The Dynamic Impact of Unilateral Environmental Policies

David Hémous*

June 15, 2014

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 polluting good generates more or less emissions depending on its relative use of a clean and a dirty input. I show that a unilateral policy combining clean research subsidies and a trade tax can ensure sustainable growth, while unilateral carbon taxes alone generally cannot. Relative to autarky and exogenous technical change respectively, the mechanisms of 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 only.

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

Keywords: climate change, environment, directed technical change, innovation, trade, unilateral policy

*INSEAD and CEPR, david.hemous@insead.edu. 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. "

1 Introduction

Despite the signature of the Kyoto protocol at the end of 1997, annual carbon dioxide (CO_2) emissions increased by 39% between 1997 and 2010. Meanwhile, climate negotiations have stalled and no global agreement is in sight. In response, 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. Partly because of the sharp increase in emissions from developing countries, numerous policymakers have suggested that these policies should now be complemented by protectionist measures both to ensure that domestic firms do not suffer a competitive disadvantage and prevent carbon leakage (i.e., an increase in emissions in non-participating countries).[1] This situation raises two questions. First, can well-designed unilateral policies achieve the necessary reduction in CO_2 emissions? Second, are the calls for protectionism justified?

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 these dynamic aspects. This paper, on the contrary, builds a dynamic framework by integrating directed technical change with a trade model that features a global pollution externality. More 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 tradeable goods with a high CO_2 emission intensity, typically energy-intensive sectors. It is be 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 polluting or the nonpolluting sector, and, within the polluting sector, at clean or dirty technologies. For most of the analysis, innovation is completely 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. (2012a; henceforth AABH), the allocation of innovation within the polluting sector exhibits path-dependence, also because

[1] Carbon tariffs for the EU ETS were discussed by the European Commission in 2008, and are championed in particular by France. The extension of the EU ETS to airline emissions originally covered flights from foreign airlines to or from Europe and was therefore a first attempt at taxing foreign production. Yet, at the moment, its application has been restricted to the European market. In the United States, the American Clean Energy and Security Act of 2009—which was supposed to set up a cap-and-trade system—planned to implement trade barriers with countries that did not have a similar system, absent an international agreement, by 2018.

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 a reallocation of some of the polluting good's production from the North 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, which combines 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 while making that sector cleaner at the same time. This ensures that emissions eventually start decreasing in both countries. If the initial environmental quality is high enough, then an environmental disaster can be averted. Importantly, directed technical change is essential for this result; if technical change were exogenous, 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 natural next step is to derive the optimal unilateral policy, which I do analytically. I also conduct a numerical exercise to illustrate the results (in accordance with the literature, the South is then identified with countries which have no binding constraints under the Kyoto protocol). 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. When the social planner values equally consumption in the North and the South, the trade tax typically takes the form of a tariff and then of an export subsidy. Its expression reflects two aims of the social planner: reducing emissions in the South and redirecting Southern innovation toward the nonpolluting sector.[2]

An important caveat for unilateral policies is that even though avoiding a disaster is possible, the welfare costs from not being able to intervene in the South may be very large, as is the case in the numerical exercise. The numerical exercise also highlights the double-edged nature of trade and directed technical change. Relative to autarky and exogenous technical change respectively, the mechanisms of trade and directed technical change accelerate environmental degradation either under laissez-faire or with unilateral carbon taxes,

[2]When the social planner values consumption in the North and the South differently, the trade tax is also used to affect the terms of trade.

yet both help reduce environmental degradation under the appropriate unilateral policy.

Finally, the model is enriched by including 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 scenario, however, 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 still specializes in the polluting good. In addition, the welfare costs from not being able to intervene in the South are much lower.

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). Here as well, a country risks specializing in the "wrong" sector, not because that sector offers poor growth prospects, but because this country cannot prevent the environmental externality associated with production in that sector. The idea that free trade may amplify comparative advantages and that a temporary trade policy could permanently reverse the trade pattern was previously touched on by Krugman (1987), and Grossman and Helpman (1991, ch. 8).[3]

The literature on trade and the environment has long recognized that, in an open world, the effectiveness of unilateral policies for reducing world pollution can be hampered by the pollution haven effect (Pethig, 1976). Empirical evidence is reported by Copeland and Taylor (2004) and more recently by Broner, Bustos and Carvalho (2012). Markusen (1975) and Hoel (1996) show that the optimal instrument for addressing the pollution haven effect is a tariff. In the specific context of global warming, where the pollutant (CO_2) enters differently at several stages of the production process, several papers use computable general equilibrium (CGE) models to track carbon through the global economy; in this way they 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 percent 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 percent from a reduction in Annex I countries—i.e., the countries with binding constraints under the Kyoto protocol—and show that border tax adjustments eliminate half of it.[4] There are comparatively few empirical

[3] Krugman's (1987) is based on learning-by-doing, and Grossman and Helpman's (1991) model features endogenous growth in one sector only. A few papers have built models with trade and directed technical change; examples include Acemoglu (2003), who studies the impact of trade on the skill bias of technological change, and Gancia and Bonfiglioli (2008), who show that trade amplifies international wage differences.

[4] Among others, Babiker and Rutherford (2005); Böhringer, Fisher and Rosendahl (2010); and Böhringer, Carbon and Rutherford (2011) find similar results.

studies. Aichele and Felbermayr (2012) find that countries which committed to the Kyoto protocol reduced domestic CO_2 emissions by about 7 percent, but that their total CO_2 consumption did not change. Against the backdrop of literature that has focused on static models, the novelty of the present paper is to incorporate dynamic aspects. This comes at the expense of a more detailed model of world trade (as in CGE models) and of a study of the strategic interactions between countries (as in Copeland and Taylor, 2005).

A growing literature has shown the importance of taking into account directed technical change when designing policies to combat climate change. On the empirical side, Popp (2002) shows that an increase in energy prices leads to more energy-saving innovation; similar results are found by Newell, Jaffe and Stavins (1999) in the air conditioner industry and by Hassler, Krusell and Olovsson (2012) using macroeconomic US data. Aghion et al. (2012) focus on the car industry and establish that *(a)* an increase in fuel prices leads to clean innovation at the expense of dirty innovation and *(b)* there is path dependence in clean versus dirty innovation—findings in line with the results reported here. Following this literature, several theoretical papers have integrated directed technical change in the study of climate change policies; here, I build on the model developed by AABH.[5] The final good in AABH and the polluting sector in this paper are produced similarly 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" as in this paper. Therefore, in AABH carbon taxes can still prevent an environmental disaster by redirecting innovation towards clean technologies, while, here, in an open economy, this is not true any more, clean research subsidies are often necessary to prevent a disaster and carbon taxes may even be counter productive. Acemoglu, Aghion and Hémous (2014) presents a two-country version of AABH in which trade occurs between two substitutable goods, the polluting tradeable good cannot become less polluting, and the South does not innovate. Here, I reverse these three assumptions which provides a more realistic and richer framework (for instance, in that paper, carbon taxes necessarily reduce the amount of emissions in the long-run). Di Maria and Smulders (2004) and Di Maria and van der Werf (2008) also tackle the issue of modeling the interaction between directed technical change and international trade. These authors study the allocation of innovation between an energy-intensive sector and a non–energy-intensive sector, but overlook that innovations within the energy-intensive sector could either reduce or increase pollution.[6]

Section 2 presents the model. Section 3 studies the laissez-faire equilibrium, identifies which policies are able to ensure sustainable growth and discusses the model's main as-

[5] Earlier work on the environment and directed technical change includes Bovenberg and Smulders (1995) Goulder and Schneider (1999), van der Zwaan et al. (2002) , Popp (2004) and Grimaud and Rouge (2008). See also Acemoglu et al. (2012b).

[6] In Di Maria and Smulders (2004), the North develops technologies that are imitated by the South, and so 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 cut in emissions in a single country.

sumptions. Section 4 solves for the second-best policy when the South is constrained to be in laissez-faire, and presents a numerical exercise which illustrates the workings of the model. Finally, Section 5 discusses how the main results generalize when knowledge flows across countries. The proofs, details on the calibration and some extensions are available in the online Appendix.

2 Model

I consider a discrete-time, infinite-horizon version of a two-country (North, N, and South, S), two-sector (E and F), three-factor (capital, labor and scientists) Heckscher–Ohlin–Ricardo model in which sector E 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, reality is more complex and most countries lay somewhere between these two extremes. As already mentioned in the introduction, the division North-South need not fall along the lines of developed versus developing countries, in particular because the United States have not signed the Kyoto protocol. Yet, in the numerical exercise, I follow the CGE literature and identify the North with the countries which were subject to binding constraints under the Kyoto protocol (Annex I countries), including the United States, and the South with the rest of the world.

2.1 Welfare

I consider two distinct problems. In the first one, 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 $\nu(S_t) C_t^X$, where S_t is the quality of the environment (identical in North and South) and C_t^X is the final good consumption in country X. The social welfare function aggregates these preferences according to:

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

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.

In the second problem, the economy admits infinitely lived representative agents in each country, whose utilities are given by $\sum_{t=0}^{\infty} \frac{1}{(1+\rho)^t} \frac{\left(v(S_t) C_t^X\right)^{1-\eta}}{1-\eta}$. The social planner maximizes a weighted sum of these utilities:

$$U=\sum_{t=0}^{\infty}\frac{1}{(1+\rho)^{t}}\frac{v\left(S_{t}\right)^{1-\eta}}{1-\eta}\left(\Psi\left(C_{t}^{N}\right)^{1-\eta}+(1-\Psi)\left(C_{t}^{S}\right)^{1-\eta}\right), \tag{2}$$

where $\Psi \in [0,1]$ is the weight on the North's representative agent. In this case, the social planner also cares about the distribution of consumption across the two countries. Since the social planner will be the only one making an intertemporal decision, the fact that the representative agents have different time-horizon under the two problems does not matter.[7]

Consumption, C_t^X, and environmental quality, S_t, are weakly positive and v is increasing in S_t. There is an upper-bound on S_t, denoted $\overline{S}$, that corresponds to a pristine environment. I define an *environmental disaster* as an instance of environmental quality reaching zero in finite time. I assume that $v(0)=0$ and $v'\left(\overline{S}\right)=0$; hence a disaster is as detrimental to welfare as zero consumption and the marginal damage of the first unit of pollution is zero.[8]

2.2 Production

Final consumption is a Cobb-Douglas aggregate of the consumption of two goods, E and F:

$$C^{X}=\left(C_{E}^{X}\right)^{\nu}\left(C_{F}^{X}\right)^{(1-\nu)}, \tag{3}$$

where C_Y^X represents the quantity of good $Y \in \{E,F\}$ consumed in country $X \in \{N,S\}$.[9] 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.[10] Goods E and F are the only goods that are traded internationally. Good E represents tradeable goods the produc-

[7] Under (2) and for $\Psi = 1/2$, the social planner has concerns for intragenerational inequality (between North and South) as long as he has concerns for intergenerational inequality ($\eta \neq 0$). It is a well-understood problem in climate economics that maximizing social welfare would then lead to recommend not only environmental policies but also major redistributional policies. The solution used by the CGE literature has been to solve (2) with time-varying Negishi weights: Ψ is time varying and chosen as the inverse of the marginal utility of consumption in laissez-faire, so that the social planner aims at keeping the distribution of income identical to the one in laissez-faire. Note that this is not a proper social welfare function if the representative agents live for more than 1 period. Instead of following this route, I use the much more transparent approach of solving both for the maximization of (1) and (2). In the case where the social planner maximizes (1), he is in fact indifferent to the distribution of income (instead of aiming at preserving it). See Stanton (2011) for a discussion of these issues.

