The Dynamic Impact of Unilateral Environmental Policies✩

David Hémous*

University of Zurich, INSEAD and CEPR

Abstract

This paper builds a two-country, two-sector (polluting, nonpolluting) trade model with directed technical change, examining whether unilateral environmental policies can ensure sustainable growth. The emission rate of the polluting sector depends on its relative use of a clean and a dirty input. A unilateral policy combining clean research subsidies and a trade tax can ensure sustainable growth, while unilateral carbon taxes alone increase innovation in the polluting sector abroad and generally cannot ensure sustainable growth. Relative to autarky and exogenous technical change respectively, trade and directed technical change accelerate environmental degradation either under laissez-faire or with unilateral carbon taxes, yet both help reduce environmental degradation under the appropriate unilateral policy. I characterize the optimal unilateral policy analytically and numerically using calibrated simulations. Knowledge spillovers have the potential to reduce the otherwise large welfare costs of restricting policy to one country.

Key words:
climate change, environment, directed technical change, innovation, trade, unilateral policy

JEL: F18, F42, F43, O32, O33, O41, Q54, Q55

✩Previous versions were circulated under the titles "Environmental Policy and Directed Technical Change in a Global Economy: Is There a Case for Carbon Tariffs?" and "Environmental Policy and Directed Technical Change in a Global Economy: The Dynamic Impact of Unilateral Environmental Policies."

*University of Zürich, Schönberggasse, 1, 8050 Zurich, Switzerland. +41 44 634 61 18
Email address: david.hemous@econ.uzh.ch (David Hémous)

URL: http://www.econ.uzh.ch/en/people/faculty/hemous.html (David Hémous)
1. Introduction

Despite the signature of the Kyoto Protocol in 1997, annual carbon dioxide (CO$_2$) emissions increased by 39% between 1997 and 2010. Meanwhile, climate negotiations have stalled: the Paris Agreement of 2015, did not set binding targets for CO$_2$ reductions (contrary to Kyoto), instead, each country is supposed to state their own target. This is line with the historical trend since the late 90s as several countries have undertaken unilateral environmental policies with varying degrees of ambition and success. For instance, the European Union implemented a cap-and-trade system (EU ETS) in 2005 which covers around 45% of the EU’s greenhouse gas emissions. However, such policies generate a pollution haven effect, as the production of tradable and polluting goods moves to countries with laxer policies, which leads to an increase in their emissions. Could unilateral policies nevertheless achieve the necessary reduction in CO$_2$ emissions? If so, how should they be designed? These questions are fundamentally about the economy’s long-run behavior. Over the time period relevant to climate change, comparative advantages evolve with innovation, which itself responds to environmental policies. Yet, the economic literature on unilateral environmental policies has largely ignored the role played by innovation. This paper builds a trade model featuring directed technical change and a global pollution externality, and thereby highlights the crucial role that innovation plays in the positive and normative analysis of unilateral climate change policies.

Formally, I consider a dynamic Ricardo–Heckscher–Ohlin model with two countries, North and South, and two sectors, polluting and nonpolluting. The North represents countries willing to implement an environmental policy, and the South, countries that are not—a division which need not fall along the lines of developed versus developing countries. The polluting good represents the tradable goods with a high CO$_2$ emission intensity, typically energy-intensive sectors. It is produced using clean inputs (e.g., renewable and nuclear energy or bioplastics) and/or dirty inputs (e.g., fossil fuel energy or traditional petroleum products). Innovation is undertaken in both countries by profit-maximizing firms that hire scientists. It can be directed at the nonpolluting sector, or, within the polluting sector, at clean or dirty technologies. For most of the analysis, innovation is local.

In laissez-faire, the allocation of innovation favors the exporting sector
and therefore reinforces comparative advantage over time. This results from a market size effect: a country exports the good that it produces relatively more, such that the market for innovation in that sector is relatively larger. As in Acemoglu et al. (2012; henceforth AABH), the allocation of innovation within the polluting sector exhibits path-dependence, also because of a market size effect (a more advanced technology has a larger market which increases the profits of subsequent innovators). If clean technologies are initially less advanced than dirty ones, the laissez-faire equilibrium leads the economy toward an environmental disaster, as the quality of the environment falls below a critical threshold. In other words, economic growth is not sustainable. The paper analyzes whether this disaster can be prevented by specific policies in the North only, and doing so, makes two important points.

First, carbon taxes are generally unable to prevent an environmental disaster and may even be counterproductive. A carbon tax in the North leads to the relocation of part of the polluting good’s production to the South (a static pollution haven effect). It cannot prevent an environmental disaster if the South initially had a comparative advantage in the polluting sector, since then, the South specializes further in the polluting sector and its emissions keep growing. Moreover, because reallocating production goes hand in hand with reallocating innovation, a Northern carbon tax actually increases dirty Southern innovations (a dynamic pollution haven effect) and thereby may accelerate environmental degradation.

Second, a temporary industrial policy, combining clean research subsidies and a trade tax, may prove to be more effective. Such a policy can help the North develop a comparative advantage in the polluting sector and make that sector cleaner. This ensures that emissions eventually start decreasing in both countries. If the initial environmental quality is high enough, an environmental disaster can be averted. Importantly, directed technical change is essential for this result; with exogenous technical change, unilateral policies in the North would still fail to prevent a disaster when the South initially has a sufficiently large comparative advantage in the polluting sector.

The optimal unilateral policy can be decentralized through a carbon tax and research subsidies in the North along with a trade tax on the polluting good. To illustrate the results, I conduct a numerical exercise where, following the literature, the North corresponds to the countries with binding constraints under the Kyoto Protocol (Annex I countries). This exercise shows that, even though avoiding a disaster is possible, the welfare costs from not being able to intervene in the South may be very large. It also highlights
the double-edged nature of trade and directed technical change. Relative to autarky and exogenous technical change respectively, trade and directed technical change accelerate environmental degradation either under laissez-faire or with unilateral carbon taxes, yet both help reduce environmental degradation under the appropriate unilateral policy.

Finally, I extend to model to include cross-country knowledge spillovers. Unilateral carbon taxes may still fail to prevent an environmental disaster; whereas a combination of clean research subsidies and a carbon tariff can do so for sufficiently high initial environmental quality. In this case, the diffusion of knowledge can ensure a switch toward clean innovation in the South; hence an environmental disaster can be prevented even though the South specializes in the polluting good. A limit of the analysis thus far is that it ignores nontradable goods which, in reality, are responsible for a large share of emissions. Without knowledge spillovers, but with nontradables, unilateral policies in the North can reduce the growth rate of emissions in the South but not sufficiently to prevent a disaster. However, with knowledge spillovers unilateral policies can prevent an environmental disaster even in the presence of a polluting nontradable sector.

This paper can be interpreted as a green version of the “infant industry argument,” which claims that trade can be detrimental to growth if it leads countries to specialize in sectors with poor development prospects (Krugman, 1981, Young, 1991, Matsuyama, 1992). Similarly, here, a country risks specializing in the “wrong” sector if it cannot prevent the associated environmental externality. Krugman (1987) and Grossman and Helpman (1991, ch. 8) already argue that free trade may amplify comparative advantages and that a temporary trade policy can permanently reverse the trade pattern.¹

It has long been recognized that, in an open world, the pollution haven effect hampers the effectiveness of unilateral environmental policies (Pethig, 1976). Empirical evidence is reported by Copeland and Taylor (2004) or Broner, Bustos and Carvalho (2012). Markusen (1975) and Hoel (1996) show that the optimal instrument to mitigate the pollution haven effect is a tariff. In the context of global warming, several papers use computable general

equilibrium (CGE) models to track carbon through the global economy in order to determine the pattern of trade and compute the carbon leakage rate (the rate at which emissions abroad increase after a domestic reduction). Developed countries are net carbon importers, which justifies the focus of the paper on the case where the South has a comparative advantage in the polluting sector: Atkinson et al. (2011) find that the net US imports of carbon from China in 2004 amounted to 244 million tons of CO$_2$ or 0.9% of total world emissions that year; the OECD STAN database estimates that for OECD countries net CO$_2$ imports represent 12.6% of CO$_2$ emissions from production. Elliott et al. (2010) compute a carbon leakage rate of 20% from a reduction in Annex I countries and show that border tax adjustments eliminate half of it.\textsuperscript{2} There are comparatively few empirical studies. Aichele and Felbermayr (2012) find that countries which committed to the Kyoto Protocol reduced domestic CO$_2$ emissions by about 7%, but did not change their total CO$_2$ consumption. While this literature has focused on static models, the novelty of the present paper is to incorporate dynamic aspects, at the expense of a more detailed model of world trade.

A growing empirical literature has shown that the direction of innovation is affected by energy prices and more generally environmental policies (see for instance Popp, 2002, Newell, Jaffe and Stavins, 1999, Hassler, Krusell and Olovsson, 2012, and Aghion et al., 2016). As a result, several theory papers have integrated directed technical change in the study of climate change policies; here, I build on AABH.\textsuperscript{3} As the polluting good in this paper, the final good in AABH is produced with a clean and a dirty input, which are substitutes for each other. Yet, AABH focus on a closed economy and does not feature a “non-polluting sector.” Therefore, in AABH carbon taxes can still prevent an environmental disaster by redirecting innovation towards clean technologies. I show that this result collapses in an open economy, as carbon taxes are often insufficient and sometimes counterproductive.

Acemoglu, Aghion and Hémous (2014, henceforth AAH) already presents a two-country extension of AABH with international trade (AAH comes from the working paper version of AABH, Acemoglu et al., 2009). In contrast to

\textsuperscript{2}Among others, Babiker and Rutherford (2005); Böhringer, Fisher and Rosendahl (2010); and Böhringer, Carbon and Rutherford (forthcoming) find similar results.

the current paper, we make the following three assumptions in AAH: (1) the polluting and non-polluting goods are substitutes (here their elasticity is one, but the analysis directly extends to the complement case); (2) the South can only imitate (here it can also innovate); and (3) the level of pollution from the polluting good is given exogenously (here it is determined endogenously and depends on the amount of clean versus dirty innovations). This leads to very different results: in AAH, carbon taxes necessarily reduce the amount of emissions in the long-run, while they may accelerate environmental degradation in the current paper; and while in AAH a carbon tax in the North may fail to prevent an environmental disaster in some but not all equilibria, here it may fail in all equilibria. More generally, the current paper goes further than AAH in several dimensions: it characterizes the optimal unilateral policy, it considers several additional policy instruments (a trade tax and a clean research subsidy), it presents a numerical exercise and it extends the model to study the role of nontradable goods.

