Clean Growth and Environmental Policies in the Global Economy *

Wei Xiang†

University of Michigan

March 3, 2025

Click Here for the Latest Version

Abstract

How effective are local environmental policies in fostering clean technology in a globalized world? To address this question, I develop a dynamic general equilibrium semi-endogenous growth model of the world economy. I focus on endogenous innovation in both clean and dirty technologies, as well as the choice of production locations for firms specializing in different technologies. Local environmental policies lead to the relocation of dirty production in the short run. Endogenous innovation in clean technology that results from such policies enhances the country's technological comparative advantage in clean technology, ultimately leading to the deployment of clean technology in foreign production locations in the long run. Counterfactual analysis shows that if both the US and EU countries increase the stringency of their environmental policies to the level of the most stringent EU country, global CO_2 emissions will decrease by 4.7 percent. The welfare implications are asymmetric across countries: in the steady state, the consumption-equivalent welfare gains are almost twice as large for EU countries as for the US.

JEL: D58, F12, F14, F18, F23, O31, O32, O33, O44

*I am indebted to Fabrizio Zilibotti, Costas Arkolakis, Sam Kortum, and Lorenzo Caliendo for their continuous guidance and support. I benefit from conversations with Sheng Cai, Zhen Huo, and Michael Peters. I thank Treb Allen, Andy Bernard, Alexander Bick, Paco Buera, Paula Bustos, Annie Chen, Xugan Chen, Maximiliano Dvorkin, Jonathan Dingel, Cecilia Fieler, Joel Flynn, Teresa Fort, Carlos Garriga, Penny Goldberg, Kenneth Gillingham, Aakash Kalyani, Oksana Leukhina, Jack Liang, Giovanni Maggi, Jeremy Majerovitz, Ricardo Marto, Shaun McRae, Giuseppe Moscarini, Martin Souchier, Juan M. Sánchez, Aleh Tsyvinski, R. Ravikumar, Andrés Rodríguez-Clare, Ana Maria Santacreu, Romain Wacziarg, Daniel Xu for their comments. I also thank participants of the Trade Lunch/Workshop and Macro Lunch at Yale University and Dissertation Fellow Workshop at St. Louis Fed for helpful comments and discussions. I am grateful to Yale Economic Growth Center Secondary Data Acquisition Grant for support. All errors are mine.

†Email: xwe@umich.edu

1 Introduction

Many governments impose pollution regulations on firms as part of a proactive approach to tackling environmental concerns. A key goal of such regulations is to catalyze the development of innovative clean technologies. Empirical evidence suggests that stringent regulations on pollution-intensive industrial activities do promote innovative clean technologies.[1] However, environmental policies are often confined to local jurisdictions, whereas industrial production operates on a global scale. This mismatch poses a challenge for policymakers. Firms, especially multinationals, can relocate pollution-intensive manufacturing operations to foreign jurisdictions, reducing their incentive to deploy or to innovate in clean technologies.[2] This strategic response potentially dampens the effect of environmental policies on global emissions reduction. In light of these considerations, a question emerges: how effective are environmental policies in promoting clean technology in a globalized landscape?

To address this question, this paper studies both the short- and long-term impact of pollution regulations on innovation in clean technology when production locations are substitutable. Leveraging a comprehensive firm-level dataset, I begin by documenting that at the firm level, a higher number of clean patents is correlated with more stringent environmental policies. However, this association diminishes when a firm's foreign production intensity increases. Motivated by these empirical findings, I develop a tractable multi-country semi-endogenous growth model in which environmental policies have immediate effects on technology adoption (clean or dirty) and production location choices. The endogenous innovation that results from such policies leads to a global rise in clean technology in the long run. Lastly, I use the model to assess the global welfare implications of increasing environmental policy stringency in the European Union and the United States.

I first use firm-level revenue and patent data to document firms' clean innovation responses to environmental policy changes.[3] To construct a firm's exposure to environmental regulation, I employ a triple-differences specification with three components. The first two are the environmental policy stringency (EPS) index compiled by the OECD, which varies across countries and years, and the country-sector level CO_2 emissions intensity reported by the World Input-Output Database. The third component of the interaction is an indicator for each firm's multinational

[1] See Aghion, Dechezleprêtre, Hemous, Martin, and Van Reenen (2016) and Calel and Dechezleprêtre (2016)

[2] See Hanna (2010) and Cai, Lu, Wu, and Yu (2016).