[8] A disaster puts the economy on an unsustainable path because the utility flow cannot be bounded away from the utility flow given by zero consumption.

[9] Whenever this does not lead to confusion, I drop the time subscript but it should be clear that allocations, technologies and policies are time-dependent (endowments are constant though).

[10] A previous version of the paper (CEPR Discussion Paper 9733) does so. On the other hand, the analysis would be different if the elasticity of substitution is greater than 1 as then both goods are not essential, and good F can be used to replace good E. This is not a very likely case considering what good E stands for in practice.

tion of which generates a lot of greenhouse gases emissions (in particular energy-intensive goods), while good F represents the other tradeable goods. When the model is calibrated, good E is identified with the sectors which have the highest CO_2 emissions over value-added ratio, namely the manufacture of chemicals and chemical products (ISIC code 24), other nonmetallic mineral products (26), and basic metals (27), good F is identified with the rest of manufacturing (see Table A.1 in Appendix A.8). Even though not all emissions can be traced to the tradeable sector, the paper initially focuses on tradeable goods, since it is because of international trade that policymakers fear that unilateral policies may have adverse consequences. Emissions for the production of tradeable goods represent a significant share of CO_2 emissions—once electricity and heat are allocated to consuming sectors, manufacturing and construction represented 36.9 % of world CO_2 emissions in 2010 according to the International Energy Agency.[11] The inclusion of nontradeables is discussed in Section 5.

Good E is produced competitively with a clean input Y_c^X and a dirty input Y_d^X following

$$Y_E^X = \left(\left(Y_c^X\right)^{\frac{\varepsilon-1}{\varepsilon}} + \left(Y_d^X\right)^{\frac{\varepsilon-1}{\varepsilon}}\right)^{\frac{\varepsilon}{\varepsilon-1}}, \tag{4}$$

where $\varepsilon > 1$ is the elasticity of substitution between the clean and the dirty input. 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.

Goods c, d and F in country X are produced competitively according to

$$Y_F^X = \left(\int_0^1 A_{Fi}^X \left(x_i^X\right)^\gamma di\right)\left(\left(K_f^X\right)^\beta \left(L_f^X\right)^{1-\beta}\right)^{1-\gamma} \text{ and} \tag{5}$$

$$Y_{zt}^X = \left(\int_0^1 A_{zi}^X \left(x_{zi}^X\right)^\gamma di\right)\left(\left(K_z^X\right)^\alpha \left(L_z^X\right)^{1-\alpha}\right)^{1-\gamma} \text{ for } z \in \{c, d\}. \tag{6}$$

K_f^X and L_f^X are the capital and labor employed in the assembly of good F in country X; x_{Fi}^X is the quantity of intermediates i employed in sector F; and A_{Fi}^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 $z \in \{c, d\}$. γ is the factor share of intermediates. Intermediates cannot be traded internationally and are produced monopolistically according to

$$x_{Fi}^X = \psi^{-1}\left(K_{fi}^X\right)^\beta \left(L_{fi}^X\right)^{1-\beta} \text{ and } x_{zi}^X = \psi^{-1}\left(K_{zi}^X\right)^\alpha \left(L_{zi}^X\right)^{1-\alpha} \text{ for } z \in \{c, d\}. \tag{7}$$

K_{fi}^X and L_{fi}^X are the capital and labor employed in the production of intermediate i for

[11] Construction is non-tradeable, but agriculture and forestry, which are tradeable activities, are not included in this figure. Using input-output tables Davis and Caldeira (2010) estimates that today, 23% of carbon emitted is attributable to the production of goods that will be exported.

good F in country X (and K_{zi}^X and L_{zi}^X are defined similarly). Since the same factor share is used in the production of intermediates and in the final assembly of the good, it follows that $\beta \in (0,1)$ is the overall factor share of capital in sector F, and $\alpha \in (0,1)$ is the overall factor share of capital in sector $z \in \{c,d\}$.[12] Therefore the production of goods c and d and the production of good F only differ in the capital share. I assume throughout that $\alpha > \beta$, which is true empirically: within tradeables, 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 $\alpha = \beta$.[13]

I use K_F^X and K_E^X to denote total employment of capital in sector F and E in country X, so that:

$$K_F^X \equiv K_f^X + \int_0^1 K_{Fi}^X di \text{ and } K_E^X \equiv K_c^X + K_d^X + \int_0^1 K_{ci}^X di + \int_0^1 K_{di}^X di, \tag{8}$$

similarly, L_F^X and L_E^X denote the total employment of labor in sector E in country X. Then, factor market clearing and good market clearing imply:

$$K_E^X + K_F^X \leq K^X \text{ and } L_E^X + L_F^X \leq L^X, \tag{9}$$

$$C_E^N + C_E^S \leq Y_E^N + Y_E^S \text{ and } C_F^N + C_F^S \leq Y_F^N + Y_F^S. \tag{10}$$

Furthermore, to simplify the exposition and focus the comparison between first-best and second-best on environmental issues, I assume throughout that the optimal subsidy to the purchase of intermediates $(1-\gamma)$ is implemented in both countries, so that all intermediates are priced at marginal cost. 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 for my results. Henceforth I abuse language by referring to the "laissez-faire" case as one where governments only implement the subsidy to the use of all intermediates.

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

[12] The Cobb–Douglas structure of production for intermediates is important because it ensures that monopolists get a constant share of the sector's revenues, which matters for the incentives to innovate. That being said, the analysis can be extended straightforwardly to production functions for which aggregation between capital and labor is not Cobb–Douglas.

[13] A Ricardian model would pose some technical difficulties for section 4 as explained below. With different factor shares in the two sectors, the analysis is not significantly more complex, the model can account for situations where both countries do not fully specialize and it will later be easy to introduce knowledge spillovers, as I will then need another reason for trade than technological differences. There is nothing special about capital and labor being the two factors here instead of high-skill and low-skill labor for instance. This is why the paper abstracts from capital accumulation.

hiring s^X_{zit} scientists, the entrepreneur holding the monopoly right on variety i in sector $z = F$ or subsectors $z \in \{c, d\}$ can increase the initial productivity $A^X_{zi(t-1)}$ of her intermediate to

$$A^X_{zit} = \left(1 + \kappa \left(s^X_{zit}\right)^{\iota} \left(\frac{A^X_{z(t-1)}}{A^X_{zi(t-1)}}\right)^{\frac{1}{1-\gamma}}\right)^{1-\gamma} A^X_{zi(t-1)} \text{ for } z \in \{c, d, F\}, \tag{11}$$

where $0 < \iota < 1$. A^X_{zt} is the average productivity of (sub)sector $z \in \{c, d, F\}$ at time t, and is defined as

$$A^X_{zt} \equiv \left(\int_0^1 \left(A^X_{zit}\right)^{\frac{1}{1-\gamma}} di\right)^{1-\gamma} \text{ for } z \in \{c, d, F\}. \tag{12}$$

The factor $\left(A^X_{zi(t-1)}\right)^{-\frac{1}{1-\gamma}}$ captures decreasing returns to scale in innovation (the more advanced is a technology, the more difficult it is to innovate further), and $\left(A^X_{z(t-1)}\right)^{\frac{1}{1-\gamma}}$ denotes knowledge spillovers from all the other intermediates in the same sector in the same country. 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). Since the mass of scientists is equal to 1 in both countries, the market clearing equation is given by

$$\int_0^1 \left(s^X_{Fit} + s^X_{cit} + s^X_{dit}\right) di \leq 1. \tag{13}$$

Because an entrepreneur has monopoly rights for one period only, she will hire scientists so as to maximize current profits instead of the entire flow of profits generated by the innovations of her scientists. The allocation of scientists across (sub)sectors is therefore myopic. One-period monopoly rights are the only inefficiency in innovation and they allow one 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, and why direct research incentives in addition to carbon taxes are welfare improving, a point made by AABH and Gerlagh et al. (2014).[14]

There are no knowledge spillovers between 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

[14] With permanent monopoly rights, infinitely lived agents, and no environmental externality, the efficient innovation allocation would be an equilibrium, but usually not the only one.

of technical change and ensures that one country does not become arbitrarily large relative to the other. This assumption is further discussed in section 3.4.

2.4 Environment

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

$$S_t = (1 + \Delta) S_{t-1} - \left(\xi^N Y_{dt}^N + \xi^S Y_{dt}^S\right). \tag{14}$$

The parameter $\xi^X > 0$ measures the rate of environmental degradation from the production of dirty inputs (which may be different in the two countries) and $\Delta > 0$ 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 then S_t reaches the disaster level.[15] The dirty input is directly responsible for environmental degradation. This specification is equivalent to one where a (cheap) fossil fuel resource can be combined with the dirty input in a Leontieff way.[16]

2.5 Policy tools

Section 4 solves the social planner's problem of maximizing (1) or (2), but Section 3 studies only whether or not an environmental disaster can be prevented with specific policy instruments, the ones that will eventually be used to decentralize the optimal policy. A policy is characterized by a sequence of *ad valorem* taxes on the dirty input τ_t^X in each country (the equivalent of a carbon tax), a sequence of *ad valorem* subsidies for scientists in each country and each subsector,[17] and a sequence of *ad valorem* trade taxes b_t on the polluting

[15] Real climate dynamics are much more complicated. In particular, 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.

[16] 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. Note that changes in the type of fossil fuel used (from coal to natural gas) can significantly affect the emission rate, yet, modeling such a possibility would not affect the propositions of the paper.

[17] In order to ensure uniqueness of the equilibrium allocation of scientists, I assume that it is possible to subsidize only a given mass of scientists; hence the social planner can use the subsidy to determine the exact allocation. If the subsidy is greater than 100 percent, then a monopolist may be willing to hire scientists even if she is not producing any good.

good (by Lerner symmetry, they could equally be on the other good). All subsidies and taxes are financed (or rebated) through lump-sum taxation at the country level.

The trade tax is implemented by the North, so that prices in the South are equal to international prices: $p^S_{Et} = p_{Et}$ and $p^S_{Ft} = p_{Ft}$, while prices in the North follow $p^N_{Ft} = p_{Ft}$ and $p^N_{Et} = p_{Et}(1 + b_t)$. A positive trade tax corresponds to a tariff (resp., export subsidy) when the North imports (resp., exports) good E.[18] When the North is the only country intervening, I assume that trade balance must be maintained every period (there is no intertemporal trade):

$$p_{Et}\left(Y^S_{Et} - C^S_{Et}\right) + p_{Ft}\left(Y^S_{Ft} - C^S_{Ft}\right) = 0. \tag{15}$$

Note that the trade tax is *not* explicitly related to the carbon content of imports. If the South does not undertake any policy, then relating the tax to the *average* carbon content of imports from a given country and in a given sector would not alter the results; since each Southern firm is atomistic, its impact on average emission is infinitesimal and so its behavior will not affect the trade tax it pays. Changing the behavior of Southern firms would require either the North to know the exact carbon content of each individual import, which seems implausible, or the South to implement a policy in response to the North's tariff.