Di Maria and Smulders (2004) and Di Maria and van der Werf (2008) also study the allocation of innovation between an energy-intensive sector and a non-energy-intensive sector in a directed technical change model with trade, but they also overlook that innovations within the energy-intensive sector could either reduce or increase pollution.\footnote{In Di Maria and Smulders (2004), the North develops technologies and the South imitates. Opening up to trade leads to a reallocation of innovation toward the sector that the North exports. Carbon leakage is reduced when the goods are substitutes and amplified otherwise. In Di Maria and van der Werf (2008), both countries innovate and carbon leakage is always reduced by the innovation response to a unilateral cut in emissions.}

Finally, this paper relates to the large “integrated assessment models” (IAMs) literature which builds dynamic models of the economy and the climate to evaluate the impact of climate change policies on welfare. This literature has been pioneered by Nordhaus (1994), who developed the DICE model for a world economy, and Nordhaus and Boyer (2000), who developed the RICE model for a multi-regions economy. It aims at deriving quantitatively the optimal policy, which generally takes the form of a carbon tax schedule (see, for instance, Golosov et al., 2014, for a recent treatment, or Krusell and Smith, 2009, for an ongoing project with highly disaggregated regions). While this literature generally ignores endogenous innovation, this paper shows that green R&D subsidies are a crucial part of the optimal climate policy (a point also made by AABH). In addition, it derives the op-
timal unilateral policy in the presence of trade in goods, which neither this literature nor the CGE trade literature mentioned above do.\(^5\)

Section 2 presents the model. Section 3 studies the equilibrium and identifies which policies can ensure sustainable growth. Section 4 solves for the second-best policy when the South is constrained to be in laissez-faire, analytically and numerically. Section 5 introduces cross-country knowledge spillovers. An online Appendix contains the proofs and further extensions.

2. Model

I consider a discrete-time, infinite-horizon two-country (North, \(N\), and South, \(S\)), two-sector (polluting \(P\) and non-polluting \(NP\)), three-factor (capital, labor and scientists) Heckscher–Ohlin–Ricardo model in which sector \(P\) is similar to the economy of AABH. Each country is endowed with a fixed amount of labor and capital, \(L^N, K^N\) and \(L^S, K^S\), and a mass \(1\) of scientists. The North is meant to represent countries which are ambitious in tackling climate change and the South countries which are not. I consider an admittedly extreme scenario in which the North is able to implement strong environmental policies and the South does not carry any policy at all—of course, in reality, most countries lay somewhere between these two extremes. As already mentioned, the division North-South need not fall along the lines of developed versus developing countries, in particular because the United States has not signed the Kyoto Protocol.

2.1. Utility

The economy admits, for each period \(t\), a representative agent in the North who lives for one period and a like representative agent in the South. The utility of time-\(t\) agent in country \(X \in \{N, S\}\) is given by \(v(S_t)C^X_t\), where \(C^X_t\) is the final good consumption in country \(X\), \(S_t\) is the quality of the environment (identical in North and South) and \(v\) is increasing in \(S_t\). \(S_t\) is bounded between \(0\) and \(S\), with \(S\) corresponding to a pristine environment. I define an \textit{environmental disaster} as an instance of environmental quality reaching zero in finite time. I assume that \(v(0) = 0\) so that a disaster is as detrimental to welfare as zero consumption.\(^6\)

---

\(^5\)Hassler and Krusell (2012) are close to doing so but their model does not feature trade in goods (nor innovation), and they do not explore the full set of policy instruments.

\(^6\)In the literature, a consumption path is said to be sustainable if the consumption flow is bounded away from \(0\). Here environmental quality directly affects utility, so a
2.2. Production

Final consumption is a Cobb-Douglas aggregate of the consumption of two goods, polluting, \( P \), and nonpolluting, \( NP \):

\[
C^X = (C^X_P)\nu \left( C^X_{NP} \right)^{1-\nu},
\]

where \( C^X_Y \) represents the quantity of good \( Y \in \{P, NP\} \) consumed in country \( X \in \{N, S\} \) (whenever this does not lead to confusion, I drop the time subscript but allocations, technologies and policies are time-dependent). The analysis extends easily to the case where the consumption aggregate between the two goods is CES with an elasticity of substitution smaller than 1.\(^7\)

Goods \( P \) and \( NP \) are the only goods that are traded internationally. Good \( P \) represents tradable goods the production of which generates a lot of greenhouse gases emissions (particularly energy-intensive goods), while good \( NP \) represents the other tradable goods. The paper initially focuses on tradables, since it is because of international trade that policymakers fear that unilateral policies may have adverse consequences. In reality, both the tradable and nontradable sectors are important in generating CO\(_2\) emissions.\(^8\) Section 5.1 extends the model to include nontradables.

Good \( P \) is produced competitively with a clean input \( Y^X_c \) and a dirty input \( Y^X_d \), with an elasticity of substitution \( \varepsilon > 1 \):

\[
Y^X_P = \left( (Y^X_c)^{\varepsilon-1} + (Y^X_d)^{\varepsilon-1} \right)^{\frac{\varepsilon}{\varepsilon-1}}.
\]

The clean input models nonpolluting inputs that could substitute for polluting inputs, for instance, renewable energies to replace fossil fuel energy or bioplastics to replace traditional petroleum products.

---

\(^7\) A previous version of the paper (CEPR Discussion Paper 9733) does so. Yet, the analysis would be different with an elasticity greater than 1 (see section 3.6).

\(^8\) According to the International Energy Agency, manufacturing and construction represented 36.9\% of world CO\(_2\) emissions in 2010 (once electricity and heat are allocated to consuming sectors). To obtain a share for the tradable sector, one would need to subtract emissions from the construction sector from that number and to add some emissions from the agriculture and forestry sectors. Relatedly, Davis and Caldeira (2010) estimates that 23\% of carbon emitted is attributable to the production of goods that will be exported. This represents a lower bound for the share of CO\(_2\) emissions coming from the tradable sector, as not all tradable goods end up being exported.
Goods $c, d$ and $NP$ in country $X$ are produced competitively following

$$Y_{NP}^X = \left( \int_0^1 A_{NP_i}^X \left( x_{NP_i}^X \right)^\gamma di \right) \left( (K_{np}^X)^\beta (L_{np}^X)^{1-\beta} \right)^{-1-\gamma} \quad \text{and} \quad (3)$$

$$Y_{zi}^X = \left( \int_0^1 A_{zi}^X \left( x_{zi}^X \right)^\gamma di \right) \left( (K_z^X)^\alpha (L_z^X)^{1-\alpha} \right)^{-1-\gamma} \quad \text{for } z \in \{c, d\}. \quad (4)$$

$K_{np}^X$ and $L_{np}^X$ are the capital and labor employed in the assembly of good $NP$ in country $X$; $x_{NP_i}^X$ is the quantity of intermediates $i$ employed in sector $NP$; and $A_{NP_i}^X$ is its productivity, which is specific to the country and the sector. $K_z^X$, $L_z^X$, $x_{zi}^X$ and $A_{zi}^X$ are defined similarly for good $z \in \{c, d\}$. $\gamma$ is the factor share of intermediates. Intermediates cannot be traded internationally and are produced monopolistically according to

$$x_{NP_i}^X = \psi^{-1} (K_{NP_i}^X)^\beta (L_{NP_i}^X)^{1-\beta} \quad \text{and} \quad x_{zi}^X = \psi^{-1} (K_{zi}^X)^\alpha (L_{zi}^X)^{1-\alpha} \quad \text{for } z \in \{c, d\}. \quad (5)$$

$K_{NP_i}^X$ and $L_{NP_i}^X$ are the capital and labor employed in the production of intermediate $i$ for good $NP$ in country $X$ (and $K_z^X$ and $L_z^X$ are defined similarly). Since the same factor share is used in the production of intermediates and in the final assembly of the good, $\alpha, \beta \in (0, 1)$ are the overall factor shares of capital in sectors $z \in \{c, d\}$ and $NP$ respectively. I assume throughout that $\alpha > \beta$, which is true empirically: within tradables, polluting sectors tend to be more capital intensive. All results hold when $\alpha < \beta$ and the analysis of this section can be extended to a pure Ricardian model with $\alpha = \beta$.\footnote{A Ricardian model would pose some technical difficulties for section 4. The Heckscher-Ohlin features are useful to introduce some curvature in the production possibility frontier, so that both countries may produce both goods. They also make it easier to introduce cross-country knowledge spillovers (as another reason for trade than technological differences is needed then). There is nothing special about capital and labor being the two factors here instead of high-skill and low-skill labor for instance, which is why the paper abstracts from capital accumulation.}

I use $K_{NP}^X$ and $K_P^X$ to denote total employment of capital in sectors $NP$ and $P$ in country $X$, so that:

$$K_{NP}^X \equiv K_{np}^X + \int_0^1 K_{NP_i}^X di \quad \text{and} \quad K_P^X \equiv K_c^X + K_d^X + \int_0^1 K_{ci}^X di + \int_0^1 K_{di}^X di. \quad (6)$$

Having similarly defined $L_{NP}^X$ and $L_P^X$, factor market clearing and good mar-
ket clearing imply:

\[
\begin{align*}
K_P^X + K_{NP}^X &\leq K_X^X \quad \text{and} \quad L_P^X + L_{NP}^X \leq L_X^X, \\
C_P^N + C_P^S &\leq Y_P^N + Y_P^S \quad \text{and} \quad C_{NP}^N + C_{NP}^S \leq Y_{NP}^N + Y_{NP}^S.
\end{align*}
\]  

(7)

(8)

Intermediates producers face an iso-elastic demand with a price elasticity of \(1/(1 - \gamma)\) and therefore charge a mark-up \(1/\gamma\) over marginal cost. This leads to a classic monopoly distortion, as too few intermediates are produced. A consumption subsidy \(1 - \gamma\) for intermediates ensures that the consumer price is equal to marginal costs, correcting for the monopoly distortion (see Appendix A.1). To simplify the exposition and focus the comparison between first-best and second-best on environmental issues, I assume throughout that this subsidy is implemented in both countries. Since the share of intermediates is the same for all sectors, the monopoly distortion only has a scale effect, and this assumption is completely innocuous. Henceforth I abuse language by referring to the "laissez-faire" case as one where governments only implement the intermediates consumption subsidy.

2.3. Innovation

At the beginning of every period, one-period monopoly rights are allocated to entrepreneurs (such that each entrepreneur holds monopoly rights on only a finite number of intermediates). Entrepreneurs can hire scientists to increase the productivity of their variety. By hiring \(s_{zt}^X\) scientists, the monopolist for intermediate \(i\) in (sub)sector \(z \in \{NP, c, d\}\) can increase the initial productivity \(A_{zt}^X(t-1)\) of her intermediate to

\[
A_{zt}^X(t) = \left(1 + \kappa \left(s_{zt}^X\right)^{\frac{1}{1-\gamma}} \frac{A_{zt}^X(t-1)}{A_{zt}^X(t-1)} \right)^{1-\gamma} A_{zt}^X(t-1),
\]

(9)

with \(\kappa \in (0,1)\). \(A_{zt}^X\) is the time-\(t\) average productivity of sector \(z \in \{c, d, NP\}\):

\[
A_{zt}^X \equiv \left(\int_0^1 \left(A_{zt}^X\right)^{\frac{1}{1-\gamma}} di \right)^{1-\gamma}.
\]

(10)

The factor \((A_{zt}^X(t-1))^{-1/(1-\gamma)}\) captures decreasing returns to scale in innovation (the more advanced is a technology, the more difficult it is to innovate further), and \((A_{zt}^X(t-1))^{1/(1-\gamma)}\) denotes knowledge spillovers from the other intermediates in the same sector and country. As \(\kappa < 1\), the innovation technology exhibits decreasing returns to scale in the mass of scientists hired
(e.g., because scientists hired for the same intermediate in the same period risk reproducing the same innovation). $\kappa$ measures the size of innovations ($\kappa$ is related to the length of a time period, a shorter time period will be associated with a lower $\kappa$).

