x_S	y_S	L_x^N	L_x^S	U_N	U_S	P^*
International Trade	0.27	0.2694	0.1556	0.27	0.9	0.9	0.2699	0.1644	2.3924
Autarky	0.3	0	0.1116	0.0872	1	0.7094	0.2312	0.1213	2.6773

Table 3.4: Value of the main economic variables in case of autarky and international trade. Parameter set: $\rho_N^x = 0.3, \rho_N^y = 0.7, \rho_S^x = 1, \rho_S^y = 0.3, \eta_N^x = 0, \eta_N^y = 0.5, \eta_S^x = 2, \eta_S^y = 0, \gamma_N^x = 0.8, \gamma_N^y = 0, \gamma_S^x = 0.4, \gamma_S^y = 0.5, \varepsilon = 0.6, \varphi = 0.1$, $\theta = 0.9$.

A comparison with Figure 3.1a allows to check that for low values of environmental degradation P, the system lies in regime R_6 and that as the environment degrades, it converges to regime R_3 going through a comparative advantage reversal. The production of both goods, the labour allocation, the utility level, and the environmental degradation for the corresponding stationary state is shown in the first row of Table 3.4. In the second row, we present the value of the same variables in case of autarky, i.e. no trade between N and S. As can be seen, in the scenario presented agents from both countries would have been better off in autarky. This result serves as an alarm on the risks of underestimating the relevance of environmental degradation on the productivity of countries and thus on the international trade patterns.

3.4 Discussion and conclusions

In this chapter, we analysed a comparative advantage model in which the productivity of two sectors depends negatively on current environmental degradation. In this framework, we highlighted that asymmetrical vulnerability of two countries to environmental degradation may result in a reversal of the comparative advantages. We showed that higher vulnerability to environmental degradation may lower the productivity in the sector in which a country has a comparative advantage, to the point that product specialisation is reversed. We sketched an intuition of the argument and then provided an illustrative example of the conditions underlying this phenomenon. In our simulation, a less developed country S goes from a complete specialisation in the production of x to a complete and opposite specialisation in the production of y, passing through periods of incomplete specialisation. The reverse applies to a more developed country N. This model suggests that forms of environmental degradation such as soil erosion, deforestation, mineralisation may in time change the patterns of international trade, affecting the comparative advantages of countries in the economic activities that are most affected. We showed that when the comparative advantage reversal drastically increases environmental degradation, it may lead to a significant fall in utility, even below the autarky level.

We designed this model to be the simplest possible representation of the specialisation reversal that may arise from environmental degradation. However, meaningful extensions may be conceived. For instance, we assumed invariance in both productivity and environmental impact coefficients. If the Environmental Kuznets Curve adagio were correct, economic growth should also bring about reductions in environmental degradation. In this model, this could be implemented by assuming that environmental impact coefficients are negatively related to the average *wealth* of agents. This could counter and even prevent the specialisation reversal phenomenon here outlined. In addition, introducing the possibility for countries to implement new policies at the national or supranational level, e.g. a cap on emission intensity, could lead to an analogous offsetting of the specialisation reversal drivers. However, the extent to which a large coalition can be maintained on environmental related matters is still uncertain, as the recent unfolding of international politics is showing. Moreover, we believe that studying this subject from a different approach might provide additional insight. For instance, the authors deem that heterogeneity of agents at the national level might prove particularly meaningful, which can be done notably by employing Agent Based Models. Indeed, taking into account that different regions are affected by environmental degradation to varying extent might prove a valuable contribution to a theoretical description of the mechanisms generating "the winners and the losers" of international trade. Indeed, researchers should not forget that international trade (and its environmental consequences) is subject to political debate when investigating the future directions of international trade patterns, lest reality may catch them off guard.

Appendix D

Study of the domains of the model

In this appendix, we show that the specialisation of the two countries N and S, i.e. their chosen allocation of labour force to the production of goods x and y, depends on the term $\Phi_i := \alpha_i^x p_x - \alpha_i^y p_y$. We recall that this term measures the comparative advantage of country i in the production of good x, as discussed in section 3.2.2.