3 Preventing an Environmental Disaster

This section presents positive results on whether certain type of policies can or cannot avert an environmental disaster. Section 3.1 details the behavior of the economy under laissez-faire. Section 3.2 explains why taxing the North's polluting sector likely fails to prevent a disaster. Section 3.3 describes how a disaster can be avoided using unilateral policies in the North, and Section 3.4 discusses some of the assumptions. For a given policy, the equilibrium is defined as follows.

Definition 1 *A* feasible allocation *is a sequence of demands for capital (K^X_{ht} K^X_{Fit}, K^X_{ct}, K^X_{cit}, K^X_{dt}, K^X_{dit}), demands for labor (L^X_{ht}, L^X_{Fit}, L^X_{ct}, L^X_{cit}, L^X_{dt}, L^X_{dit}), demands for intermediates (x^X_{zit} for $z \in \{c, d\}, F$), demands for inputs $\left(Y^X_{ct}, Y^X_{dt}\right)$, goods production $\left(Y^X_{Et}, Y^X_{Ft}\right)$, demands for goods $\left(C^X_{Et}, C^X_{Ft}\right)$, research allocations $\left(s^X_{zit}\text{ for } z \in \{c, d\}, F\right)$, and quality of the environment S_t such that, in each period t and in each country $X \in \{N, S\}$, factor and good markets clear (i.e.,* (9), (10)*, and* (13) *hold).*

Definition 2 *For a given policy, an* equilibrium *is given by a feasible allocation and sequences of wages of workers (w^X_t), returns to capital (r^X_t), wages of scientists (υ^X_t), consumer prices for intermediates $\left(\varphi^X_{zit}\text{ for } z \in \{c, d\}, F\right)$, producer prices for clean and dirty*

[18] Starting from a situation where the North imports the polluting good 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 the polluting good and the trade tax is a positive export subsidy.

inputs (p^X_{ct}, p^X_{dt})*, and international prices of goods* (p_{Et}, p_{Ft}) *for* $X \in \{N, S\}$ *such that: (i)* $\left(\varphi^X_{zit}, x^X_{zit}, s^X_{zit}, K^X_{zit}, L^X_{zit}\right)$ *maximizes profits by the producer of intermediate* i *in sector* $z \in \{c, d, F\}$ *in country* X*; (ii)* L^X_{zt}*, and* K^X_{zt} *maximize the profits of the producer of good* $z \in \{c, d, F\}$*; (iii)* Y^X_{ct} *and* Y^X_{dt} *maximize the profits of producer of good* E*; (iv)* C^X_{Et} *and* C^X_{Ft} *maximize consumers' utility under the trade balance constraint (15).*

3.1 Laissez-Faire

Trade pattern. Here I analyze the laissez-faire equilibrium; the results are derived and generalized in Appendix A.1. In each country, aggregate production in each sector can be written as

$$Y^X_{Et} = \zeta A^X_{Et} \left(K^X_{Et}\right)^{\alpha} \left(L^X_{Et}\right)^{1-\alpha} \text{ and } Y^X_{Ft} = \zeta A^X_{Ft} \left(K^X_{Ft}\right)^{\beta} \left(L^X_{Ft}\right)^{1-\beta}, \tag{16}$$

where $\zeta \equiv \gamma^{\gamma} (1-\gamma)^{1-\gamma} \psi^{-\gamma}$ and $A^X_{Et} \equiv \left(\left(A^X_{ct}\right)^{\varepsilon-1} + \left(A^X_{dt}\right)^{\varepsilon-1}\right)^{\frac{1}{\varepsilon-1}}$ is the average productivity of sector E. This formulation highlights that, in a given period, the model collapses to a Heckscher–Ohlin model with varying productivity across countries. The South has the comparative advantage in the polluting good E if and only if

$$\left(\frac{A^S_{Et}}{A^S_{Ft}}\right)^{\frac{1}{\alpha-\beta}} \frac{K^S}{L^S} > \left(\frac{A^N_{Et}}{A^N_{Ft}}\right)^{\frac{1}{\alpha-\beta}} \frac{K^N}{L^N}. \tag{17}$$

Trade results from Ricardian forces (relative productivity) as well as Heckscher–Ohlin forces (relative factors endowment). Provided the difference in comparative advantage is not too large, both countries produce both goods. When the difference in comparative advantage is larger, one and eventually both countries fully specialize.

Emissions are given by $E^X_t = \xi^X \left(\frac{A^X_{dt}}{A^X_{Et}}\right)^{\varepsilon} Y^X_{Et}$. Thus the emission rate in the polluting sector is increasing in the ratio of dirty to clean productivities A^X_{dt}/A^X_{ct}. Over time, innovation changes the comparative advantage and the emission rate.

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

$$\pi^X_{zit} = (1-\gamma) \left(\frac{A^X_{zit}}{A^X_{zt}}\right)^{\frac{1}{1-\gamma}} p^X_{zt} Y^X_{zt}. \tag{18}$$

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^X_{zit}. 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 according to

$$A_{zt}^{X} = \left(1 + \kappa \left(s_{zt}^{X}\right)^{\iota}\right)^{1-\gamma} A_{z(t-1)}^{X} \text{ for } z \in \{c, d, F\}.$$

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

$$\frac{\left(s_{ct}^{X}\right)^{1-\iota}\left(1 + \kappa \left(s_{ct}^{X}\right)^{\iota}\right)}{\left(s_{dt}^{X}\right)^{1-\iota}\left(1 + \kappa \left(s_{dt}^{X}\right)^{\iota}\right)} = \frac{p_{ct}^{X} Y_{ct}^{X}}{p_{dt}^{X} Y_{dt}^{X}} = \frac{\left(A_{ct}^{X}\right)^{\varepsilon-1}}{\left(A_{dt}^{X}\right)^{\varepsilon-1}}. \tag{19}$$

The second equality follows from the demand equation for both inputs in sector E (knowing that the production technologies differ only by their productivity level). The ratio of revenues in the clean sector to those in the dirty sector increases with the ratio of clean to dirty technologies. This association reflects 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. Thus, for a sufficiently small innovation size κ, more scientists are allocated to the dirty than to the clean subsector if and only if $A_{d(t-1)}^{X} > A_{c(t-1)}^{X}$ (if κ is too large then there may be multiple equilibria when $A_{d(t-1)}^{X}$ and $A_{c(t-1)}^{X}$ are close to each other; see Appendix A.2). So, within the polluting sector, under laissez-faire, innovation tends to be allocated to the sector that is already the most advanced: there is path dependence.

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 F, c and d, I obtain

$$\frac{\left(s_{ct}^{X}\right)^{1-\iota}\left(1 + \kappa \left(s_{ct}^{X}\right)^{\iota}\right) + \left(s_{dt}^{X}\right)^{1-\iota}\left(1 + \kappa \left(s_{dt}^{X}\right)^{\iota}\right)}{\left(s_{Ft}^{X}\right)^{1-\iota}\left(1 + \kappa \left(s_{Ft}^{X}\right)^{\iota}\right)} = \frac{p_{Et}^{X} Y_{Et}^{X}}{p_{Ft}^{X} Y_{Ft}^{X}}. \tag{20}$$

Therefore, for a given ratio $A_{d(t-1)}^{X}/A_{c(t-1)}^{X}$ of initial productivities within sector E, the number of scientists allocated to sector E is increasing in the ratio of sector E to sector F revenues. Under free trade, prices are equalized in North and South; hence each country tends to innovate relatively more in the sector it exports, and does so at an equal ratio of initial productivities within sector E. Comparative advantages are typically amplified over time, so that one and eventually both countries fully specialize.

This contrasts with the autarky case. Consumer demand implies that

$$\frac{p_{Ft}^{X} Y_{Ft}^{X}}{p_{Et}^{X} Y_{Et}^{X}} = \frac{1-\nu}{\nu}, \tag{21}$$

so that innovation always occurs in both sectors.

Equilibrium uniqueness. As innovating more in a sector increases a country's comparative advantage in that sector, which, in turn, prompts more innovation in the same sector, multiple equilibria could arise. The results of this section can be extended to such a case, but focusing on a unique equilibrium simplifies the exposition. Henceforth, I assume that the conditions of the following lemma are satisfied, so that the equilibrium is unique.[19]

Lemma 1 *If κ is small enough and $\iota \geq 1/2$, the laissez-faire equilibrium is unique.*

Proof. See Appendix A.2. ■

Environmental disaster. Under laissez-faire, as long as dirty technologies are more advanced than clean ones in both countries, innovation in the polluting sector remain directed primarily toward dirty technologies. Since innovation in the polluting sector does not asymptotically vanish (the exporting country innovates more in the polluting good than it would under autarky), the production of good E grows unboundedly and so do emissions. At some point, the regenerative capacity of the environment becomes overwhelmed and the economy reaches an environmental disaster.

As in AABH, a global government could use clean research subsidies, taxes on dirty research and/or carbon taxes to redirect innovation from the dirty toward the clean subsector in countries that produce the polluting good. Once clean technologies acquire a sufficient lead over dirty intermediates, market forces will ensure that most research is directed toward the clean subsector, which is now the most advanced. Eventually, the emission rate of the polluting good approaches zero—sufficiently fast to offset the growth in the polluting good's production—and a disaster can be avoided for sufficiently high initial environmental quality (see Appendix A.3).

3.2 Taxes on the Polluting Good in the North only

Assume now that only the North implements some policy (ruling out the case where the North pays the South to implement some policy). Is this alone enough to avoid environmental disaster? Observe that, in autarky and without knowledge spillovers, no policy restricted to the North can prevent a disaster because Southern emissions grow unboundedly regardless of what the North does. Absent international cooperation, trade is necessary to avoid an environmental disaster. Now, the key to avoid environmental disaster with Northern policies only is ensuring that the South asymptotically fully specializes in the nonpolluting sector. Otherwise, innovation in the polluting sector always occurs in the

[19] A sufficiently small size of innovation κ ensures that changes in productivities during one period remain sufficiently small. The technical assumption $\iota \geq 1/2$ is further necessary to ensure that the equilibrium is unique when one country is close to a corner of specialization (i.e., to a point at which a producer of the imported good would break even only if he produces an infinitesimal amount of the good). The lemma does not extend to the Ricardian case where $\alpha = \beta$: in that case, no matter how small κ is, there are multiple equilibria when the initial comparative advantage is small.

South, and the production of the polluting good and therefore emissions grow unboundedly (see Appendix A.4 for a formal proof).

I first focus on the case where the North can implement a positive carbon tax and/or a positive tax on dirty research. Both instruments have no protectionist aspects, can reduce emissions in the North, and prompt clean innovation there; and both could prevent an environmental disaster if the North were the only country or if the South undertook the same policy. However, such policies may be incompatible with a South specializing in the nonpolluting sector and thus may fail to prevent an environmental disaster.