With a unit mass of scientists in both countries, market clearing gives:

$$\int_0^1 (s_{NPit}^X + s_{cit}^X + s_{dit}^X) \, di = 1.$$  \hspace{1cm} (11)

Since an entrepreneur has monopoly rights for one period only, she hires scientists so as to maximize current profits instead of the entire flow of profits generated by their innovations. Therefore the allocation of scientists across (sub)sectors is myopic. One-period monopoly rights allow to model as simply as possible the “building on the shoulder of giants” externality, the existence of which has long been recognized by the endogenous growth literature. In the specific context of climate change, this externality plays a crucial role in explaining why clean technologies have so far failed to really take off (e.g., there are still more dirty than clean patents in the car industry, Aghion et al., 2016), and why direct research incentives in addition to carbon taxes are welfare improving, a point made by AABH and Gerlagh et al. (2014).

There are no knowledge spillovers across sectors. Cross-country spillovers are absent for the moment but introduced in Section 5. A fixed mass of scientists in both countries implies that the amount of resources devoted to productivity improvements (in particular R&D) remains the same in both countries and over time. It allows to focus on the direction of technical change and ensures that one country does not become arbitrarily large relative to the other. This assumption is further discussed in section 3.6.

2.4. Environment

Within the bounds $0$ and $\bar{S}$, environmental quality evolves according to

$$S_t = (1 + \Delta) S_{t-1} - (\xi_Y^{N} Y_{dt}^{N} + \xi_S Y_{dt}^{S}).$$  \hspace{1cm} (12)

---

10If North and South are groups of countries then (9) implies that technically there are knowledge spillovers within these groups. This could be interpreted in at least three ways: 1) North and South are really two (close groups of) countries such as the EU versus China, the EU versus the US or the US versus China; 2) North countries are so different from South countries that technologies developed in the first group are not transferable to the second; 3) Within each group, countries are similar and therefore would develop similar technologies even without knowledge spillovers between them.
\(\xi^X Y_{dt}^X\) represents emissions in country \(X\), with \(\xi^X > 0\), the emission rate of dirty input production (which may differ across countries). \(\Delta > 0\) is the regeneration rate of the environment. Without loss of generality, I assume that \(S_0 = \overline{S}\). Such a law of motion captures the idea that the environment’s regeneration capacity decreases with greater environmental degradation—the type of negative feedback that climatologists worry about, e.g., the change in Earth’s albedo and the release of captured greenhouse gases which may occur as the polar ice cap melts. It is adopted for simplicity’s sake but, unless explicitly mentioned, the analytical results do not depend on it. The only important assumption is that if emissions become too large, \(S_t\) reaches the disaster level.\(^{11}\) The dirty input is directly responsible for environmental degradation, which is equivalent to a situation where it can be combined with a (cheap) fossil fuel resource in a Leontief way.\(^{12}\)

2.5. Policy tools

Section 3 considers policies which are characterized by a sequence of \(ad\ valorem\) taxes on the dirty input \(\tau_t^X\) in each country (equivalent to a carbon tax), a sequence of subsidies for scientists in each country and each subsector, and a sequence of \(ad\ valorem\) trade taxes \(b_t\) on the polluting good (by Lerner symmetry, they could equally be on the other good). All subsidies and taxes are financed or rebated lump-sum at the country level. Section 4 shows that the optimal policy can be decentralized with these instruments.

The trade tax is implemented by the North, so that prices in the South are equal to international prices: \(p_{NPt}^S = p_{NPt}\) and \(p_{Pt}^S = p_{Pt}\), while prices in the North follow \(p_{NPt}^N = p_{NPt}\) and \(p_{Pt}^N = p_{Pt}(1 + b_t)\).\(^{13}\) A positive trade tax corresponds to a tariff (resp., export subsidy) when the North imports

\(^{11}\)Real climate dynamics are much more complicated: emissions have a lagged impact on temperature, part of their impact is essentially infinitely-lived and there is a lot of uncertainty in the magnitude of the impact of CO\(_2\) on temperatures. This matters for the numerical exercise but not the results of section 3.

\(^{12}\)Therefore, we abstract from resource exhaustion. This is not a bad assumption since oil does not play a major role in emissions for the manufacturing sector, while reserves of coal, natural gas and non-traditional fossil fuels are in large supply relative to the time scale of critical environmental degradation.

\(^{13}\)The trade tax is not explicitly related to the carbon content of imports. A carbon content specific trade tax could alter Southern firms’ behavior only if the South were to implement its own policy or if the North could observe the carbon content of individual imports, which seems implausible.
(resp., exports) good $P$.$^{14}$ When the North is the only country intervening, trade balance must be maintained every period:

$$p_{Pt} (Y^S_{Pt} - C^S_{Pt}) + p_{NPt} (Y^S_{NPt} - C^S_{NPt}) = 0.$$  \hfill (13)

3. A positive analysis of unilateral environmental policies

This section presents a positive analysis of unilateral environmental policies. It first solves for the allocation of innovation and shows how directed technical change reinforces the pollution haven effect. Then, it studies whether an environmental disaster can be prevented under different scenarii.

3.1. Trade and innovation allocation

**Trade pattern.** Here I analyze the equilibrium when the only policy implemented is a carbon tax in the North $\tau^N_t \geq 0$; the results are derived and generalized in Appendix A.1. In each country, aggregate production in each sector can be written as

$$Y^X_{Pt} = \frac{\zeta A^X_{Pt} (K^X_{Pt})^{\alpha} (L^X_{Pt})^{1-\alpha}}{1 - \delta^X_t} \text{ and } Y^X_{NPt} = \zeta A^X_{NPt} (K^X_{NPt})^\beta (L^X_{NPt})^{1-\beta},$$ \hfill (14)

where $\zeta \equiv \gamma (1 - \gamma)^{1-\gamma} \psi^{-\gamma}$, $A^X_{Pt} \equiv \left( (A^X_{Pt})^{\varepsilon-1} + (1 + \tau^X_t)^{1-\varepsilon} (A^X_{Pt})^{\varepsilon-1} \right)^{\frac{1}{\varepsilon-1}}$ and $\delta^X_t \equiv \tau^X_t (A^X_{Pt}/A^X_{Pt})^{\varepsilon-1} (1 + \tau^X_t)^{-\varepsilon} \in [0, 1)$. $A^X_{Pt}/(1 - \delta^X_t)$ decreases in $\tau^X_t$ and measures the effective average productivity of sector $P$ in country $X$. $\delta^X_t$ is introduced for convenience and is equal to the ratio of carbon taxes paid over the revenues of the polluting sector: $\delta^S_t = \tau^S_t p^S_{Pt} Y^S_{Pt} / (p_{Pt} Y^S_{Pt})$ (as $\tau^S_t = 0$, $\delta^S_t = 0$). This formulation highlights that, in a given period, the model collapses to a Heckscher–Ohlin model with varying productivity across countries. In laissez-faire, the South has the comparative advantage in the polluting good $P$ if and only if

$$\left( \frac{A^S_{Pt}}{A^S_{NPt}} \right)^{\frac{1}{\varepsilon-1}} K^S_t / L^S_t > \left( \frac{A^N_{Pt}}{A^N_{NPt}} \right)^{\frac{1}{\varepsilon-1}} K^N_t / L^N_t.$$ \hfill (15)

$^{14}$If the North imports good $P$ under free trade, an increasingly higher trade tax corresponds to a positive tariff up to the point where it implements autarky. Beyond that point, the North begins to export good $P$ and the trade tax is a positive export subsidy.
Trade results from Ricardian (relative productivity) and Heckscher–Ohlin forces (relative factors endowment). If the difference in comparative advantage is not too large, both countries produce both goods. As it grows, one and eventually both countries fully specialize. Besides, a positive carbon tax in the North reduces the productivity of sector $P$ there and increases the chance that the South has a comparative advantage in that sector.

**Emissions.** Emissions are given by $E_t^X = \xi^X (1 + \tau_t^X)^{-\varepsilon} \left( A_{dt}^X/A_{Pt}^X \right)^\varepsilon Y_t^X$. Thus the emission rate in sector $P$ is increasing in the ratio of dirty to clean productivities $A_{dt}^X/A_{ct}^X$ and decreasing in the carbon tax $\tau_t^X$.

**Allocation of innovation.** Entrepreneurs face a two-stage problem. In the second stage, they choose prices to maximize profits given their productivities. Post-innovation profits in sector $z \in \{c, d, NP\}$ are given by:

$$\pi_{zit}^X = (1 - \gamma) \left( A_{zit}^X/A_{cit}^X \right)^{1-\gamma} p_{zt}^Y Y_{zt}^X.$$ 

(16)

These profits are proportional to the revenues of the intermediate’s (sub)sector (because of the Cobb-Douglas specification) and are increasing in the productivity of the intermediate, $A_{zit}^X$. In the first stage, entrepreneurs hire scientists to increase the productivity of their intermediate. Thanks to the knowledge spillovers across varieties, all monopolists in a given (sub)sector hire the same number of scientists and average productivity grows following

$$A_{zt}^X = (1 + \kappa (s_{zt}^X)^{\gamma})^{1-\gamma} A_{zt(t-1)}^X$$

for $z \in \{c, d, NP\}$. Therefore, regardless of the technology and its level of development, the same amount of innovation resources (scientists) is required for a given proportional increase in productivity. Such formulation is common in endogenous growth models as it is consistent with steady-state growth.

**Path dependence in clean versus dirty technologies.** Assume that country $X$ produces good $P$ (otherwise, $s_{ct}^X = s_{st}^X = 0$). Combining the first-order conditions with respect to the number of scientists in the clean and dirty subsector yields the following equation:

$$\left( s_{ct}^{X(1-\gamma)} / s_{dt}^{X(1-\gamma)} \right) (1 + \kappa (s_{ct}^X)^{\gamma})^{1-\gamma} Y_{ct}^X = Y_{ct}^X \left( A_{ct}^X/A_{dt}^X \right)^{1-\gamma}.$$

(17)

The second equality follows from the demand equation for both inputs in sector $P$. As the production technologies differ only by their productivity levels, the ratio of clean to dirty sector revenues increases with the ratio
of clean to dirty productivities. This results from two counteracting forces: a larger technology ratio leads to a larger market share ratio but also to a lower price ratio; the former effect dominates when the inputs are substitutes ($\varepsilon > 1$). Thus, in laissez-faire ($\tau^X_t = 0$), for a sufficiently small innovation size $\kappa$, more scientists are allocated to the dirty than to the clean subsector if and only if the dirty sector is already the most advanced $A^X_{d(t-1)} > A^X_{c(t-1)}$; there is path dependence (see Appendix A.2).\textsuperscript{15} A carbon tax reduces demand for the dirty input. Therefore, for a given mass of scientists in sector $P$, it leads to a reallocation of innovation within that sector towards clean technologies, which is stronger the larger the elasticity of substitution $\varepsilon$ is.