In the free trade case, the maximisation problem is:

$$\max_{L_i^x, x_i^C, y_i^C} (x_i^C)^\theta (y_i^C)^{(1-\theta)} \tag{D.1}$$

subject to the budget constraint (D.2):

$$p_x \alpha_i^x L_i^x + p_y \alpha_i^y (1 - L_i^x) = p_x x_i^C + p_y y_i^C \tag{D.2}$$

and to the market clearing conditions (D.3):

$$x_N + x_S = x_N^C + x_S^C \tag{D.3a}$$

$$y_N + y_S = y_N^C + y_S^C \tag{D.3b}$$

In order to solve this maximisation problem, we firstly maximise the Lagrangian function (3.10) with respect to L_i^x, only taking into account the budget constraint (D.2). This will yield the optimal value of $L^{\circ}{}_i^x$ for countries $i = N, S$ and the related Lagrangian function $\mathscr{L}^{\circ}{}_i$. We employ this notation in the next section of this appendix.

We may rewrite the Lagrangian function to highlight its relation with the term Φ_i:

$$\mathscr{L}_i(x_i^C, y_i^C, L_i^x; \lambda) = (x_i^C)^\theta \ (y_i^C)^{1-\theta} + \lambda \left(\Phi_i L_i^x + p_y \alpha_i^y - p_x x_i^C - p_y y_i^C\right) \tag{D.4}$$

It is thus apparent that the sign of Φ_i is highly informative with respect to optimal labour allocation. A positive value of the term Φ_i is a signal that the representative agent from country i is able to gain more in the international market by trading off units of x in exchange for y. In this case, the optimal labour share $L^{\circ x}_i$ is equal to its upper limit of 1. The reverse applies when $\Phi_i < 0$: the representative agent of country i devotes all her labour to the production of good y, so that $L^{\circ x}_i = 0$. Formally, we write this relation in the following way:

$$L^{\circ x}_i = \begin{cases} 0, & \text{if } P \in \Delta^1_i = \{P : P \in \Phi_i < 0\} \\ L^{*x}_i \in (0,1), & \text{if } P \in \Delta^2_i = \{P : P \in \Phi_i = 0\} \\ 1, & \text{if } P \in \Delta^3_i = \{P : P \in \Phi_i > 0\} \end{cases} \tag{D.5}$$

where we emphasize that Φ_i depends on the current level of environmental degradation P. We also note that in case $\Phi_i = 0$, the allocation of labour is equal to the share L^{*x}_i from the autarky case. Since these three domains are defined for both countries, our two-country model has a total of nine domains over which the values L^x_N and L^x_S are defined:

$$\Delta^{(s,k)}_{NS} = \{P : P \in \Delta^s_N \cap \Delta^k_S\}, \qquad s,k = 1,2,3 \tag{D.6}$$

D.1 The domains $\Delta^{(1,1)}_{NS}$, $\Delta^{(2,2)}_{NS}$, $\Delta^{(3,3)}_{NS}$, $\Delta^{(3,1)}_{NS}$

Since a full study of all domains is beyond the scope of this appendix, we here present the study of the domains relating to the most salient characterisations: $\Delta^{(1,1)}_{NS}$, $\Delta^{(2,2)}_{NS}$, $\Delta^{(3,3)}_{NS}$, $\Delta^{(3,1)}_{NS}$. In order to illustrate our choice of domains, we anticipate their interpretation. Firstly, the domains $\Delta^{(1,1)}_{NS}$ and $\Delta^{(3,3)}_{NS}$ represent the cases in which both countries specialise on the production of the same good: y in the former case and x in the latter. Secondly, the domain $\Delta^{(2,2)}_{NS}$ represents the case in which both countries have the same opportunity costs and thus do not benefit from international trade. We consider this scenario non-robust and, most importantly, not descriptive of the dichotomy between more developed countries and less developed ones presented in this work. Finally, the domain $\Delta^{(3,1)}_{NS}$ represents a case where countries have comparative advantages in different goods and may thus benefit from international trade. We remark that this case is symmetrical with respect to the one of domain $\Delta^{(1,3)}_{NS}$ whose analysis would thus yield analogous results. We underline that the other domains may be studied in a similar way to the ones here analysed. On a methodological note, we recall that the Lagrangian functions in the following analysis are the result of a precedent maximisation in L^x_i, with $L^{\circ x}_i$ representing such maximum point.

1. **Domain** $\Delta_{NS}^{(1,1)}$: $\Phi_N < 0$, $\Phi_S < 0$ and $L^{\circ x}_N = 0$, $L^{\circ x}_S = 0$.