Proposition 1 *If innovation size κ is small enough then, no matter how high $\overline{S}$ 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 the North ($A_{c0}^N/A_{d0}^N \leq 1$); (ii) clean technologies are sufficiently less developed than dirty ones in the South (A_{c0}^S/A_{d0}^S is sufficiently small); and (iii) the South has a weak initial comparative advantage in the polluting sector (i.e., $\left(A_{E0}^S/A_{F0}^S\right)^{\frac{1}{\alpha-\beta}} K^S/L^S \geq \left(A_{E0}^N/A_{F0}^N\right)^{\frac{1}{\alpha-\beta}} K^N/L^N$).*

Proof. See Appendix A.5. ■

Under laissez-faire and with the assumptions of the proposition, the South keeps its comparative advantage in the polluting sector, eventually specializing in that sector. The North government cannot reverse this pattern simply by using a positive tax on dirty research or a positive carbon tax. On the contrary, a positive tax on dirty innovation drives scientists away from the polluting sector E toward the nonpolluting sector F; moreover, within the polluting sector it allocates innovation toward the initially backward clean subsector, which further reduces the growth rate of average productivity A_{Et}^N. A positive carbon tax has the same effect on innovation and also directly reduces the productivity of the polluting sector in the North. Since both instruments increase the costs of producing the polluting good in the North, they lead to an increase in its world relative price. This induces an increase in production of the polluting good E in the South and hence more emissions there, which is the classic pollution haven effect. As the relative revenues of the polluting sector increase in the South, and following equation (20), Southern innovation is further tilted toward the polluting sector, where it is mostly directed at the dirty technologies. Accordingly, positive Northern taxes on the polluting good can only accelerate the Southern specialization in the polluting sector.

In fact, such policies are likely to *accelerate* environmental degradation because of the reallocation of innovation in the South. Indeed, the economy tends to grow faster when countries are more specialized since there is less overlap in the type of innovations being undertaken by both countries. In addition, the gap between clean and dirty technologies in the South grows faster, which increases the South's emissions rate. Both effects work towards an increase in emissions. Furthermore, although carbon taxes and taxes on dirty research can tilt innovation within the polluting sector toward clean technologies, they typically fail to ensure that such technologies get significantly developed. As production of the polluting sector moves to the South, the market size for clean technologies in the North becomes too small to attract much innovation.

In Proposition 1, the condition that A^S_{c0}/A^S_{d0} be sufficiently small (and not simply less than 1) 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 the polluting sector might take place in the former even if the latter exports the polluting good.[20] The condition could be dispensed with if the initial comparative advantage were sufficiently large.

The crucial hypothesis of Proposition 1 is that the South has a comparative advantage in the polluting sector. 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 the polluting sector initially in the numerical exercise in section 4.1). Yet, with a different definition of the North, this hypothesis may not hold, in which case, 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 work in opposite directions.

3.3 Introducing Clean Research Subsidies and Trade Taxes

I now allow the North to use clean research subsidies and a trade tax. Both policies have protectionist aspects, in that the clean research subsidy is a conditional subsidy granted to the polluting sector, which is the sector facing import competition (when the North has a comparative advantage in the non-polluting sector).

Proposition 2 *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 $\overline{S}$ is sufficiently high.*

The key difference between clean research subsidies and the carbon tax or the tax on dirty research is that the former can also reallocate scientists who were working in the nonpolluting sector F toward the clean subsector. This boosts innovation in clean technologies in the North, even when the North does not have the comparative advantage in the polluting sector. Increasing innovation in clean technologies makes the polluting sector less polluting and helps build a comparative advantage in the polluting sector. In the meantime, a positive trade tax reduces production and therefore innovation in the polluting sector in the South, which also helps reverse the pattern of comparative advantage. For sufficiently

[20] More specifically: the incentive to innovate in sector G is, *ceteris paribus*, lower when the revenues in the clean and dirty subsectors are close to each other — that is when $\left(A^X_{c(t-1)}\right)^{\varepsilon-1}$ and $(1+\tau_{Xt})^{-\varepsilon}\left(A^X_{d(t-1)}\right)^{\varepsilon-1}$ are comparable. Given carbon taxes that are high enough or taxes on dirty research that are of sufficient duration, the ratio of clean to dirty revenues may become farther from unity in the North than in the South. In that event, the assumption on A^S_{c0}/A^S_{d0} ensures that, when this occurs, the difference in comparative advantages is large enough to ensure that there is more polluting innovation in the South.

high initial environmental quality, a policy combining these two instruments can prevent a disaster. To see this, consider the following two-phase approach (this is not the optimal policy, which is derived in Section 4). First, a social planner implements a tariff large enough to shut down trade, so that innovation in the South must be balanced between the polluting and nonpolluting sectors. Simultaneously, she implements large clean research subsidies so that nearly all Northern scientists innovate in the clean subsector, and the North innovates more in the polluting sector E than the South. Once the North has acquired the comparative advantage in the polluting sector and $A^N_{c(t-1)}/A^N_{d(t-1)}$ is sufficiently large, the social planner can discontinue all policies and re-open up to trade. Market forces then ensure that the production of the polluting good eventually moves entirely to the North where it relies essentially on clean technologies,[21] emissions go down to zero in both countries, and a disaster can be avoided.

From this discussion 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 as the following remark stipulates, it does not always hold.

Remark 1 There exist initial factor endowments and technologies, such that no matter how high $\overline{S}$ is, no combination of a carbon tax, a tax on dirty research, and a subsidy for clean research can prevent a disaster.

Proof. See Appendix A.6. ∎

Clean research subsidies alone cannot prevent a disaster when the South fully specializes in the polluting sector and clean technologies in the South are sufficiently less advanced than dirty ones. In that case, all Southern scientists are allocated to the polluting sector and, asymptotically, to dirty technologies. So even if the North were to allocate all its scientists to clean technologies, A^S_{Et} would grow as fast as A^N_{Et}. 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 the polluting sector is sufficiently large or when clean technologies are sufficiently backward in the North, as the average productivity of the polluting sector in the North, A^N_{Et}, grows slowly during the period when clean technologies are catching up with dirty ones.[22]

3.4 Discussion

Here, I discuss some of the assumptions of the model and present additional results regarding alternative instruments.

[21] This follows lemma A.3, applied to the case where the North now has the comparative advantage in the polluting sector at some date τ, with $A^N_{d\tau}/A^N_{c\tau} < A^S_{c\tau}/A^S_{d\tau} < 1$.

[22] This is the only result of this section that would not hold if goods E and F were strict complements (instead of Cobb-Douglas): in this case the South could not stay fully specialized in the polluting sector if both countries innovate only in that sector.

Size. The relative size of the two countries in terms of capital and labor endowments plays a role quantitatively: the larger the North is, the easier it is to reverse comparative advantages. In the long-run, the relative size of the two economies depends crucially on the mass of scientists, which is a proxy for the amount of resources spent on innovation. Here we have considered the case where the two countries have the same mass of scientists, so that our analysis is implicitly restricted to a situation where the two countries (or groups of countries) are of similar size. Appendix A.7 considers alternative scenarios, and shows generally that the larger the mass of scientists in the North relative to the South, the easier it is for unilateral policies to prevent an environmental disaster. Unsurpisingly, this implies that the inclusion of the United States in the North is crucial. In addition, one may think that the mass of scientists in the South is bound to increase, making it harder and harder for the North to intervene decisively as time passes.

Other instruments. It is clear that the reasoning behind Proposition 2 extends to the case where the North uses a combination of clean research subsidies and subsidies to the polluting good (which is relevant if trade taxes are impossible), a combination of carbon taxes and trade taxes (relevant if targeted research subsidies are impossible to implement),[23] or subsidies to the production of the clean input alone (relevant if both research subsidies and trade taxes are impossible). Paradoxically, a negative carbon tax combined with a positive tax on dirty research might also avert a disaster when a positive carbon tax could not: the negative carbon tax can be used to reverse the pattern of trade while the tax on dirty research can ensure that innovation occurs in clean technologies.[24] So far I have assumed that the North cannot find the carbon content of imports at the firm level. Under the opposite assumption, trade taxes related to the emission content of imports could directly influence the behavior of Southern firms. In some cases this instrument (combined, for instance, with a carbon tax in the North, which is what a carbon consumption tax would look like) can prevent a disaster by inducing a switch to clean innovation in the South; but, this requires that clean technologies are not too backward in the South and that the export market is large.

Three sectors. The results of the paper crucially depend on the assumption that innovation may occur in all three (sub)sectors (clean, dirty and non-polluting). If innovation were limited to clean and dirty technologies within the polluting sector, then the North could not build a comparative advantage in a specific sector. With clean innovation in the polluting sector only (as in Di Maria and Smulders, 2004; and Di Maria and van der Werf, 2008), the model would falsely assume that all innovations in the polluting sector decrease

[23] For this combination to work, it is important that trade taxes large enough to reverse the pattern of trade immediately are allowed. A trade tax that implements autarky is generally insufficient to prevent a disaster when combined with carbon taxes (while it is sufficient when combined with clean research subsidies).

[24] This scenario is not very realistic: achieving the right combination of a large negative carbon tax and a large positive tax on dirty research seems difficult, moreover, since emissions are likely to increase considerably in the short-run, the initial level of environmental quality necessary to avert a disaster can be very high.

emissions. On the contrary, 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.[25]

Assumptions on ε and the elasticity of substitution between goods E and F are also crucial. If $\varepsilon < 1$, avoiding a disaster with unilateral policies is not possible: it requires positive growth in the polluting sector in the North to ensure that Southern emissions do not grow unboundedly, but positive growth in the polluting sector is not possible without positive growth in dirty input production. As already mentioned, the analysis extends to the case where the elasticity of substitution between goods E and F is strictly lower than 1, but not the the case where it is strictly greater. In the later case goods E and F are not both essential. Whether an environmental disaster can be avoided or not depends crucially on how large A_d^S is relative to A_F^S: a large A_d^S may push the South towards producing the dirty input rather than the non-polluting good, regardless of the policy in the North.

Importantly, note that dirty innovations generally include not only 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,[26] or the introduction of new goods or inputs that rely on fossil fuel energy. In practice, some innovations in the polluting sector may complement both fossil fuel energy and alternative forms of energy; one could represent such innovations in this model as improving the productivity of an additional input in the polluting sector complement to both the clean and dirty inputs. This would not affect the economic intuitions developed and my results could be extended to this scenario.

The South's behavior. The paper assumes that the South does not implement any policy. Regarding environmental policy today, this seems a reasonable assumption: several countries seem willing to move forward, while others are opposing a global agreement while often undertaking very limited domestic policies (Barrett, 1994, explains why designing a self-enforcing international agreement on climate change is difficult). A reason why these divisions may persist in reality is the significant delay between emissions and damages that climate models predict, an aspect that I abstract from here: as a result, it may be too late before skeptic countries get convinced that they should start undertaking significant policy actions. Even if one expects that these divisions will eventually end, the results of the paper are still useful for countries who are willing to intervene before the rest of the world.