**Amplification of comparative advantage.** Assume that production occurs in both sectors (otherwise, innovation occurs only in the active sector). By combining the first-order conditions with respect to the number of scientists in (sub)sectors $NP$, $c$ and $d$, I obtain

$$\frac{(s^X_{ct})^{1-\varepsilon} (1 + \kappa (s^X_{ct})^\varepsilon) + (s^X_{dt})^{1-\varepsilon} (1 + \kappa (s^X_{dt})^\varepsilon)}{(s^X_{NPt})^{1-\varepsilon} (1 + \kappa (s^X_{NPt})^\varepsilon)} = \frac{(1 - \delta^X_t) p^X_{Pt} Y^X_{Pt}}{p^X_{NPt} Y^X_{NPt}}. \quad (18)$$

$(1 - \delta^X_t) p^X_{Pt} Y^X_{Pt}$ are the net of tax revenues of the polluting sector. Therefore, for a given ratio $A^X_{d(t-1)}/A^X_{c(t-1)}$ of initial productivities within sector $P$ and given carbon tax $\tau^X_t$, the number of scientists allocated to sector $P$ is increasing in the ratio of sector $P$ to sector $NP$ revenues. Under free trade, prices are equalized in both countries, and each tends to innovate relatively more in its exporting sector (it necessarily does so when $A^N_{ct}/A^N_{dt} = A^S_{ct}/A^S_{dt}$ and $\tau^N_t = 0$). In laissez-faire, comparative advantages are typically amplified over time, so that one and eventually both countries fully specialize.\textsuperscript{17}

---

\textsuperscript{15}If $\kappa$ is large and $A^X_{d(t-1)}/A^X_{c(t-1)}$ is close to 1, multiple equilibria may exist (as in AABH): if more scientists innovate in clean than in dirty technologies at time $t$, clean technologies may become more developed than dirty ones (even though they were not at time $t-1$), justifying in return the higher rate of clean innovation.

\textsuperscript{16}One may think that innovation could be easier for new technologies (such as clean ones) where low-hanging fruits are more common. For instance we could have $A^X_{ct} = \left(1 + \kappa \max \left(1, (\overline{A}/A^X_{ct})^\eta \right) (s^X_{ct})^\gamma \right)^{1-\gamma} A^X_{c(t-1)}$ for some $\overline{A} > 0$ and $\eta > 0$. This would create a temporary force going against path dependence in clean versus dirty technologies. Yet, Aghion et al. (2016) find evidence of path dependence in the car industry.

\textsuperscript{17}The amplification of comparative advantage mechanism bears some similarities with the one that arises in models of external learning by doing like Lucas (1988): There, as an
By contrast, in autarky, innovation always occurs in both sectors since consumer demand implies that $p^X_{NPt}Y^X_{NPt} / (p^X_{Pt}Y^X_{Pt}) = (1 - \nu)/\nu$.

**Equilibrium uniqueness.** Although the main results of this section could be extended to a case with multiple equilibria, focusing on a unique equilibrium simplifies the exposition. I obtain (proof in Appendix A.2):

**Lemma 1.** If i) either $\nu (\varepsilon - 1)(1 - \gamma) \leq 1$ or $\kappa < (1 - \nu) / [(1 - \varepsilon - 1)(1 - \gamma) - 1]$; ii) $\nu \geq 1/2$; and iii) $\kappa \leq (1 - \nu) (\alpha - \beta)^2 / [\nu (1 - \gamma) L]$ where $L \equiv \beta (1 - \alpha) + \alpha (1 - \beta) + \max((1 - \beta) \beta (2 + \kappa), \alpha (1 - \alpha) (1 + 2^\alpha + \kappa))$, then the equilibrium is unique.

Condition i) ensures that the allocation of innovation within the polluting sector is unique, condition ii) that the equilibrium is unique when one country is close to a corner of specialization (i.e. when small changes in productivity could push toward full specialization or diversification). Moreover, innovating more in a sector increases a country’s comparative advantage in that sector, which, in turn, prompts more innovation in the same sector. This could also lead to multiple equilibria, but condition iii) stipulates that the innovation size is sufficiently small to ensure uniqueness. This condition is far from necessary and Appendix A.2 provides a laxer, but more complicated, condition (from then on, when I refer to the assumptions of lemma 1 being met, iii) can always be replaced by this laxer condition). The lemma does not extend to the Ricardian case ($\alpha = \beta$) since then, no matter how small $\kappa$ is, there are multiple equilibria if the initial comparative advantage is small.

### 3.2. The dynamic pollution haven effect

The introduction of a positive carbon tax $\tau^N_i$ in the North increases the production cost of the polluting good $P$ there and therefore reduces its production. This raises its price which leads to an increase in its production in the South, resulting in more Southern emissions (this is the classic static pollution haven effect). Following (18), sector $P$ innovation increases in the South because the relative revenues of sector $P$ increase, that is the mass of scientists allocated to the polluting sector, $s^S_{Pt} \equiv s^S_{ct} + s^S_{at}$, increases. Meanwhile in the North, a decrease in the relative revenues of sector $P$ leads to a decrease in sector $P$ innovation ($s^N_{Pt}$ decreases). This changes technologies economy develops, it accumulates human capital, and a higher future level of aggregate human capital encourages each agent to invest more in human capital.
in a way which further favors production of good $P$ by the South and so further increases Southern emissions, creating a dynamic pollution haven effect. Formally, I get (proof in Appendix A.3):

**Proposition 1.** Assume that the assumptions of Lemma 1 are met, that $A_d^{N(t-1)} \geq A_c^{N(t-1)}$ and $(1 + \tau_i^N) \leq (A_d^{N(t-1)}/A_c^{N(t-1)})^2$. Then, the introduction at time $t$ of a positive carbon tax $N_t > 0$ in the North increases innovation in the polluting sector in the South, $s_{Pt}^S$, and decreases innovation in the polluting sector in the North, $s_{Pt}^N$. These changes in the innovation allocation further increase emissions in the South, provided that the Southern dirty technologies are more advanced than clean ones ($A_S^d(t-1) \geq A_S^c(t-1)$) and that $A_{Pt}^N/(1 - \delta_i^N)$ increases in $s_{Pt}^N$.

Proposition 1 adds a few assumptions for the logic developed above to reach its conclusion. The assumption that $A_d^{N(t-1)} \geq A_c^{N(t-1)}$ and $1 + \tau_i^N \leq (A_d^{N(t-1)}/A_c^{N(t-1)})^2$ ensures that the ratio of clean to dirty revenues is closer to unity post-tax than it was pre-tax. The further is this ratio from unity, the more sector $P$ innovation is attractive, so that without this assumption, there would be a force pushing towards an increase in the the total mass of scientists in the polluting sector $s_{Pt}^N$ for a high carbon tax.\(^\text{18}\)

Second, more innovation in sector $P$ in the South increases the emission rate if and only if $A_d^{S(t-1)} > A_c^{S(t-1)}$. Third, because of the distortion created by the carbon tax, for some very specific combination of parameters, a decrease in $s_{Pt}^N$ might increase $A_{Pt}^N/(1 - \delta_i^N)$ and thereby $Y_{Pt}^N$, which would push towards a decrease of Southern emissions (this case is ruled out if $(1 + \tau_i^N)\leq (A_d^{N(t-1)}/A_c^{N(t-1)})^{\varepsilon - 1}$ or if $(1 + \tau_i^N)\geq (A_d^{N(t-1)}/A_c^{N(t-1)})^{\varepsilon - 1} (\tau_i^N (\varepsilon - 1) - 1)$).

Whether world emissions are more likely to increase when innovation responds to the policy change depends on the pattern of comparative advantage and the emission rates. The dynamic pollution haven effect increases Southern emissions but it also further decreases Northern emissions. Moreover, the dynamic pollution haven effect has permanent consequences: in all subsequent periods, the relative productivity of the South in sector $P$ will have

\(^\text{18}\)This force is still unlikely to dominate because of the direct effect of the carbon tax on the revenue ratio between sectors $P$ and $NP$. Moreover, even if this assumption is violated and $s_{Pt}^N$ were to increase, as long as the direct effect of the carbon tax on Northern production dominates ($Y_{Pt}^N$ decreases and $Y_{NPt}^N$ increases, which is true for $\kappa$ sufficiently small), then $s_{Pt}^N$ must increase and $s_{d(t-1)}^N$ decrease.
increased, which favors innovation in that sector itself and therefore tends to 
increase Southern emissions. Proposition 2 below further characterizes the 
dynamic consequences of a Northern carbon tax.

A carbon tax has an ambiguous effect on clean innovation in the North: 
as the market for good $P$ shrinks, overall sector $P$ innovation is reduced, but 
within that sector, it gets reallocated towards clean technologies.

3.3. Laissez-faire, a global social planner and autarky

Under laissez-faire, if dirty technologies are more advanced than clean 
one in both countries, innovation in sector $P$ remains directed primarily 
toward the former. Since sector $P$ innovation does not asymptotically vanish 
(the exporting country innovates more in that sector than it would under 
autarky), the production of good $P$ grows unboundedly and so do emissions. 
Eventually, the economy reaches an environmental disaster (see (12)).

With only one country, the logic of AABH applies. A social planner 
can use clean research subsidies, taxes on dirty research or carbon taxes to 
redirect innovation from dirty to clean. Once clean technologies acquire a 
sufficient lead over dirty ones, market forces ensure that most research is 
directed toward the clean subsector, which is now the most advanced. Eventually, 
the emission rate of good $P$ approaches zero—fast enough to offset 
growth in good $P$’s production—and a disaster can be avoided for sufficiently 
high initial environmental quality. A social planner who can intervene in both 
countries can use the same instruments and avert a disaster (for high enough 
$S_0$) by redirecting sector $P$ innovation towards clean technologies in countries 
which produce good $P$ (proof in Appendix A.4).

In autarky (and without knowledge spillovers), a disaster cannot be avoided 
through Northern policies alone because Southern emissions necessarily grow 
unboundedly: absent international cooperation, trade is necessary to avoid 
an environmental disaster.

3.4. Taxes on the polluting good in the North only

I now assume that trade is possible and that only the North can im-
plement some policy (and the North cannot pay the South to implement a 
policy). To avoid an environmental disaster, the South must asymptotically 
fully specialize in the nonpolluting sector $NP$. Otherwise, there is always 
innovation in the polluting sector $P$ in the South, and the production of good 
$P$ and therefore emissions grow unboundedly (see Appendix A.5 for a for-
mal proof). I first consider that the North can implement a positive carbon
tax and/or a positive tax on dirty research. Both instruments can reduce emissions in the North, and prompt clean innovation there.