   In this scenario, representative agents from both countries N, S have a comparative advantage in the production of good y, leading them to allocate all their labour force to such sector. We derive from the quantities produced of the two goods from the production functions (3.7) and (3.8): $x_i = 0$, $y_i = \alpha_i^y$ with $i = N, S$. The resulting Lagrangian function is:

   $$\mathscr{L}_i = (x_i^C)^\theta \; (y_i^C)^{1-\theta} + \lambda \left(p_y \alpha_i^y - p_x x_i^C - p_y y_i^C \right) \tag{D.7}$$

   We may thus find the optimal consumption values by applying the first order conditions, from which we obtain that $x_i^{C*} = \theta \dfrac{p_y}{p_x} \alpha_i^y$ and $y_i^{C*} = (1-\theta)\alpha_i^y$. Replacing them in the Market Clearing Conditions (3.3) we derive that $\theta \dfrac{p_y}{p_x}(\alpha_N^y + \alpha_S^y) = 0$. Such equation cannot be satisfied since all parameters are assumed to be strictly positive. In other terms, under this domain the MCC cannot be satisfied, as both countries are specialised in the production of the same good and no unit of x is produced.

2. **Domain** $\Delta_{NS}^{(3,3)}$: $\Phi_N > 0$, $\Phi_S > 0$ and $L^{\circ x}_N = 1$, $L^{\circ x}_S = 1$.

   This scenario is specular to the previous one and represents the case in which both countries have a comparative advantage in the production of good x. Again, we derive the quantities produced of the two goods from the production functions (3.7) and (3.8): $x_i = \alpha_i^x$, $y_i = 0$ with $i = N, S$. The corresponding Lagrangian function is given by:

   $$\mathscr{L}_i = (x_i^C)^\theta \; (y_i^C)^{1-\theta} + \lambda \left(p_x \alpha_i^x - p_x x_i^C - p_y y_i^C \right) \tag{D.8}$$

   First order conditions yield that $x_i^{C*} = \theta \alpha_i^x$ and $y_i^{C*} = (1-\theta)\dfrac{p_x}{p_y}\alpha_i^y$, which we replace in the Market Clearing Conditions (3.3) to derive that $(1-\theta)\dfrac{p_x}{p_y}(\alpha_N + \alpha_S) = 0$. Also this equation cannot be satisfied, since all parameters are assumed to be strictly positive. Analogously to the previous case, this domain is impossible as it describes a situation in which both countries have a comparative advantage in the production of good x.

3. **Domain** $\Delta_{NS}^{(2,2)}$: $\Phi_N = 0$, $\Phi_S = 0$ and $L^{\circ x}_N = L_N^{x}{}^*$, $L^{\circ x}_S = L_S^{x*}$.

   In this case both countries have no comparative advantage, reflecting a situation in which the representative agents face the same opportunity cost when choosing where to allocate their labour force. From $\Phi_N = 0$ and $\Phi_S = 0$, we obtain that $\dfrac{p_x}{p_y} = \dfrac{\alpha_N^y}{\alpha_N} = \dfrac{\alpha_S^y}{\alpha_S}$, which implies that $\delta := \alpha_S^x \alpha_N^y - \alpha_N^x \alpha_S^y = 0$. We have seen in section 3.2.1 that an analogous condition applies in the autarky case. We can thus apply the analysis performed on the autarky case. We thus find the consumption values x_i^C, y_i^C are:

$$x_N^C = \theta\alpha_N^x \qquad y_N^C = (1-\theta)\alpha_N^y \tag{D.9a}$$
$$x_S^C = \theta\alpha_S^x \qquad y_S^C = (1-\theta)\alpha_S^y \tag{D.9b}$$

   By substituting (D.9) into the MCCs (3.3) we can rewrite the latter as:

$$\alpha_N^x L_N^{x}{}^* + \alpha_S^x L_S^{x*} = \theta(\alpha_N^x + \alpha_S^x) \tag{D.10a}$$
$$\alpha_N^y L_N^{x}{}^* + \alpha_S^y L_S^{x*} = \theta(\alpha_N^y + \alpha_S^y) \tag{D.10b}$$

   From $\delta = 0$, and $L_i^{x*} \in (0,1)$, by means of straightforward calculations, we derive the solution of the previous system as:

$$\max\left(0, \theta\frac{\alpha_N^x + \alpha_S^x}{\alpha_N^x} - \frac{\alpha_S^x}{\alpha_N^x}\right) < L_S^{x*} < min\left(1, \theta\frac{\alpha_N^x + \alpha_S^x}{\alpha_N^x}\right) \tag{D.11a}$$
$$L_N^{x}{}^* = \theta\frac{\alpha_N^x + \alpha_S^x}{\alpha_N^x} - \frac{\alpha_S^x}{\alpha_N^x} L_S^{x*} \tag{D.11b}$$

   These results are then summarised in Table 3.3. In this scenario, representative agents from the two countries are indifferent toward all sets of consumption bundles satisfying the above conditions.