Even if the South does not implement any environmental policy, it may still want

[25] Here clean innovations allow to develop an input which substites for the dirty one, and the polluting sector's productivity can grow at the same rate whether it relies mostly on the clean or the dirty inputs. If the clean alternative had some growth's costs, then preventing a disaster with unilateral policies would be more difficult (this would be the case in a different model where the clean alternative refers to energy efficiency improvements, and where energy is complement to the other inputs in the polluting sector).

[26] Aghion et al. (2012) for instance show that the majority of innovation for fossil fuel engines in the automotive industry are of this type.

to implement trade policies, particularly if the North's trade policy hurts the South. Yet, South's consumption is not necessarily negatively affected by the North's unilateral policies, and the South benefits from better environmental quality. For instance, if the North's temporary policy reverses the pattern of comparative advantages, both countries fully specialize in the long run. Income shares are linked to the consumption share of the good that the country exports; therefore, if the income share for the polluting good is smaller than for the nonpolluting one ($\nu < 1 - \nu$), then the South's income share will be larger under the North unilateral policy than under laissez-faire.[27]

Although a full analysis of the strategic interactions between two governments is beyond the scope of this paper, one can consider the case where the South government is myopic and maximizes current consumption. This government implements its own trade tax to improve its terms of trade. As long as the South retains an initial comparative advantage in the polluting good, this trade tax moves both countries closer to autarky and thus does not prevent the North from reversing the pattern of comparative advantage. Once the North exports the polluting good, the South implements its own tariff. This tariff slows down the South's specialization in the nonpolluting sector. Yet, once the North has acquired a sufficiently large comparative advantage, it does not prevent the South from fully specializing in the nonpolluting sector. Therefore, a disaster can still be avoided for sufficiently high initial environmental quality.

4 Optimal policy and numerical illustration

I now turn to the normative part of the paper, characterizing the first-best policy and the second-best policy under the constraint that the social planner cannot intervene in the South. I use a numerical example to illustrate both policies and compute their welfare costs, and to show that both trade and directed technical change act as double-edge swords.

4.1 Parameter Choices

This subsection briefly describes the parametrization; details are given in Appendix A.8. A period corresponds to 5 years, and initial values are based on the 2003–2007 world economy while 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[28] and the South to 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, here using data on sectoral emissions of CO_2

[27] Even in the short run, the South might benefit: a tariff implemented by the North hurts the South when the South exports the polluting good, but a trade tax high enough to reverse the pattern of trade immediately may benefit the South (this trade tax is then an export subsidy).

[28] This includes the US even though the US government has not ratified the treaty.

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 E—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 F.[29] As already mentioned, Southern production is tilted toward sector E relative to Northern production ($Y_{E0}^N/Y_{E0}^S \times Y_{F0}^S/Y_{F0}^N = 0.77$), so that the South has a small initial comparative advantage in the polluting sector E.

The consumption share of good E is computed using world production of both sectors: $\nu = 0.257$. The capital shares are $\alpha = 0.5$ for sector E and $\beta = 0.3$ for sector F, here using the ratio of capital to labor compensations in both sectors in the United States according to the EU KLEMS dataset, Timmer et al. (2008), and the share of intermediates $\gamma = 1/3$, a common value in endogenous growth models. The elasticity of substitution between the clean and the dirty input, ε is fixed at 5, but Appendix A.14 considers the cases of $\varepsilon = 3$ and 10. The innovation size κ is adjusted so that the long-run annual growth rate is 2 percent, and the concavity of the innovation function is fixed by choosing $\iota = 0.55$ (≥ 0.5 so that the equilibrium is unique for a small κ).

The quality of the environment S_t is linearly and negatively related to the atmospheric concentration of CO_2; the assumption that $S_0 = \overline{S}$ is relaxed, and the initial environmental quality S_0 corresponds to the atmospheric concentration in 2003-2007 (379 ppm). Δ 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_{c0}^X/Y_{d0}^X 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_{c0}^X/A_{d0}^X. This, together with the emission rates in sector E in both countries, gives me the emission rates per unit of dirty input ξ^X.[30]

Nevertheless, the model is still very stylized and, at this stage, 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 production and imports are not perfect substitutes, a nontradeable sector, and some technologies in sector E common to both the clean and the dirty inputs. Such an exercise would deserve a separate paper and is left for future research.

[29] According to the model, I ignore emissions from sector F. Sector F corresponds to the other sectors in manufacturing except 23, 25, 33, 36, and 37, for which data are not available.

[30] Overall the emission rates in the polluting sector in the South is nearly 4 times that of the North's, so that, even though $A_{d0}^N/A_{c0}^N < A_{d0}^S/A_{c0}^S$, I have $\xi^S > \xi^N$.

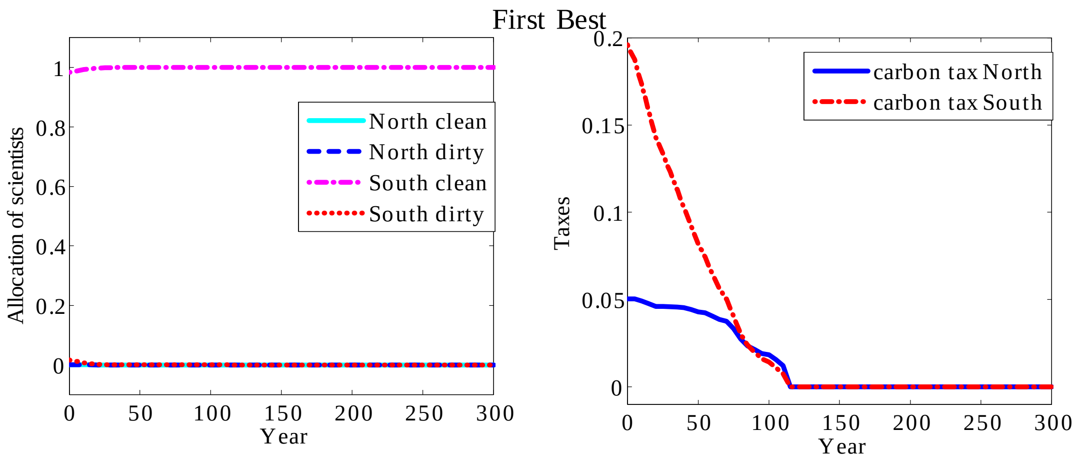

Figure 1: First-best policy. From left to right, figures: 1.A and 1.B.

4.2 First-Best

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 (1) or (2) subject to the following constraints: the production function equations (3), (4), (5), (6), (7) and; the factor market–clearing equations (9) and (13); the goods market–clearing equation (10); the environmental degradation equation (14); and the knowledge accumulation equation (11).

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 *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. When the social planner cares about the cross-country distribution of consumption, transfers are used to equalize the marginal social value of consumption in each country (i.e., $\Psi\left(C_t^N\right)^{-\eta} = (1-\Psi)\left(C_t^S\right)^{-\eta}$). Since utility flow is minimized during a disaster and since the social planner can always reduce world emissions, the optimal policy always avoids a disaster.

Figure 1 describes the first-best policy in our numerical example. Figure 1.A shows that sector-E innovation switches to clean technologies (here immediately), and is rapidly only carried out in the South, since both countries rapidly fully specialize (in the figure the North clean, North dirty and South dirty lines are indistinguishable from the x-axis and the scientists allocated to sector F 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. Either imperfect mobility of factors, cross-sector or cross-country knowledge spillovers, or imperfect substitutability between domestic and foreign goods would have the effect of slowing down the specialization process. Figure 1.B shows the *ad valorem* carbon taxes in both countries (the tax is expressed as a fraction of the dirty good's price), they decline and eventually reach 0 as the environment recovers; it declines faster in the South where clean technologies catch up with dirty ones. More generally, Appendix A.10 demonstrates that if the discount rate ρ is sufficiently small and the inverse elasticity of intertemporal substitution $\eta \leq 1$, then, as in this example, both countries specialize in finite time and innovation in sector E switches to mostly clean.[31] In this case, emissions eventually vanish. With the law of motion (14), the quality of the environment reverts to $\overline{S}$—and the carbon tax reaches zero—in finite time.[32]

4.3 Second-Best

I now turn to the case where the social planner cannot implement any policy in the South, whose economy is in laissez-faire, and cannot transfer income from one country to another. Trade balance must be maintained at every point in time. The second-best policy is defined by the social planner maximizing (1) or (2) subject to the following constraints: (3) for the North and the South; constraints (4), (5), (6), (7), (9), (13) and (11) for the North only; the environmental degradation constraint (14); the goods market–clearing constraints in both countries, which are now written as

$$C^N_{Yt} = Y^N_{Yt} + M_{Yt} \text{ and } C^S_{Yt} = Y^S_{Yt} - M_{Yt}, \text{ for } Y \in \{E, F\}, \tag{22}$$

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

$$p_t M_{Et} + M_{Ft} = 0, \tag{23}$$

where $p_t \equiv p_{Et}/p_{Ft}$ is the international price ratio; and constraints describing the South's laissez-faire economy. These latter constraints (detailed in Appendix A.11) are: a consumer

[31] These are only sufficient conditions, and the optimal policy is likely to feature a switch to clean innovations also when $\eta > 1$.

[32] For an alternative law of motion where environmental regeneration decreases as the quality of the environment S_t approaches $\overline{S}$, or where some emissions are permanent, then S_t may not reach $\overline{S}$ asymptotically. The optimal carbon tax may then not converge to 0 but it becomes irrelevant in the sense that a 0 carbon tax would only have a negligible effect on welfare.

demand equation

$$\left(\frac{\partial C^S}{\partial C_E^S}\right)\left(\frac{\partial C^S}{\partial C_F^S}\right)^{-1}=\frac{\nu}{1-\nu}\frac{C_{Ft}^S}{C_{Et}^S}=p_t; \tag{24}$$

offers equations in the South of the type

$$Y_{Et}^S=y_E^S\left(p_t,A_{Et}^S,A_{Ft}^S\right) \text{ and } Y_{Ft}^S=y_F^S\left(p_t,A_{Et}^S,A_{Ft}^S\right); \tag{25}$$

an emissions equation $Y_{dt}^S=\left(A_{dt}^S/A_{Et}^S\right)^{\varepsilon}Y_{Et}^S$; an equation that specifies the mass of scientists allocated to sector E,

$$s_{Et}^S=s_E^S\left(p_t,A_{dt}^S,A_{ct}^S,A_{Ft}^S\right); \tag{26}$$

and the resulting law of motion of aggregate productivity in the South:

$$\begin{aligned} A_{Ft}^S &= \left(1+\kappa\left(1-s_{Et}^S\right)^{\iota}\right)^{1-\gamma}A_{F(t-1)}^S, \\ A_{zt}^S &= \left(1+\kappa\left(s_{zt}^S\left(s_{Et}^S,a_{t-1}^S\right)\right)^{\iota}\right)^{1-\gamma}A_{z(t-1)}^S, \text{ for } z\in\{c,d\}. \end{aligned} \tag{27}$$

The allocation between clean and dirty innovation s_{ct}^S, s_{dt}^S is uniquely determined by the total mass s_{Et}^S and the ratio $a_{t-1}^S\equiv\left(A_{c(t-1)}^S/A_{d(t-1)}^S\right)^{\varepsilon-1}$. For the problem to be well-defined, the South's equilibrium must be unique given the North's allocation. An argument similar to that of Appendix A.2 shows that it is the case when κ is sufficiently small and $\iota\geq 1/2$. (This is where the Ricardian case would pose a technical difficulty, with $\alpha=\beta$, even for a small κ, the South's equilibrium may not be uniquely defined.) This leads to the following result.