Under laissez-faire and if the South has initially a comparative advantage in the polluting sector, \( P \), then it tends to innovate more than the North in that sector, which reinforces its comparative advantage over time, so that the South eventually specializes in \( P \). This is the “amplification of comparative advantage effect” described above. The North government cannot reverse this pattern by using a positive tax on dirty research or a positive carbon tax. In contrast, a carbon tax reduces the productivity of sector \( P \) in the North, which leads to an increase in sector \( P \) innovation in the South and a decrease in the North (as specified in Proposition 1). This further strengthens the comparative advantage of the South in sector \( P \). A positive tax on dirty innovation in the North has similar effects to a carbon tax: it drives scientists away from sector \( P \) toward the nonpolluting sector \( NP \); and distorts innovation within sector \( P \) which further reduces the growth rate of average productivity \( A_N^P \)—this results in an increase in the price of good \( P \) and thereby more sector-\( P \) innovation in the South. Accordingly, positive Northern taxes on good \( P \) can only accelerate the South’s specialization in that sector. In fact, the economy typically grows faster since more specialization entails less overlap in the type of innovations being undertaken by both countries, and, as a result, such policies are likely to accelerate environmental degradation.\(^{19}\) Formally, I get (proof in Appendix A.6):

**Proposition 2.** Assume that the assumptions of Lemma 1 are met. Then, no matter how high initial environmental quality \( S_0 \) is, no combination of a positive carbon tax and a positive tax on dirty research can prevent an environmental disaster if: (i) clean technologies are less developed than dirty ones in both countries (\( A_{c0}^N/A_{d0}^N \leq 1 \) and \( A_{c0}^S/A_{d0}^S \leq 1 \)), (ii) the South has a weak initial comparative advantage in the polluting sector \( P \) (i.e., \( (A_{P0}^S/A_{NP0}^S)^{1/\gamma} K^N/L^N \geq (A_{P0}^N/A_{NP0}^N)^{1/\gamma} K^N/L^N \)), and (iii) either clean technologies are sufficiently less developed than dirty ones in the South (\( A_{c0}^S/A_{d0}^S \)

---

\(^{19}\)The extreme version of this argument is illustrated by the knife-edge case where \( A_{c0}^N/A_{d0}^N = A_{c0}^S/A_{d0}^S < 1 \) and \( (A_{P0}^S/A_{NP0}^S)^{1/\gamma} K^N/L^N = (A_{P0}^N/A_{NP0}^N)^{1/\gamma} K^N/L^N \), with no carbon tax, there would be no trade in equilibrium, and emissions and the economy would grow at rate \( (1 + \kappa 2^{-\nu})^{1-\gamma} - 1 \) (for \( \nu = 1/2 \)). A small carbon tax \( \tau_N^N \) is enough to ensure that both countries eventually specialize so that emissions (and the economy) asymptotically grow at rate \( (1 + \kappa)^{1-\gamma} - 1 \).
is sufficiently small) or the South has a sufficiently strong initial comparative advantage in $P$.

Assumption iii) in Proposition 2 is necessary because when the ratio of clean to dirty revenues is farther from unity in the North than in the South, more innovation in sector $P$ might take place in the North even if the South exports good $P$. The assumption of equilibrium uniqueness could also be relaxed if $A_{d0}^S / A_{d0}^S$ is sufficiently small and the South has a sufficiently strong initial comparative advantage in $P$: in this case, all equilibria would feature the South specializing in sector $P$ leading to an environmental disaster.

The crucial hypothesis of Proposition 2 is that the South has a comparative advantage in sector $P$. When the North is identified with Annex I countries, this hypothesis seems to hold since the CGE literature systematically finds that developed countries are net carbon importers as mentioned in the introduction (and I also find that the South has a comparative advantage in sector $P$ initially in the numerical exercise in section 4.3). Yet, with a different definition of the North, this hypothesis may not hold, and the North might be able to prevent an environmental disaster with a carbon tax only, as the pollution haven and the amplification of initial comparative advantage effects would work in opposite directions.

3.5. Introducing clean research subsidies and the trade tax

I now assume that the North can use clean research subsidies and a trade tax. Contrary to the previous policies, clean research subsidies can also reallocate scientists who were working in the nonpolluting sector $NP$ toward the clean subsector. This boosts innovation in clean technologies in the North, even when it does not have the comparative advantage in sector $P$. Increasing innovation in clean technologies makes sector $P$ less polluting and helps build a comparative advantage in that sector. In the meantime, a positive trade tax on good $P$ can reduce production and therefore innovation in sector $P$ in the South, which also helps reverse the pattern of comparative advantage. For sufficiently high initial environmental quality, a policy combining these two instruments can prevent a disaster.

\footnote{Assumption iii) ensures that this effect is dominated. It plays a role similar to that of the assumption $A_{d(t-1)}^N \geq A_{c(t-1)}^N$ and $1 + \tau_t^N \leq (A_{d(t-1)}^N / A_{c(t-1)}^N)^2$ in Proposition 1.}
Proposition 3. A combination of a temporary trade tax and a temporary clean research subsidy in the North can prevent an environmental disaster provided that the initial environmental quality $S_0$ is sufficiently high.

Proof. For instance, consider the following two-phase approach. First, the social planner implements a tariff large enough to shut down trade, so that the South must innovate in both sectors $P$ and $NP$. Simultaneously, she implements large clean research subsidies so that nearly all Northern scientists innovate in the clean subsector, resulting in more sector $P$ innovation than in the South. Once the North has acquired the comparative advantage in that sector and $A^N_{ct(t-1)}/A^N_{dt(t-1)}$ is sufficiently large, the social planner can discontinue all policies and re-open up to trade. Market forces then ensure that the production of good $P$ eventually moves entirely to the North where it relies essentially on clean technologies.\footnote{If the equilibrium is unique, this follows lemma A.3, when the North has the comparative advantage in sector $P$ at some date $t$, with $A^N_{ct}/A^N_{st} < A^S_{ct}/A^S_{st} < 1$. With multiple equilibria, the same logic applies if the temporary policy is maintained until the North has a comparative advantage in all possible equilibria.} Emissions go down to zero in both countries, and a disaster can be avoided. 

With a nontradable polluting sector, such policy could reduce innovation in the polluting sector in the South, decreasing the growth rate of emissions. Yet, it could not prevent an environmental disaster because the South would still produce and therefore innovate in the nontradable polluting sector.

One might think that clean research subsidies alone should be enough to prevent an environmental disaster. This is true if the initial comparative advantage of the South is not too large, but not if the South fully specializes in sector $P$ and clean technologies in the South are sufficiently less advanced than dirty ones. In that case, all Southern scientists are allocated to sector $P$ and, asymptotically, to dirty technologies. Even if the North were to allocate all its scientists to clean technologies, $A^S_{Pt}$ would grow as fast as $A^N_{Pt}$. That situation is irreversible and an environmental disaster cannot be avoided. Full specialization in the South occurs in the first place when its initial comparative advantage in sector $P$ is sufficiently large or when clean technologies are sufficiently backward in the North, as $A^N_{Pt}$ grows slowly when clean technologies are catching up with dirty ones. The following remark summarizes this discussion (proof in Appendix A.7).\footnote{This is the only result of this section that would not hold if goods $P$ and $NP$ were strict complements (instead of Cobb-Douglas): then the South could not stay fully specialized}
Remark 1. There exist initial factor endowments and technologies, such that no matter how high $S_0$ is, no combination of a carbon tax, a tax on dirty research, and a subsidy for clean research can prevent a disaster.

3.6. Discussion

Here, I discuss some assumptions of the model. Appendix A.8 presents additional results regarding alternative instruments.

Size and mass of scientists. The relative capital and labor endowments of the two countries play a role quantitatively: the larger the North is, the easier it is to reverse comparative advantages. In the long-run, the size of the two economies depends on the mass of scientists, which is a proxy for the amount of resources spent on innovation. As North and South have the same fixed mass of scientists, the analysis is implicitly restricted to a case where the two (groups of) countries are of similar size.

If the mass of scientists in the North ($s^N$) is much smaller than in the South ($s^S$), the North eventually becomes a small economy relative to the South, and the South behaves as if it were in autarky: a disaster is unavoidable. In contrast, if $s^N$ is larger than $s^S$, a disaster can be avoided using only clean research subsidies—as $A^N_{Pt}$ could grow faster than $A^S_{Pt}$ even if the South had fully specialized in sector $P$. Depending on parameters, a disaster may also be avoided using taxes on dirty research or a carbon tax under the assumptions of Proposition 2.\footnote{For instance, if the consumption share of good $P$ ($\nu$) is close to 1 and $s^N$ is large enough, $A^N_{Pt}/A^N_{NPt}$ can grow faster than $A^S_{Pt}/A^S_{NPt}$, leading to a reversal of comparative advantage; if $\nu = 1/2$, on the contrary, the previous analysis still holds.}

This suggests that the inclusion of the United States in the North is crucial. However, one may think that $s^S$ is bound to increase, making it harder over time for the North to intervene decisively.

The assumption that the mass of scientists is exogenous is not innocuous either. With an endogenous mass of scientists, clean research subsidies become an even more potent instrument, as they can ensure that the amount of resources spent in R&D in the North becomes greater than that in the South (so that a disaster becomes avoidable with clean research subsidies only). The impact of a carbon tax depends on the specific way in which innovation is endogenized, but in models where the cost of innovation moves with GDP, a carbon tax is likely to reduce overall innovation since innovation benefits are scaled by the revenues net of taxes (this would reinforce Proposition 2).
Three sectors. The results of the paper crucially depend on the assumption that innovation may occur in three (sub)sectors: clean, dirty and non-polluting (in contrast with Di Maria and Smulders, 2004; Di Maria and van der Werf, 2008 and AAH). If innovation were limited to clean and dirty technologies within the polluting sector, the North could not build a comparative advantage in a specific sector. With only dirty innovations in the polluting sector, no innovations could replace existing polluting technologies since the final good is a Cobb-Douglas aggregate of the polluting and the nonpolluting goods. And only clean innovation in the polluting sector would falsely assume that all innovations in that sector decrease emissions.\textsuperscript{24}

The assumption that the clean and dirty inputs are substitute ($\varepsilon > 1$) is crucial, as otherwise avoiding a disaster with unilateral policies is impossible.\textsuperscript{25} Yet, this is a very natural assumption, first if $\varepsilon \leq 1$, both inputs would be essential in the production of the polluting good and therefore clean inputs would not really represent an alternative to dirty inputs; second, Papageorgiou, Saam and Schulte (2013) show empirically that the elasticity of substitution between clean and dirty inputs (notably in energy) is significantly greater than 1. The assumption of a unit elasticity between the polluting and non-polluting good is very common in the literature, and, as already mentioned, the analysis extends to the case where goods $P$ and $NP$ are complements.\textsuperscript{26} It does not extend to the case where $P$ and $NP$ are substitutes (and therefore not essential). Then whether an environmental disaster can be avoided or not depends crucially on how large $A_{d}^{S}$ is relative to $A_{NP}^{S}$: a large $A_{d}^{S}$ may push the South towards producing the dirty input.

\textsuperscript{24}Dirty innovations include innovations in the energy sector that make fossil fuel energy cheaper (for instance by allowing the use of shale gas or bituminous sands), but also innovations in components that are complements to fossil fuel energy and thus increase its demand, or the introduction of new goods or inputs that rely on fossil fuel energy.

\textsuperscript{25}To do so, the South must specialize in sector $NP$. This implies positive growth in sector $P$ in the North, which is then only possible with positive growth in the dirty input and therefore in emissions.