4. **Domain** $\Delta_{NS}^{(3,1)}$: $\Phi_N > 0$, $\Phi_S < 0$ and $L^{\circ x}_N = 1$, $L^{\circ x}_S = 0$.

   In the last domain we present, the two country have a comparative advantage in different goods. In this specific case, N has a comparative advantage on the production of good x, while S has a comparative advantage in the production of y. We underline that the domain $\Delta_{NS}^{(1,3)}$ represents a specular case in which the comparative advantages of the countries are reversed. From the optimal labour allocation of this domain and from production functions (3.7) and (3.8), it follows that the quantities produced are $x_N = \alpha_N^x$, $y_N = 0$, $x_S = 0$, and $y_S = \alpha_S^y$. Therefore, the quantities consumed in each countries are given by:

$$x_N^C = \theta\alpha_N^x, \qquad y_N^C = (1-\theta)\frac{p_x}{p_y}\alpha_N^x \tag{D.12a}$$

$$x_S^C = \theta\frac{p_y}{p_x}\alpha_S^y, \qquad y_S^C = (1-\theta)\alpha_S^y \tag{D.12b}$$

Therefore, the MCCs (3.3) can be rewritten as:

$$(1-\theta)\alpha_N^x = \theta\frac{p_y}{p_x}\alpha_S^y \tag{D.13a}$$

$$\theta\alpha_S^y = (1-\theta)\frac{p_x}{p_y}\alpha_N^x \tag{D.13b}$$

from which we derive that:

$$\frac{p_x}{p_y} = \frac{\theta}{1-\theta}\frac{\alpha_S^y}{\alpha_N^x} \tag{D.14}$$

Since in this domain $\Phi_N > 0$ and $\Phi_S < 0$, we state that the following relation must hold:

$$\frac{\alpha_N^y}{\alpha_S^y} < \frac{\theta}{1-\theta} < \frac{\alpha_N^x}{\alpha_S^x} \tag{D.15}$$

which is satisfied only if $\delta \coloneqq \alpha_S\alpha_N^y - \alpha_N\alpha_S^y < 0$. Rearranging the terms, we obtain:

$$\frac{\alpha_N^y}{\alpha_N^y + \alpha_S^y} < \theta < \frac{\alpha_N^x}{\alpha_N^x + \alpha_S^x}$$

We summarise the results from the study of this domain in Table 3.1, whereas the specular case in which $\Phi_N < 0$ and $\Phi_S > 0$ leads to the results presented in Table 3.2.

Bibliography

Antweiler, Werner (1996) "The pollution terms of trade," *Economic Systems Research*, Vol. 8, pp. 361–366.

Asako, Kazumi (1979) "Environmental pollution in an open economy," *Economic Record*, Vol. 55, pp. 359–367.

Barbier, Edward B (2010) "Poverty, development, and environment," *Environment and Development Economics*, Vol. 15, pp. 635–660.

Broner, Fernando, Paula Bustos, and Vasco M Carvalho (2012) "Sources of comparative advantage in polluting industries,"Technical report, National Bureau of Economic Research.

Bruckner, Martin, Stefan Giljum, Christian Lutz, and Kirsten Svenja Wiebe (2012) "Materials embodied in international trade–Global material extraction and consumption between 1995 and 2005," *Global Environmental Change*, Vol. 22, pp. 568–576.

Candau, Fabien and Elisa Dienesch (2017) "Pollution haven and corruption paradise," *Journal of Environmental Economics and Management*, Vol. 85, pp. 171–192.

Cole, Matthew A (2004) "Trade, the pollution haven hypothesis and the environmental Kuznets curve: examining the linkages," *Ecological Economics*, Vol. 48, pp. 71–81.

Copeland, Brian R and M Scott Taylor (2004) "Trade, growth, and the environment," *Journal of Economic Literature*, Vol. 42, pp. 7–71.