Proposition 3 *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.*

Proof. See Appendix A.11. ■

In this second-best scenario, 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, b_t, allows the social planner to distort prices in the South thereby affecting the allocation of factors there. When the social planner maximizes (1), the optimal allocation satisfies:

$$\begin{aligned} &\nu^{\nu}(1-\nu)^{1-\nu}p_t^{-\nu}\left((1+b_t)^{-\nu}b_tp_t\frac{\partial y_E^S}{\partial p_t}+\left((1+b_t)^{1-\nu}-1\right)\frac{\nu Y_{Ft}^S}{p_t}+\left(1-(1+b_t)^{-\nu}\right)(1-\nu)Y_{Et}^S\right) \\ =\ &\widehat{\omega}_t\xi^S\left(\frac{A_{dt}^S}{A_{Et}^S}\right)^{\varepsilon}\frac{\partial y_E^S}{\partial p_t}-\widehat{\phi}_t\frac{\partial s_{Et}^S}{\partial p_t} \end{aligned} \tag{28}$$

where $\widehat{\omega}_t$ is the shadow value of a unit of environmental quality at time t (in units of consumption at time t) and $\widehat{\phi}_t$ is the shadow value of moving an additional scientist in the South from sector F to sector E. In this expression, the left-hand side has the sign of b_t, which shows that the social planner imposes a wedge between relative prices in the North and in the South. This wedge is generated by an environmental motive (the first term on the right-hand side) and an innovation motive (the second term). The first term is always positive. A positive trade tax on the polluting good E imposed by the North reduces its relative price in the South, which decreases its production there and emissions. The second term is generally also positive as there is typically too much innovation in the polluting sector in the South ($\widehat{\phi}_t < 0$) for two reasons. First, more innovation in the polluting sector in the South leads to more emissions. Second, to avoid a disaster—which the social planner typically does—the South must at least asymptotically fully specialize in the nonpolluting sector, so that current innovations in the polluting sector will be of little use in the future. Because of their myopia, Southern innovators do not internalize this and their innovation efforts are tilted too much toward the polluting good. By reducing the production of the polluting good in the South, a positive trade tax moves Southern scientists from sector E to sector F. Therefore, the trade tax is generally positive; it takes the form of a tariff when the North imports the polluting good and of an export subsidy otherwise.

For the maximization of (2), terms-of-trade matter and the optimal trade tax is modified in order to favor the country with the largest social marginal value of consumption. If the social planner cares only about the North ($\Psi = 1$), then this motive pushes toward a tariff when the North imports the polluting good and toward an export tax otherwise. If the social planner cares equally about both countries ($\Psi = 1/2$) but the South is poorer, then it pushes toward an import or an export subsidy.[33]

The next proposition further characterizes the optimal policy.

Proposition 4 *(i) Whenever doing so is feasible, the social planner avoids a disaster if the inverse elasticity of intertemporal substitution* $\eta \geq 1$*; or if* $\eta < 1$ *and the discount rate* ρ *is sufficiently low. The South must asymptotically be fully specialized in the nonpolluting sector* F *if initially clean technologies are less developed than dirty ones there* ($A^S_{c0} \leq A^S_{d0}$)*. (ii) If* $A^S_{c0} \leq A^S_{d0}$*, avoiding a disaster is feasible,* ρ *is sufficiently small, and the inverse elasticity of intertemporal substitution* $\eta \leq 1$*, then the mass of scientists allocated to clean technologies in the North is asymptotically* 1*.*

Proof. See Appendix A.13. ■

[33] Equation (28) is modified in the following way: $\widehat{\omega}_t$ and $\widehat{\phi}_t$ are shadow values in units of consumption in the North and the right-hand side is followed by the term $\nu^{\nu}(1-\nu)^{1-\nu} p_t^{-\nu} \left(1 - \frac{\lambda_t^S}{\lambda_t^N}\right) \left(\frac{C^S_{Ft}}{C^S_{Et}} \frac{Y^S_{Et}}{Y^S_{Ft}} - 1\right) \frac{\nu Y^S_{Ft}}{p_t}$, where λ_t^X is the shadow value of a unit of consumption at time t in country X. The additional term indicates the terms of trade effects. It is then less straightforward to sign $\widehat{\phi}_t$, because the social value of moving a Southern scientist from one sector to the other also reflects how it affects terms of trade.

Since the North cannot fully control the Southern economy, avoiding a disaster may not be feasible when 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 the nonpolluting sector, and production of the polluting good in the North must be limited or a switch to clean innovations must occur. Statement (ii) specifies *sufficient* conditions under which innovation does indeed switch to mostly clean innovation in the North: long-run growth is maximized if the North asymptotically innovates only in clean technologies, and the optimal policy maximizes long-run growth when $\eta \leq 1$ and the discount rate is low enough.[34]

These results are illustrated in Figure 2 which shows the second-best policies for the cases where the social planner maximizes (1) and (2) with $\Psi = 1$ (so that the North only cares about the welfare of its representative agent). Contrary to the first-best case, the North must now export the polluting good E in the long run. For these parameter values, a large trade tax on good E (see Figures 2.B and 2.D) ensures that, right from the first period, the South specializes in the nonpolluting sector F, and thus does not innovate at all in the polluting sector (see Figures 2.A and 2.C: the South clean and South dirty lines are indistinguishable from the x-axis, all Southern innovation is in sector F). Several factors explain this feature: the high emission rate in the South means that the South should specialize rapidly in sector F, the low initial comparative advantage of the South that the pattern of trade is easily reversed, and the smaller size of the South together with a small difference $\alpha - \beta$ in factor shares between the two sectors imply that full specialization in the South is reached quickly.

With no redistributive motive, the switch from predominantly dirty to clean innovation in the polluting sector occurs after 65 years (Figure 2.A). The switch is delayed relative to the first-best because the North starts with a lower emission rate in the polluting sector, so that the initial temperature increase is lower, and because continuing to invest in dirty technologies helps the North build a large comparative advantage in the polluting sector. It occurs even later (after 215 years) when the North cares only about its own consumption: less innovation in the polluting sector improves the North's terms of trade, as it exports the polluting good, and in return allows for a delayed switch toward clean innovation. The amount of clean innovation increases over time and, beyond the time frame of the simulation, eventually reaches one when the North fully specializes in the polluting sector (in line with Proposition 4).

In the optimum for the North case, the trade tax eventually becomes negative (it reaches -1 asymptotically) as the North eventually acquires the comparative advantage in

[34] For the maximization of (1), one can further shows that, in finite time, both countries fully specialize and the optimal trade tax reaches 0. Since environmental quality can fully recover, the optimal carbon tax reach 0. As in the First Best, the optimal carbon tax would still become irrelevant if the environmental quality were not to recover fully.

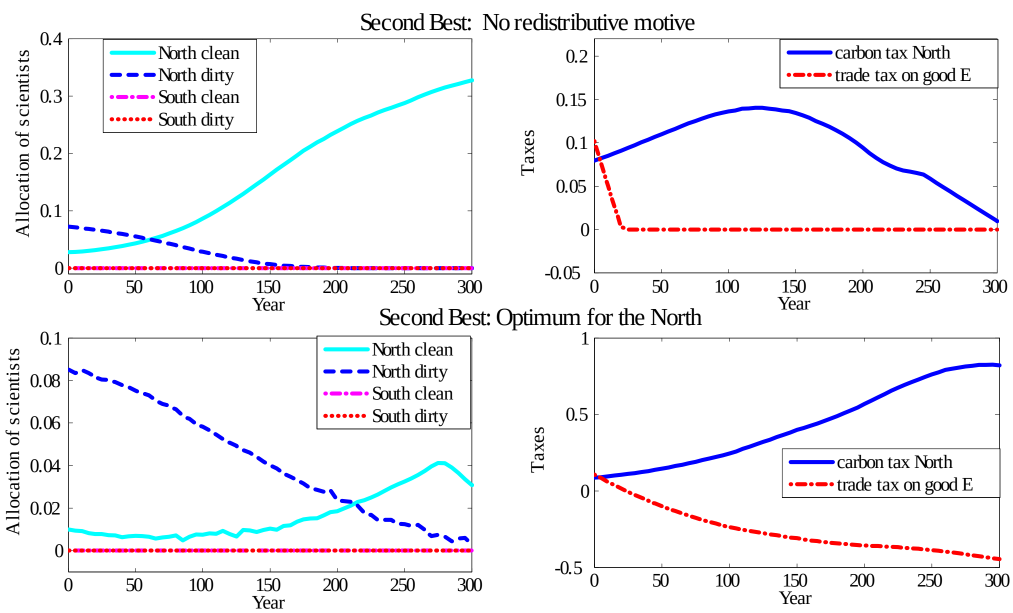

Figure 2: Second-best policies, when the social planner has no redistributive motive and when she cares only about the North. From left to right, top to bottom, figures 2.A, 2.B, 2.C and 2.D.

the polluting sector, and a negative trade tax—an export tax here—improves the North's terms of trade. The later the switch to clean technologies, the more temperature eventually increases, so that the carbon tax is higher than in the first-best and even higher when the social planner cares more about consumption in the North than in the South (see Figures 2.B and 2.D).

4.4 Welfare costs

The unilateral policies which are able to avoid a disaster are quite radical since they involve a reversal in the pattern of trade, so the welfare costs from not being able to intervene in the South may be very large. 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. $\xi^N = \xi^S = 0$ from the first period). As already emphasized, the numbers should not be taken literally considering the limits of the numerical exercise. Yet, comparing the costs of the first-best and second-best policies is still of interest: in this simple model, not being able to intervene in the South increases the welfare costs by a factor 4. The reason is that 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 here but a global one is much preferred.[35]

Table 1: Disaster and welfare cost (with no redistributive motive)

	First-best	Second-best	Third-best
Welfare cost (%)	6.36	24.64	24.75

Table 1 also presents the case of a "third" best in which the North can implement a positive carbon tax and research subsidies/taxes but cannot implement trade, consumption, or production taxes. With the calibrated parameter values it is still possible to avoid disaster under such a policy (which is not always the case, see Remark 1). In fact, the welfare costs of dispensing with the trade tax are not large. With these parameters, the difference in initial comparative advantages is small and innovation is very effective in affecting technological levels. As a result, the North can quickly acquire a comparative advantage in the polluting sector without the help of a trade tax by innovating more in the polluting sector than in the second-best and implementing a low carbon tax initially. For the reasons explained above, the South quickly specializes in the non-polluting sector. Importantly though, this third-best policy still bears protectionist aspects since it indirectly subsidizes the production of the polluting good, which the North initially imported. The distributional impacts of the policy are also interesting and discussed in Appendix A.14.