\textsuperscript{26}The CGE (e.g. Babiker and Rutherford, 2005 or Boehringer et al., 2010) literature often assumes a unit elasticity of substitution between manufacturing sectors, which include a non-polluting aggregate and a detailed representation of polluting sectors. Copeland and Taylor (2005) assume that polluting and non-polluting goods are both essential (which rules out the substitute case). This also seems reasonable given what the two sectors stand for: for instance the polluting good includes the manufacture of basic metals while the non-polluting good includes the manufacture of machinery and equipment.
rather than the non-polluting good, regardless of the policy in the North.

4. Optimal policy and numerical illustration

I now turn to the normative part of the paper. I consider a social welfare function which aggregates the preferences of the time-\(t\) representative agents in the North and the South according to:

\[
U = \sum_{t=0}^{\infty} \frac{1}{(1 + \rho)^t} \left(1 + v(S_t) \left(C_t^N + C_t^S\right)\right)^{1-\eta};
\]

(19)

where \(\rho > 0\) is the discount rate and \(\eta \geq 0\) is the inverse elasticity of intertemporal substitution (\(\eta = 1\) corresponds to a logarithmic utility). Therefore, the social planner cares only about the time profile of world consumption and environmental quality, which allows to separate climate issues from redistribution issues (see Appendix A.13 for an alternative set-up where the social planner cares about consumption distribution).

4.1. First-Best: North and South Policy

Before solving for the optimal unilateral policy, the focus of this paper, I briefly present the first-best which is a useful benchmark—the solution is derived in Appendix A.9. In the first-best, the social planner maximizes (19) subject to the following constraints: the production function equations (1), (2), (3), (4), (5); the factor market clearing equations (7) and (11); the goods market clearing equation (8); the environmental degradation equation (12); and the knowledge accumulation equation (9).

The first-best policy can be decentralized in the following way. As already mentioned, a subsidy \(1 - \gamma\) to all intermediates corrects for the monopoly distortion. The environmental externality is corrected by a carbon tax in both countries that equalizes the marginal cost of the tax (lower current consumption) with the marginal benefit (higher environmental quality in all subsequent periods). Carbon taxes in the North and the South differ in \textit{ad valorem} values across countries but are identical as a tax per unit of CO\(_2\). The social planner corrects for the myopia of monopolists in their innovation decisions by allocating scientists in accordance with the discounted value of the entire stream of additional revenues generated by their innovation.

As the utility flow is minimized during a disaster and the social planner can directly control emissions, the optimal policy always avoids a disaster.
In addition (as shown in Appendix A.10), if the discount rate $\rho$ is sufficiently small and the inverse elasticity of intertemporal substitution $\eta \leq 1$, then both countries specialize in finite time and innovation in sector $P$ switches to mostly clean, so that emissions eventually vanish. With the law of motion (12), the quality of the environment reverts to $S$ and the carbon tax reaches zero in finite time.\textsuperscript{27} These are only sufficient conditions, and the optimal policy is likely to feature a switch to clean innovations also when $\eta > 1$.

4.2. Second-Best: Policy only in the North
Assume now that the social planner cannot implement any policy in the South, whose economy is in laissez-faire. Trade balance must be maintained at every point in time and international income transfers are impossible. In the second-best policy, the social planner maximizes (19) subject to the following constraints: (1) for the North and the South; constraints (2), (3), (4), (5), (7), (11) and (9) for the North only; the environmental degradation constraint (12); the goods market-clearing constraints in both countries:

$$C^N_{Yt} = Y^N_{Yt} + M_{Yt} \quad \text{and} \quad C^S_{Yt} = Y^S_{Yt} - M_{Yt}, \quad \text{for } Y \in \{P, NP\},$$

(20)

where $M_{Yt}$ denotes net imports of the North of good $Y$; the trade balance constraint

$$p_t M_{Pt} + M_{NPt} = 0,$$

(21)

where $p_t \equiv p_{Pt}/p_{NPt}$ is the international price ratio; and constraints describing the South’s laissez-faire economy (a consumer demand equation, offers equation, emissions equation, equilibrium allocation of scientists equations and laws of motion for aggregate productivity). The South’s allocation is unique under the same assumptions as those of lemma 1. This leads to the following result (proof in Appendix A.11).

**Proposition 4.** The second-best policy can be decentralized through a carbon tax in the North, research subsidies/taxes in the North, a subsidy for the use of all intermediates, and a trade tax.

\textsuperscript{27}With a different law of motion, $S_t$ may not approach $S$ and the optimal carbon tax may not converge to 0. Yet, it becomes irrelevant in the sense that a 0 carbon tax would only have a negligible effect on welfare.
The social planner uses the same instruments as before to address the inefficiencies in the North’s economy: the environmental externality, the knowledge externality and the monopoly distortion. The trade tax allows the social planner to distort prices and therefore to affect the allocation of factors in the South. The trade tax results from both an environmental motive and an innovation one. First, emissions in the South are not directly taxed, this favors a positive trade tax which decreases the relative price of the polluting good $P$ in the South and therefore its production and emissions. Second, there is typically too much sector-$P$ innovation in the South, both because sector-$P$ innovations lead to more untaxed emissions in the future and because to avoid a disaster, the South must at least asymptotically fully specialize in sector $NP$. As a result, current sector-$P$ innovations will be of little use in the future, which Southern innovators do not internalize because of their myopia. A positive trade tax can be used to reduce sector-$P$ innovations in the South. Therefore, the trade tax is generally positive; it takes the form of a tariff when the North imports good $P$ and of an export subsidy otherwise (see Appendix A.11.3 for details). The next proposition further characterizes the optimal policy (proof in Appendix A.12).

**Proposition 5.** (i) For a sufficiently high initial environmental quality $S_0$, the social planner avoids a disaster if the inverse elasticity of intertemporal substitution $\eta \geq 1$; or if $\eta < 1$ and the discount rate $\rho$ is sufficiently low. The South asymptotically fully specializes in the nonpolluting sector if initially clean technologies are less developed than dirty ones there ($A_{00}^S \leq A_{00}^D$).

(ii) If $A_{00}^S \leq A_{00}^D$, $S_0$ is sufficiently high, $\rho$ is sufficiently small, and $\eta \leq 1$, then the mass of scientists allocated to clean technologies in the North is asymptotically 1, both countries fully specialize and the optimal trade tax reaches 0 in finite time.

As the North cannot fully control the Southern economy, avoiding a disaster may not be feasible if $S_0$ is low. Yet, when it is feasible, a social planner will do so if the elasticity of intertemporal substitution $\eta \geq 1$ (as then a disaster brings a utility of $-\infty$), or if $\eta < 1$ and the discount rate is sufficiently low (as then the social planner maximizes long-run utility growth and $S = 0$ is an absorbing state). To avoid a disaster, the South must asymptotically fully specialize in sector $NP$. In the North, either good $P$ production must be bounded, or sector $P$ innovation must switch to mostly clean. Statement (ii) then specifies sufficient conditions under which the latter occurs. The
optimal policy maximizes long-run growth when $\eta \leq 1$ and the discount rate is low enough. Long-run growth, in return, is maximized if the North asymptotically innovates only in clean technologies and the South in sector $NP$. Then, both countries fully specialize in finite time and the trade tax reaches 0 (as environmental quality fully recovers, the carbon tax also reaches 0).

4.3. Parameter Choices

Here, I briefly describe the parametrization with details in Appendix A.14. A period corresponds to 5 years, and initial values are based on the 2003–2007 world economy, assuming laissez-faire in both countries. The elasticity of intertemporal substitution is unity ($\eta = 1$). The annual time discount rate is 0.015, as in Nordhaus (2008). In line with the CGE literature, the North comprises 33 countries in Annex I of the Kyoto Protocol (including the United States) and the South 18 major countries in the rest of the world. Restricting attention to manufacturing, I compute the world rate of emissions per dollar of value-added in each sector at the available aggregation level, using data on sectoral emissions of CO$_2$ from fossil fuel combustion given by the International Energy Agency (IEA, 2010a), and data on sectoral value added by the United Nations Industrial Development Organization (UNIDO, 2011). The sectors with the highest rate are identified with sector $P$—namely the manufacture of chemicals and chemical products, ISIC code 24, of other nonmetallic mineral products, 26, and of basic metals, 27—and the others with sector $NP$. As already mentioned, the South has a small initial comparative advantage in sector $P$ ($Y^N_{P0}/Y^S_{P0} \times Y^S_{NP0}/Y^N_{NP0} = 0.77$).

I compute the consumption shares using world production of both sectors and find $\nu = 0.257$. I use the ratio of capital to labor compensations in both sectors in the United States according to the EU KLEMS dataset to compute the capital shares: $\alpha = 0.5$ for sector $P$ and $\beta = 0.3$ for sector $NP$. The share of intermediates is $\gamma = 1/3$, a common value in endogenous growth models. The elasticity of substitution between the clean and the dirty input, $\varepsilon$ is fixed at 5, but Appendix A.15 considers the cases of $\varepsilon = 3$ and 10. The innovation size $\kappa$ is adjusted so that the long-run annual growth rate is 2%, and the concavity of the innovation function is fixed by choosing $\iota = 0.55$.

---

28See Table A.1 in Appendix A.14. According to the model, I ignore emissions from sector $NP$, which corresponds to the other sectors in manufacturing (except 23, 25, 33, 36 and 37, for which data are not available).
The quality of the environment $S_t$ is linearly and negatively related to the atmospheric concentration of CO$_2$; the assumption that $S_0 = S$ is relaxed, and the initial environmental quality $S_0$ corresponds to the atmospheric concentration in 2003-2007 (379 ppm). $\Delta$ is chosen such that, for $S_t = S_0$, half of CO$_2$ emissions are absorbed and do not add to atmospheric concentrations. Changes in atmospheric CO$_2$ concentrations are mapped against changes in temperature, and $S = 0$ is chosen to correspond to a disaster temperature level of 6°C. The function $\nu(S_t)$ is the same as in AABH and mimics the cost function of Nordhaus (2008) for temperature increases up to 3°C. I identify the ratio $Y_d^X/Y_d^0$ with the ratio of nonfossil to fossil fuel energy produced for country $X$’s primary energy supply (following IEA, 2010b). From this, I derive the ratio $A_d^X/A_d^0$. This, together with the emission rates in sector $P$ in both countries, gives the emission rates per unit of dirty input $\xi^X$.

Yet, the model is still very stylized and the numerical exercise should not be taken too literally. A more complete calibration would feature a more realistic carbon cycle, a more detailed trade model where domestic goods and imports are not perfect substitutes, a nontradable sector, and some technologies in sector $P$ common to the clean and the dirty inputs. Such an exercise would deserve a separate paper and is left for future research.

4.4. Simulation results

Figure 1 describes the first-best policy. Sector $P$ innovation immediately switches to clean technologies, and is rapidly only carried out in the South, as both countries quickly fully specialize (see Figure 1.A, the North clean, North dirty and South dirty lines are indistinguishable from the x-axis and the scientists allocated to sector $NP$ are not represented). This rapid full specialization results from a relatively large growth rate (2% a year), combined with a small difference in capital shares between the two sectors ($\alpha - \beta = 0.2$) and a small initial comparative advantage. In reality, the specialization process likely occurs slower because of imperfect mobility of factors, cross-sector or cross-country knowledge spillovers, or imperfect substitutability between domestic and foreign goods. Ad valorem carbon taxes decline and eventually reach 0 as the environment recovers (Figure 1.B). This occurs faster in the South as clean technologies catch up with dirty ones.