Duan, Yuwan and Xuemei Jiang (2017) "Temporal change of China's pollution terms of trade and its determinants," *Ecological Economics*, Vol. 132, pp. 31–44.

Gallagher, Kevin (1999) "Pollution intensive industry in Mexico under NAFTA: model and empirical evidence," *Praxis*, Vol. 15, pp. 1–15.

——— (2000) *Trade Liberalization and Industrial Pollution in Mexico: Lessons of the FTAA*, Vol. 00-07: Tufts University, Boston, Massachussetts, USA.

Garsous, Grégoire and Tomasz Kozluk (2017) "Foreign Direct Investment and The Pollution Haven Hypothesis: Evidence from Listed Firms," *OECD Economic Department Working Papers*, p. 0_1.

Grether, Jean-Marie and Nicole A Mathys (2013) "The pollution terms of trade and its five components," *Journal of Development Economics*, Vol. 100, pp. 19–31.

Hazell, Peter and Stanley Wood (2008) "Drivers of change in global agriculture," *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, Vol. 363, pp. 495–515.

Kellenberg, Derek K (2009) "An empirical investigation of the pollution haven effect with strategic environment and trade policy," *Journal of International Economics*, Vol. 78, pp. 242–255.

Kozluk, Tomasz and Christina Timiliotis (2016) "Do environmental policies affect global value chains?: A new perspective on the pollution haven hypothesis," *OECD Economic Department Working Papers*, p. 0_1.

Millimet, Daniel L and Jayjit Roy (2016) "Empirical tests of the pollution haven hypothesis when environmental regulation is endogenous," *Journal of Applied Econometrics*, Vol. 31, pp. 652–677.

Muradian, Roldan, Martin O'Connor, and Joan Martinez-Alier (2002) "Embodied pollution in trade: estimating the 'environmental load displacement'of industrialised countries," *Ecological Economics*, Vol. 41, pp. 51–67.

Pelletier, Nathan and Peter Tyedmers (2010) "Forecasting potential global environmental costs of livestock production 2000–2050," *Proceedings of the National Academy of Sciences*, Vol. 107, pp. 18371–18374.

Peng, Shuijun, Wencheng Zhang, and Chuanwang Sun (2016) "'Environmental load displacement' from the North to the South: A consumption-based perspective with a focus on China," *Ecological Economics*, Vol. 128, pp. 147 – 158.

Phélinas, Pascale and Johanna Choumert (2017) "Is GM soybean cultivation in Argentina sustainable?" *World Development*, Vol. 99, pp. 452–462.

Sapkota, Pratikshya and Umesh Bastola (2017) "Foreign direct investment, income, and environmental pollution in developing countries: Panel data analysis of Latin America," *Energy Economics*, Vol. 64, pp. 206–212.

Scherr, Sara J (1999) *Soil degradation: a threat to developing-country food security by 2020?*, Vol. 27: Intl Food Policy Res Inst.

Scoppola, M. and V. Raimondi (2018) "Foreign Land Acquisitions and Environmental Regulations: does the Pollution-Haven Hypothesis hold?," 2018 Conference, July 28-August 2, 2018, Vancouver, British Columbia 277098, International Association of Agricultural Economists.

Strömberg, David (2007) "Natural disasters, economic development, and humanitarian aid," *Journal of Economic Perspectives*, Vol. 21, pp. 199–222.

Tang, Jitao (2015) "Testing the pollution haven effect: Does the type of FDI matter?" *Environmental and Resource Economics*, Vol. 60, pp. 549–578.

Tisdell, Clement A (1999) "Economics, aspects of ecology and sustainable agricultural production," in *Sustainable Agriculture and Environment*: Edward Elgar, Cheltenham, GB, pp. 37–56.

Toya, Hideki and Mark Skidmore (2007) "Economic development and the impacts of natural disasters," *Economics Letters*, Vol. 94, pp. 20–25.

Wilson, Clevo and Clem Tisdell (2001) "Why farmers continue to use pesticides despite environmental, health and sustainability costs," *Ecological Economics*, Vol. 39, pp. 449–462.

World Bank (2003) *World Development Report 2003: Sustainable Development in a Dynamic World: Transforming Institutions, Growth, and Quality of Life*: Oxford University Press, New York, DC, USA.

Zugravu-Soilita, Natalia (2017) "How does foreign direct investment affect pollution? Toward a better understanding of the direct and conditional effects," *Environmental and Resource Economics*, Vol. 66, pp. 293–338.