[3] In what follows, I use "clean/dirty innovation" to refer to innovation in clean/dirty technology.

production status from the Orbis Historical database. The main outcome variable is the number of clean patents at the firm level, obtained from the PATSTAT database by the European Patent Office (EPO). I show that the number of clean patents is positively correlated with the environmental policy stringency for those in dirty industries, but the correlation is weaker for firms with more foreign production. This evidence suggests that multinational production attenuates the effectiveness of local environmental regulation in fostering clean innovation.

Motivated by this evidence and by the central question of the paper, I develop a dynamic general equilibrium semi-endogenous growth model of the world economy, in which I focus on (1) endogenous investments in innovation and (2) firm decisions on production locations and technology type (clean and dirty). Dirty production generates negative externalities for consumers, while clean production does not.

I model innovation as a process of generating new insights. Insights arrive randomly but their arrival intensity is enhanced by R&D efforts, as modeled in Eaton and Kortum (2001). The model departs from Eaton and Kortum in two key ways. First, it allows firms to utilize insights for production outside of the home market. Second, it incorporates the choice of innovation between clean and dirty technologies. Consequently, endogenous innovation in this model leads to improvements in the technology frontier across different production locations for two technology types.

I establish conditions under which the frontier productivity distribution induced by innovation in each country follows a multivariate Fréchet distribution for each technology type. As a result, the state of each technology type in a country can be effectively summarized by the scale parameter of the Fréchet distribution, which represents the knowledge stock of this technology type. These knowledge stocks accumulate over time due to endogenous directed R&D investments, which are contingent upon several factors, including the existing knowledge stock of clean technology, strictness of environmental policies, and exposure to policies.

The Fréchet-distributed technology frontier allows me to characterize technology and production location choices analytically. In this model, firms utilize technologies stemming from the knowledge stock in their headquarter countries for production. Firms can produce in different locations and deploy different technologies. In equilibrium, when firms deploy technology for both domestic and foreign production, they take into account their technological comparative advantage and the environmental policies in the production location. To elaborate, firms tend to use dirty (clean) technology in countries with less (more) stringent environmental regulations, conditional on their technological comparative advantage. When faced with the same level of environmental policy stringency, firms deploy the type of technology in which they possess a comparative advantage.

In this model, environmental policies affect the optimal choice in both production and innovation. More specifically, when the technology level is held constant, more stringent environmental policies result in two short-run consequences: (1) increased adoption of clean technology, as firms respond to regulatory pressures by shifting towards cleaner alternatives; (2) relocation of dirty production, as firms seek locations with less stringent environmental regulations to mitigate the additional costs associated with dirty technologies. The latter case illustrates a weakening of the immediate impact of local environmental policies on global pollution reduction. However, multinational production possibilities can strengthen the long-term efficacy of these policies. Over time, innovation plays a pivotal role in adjusting the relative levels of clean and dirty technology in response to the tightening of environmental regulations. Investments in R&D for clean technology within a country enhance its comparative advantage in clean technology in the long run. Consequently, firms from such countries tend to deploy clean technology for global production. Thus, in the long run, the interplay between local environmental policies and multinational production possibilities reduces global pollution emissions through the global deployment of cleaner technology.

I calibrate the model using 2016 data on trade, multinational production (MP), environment-related taxes, patents, and CO_2 emissions for fifty-two countries. First, I use aggregate bilateral trade/MP flow and environment-related tax data to recover the full set of trade and MP costs between countries as well as environmental policy stringency in each country. Second, I use patent application and citation data from the PATSTAT Global database to construct a model-consistent measure for knowledge stocks related to clean and dirty technologies. Third, I employ CO_2 emissions data to recover the emissions intensities for dirty production carried out by firms originating in various countries. Lastly, I calibrate the environmental damage in each country to match the social cost of carbon.

The calibrated model enables me to study the effect of a variety of policies. First, I show that environmental policies could lead to opposing effects in the short and long run. For instance, a 10 percent increase in Germany's environmental policy stringency leads to a short-term increase in CO_2 emissions in Hungary and Slovakia due to the relocation of dirty production. However, the endogenous response of innovation in Germany enhances its comparative advantage in clean technology, ultimately leading to the deployment of clean technology in Hungary and Slovakia. As a result, CO_2 emissions in these two countries decline in the long run. In another counterfactual analysis, I illustrate that if both the US and EU countries increase the stringency of their environmental policies to the level of the most stringent EU country (Slovenia), global CO_2 emissions will decrease by 4.7 percent. While environmental regulation reduces real consumption, the reduction in global carbon emissions increases welfare. The welfare implications are asymmetric across countries: in the steady state, the consumption-equivalent welfare gains are almost twice as large for EU countries compared to the US.