---

$^{29}$ Overall the emission rate in sector $P$ in the South is nearly 4 times that of the North’s, so that $\xi^S > \xi^N$, even though $A_d^N/A_d^0 < A_d^S/A_d^0$. 

28
Figure 2 shows the second-best policy. The North must now export good $P$ in the long run. For these parameter values, a large trade tax on good $P$ (Figure 2.B) ensures that from the first period, the South specializes in sector $NP$, and thus does not innovate in sector $P$ (in Figure 2.A, the South clean and dirty lines are indistinguishable from the x-axis, all Southern innovation is in sector $NP$). Several factors explain this feature: the South should specialize rapidly in sector $NP$ due to its high emission rate, the pattern of trade is easily reversed because the initial comparative advantage is small and full specialization is reached quickly because the South is relatively small and the factor share difference $(\alpha - \beta)$ is small. The switch from predominantly dirty to clean innovation in sector $P$ occurs after 65 years. It is delayed relative to the first-best because the initial sector-$P$ emission rate is lower in the North, which reduces the initial temperature increase, and because investing in dirty technologies helps the North build a large comparative advantage in sector $P$. Clean innovation increases over time and $s_{ct}^N$ asymptotes 1 as the North eventually fully specializes in sector $P$ (in line with Proposition 5).

The optimal unilateral policy avoids a disaster but involves a reversal in the pattern of trade, which is quite costly. Table 1 reports the welfare costs of the first-best and second best-policies, computed as the equivalent percentage loss of world consumption every period relative to the first-best case in a “miracle” scenario under which the dirty input would cease to pollute (i.e. $\xi_N = \xi_S = 0$). As already emphasized, the numbers should not be taken
literally considering the limits of the numerical exercise. Yet, comparing the relative costs of the two policies is still of interest: not being able to intervene in the South increases the welfare costs by a factor 4. Indeed, reversing the pattern of comparative advantages leads to significant static costs in the first periods and to lower productivity levels in subsequent periods. Therefore unilateral intervention is possible but a global one is much preferred.\footnote{This increase in cost is almost entirely due to the environmental externality. In the miracle case, the inability to intervene in the South generates welfare cost since innovation there is not allocated optimally, but these costs are very small: 0.03\%.}

Table 1: Disaster and welfare cost

<table>
<thead>
<tr>
<th></th>
<th>First-best</th>
<th>Second-best</th>
<th>Third-best</th>
</tr>
</thead>
<tbody>
<tr>
<td>Welfare cost (%)</td>
<td>6.36</td>
<td>24.64</td>
<td>24.75</td>
</tr>
</tbody>
</table>

Table 1 also presents the case of a “third” best in which the North can implement (clean) research subsidies and a positive carbon tax but cannot implement trade, consumption, or production taxes. Because of the small difference in initial comparative advantages, the calibrated parameters violate the assumptions of Remark 1 so that a disaster can be avoided with such a policy. In fact the welfare costs of dispensing with the trade tax are small, because the North can quickly acquire a comparative advantage in sector $P$ by innovating even more in that sector than in the second-best. Appendix A.15 further discusses the distributional consequences. Perhaps surprisingly most of the economic costs of the unilateral policy are born by the North:
because of this policy, the South eventually specializes in the sector with the larger consumption share, which increases its share of world income.

4.5. Trade and Directed Technical Change, Two Double-Edged Swords

Figure 3 shows the temperature increase for different policies with and without trade, in laissez-faire and under various policies. Laissez-faire leads to an environmental disaster after 50 years in the open economy case but occurs later in autarky (as economic growth is lower). Under free-trade, no combination of a positive carbon tax or a tax on dirty research in the North can prevent an environmental disaster as in Proposition 2. Figures 3.A depicts the combination that minimizes CO$_2$ emissions (“Taxes on Good $P$ in the North Only”), the curve is indistinguishable from the laissez-faire one, as it is impossible to even delay a disaster with such a policy when trade is allowed. In contrast, in autarky, such a policy can postpone the disaster for 85 years, as there is no pollution haven effect. The second-best curve in Figure 3.A shows how the appropriate unilateral intervention avoids an environmental disaster, while the same instruments cannot affect emissions much in autarky (in Figure 3.A, the second-best refers to the maximization of (19), while in Figure 3.B, it is the combination of research subsidies and positive carbon tax which minimizes CO$_2$ emissions). Even in the first-best, temperature increases more in autarky than with trade because clean technologies grow slower.$^{31}$ Overall, Figure 2 illustrates the double-edged nature of trade: without it, unilateral policies cannot prevent a disaster; but opening up to trade accelerates environmental degradation if the North does not undertake the appropriate policy (this relates this paper to the literature on the impact of trade on the environment, e.g. Copeland and Taylor, 1995).

Directed technical change (DTC) plays a similar role. To study it, I compare the current scenario with DTC to one in which innovation is exogenous and equal in all subsectors ($s^X_{ct} = s^X_{dt} = s^X_{NPt} = 1/3$). With the calibrated values, Northern taxes on the polluting good cannot postpone the disaster even in the exogenous growth case. So as to better illustrate the impact of DTC, I change the parameter values to $\alpha = 0.7$ and $\beta = 0.1$. (A larger difference in capital shares limits the pollution haven effect in a static model and therefore better illustrates how it is amplified by the innovation response.)

$^{31}$In the open economy case, the temperature is initially higher in the first-best than in the second-best because the South’s emission rate is higher. Yet, since the switch to clean innovation occurs sooner, the temperature also starts decreasing sooner.
Figure 4 shows that DTC accelerates the disaster under laissez-faire because it accelerates the economy’s growth rate. A combination of positive carbon tax and tax on dirty research in the North cannot postpone a disaster with DTC, but it can delay it for up to 30 years without DTC as the dynamic pollution haven effect is then absent. Here, the second-best policy can avoid a disaster both with and without DTC, but without DTC, the increase in temperature is much larger—despite a lower growth rate—and a large trade tax must be permanently maintained to reverse the pattern of trade.

In fact, there are parameters for which unilateral policies cannot prevent a disaster without DTC, regardless of initial environmental quality. To avoid a disaster, the North must use mostly the clean input to produce good $P$ and force the South to asymptotically fully specialize in sector $NP$. In the most extreme scenario, the North would only produce good $P$ (mostly with the clean input) and give it for free to the South. Yet, without DTC, the ratio of relative productivities stays constant, so even that would not be enough to push the South towards full specialization and to avoid a disaster if initially the South has a large comparative advantage in sector $P$, or if clean technologies in the North are sufficiently backward. This thought experiment demonstrates that innovation’s ability to affect comparative advantage is essential in deriving the previous results.

5. Knowledge Diffusion

So far the paper has ruled out cross-country knowledge spillovers. In reality, some productivity improvements cross borders, mitigating the am-
Figure 4: Temperature increase with and without directed technical change (capital shares modified to: $\alpha = 0.7$, $\beta = 0.1$). From left to right: figures 3.A and 3.B.

The augmentation of comparative advantage effect, which partly drove the previous results.\textsuperscript{32} This brings into question the robustness of the previous analysis. I now assume that the lagging country can benefit from the diffusion of innovations produced in the leading country (Appendix A.18 considers another extension where innovation is undertaken by multinational firms).

Specifically, before any innovation occurs, the producer of intermediate $i$ in sector $z \in \{c, d, NP\}$ gains access to the technology $A_{zit}^{X} = \max (A_{z(t-1)}^{-X}/A_{z(t-1)}^{X}, 1)^{\Omega} A_{z(t-1)}^{X}$, where $\Omega \in [0, 1]$ measures the strength of technological diffusion. Aggregate productivity then obeys the law of motion:

$$A_{zt}^{X} = (1 + \kappa (s_{zt}^{X})^\gamma) 1^{-\gamma} \max (A_{z(t-1)}^{-X}/A_{z(t-1)}^{X}, 1)^{\Omega} A_{z(t-1)}^{X}$$

for $z \in \{c, d, NP\}$. When $\Omega > 0$, the technology ratio across countries cannot diverge: if a country acquires a strong lead, the other partly catches up in the next period. Yet, the main intuitions of the baseline model carry through.

Northern policies that foster clean innovation in the North now also increase the productivity of clean Southern technologies. They may even put the South on a clean innovation track. If, in some period, pre-innovation

\textsuperscript{32}One should not expect all productivity improvement to cross borders easily, because some may be embedded in capital or depend on local know-how. Dechezleprêtre et al. (2011) suggest that clean technology transfers between developing and developed countries exist but are limited: only 15 percent of clean innovations in 2000-2005 were patented in more than one country; compared to 17 percent for all innovations.
clean Southern technologies become more advanced than dirty ones (i.e., for some $t$, $A_{ct}^S > A_{dt}^S$), market forces will induce more clean than dirty innovations in the South from that period onwards. Preventing a disaster does not necessarily involve pushing the South toward specializing in sector $NP$ any more; it can also be achieved by ensuring a switch to clean innovation there. That transition will occur if more scientists are allocated to clean technologies in the North than to dirty technologies in the South for a sufficient amount of time. Clean innovation in the North and dirty innovation in the South enter a horse race, which determines whether or not good $P$ will be produced in a clean way in the long-run. Which wins depends on the policies that the North allows for and on the pattern of comparative advantage, as in Section 3. This leads to the following result (proof in Appendix A.16).

**Proposition 6.** Assume that i) the assumptions of Lemma 1 are met; ii) technologies are sufficiently close to each other across countries; iii) the South is relatively well-endowed in capital, $K^S/L^S > K^N/L^N$; and iv) clean technologies are sufficiently less advanced than dirty ones ($A_{c0}^S/A_{d0}^S$ sufficiently small). Then, no combination of a carbon tax and a tax on dirty research in the North can prevent a disaster, irrespective of how high initial environmental quality $S_0$ is.

This proposition shows that a result akin to Proposition 2 exists even with large spillovers. The intuition is similar, if in laissez-faire there is more dirty innovation in the South than clean in the North, then a Northern carbon tax cannot reverse this pattern (since on the contrary it increases sector $P$ innovation in the South, as per Proposition 1). Because the market for good $P$ innovation is too small in the North, clean innovation loses the horse race and a disaster cannot be avoided. Specifically, assumptions ii) and iii) ensure that the South keeps a comparative advantage in sector $P$ in laissez-faire and assumption iv) ensures that in this case, the South innovates more in sector $P$ than the North. Instead of assumptions i) and ii) one may assume that $K^S/L^S/(K^N/L^N)$ is large enough.