[35] 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 percent.

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

Figure 3 shows the temperature increase for different policies when trade is allowed for and when the two countries are in autarky in laissez-faire and under various policies. Laissez-faire leads to an environmental disaster after 50 years for the open economy case but occurs later in autarky, since economic growth is lower in that case. Under free-trade and following Proposition 1, no combination of a positive carbon tax or a tax on dirty research in the North can prevent an environmental disaster. Figures 3.A depicts the combination that minimizes CO_2 emissions ("Taxes on Good E in the North Only"), the curve is indistinguishable from the laissez-faire one, as it is not even possible to delay a disaster with such a policy when trade is allowed.[36] On the contrary, 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 adding the same instrument (research subsidies) does not affect emissions much in autarky (in Figure 3.A, the second-best refers to the maximization of (1), 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 case temperature increases more in autarky because the growth rate of clean technologies is lower than in the open economy scenario.[37] 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 the allocation of innovation is exogenous and equal in all subsectors ($s_{ct}^X = s_{dt}^X = s_{Ft}^X = 1/3$). With the calibrated values, however, 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 perform the same exercise but now assume that $\alpha = 0.7$ 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.) Figure 4 shows that DTC accelerates the disaster under laissez-faire because it accelerates the economy's growth rate. With DTC, a disaster cannot be postponed with a combination of positive carbon tax and tax on dirty research in the North: in fact the combination that minimizes CO_2 emissions is no taxes. Without DTC, it is possible to delay an environmental disaster for up to 30 years with this policy because the dynamic pollution haven effect that this paper emphasizes is absent. Here, the second-best policy can avoid a disaster both with and without DTC, but without DTC, the increase

[36] The dynamic aspect is key in obtaining this result, since it also holds when the North and the South initially have the same emission rate (when $A_{c0}^S/A_{d0}^S = A_{c0}^N/A_{d0}^N$ and $\xi^S = \xi^N$):

[37] Comparing the increase in temperature between the first-best and the second-best in the open economy case is interesting. The temperature is initially higher in the first-best because the South's emission rate is higher, but since the switch to clean innovation occurs sooner, temperatures decrease faster.

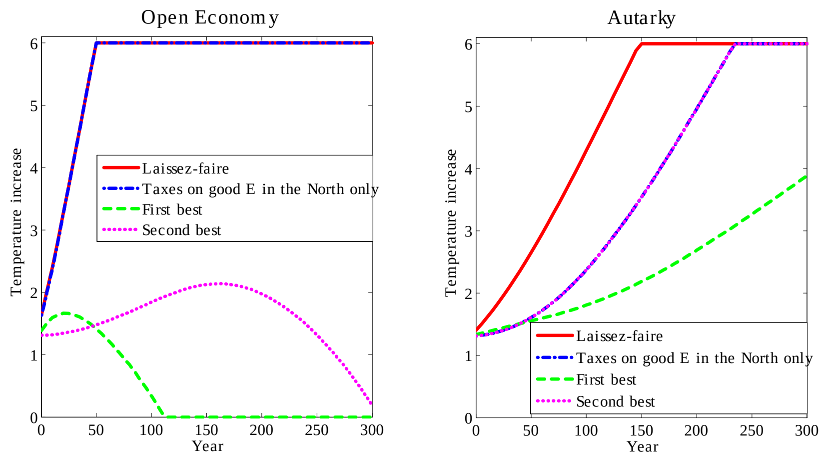

Figure 3: Temperature increase in open economy and in autarky (no redistribution concerns for the social planner). From left to right: Figures 2.A and 2.B.

in temperature is much larger—despite a much lower growth rate—and a large trade tax must be permanently maintained in order 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 should be able to produce the polluting good relying mostly on clean technologies and to force the South to asymptotically fully specialize in the non-polluting sector. The most extreme way for the North to do this is to produce only the non-polluting good (with nearly only the clean input) and to give it for free to the South. Yet, without DTC, the ratio of relative productivities stay the same over time, so if initially the South has a large comparative advantage in the non-polluting sector, or if clean technologies in the North are sufficiently backward, this is not enough to push the South towards full specialization and to avoid a disaster. This thought experiment demonstrates that innovation's ability to affect comparative advantage is essential to deriving the previous results.

5 Knowledge Diffusion

I now relax the assumption that productivity improvements are entirely country specific. In reality, some productivity improvements cross borders, mitigating the amplification of comparative advantage effect, which partly drove the previous results.[38] This brings into

[38] One should not expect all productivity improvement to cross borders easily, because some may be embedded in capital or may depend on local know-how. Dechezleprêtre et al. (2011) suggest that clean technology transfers between developing and developed countries exist but are limited: for the period 2000–

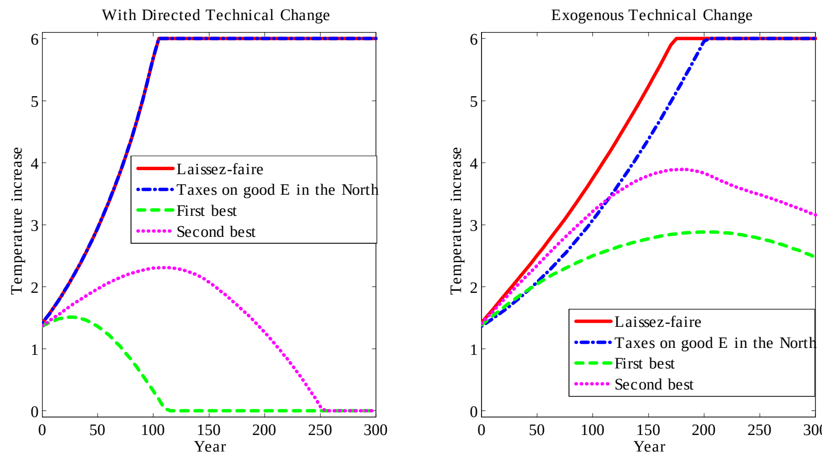

Figure 4: Temperature increase with and without directed technical change (no redistribution concerns for the social planner, different capital shares than in the baseline scenario: $\alpha = 0.7$, $\beta = 0.1$). From left to right: figures 3.A and 3.B.

question the robustness of the previous analysis. Here I consider an extension of the original model whereby the lagging country can benefit from the diffusion of innovations produced in the leading country, while Appendix A.16 considers a different extension where innovation is undertaken by multinational firms so that technologies are the same in both countries.

To model knowledge diffusion in a simple way, I assume that, at the beginning of every period, the country with the less advanced average productivity in a given sector can partially catch up exogenously. That is, before any innovation occurs, the producer of intermediate i in sector $z \in \{c, d, F\}$ gains access to the technology:

$$\overline{A_{zit}^X} = \max\left(\left(\frac{A_{z(t-1)}^{(-X)}}{A_{z(t-1)}^X}\right)^\delta, 1\right) A_{zi(t-1)}^X,$$

where $\delta \in [0, 1]$ measures the strength of the technological diffusion. This equality then

2005, only 15 percent of the clean innovations were patented in more than one country; this is slightly less than the share (17 percent) of all innovations patented in more than one country.

delivers the following law of motion for aggregate productivity:

$$A_{zt}^{X}=\left(1+\kappa\left(s_{zt}^{X}\right)^{\iota}\right)^{1-\gamma}\max\left(\left(\frac{A_{z(t-1)}^{(-X)}}{A_{z(t-1)}^{X}}\right)^{\delta},1\right)A_{z(t-1)}^{X}$$

for $z\in\{c,d,F\}$. Under this formulation, the ratio of the technological levels across countries cannot diverge: as soon as one country acquires a strong advantage over the other one, the catching-up process ensures that this difference is reduced in the next period. Unless factor endowments are significantly different across countries, this limits considerably the scope for full specialization in the long-run. 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 clean Southern technologies become more advanced than dirty ones (i.e., for some t, $\overline{A_{ct}^{S}}>\overline{A_{dt}^{S}}$), 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 the nonpolluting sector 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 the polluting sector will be produced in a clean way in the long-run. Who wins depends on the policies that the North allows for and on the pattern of comparative advantage, much as in Section 3, which leads to the following result.

Proposition 5 *Assume that initially: (i) technologies are sufficiently close to each other across countries, that κ is sufficiently small, and that the spillovers δ are sufficiently strong; (ii) the South is relatively well-endowed in capital, $K^S/L^S>K^N/L^N$; and (iii) 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 $\overline{S}$ is.*

Proof. See Appendix A.15. ∎

This Proposition mirrors Proposition 1. Assumptions (i) imply that technological levels remain sufficiently close to each other across countries. When combined with assumption (ii), this ensures that the South maintains its comparative advantage in the polluting sector. Assumption (iii) plays the same role as in Proposition 1, ensuring that, when the South has the comparative advantage in the polluting sector, it innovates there more than the North does. As a result, the South keeps its comparative advantage in the polluting sector, and since a carbon tax in the North can only reinforce this comparative advantage, there are more Southern scientists innovating in dirty technologies than Northern scientists innovating in clean ones. Hence Southern clean productivity $\overline{A_{ct}^S}$ never catches up, so a switch in the South to clean innovation never occurs. Intuitively, the Northern market for the polluting good is too small to generate enough clean innovations.

When the South is identified with non-Annex I countries, assumption (ii) seems less likely to hold relative to its counterpart in Proposition 1, which only stipulates that the South has a comparative advantage in the polluting sector.[39] In fact, if the North has a large "endowments"-comparative advantage in the polluting sector (that is for $\left(K^N/L^N\right) / \left(K^S/L^S\right)$ large enough) and knowledge spillovers are strong enough, it can prevent a disaster using a combination of a carbon tax and a tax on dirty research for sufficiently high initial environmental quality. Therefore, a possible interpretation of this analysis is that knowledge spillovers weaken the conclusion that Annex I countries could not prevent worlwide emissions from growing using a carbon tax only. Yet assumption (ii) could also be more generally interpreted as assuming that the South has a comparative advantage in the polluting sector for reasons beyond imitable technological factors, which could include factor endowments (capital, labor but also natural resources), policies, different market distortions, etc... This broader interpretation is more likely to hold, in which case the conclusion that Annex I countries could not prevent worldwide emissions from growing without research subsidies would be reinforced.

Indeed, as before, a temporary combination of clean research subsidies and a tariff can prevent a disaster for sufficiently large initial environmental quality (i.e., Proposition 2 still holds). Clean research subsidies can reallocate Northern innovation to clean technologies, and a tariff can limit Southern innovation in dirty technologies. Then $\overline{A_{ct}^S}$ grows faster than $\overline{A_{dt}^S}$, and a switch to clean innovation eventually occurs in the South.[40]

Table 2: Welfare cost in the presence of knowledge spillovers

	First-best	Second-best
$\delta = 0.4$ (%)	5.71	6.92
$\delta = 0.8$ (%)	5.95	6.58

The structures of the first-best and second-best policies are broadly similar, but the trade tax and subsidies for research must take knowledge spillovers into account, and the second-best policy may prevent a disaster with a South exporting the polluting good in the long-run. In addition, the welfare costs of unilateral intervention are typically lower than

[39] It may not be rejected since L^X should stand for human capital and not simply labor.

[40] Remark 1 no longer holds when clean and dirty inputs are imperfect substitutes. Because of the knowledge spillvers the ratio A_{ct}^S/A_{dt}^S cannot approach zero if the North allocates all its scientists to clean technologies, so the South always allocate some scientists to clean technologies. $\overline{A_{ct}^S}$ becomes greater than $\overline{A_{dt}^S}$ at some t, after which a switch to clean innovation occurs in the South. Remark 1 still holds if $\varepsilon = \infty$, or with a different innovation function which does not satisfy the Inada condition (such as $\kappa\left((s+\Upsilon)^\iota - \Upsilon^\iota\right)$ with $\Upsilon > 0$).

in the absence of knowledge spillovers. Indeed, the reversal in comparative advantages, 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 the welfare costs in the first-best and the second-best cases in the presence of knowledge spillovers ($\delta = 0.4$ and $\delta = 0.8$, and the social planner maximizes (1)): the welfare costs of the first-best policy are very similar to those in Table 1, but those of the second-best policy are now much lower.[41]

To some extent, technological diffusion itself is a parameter that can be affected by policy: laxer intellectual property rights, direct financing of projects abroad, or migrations of skilled workers could all contribute to a faster diffusion of technology. Therefore, according to the analysis presented here, the diffusion of clean technologies from North to South renders a tariff less necessary, and significantly reduces the costs of a unilateral policy intervention.

With the inclusion of knowledge spillovers, one can now add a nontradeable sector to the economy without changing the results. Assume that final consumption is a Cobb-Douglas aggregate of nontradeable and tradeable goods. Both are produced according to (3), with the associated goods E and F (and the associated subsectors c and d), but for the nontradeable good, the polluting and non-polluting inputs must be sourced locally. The same intermediates are used whether the good is produced for the tradeable or nontradeable sector. In the no-spillovers case, it is impossible to prevent a disaster because Southern emissions from nontradeables will increase unboundedly regardless of Northern policy. In the spillover case, however, the same results as before still apply: if Northern clean technologies win the horse race over Southern dirty technologies, then nontradeables in the South will also begin using clean inputs more intensively, so that emissions can decrease in both countries. Similarly, with knowledge spillovers, the results would carry through if production of good F also relied partly on the clean and dirty inputs.

6 Conclusion

On the backdrop of a literature on trade and the environment, which has largely ignored innovation, this paper presents a simple model which puts innovation at the center. It 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 of course 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 placing the economy on a path that leads to the South having a comparative advantage in the polluting sector. This leads to

[41] Here, the reversal of comparative advantage still takes place in the presence of knowledge spillovers because the difference in factor endowments is small.

the relocation of not only the production of the polluting good but also of innovation in the polluting sector, which dramatically hampers the benefits of such a policy on worldwide emissions. The South innovates more in dirty technologies, while innovation in clean technologies in the North does not take off because the market share for the polluting sector is reduced. Second, sustainable growth can be achieved without cooperation from the South, but this requires a somewhat protectionist 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.

Therefore, in practice, the paper argues that unilateral environmental policies should be devoted to 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. The paper aims at analyzing what "well-intentioned" countries should do until then, and therefore, as a first step, it has taken as given the absence of such an agreement. The next logical 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. This is, however, a complex issue as the incentive to sign a global agreement depends on the benefit that the reluctant country would get from it. Unilateral policies can affect this potential 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).

Another aspect left for future research is to study policies that directly boost technological diffusion. Such policies (e.g., the clean development mechanism) are already part of climate negotiations. Studying technological diffusion would, however, require a proper model of intellectual property rights (IPR), whose impact on emissions is ambiguous. On the one hand, laxer IPR could lead to more rapid diffusion of clean technologies to the South, which would facilitate the switch to a clean path there. On the other hand, they might reduce the incentives to develop Northern clean technologies in the first place. Finally. the paper's results suggest that directed technical change 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 models. To properly evaluate the impact of local carbon taxes and carbon tariffs, numerical models of the world economy should incorporate directed technical change.

7 Acknowledgments

I am very grateful to my advisors Daron Acemoglu, Philippe Aghion, Pol Antràs and Elhanan Helpman for their invaluable guidance. I thank David Atkin, Steve Cicala, Richard Cooper, Rafael Dix-Carneiro, Vasco Carvalho, Matt Darnell, 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 and the Oxcarre conference.

References

Acemoglu, D., 2003. Patterns of skill premia. Review of Economic Studies 70, 199–230.

Acemoglu, D., Aghion, P., Bursztyn, L., Hémous, D., 2012a. The environment and directed technical change. The American Economic Review 102 (1), 131–166.

Acemoglu, D., Aghion, P., Hémous, D., 2013. The environment and directed technical change in a North-South model.

Acemoglu, D., Akcigit, U., Hanley, D., Kerr, W., 2012b. The transition to clean technology, mimeo.

Aghion, P., Dechezleprêtre, A., Hémous, D., Martin, R., Reenen, J. V., 2012. Carbon taxes, path dependency and directed technical change: Evidence from the auto industry, NBER Working Paper 18596.

Aichele, R., Felbermayr, G., 2012. Kyoto and the carbon footprint of nations. Journal of Environmental Economics and Management (forthcoming).

Atkinson, G., Hamilton, K., Ruta, G., Mensbrugghe, D. V. D., 2011. Trade in "virtual carbon": Empirical results and implications for policy. Global Enviromental Change 21, 563–574.

Babiker, M. H., Rutherford, T. F., 2005. The economic effects of border measures in subglobal climate agreements. The Energy Journal 26, 99–126.

Barrett, S., 1994. Self-enforcing international environmental agreements. Oxford Economic Papers 46, 878–894.

Bovenberg, A. L., Smulders, S., 1995. Environmental quality and pollution-augmenting technological change in a two-sector endogenous growth model. Journal of Public Economics 57 (3), 369–391.

Broner, F., Bustos, P., Carvalho, V., 2012. Sources of comparative advantages in polluting industries, Working Paper 655, Barcelona Graduate School of Economics.

Böhringer, C., Carbon, J. C., Rutherford, T. F., 2011. Embodied carbon tariffs, NBER Working Papers 17376.

Böhringer, C., Fischer, C., Rosendahl, K. E., 2010. The global effects of subglobal climate policies. The B.E. Journal of Economic Analysis and Policy 10(2), article 13.

Copeland, B. R., Taylor, M. S., 1995. Trade and transboundary pollution. The American Economic Review 85(4), 716–737.

Copeland, B. R., Taylor, M. S., 2004. Trade, growth and the environment. Journal of Economic Literature 42(1), 7–71.

Copeland, B. R., Taylor, M. S., 2005. Free trade and global warming: a trade theory view of the kyoto protocol. Journal of Environmental Economics and Management 49, 205 – 234.

Davis, S., Caldeira, K., 2010. Consumption-based accounting of CO2 emissions. Proceedings of the National Academy of Sciences 107(12), 5867–5693.

Dechezleprêtre, A., Glachant, M., Hascic, I., Johnstone, N., Ménière, Y., 2011. Invention and transfer of climate change mitigation technologies on a global scale: A study drawing on patent data. Review of Environmental Economics and Policy 5(1), 109–130.

Elliott, J., Foster, I., Kortum, S., Munson, T., Cervantes, F. P., Weisbach, D., 2010. Trade and carbon taxes. American Economic Review: Papers & Proceedings 100, 465–469.

Gancia, G., Bonfiglioli, A., 2008. North-South trade and directed technical change. Journal of International Economics 76, 276–295.

Gerlagh, R., Snorre, K., Rosendahl, K. E., 2014. The optimal time path of clean energy R&D policy when patents have finite lifetime. Journal of Environmental Economics and Management 67 (1), 2–19.

Goulder, L. H., Schneider, S. H., 1999. Induced technological change and the attractiveness of CO2 abatement policies. Resource and Energy Economics 21 (3-4), 211–253.

Grimaud, A., Rouge, L., 2008. Environment, directed technical change and economic policy. Environmental and Resource Economics 41 (4), 439–463.

Grossman, G. M., Helpman, E., 1991. Innovation and Growth in the Global Economy. MIT Press. Cambridge, Massachusetts.

Hassler, J., Krusell, P., Olovsson, C., 2012. Energy-saving technical change, NBER Working Paper No. 18456.

Hoel, M., 1996. Should a carbon tax be differentiated across sectors? Journal of Public Economics 59, 17–32.

IEA, 2010a. Detailed CO2 estimates. In: IEA CO2 Emissions from Fuel Combustion Statistics (database).

IEA, 2010b. World energy balances. In: IEA World Energy Statistics and Balances.

Krugman, P., 1981. Trade, accumulation, and uneven development. Journal of Development Economics 8, 149–161.

Krugman, P., 1987. The narrow moving band, the Dutch disease, and the competitive consequences of Mrs Thatcher. Journal of Development Economics 27, 41–55.

Maria, C. D., Smulders, S. A., 2004. Trade pessimists vs technology optimists: Induced technical change and pollution havens. The B.E. Journal of Economic Analysis and Policy 4(2), Article 7.

Maria, C. D., van der Werf, E., 2008. Carbon leakage revisited: Unilateral climate policy with directed technical change. Environmental and Resource Economics 39, 55–74.

Markusen, J. R., 1975. International externalities and optimal tax structures. Journal of International Economics 5, 15–29.

Matsuyama, K., 1992. Agricultural productivity, comparative advantage and economic growth. Journal of Economic Theory 58, 317–334.

Newell, R. G., Jaffe, A. B., Stavins, R. N., 1999. The induced innovation hypothesis and energy-saving technological change. The Quarterly Journal of Economics 114 (3), 941–975.

Nordhaus, W. D., 2008. A Question of Balance: Economic Modeling of Global Warming. Yale University Press, New Haven, CT.

OECD, 2013. Stan, database.

Pethig, R., 1976. Pollution, welfare, and environmental policy in the theory of comparative advantage. Journal of Environmental Economics and Management 2, 160–169.

Popp, D., 2002. Induced innovation and energy prices. The American Economic Review 92 (1), 160–180.

Popp, D., 2004. Entice: Endogenouse technological change in the DICE model of global warming. Journal of Environmental Economics and Management 24 (1), 742–768.

Stanton, E., 2011. Negishi welfare weights in integrated assessment models: the mathematics of global inequality. Climatic Change 107, 417–432.

Timmer, M., O'Mahony, M., van Ark, B., March 2008. EU KLEMS database. In: The EU KLEMS Growth and Productivity Accounts: An Overview. University of Groningen & University of Birmingham, downloadable at www.euklems.net.

UNIDO, 2011. Industrial demand supply balance database. Mimas, University of Manchester.

van der Zwaan, B. C. C., Gerlagh, R., Klaassen, G., Schrattenholzer, L., 2002. Endogenous technological change in climate change modelling. Energy Economics 24 (1), 1–19.

Young, A., 1991. Learning by doing and the dynamic effects of international trade. Quarterly Journal of Economics 106, 369–405.

.