If the North corresponds to Annex I countries, assumption iii) taken literally seems less likely to hold than its counterpart in Proposition 2, which only stipulates that the South has a comparative advantage in sector $P$. In fact, if the North has a large “endowments”-comparative advantage in sector $P$ (that is for $(K^N/L^N)/(K^S/L^S)$ large enough) and knowledge spillovers are strong enough, it can prevent a disaster using a carbon tax for sufficiently
high $S_0$. This suggests that knowledge spillovers may weaken the conclusion
that Annex I countries cannot prevent worldwide emissions from growing
using a carbon tax only. Yet this conclusion may be premature: a broader
and more likely to hold interpretation of Assumption iii) is that the South
has a comparative advantage in $P$ for reasons beyond imitable technological
factors, such as factor endowments (capital, labor but also human capital
and natural resources), policies, market distortions, etc ...

As before, a temporary combination of clean research subsidies and a
tariff can prevent a disaster for sufficiently large initial environmental quality:
Proposition 3 still holds. Clean research subsidies can reallocate Northern
innovation to clean technologies, and a tariff can limit Southern innovation
in dirty technologies. Then $\bar{A}_{ct}^S$ grows faster than $\bar{A}_{dt}^S$, and a switch to clean
innovation eventually occurs in the South.$^{33}$

<table>
<thead>
<tr>
<th>Table 2: Welfare cost in the presence of knowledge spillovers</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Omega = 0.4$ (%)</td>
</tr>
<tr>
<td>$\Omega = 0.8$ (%)</td>
</tr>
</tbody>
</table>

The first-best and second-best policies are broadly similar to before, but
the trade tax and research subsidies must take knowledge spillovers into ac-
count, and the second-best policy may prevent a disaster with the South
exporting good $P$ in the long-run. In addition, the welfare costs of unilateral
intervention are typically lower. Indeed, the reversal in comparative advan-
tages, which generated the large welfare cost in the no-spillover case, may
not happen, and even if it does, is much less costly since the South ends up
benefiting from the technologies that the North had developed. Accordingly,
Table 2 shows that with knowledge spillovers ($\Omega = 0.4$ and $\Omega = 0.8$), the
welfare costs of the first-best policy are very similar to those in Table 1, but
those of the second-best policy are much lower.

To some extent, (clean) technology diffusion itself can be enhanced by
policies such as laxer intellectual property rights, the clean development
mechanism, direct financing of projects abroad, or migrations of skilled work-
ers. Therefore, this analysis shows that such policies render a tariff less
necessary, and significantly reduce the costs of a unilateral intervention.

$^{33}$If the clean and dirty inputs are imperfect substitutes, Remark 1 no longer holds and
clean research subsidies are enough to prevent a disaster if $S_0$ is high enough.
5.1. An extension with nontradables and different endowments of scientists

Finally, I extend the model to include nontradables. Final consumption is now given by a Cobb-Douglas aggregate of nontradable \((C^X_{NT})\) and tradable goods \((C^X_T)\), where the consumption share of tradables is given by \(a\), that is:

\[
C^X = (C^X_T)^a (C^X_{NT})^{1-a}.
\]

Both goods are produced according to (1), with the associated goods \(P\) and \(NP\) (and the associated subsectors \(c\) and \(d\)), but for the nontradable good, the polluting and non-polluting inputs must be sourced locally. The same intermediates are used whether the good is produced for the tradable or nontradable sector. Without spillovers \((\Omega = 0)\), Northern policies cannot prevent a disaster because Southern emissions will grow at a positive rate from nontradables. In addition, I allow for the mass of scientists in the South \((s^S)\) and the North to differ \((s^N)\). For simplicity, I focus on the case where the clean and dirty inputs are perfect substitutes, \(\varepsilon = \infty\)—the results can be generalized if the initial ratio \(A^S_{c0}/A^S_{d0}\) is small enough.

Since both goods must be produced in each country, innovation must occur in both sectors in laissez-faire regardless of factor endowments or technology levels. The minimal amount of sector-\(P\) innovation in laissez faire, \(s^X_P\), and the maximal amount, are defined as the unique solutions to

\[
\frac{(s^X_P)^{1-\nu}}{(s^X - s^X_P)^{1-\nu}} \left(1 + \kappa \left(\frac{s^X_P}{s^X}ight)\right) = \frac{\nu (1-a)}{1 - \nu (1-a)}
\]

and

\[
\frac{(s^X)^{1-\nu}}{(s^X - s^X_P)^{1-\nu}} \left(1 + \kappa \left(\frac{s^X_P}{s^X}ight)\right) = \frac{1 - (1-\nu)(1-a)}{(1-\nu)(1-a)}.
\]

As shown in Appendix A.17, one then gets:

**Proposition 7.** Part 1) Assume that dirty technologies are initially sufficiently more advanced than clean in both countries so that \(A^X_{d0}/A^X_{c0} > \left(1 + \kappa \left(\frac{s^S}{s^P}\right)\right)^{1-\gamma}\) and that knowledge spillovers \(\Omega\) are strong enough. Then, a) unilateral policies can never prevent a disaster if \(s^N < s^S_P\). b) No combination of a carbon tax and a tax on dirty research in the North can prevent a disaster provided that either i) the South is sufficiently relatively well-endowed in capital.
(so that $K^S/L^S > \left( (1 + \kappa (s^S)^{\lambda}) (1 + \kappa (s^N)^{\lambda}) \right)^{\frac{1-\kappa}{1-\beta}} K^N/L^N$) and $s^S \geq s^N$; or ii) $(K^N/K^S)^{\alpha} (L^N/L^S)^{1-\alpha}$ is small enough, $(K^N/K^S)^{\beta} (L^N/L^S)^{1-\beta}$ is large enough, and $\frac{s^N}{L^P} < \frac{s^S}{L^P}$.

Part 2) If $\Omega > 0$ and $S_0$ is high enough, then the North can prevent a disaster using a) only research subsidies alone if $s^N > \frac{s^S}{L^P}$; or b) only a carbon tax (or a tax on dirty research) if $s^N > \frac{s^S}{L^P}$.

In line with Proposition 2, a unilateral carbon tax is generally unable to prevent a disaster particularly when the South has a large comparative advantage in sector $P$ even if knowledge spillovers are very large (part 1b)).$^{34}$ As $s^X_P$ and $\frac{L^X_P}{X^P}$ increase in $s^X$, it is only if the North has much more innovation resources than the South that such a policy may succeed regardless of the pattern of comparative advantage (part 2b)). Further, since $\frac{L^X_P}{X^P}$ increases in $a$ and $s^X_P$ decreases in $a$, a larger share of tradables makes it more likely that it fails. The North can avoid a disaster using research subsidies (Proposition 3 survives) as long as it has enough innovation resources (part 2a)). Otherwise, the South may still innovate more in dirty technologies for its nontradable sector than the North can innovate in clean making it impossible to prevent a disaster with any unilateral policy (part 1a)).$^{35}$

6. Conclusion

This paper shows that when evaluating the long-term consequences of unilateral environmental policies, it is essential to consider their impact on the allocation of innovation within the polluting sector between technologies (clean/dirty) and between countries (intervening/non-intervening). The propositions in the text are model-specific but they allow to illustrate fundamental intuitions. First, the pollution haven effect becomes worse in a dynamic setting. Positive taxes on the polluting sector in the North risk

$^{34}$Part 1) assumes a high level of knowledge spillovers in order to give simple conditions on factor endowments and initial productivity ratios $A^X_{0}/A^X_{0}$ under which Northern policies cannot prevent a disaster. Yet, in general a high $\Omega$ does not make it harder to prevent an environmental disaster.

$^{35}$Part 2a) characterizes a range of parameters for which a disaster can be avoided with research subsidies only, allowing for a trade tax expands this range.
placing the economy on a path where the South has a comparative advantage in that sector. This leads to the relocation of not only the production of the polluting good but also of innovation in that sector, which hampers the benefits of such a policy on worldwide emissions. The South innovates more in dirty technologies, while the market for clean innovation in the North is reduced.\textsuperscript{36} Second, sustainable growth can be achieved without cooperation from the South, but this requires a “green industrial policy” (with clean research subsidies and perhaps a trade tax) in order to ensure that there is more clean than dirty innovation worldwide. Such a policy can guarantee that either the North acquires a long-run comparative advantage in the polluting sector, or, with knowledge spillovers, that a switch towards clean innovation occurs in the South.

The paper argues that unilateral environmental policies should aim at developing clean technologies, which have the potential to reduce emissions in the North, but also in the South either through technology diffusion or by slowing down the move of polluting industries there. These policies should be thought of as transitory until a satisfactory global agreement is reached. I analyzed what “well-intentioned” countries should do until then, and therefore, as a first step, I took as given the absence of such an agreement.\textsuperscript{37}

The next step is to analyze why some countries are willing to participate and others are not, and how unilateral policies shape their intentions in the long-run. One hypothesis is that richer countries are more likely to participate because their citizens are more willing to sacrifice current consumption for environmental quality or future generations (although of course, the non-signing of the Kyoto protocol by the United States challenges this view). In that case, economic development would naturally push more and more countries in the participating coalition. A second hypothesis emphasizes the benefit that the reluctant country could obtain from joining the coalition.

\textsuperscript{36}Therefore, DTC renders Southern emissions much more responsive to Northern policies in the long run. This finding calls into question existing estimates of the carbon leakage rate obtained from static CGE models.

\textsuperscript{37}Within the framework of this paper, one can study the case of a myopic South government which maximizes current consumption, and therefore implements its own trade tax to improve its terms of trade. When the South has a comparative advantage in the polluting good, this trade tax moves both countries closer to autarky. Clean research subsidies in the North can still ensure that more clean innovation occurs there than dirty innovation in the South, and therefore the North can still prevent an environmental disaster for sufficiently high initial environmental quality.
Unilateral policies can affect this benefit in at least three dimensions: by decreasing environmental damages which discourages a reluctant country from joining (the free-rider problem), by developing clean technologies which can diffuse and therefore reduce the costs of an environmental policy for the reluctant country, and by affecting comparative advantages and therefore the impact of a potential environmental policy on the reluctant country’s terms of trade (as analyzed in a static framework by Copeland and Taylor, 2005).

7. Acknowledgments

I am very grateful to my advisors Philippe Aghion, Daron Acemoglu, Pol Antràs and Elhanan Helpman for their invaluable guidance. I thank Andres Rodriguez-Clare and two anonymous referees for their suggestions. I thank David Atkin, Steve Cicala, Richard Cooper, Rafael Dix-Carneiro, Vasco Carvalho, Dave Donaldson, Emmanuel Farhi, Gita Gopinath, Adam Guren, Christian Hellwig, Per Krussell, Marc Melitz, Nathan Nunn, Morten Olsen, Jennifer Page, Torsten Persson, Dorothée Rouzet, Robert Stavins, Vania Stavrakeva, Jean Tirole, Anthony Venables, and Martin Weitzman for their helpful comments. I also thank seminar and conference participants at Harvard University, Harvard University Kennedy School of Government, Toulouse School of Economics, IIES Stockholm, INSEAD, University of Maryland, Ohio State University, Pennsylvania State University, the conference on macroeconomics and climate change at Yale University, the conference on climate and the economy at IIES Stockholm, Ecole Polytechnique, the EAERE 2013 conference, ETH Zurich, CIRED, the Oxcarre conference and Paris School of Economics.

References