Related Literature. This paper bridges the gap between two bodies of literature on environmental policies and their impact. The first pertains to directed technical change and the environment, which examines how environmental policies influence the direction of innovation.[4] Aghion et al. (2016) provides evidence that stricter environmental policies result in more green patents in the automobile industry. Optimal environmental regulation and R&D subsidies have thus been proposed in light of this fact (Acemoglu, Aghion, Bursztyn, and Hemous, 2012). Donald (2025) studies how cross-technology knowledge spillovers influence the transition to clean technology. However, these papers focus on a closed economy and ignore the possibility of carbon leakage, where firms escape environmental regulation by shifting dirty production abroad instead of developing clean technology. The second strand of literature investigates whether domestic environmental policies lead to relocation of production activities (Hanna, 2010; Broner, Bustos, and Carvalho, 2012) but typically ignores clean innovation and clean growth, hence providing little insight into the long-term effects of environmental policies. This paper combines insights from both literatures to study the long-term growth and innovation outcomes of environmental policies in a globalized world. It complements the first literature by considering multinational production possibilities and contributes to the second literature by analyzing dynamic consequences. By bridging the two strands of literature mentioned above, this paper provides a consistent explanation for the mixed empirical evidence on pollution haven effects in the existing literature (Eskeland and Harrison, 2003; Copeland and Taylor, 2004; Levinson and Taylor, 2008).[5] The relocation of production may

[4]For a comprehensive review of this literature, see Hémous and Olsen (2021). There are also related papers about directed technical change and natural resources. For example, Acemoglu, Aghion, Barrage, and Hémous (2023) study how shale gas booms shape energy transition through their impact on directed innovation in clean energy; Hassler, Krusell, and Olovsson (2021) study how scarce natural resources influence input-saving technical change, with an application to fossil fuel.

[5]I would like to clarify the distinction between two similar but different concepts in the environmental economics literature: pollution haven hypothesis and pollution haven effect, following Copeland and Taylor (2004). The pollution

introduce pollution to a foreign country in the short run. However, in the long run, the increased comparative advantage in clean technology due to innovation makes firms more inclined to deploy cleaner technology for global production. Consequently, the prominence of the pollution haven effect dissipates in the long run, possibly giving way to the pollution halo effect.[6]

This paper also expands the literature on environment and international trade, as recently reviewed by Cherniwchan, Copeland, and Taylor (2017) and Copeland, Shapiro, and Scott Taylor (2022). It is most closely related to Garcia-Lembergman, Ramondo, Rodriguez-Clare, and Shapiro (2023), who investigate how multinational firms use clean technology in their production locations. However, this paper differs from their study in several ways. First, it presents a dynamic model that complements their static framework. Second, while still incorporating the global technology deployment mechanism of multinational firms, this paper focuses on the innovation response and technological progress of these firms. There is a growing body of literature on the static impact of carbon policies in the presence of trade such as Farrokhi and Lashkaripour (2021), Kortum and Weisbach (2021), and Weisbach, Kortum, Wang, and Yao (2023). My paper complements this literature by providing a dynamic framework to understand the long-run impact of environmental policies.

Empirical studies on technology, international trade, and pollution include Levinson (2009) and Shapiro and Walker (2018), among others. The focus of this paper, instead, is to propose a theoretical and quantitative framework. There is also a growing body of literature on climate change, the environment, and economic geography (Desmet and Rossi-Hansberg, 2015; Conte, Desmet, Nagy, and Rossi-Hansberg, 2021; Cruz, 2021; Desmet, Kopp, Kulp, Nagy, Oppenheimer, Rossi-Hansberg, and Strauss, 2021; Cruz and Rossi-Hansberg, 2022; Arkolakis and Walsh, 2023; Desmet and Rossi-Hansberg, 2023; Cruz and Rossi-Hansberg, forthcoming). This paper, instead, focuses on the interaction between innovation and multinational production in response to local environmental policies. Discussions on the macroeconomic implications of carbon emission in this paper relate to the literature on environmental macroeconomics (Golosov, Hassler, Krusell, and Tsyvinski, 2014; Krusell and Smith Jr, 2022).[7]

haven effect means that a tightening up of pollution regulation will, at the margin, have an effect on plant location decisions and trade flows, while the pollution haven hypothesis states a reduction in trade barriers will lead to a shifting of pollution-intensive industry from countries with stringent regulations to countries with weaker regulations. In this paper, I focus on pollution haven effect.

[6] As Albornoz et al. (2014) point out, "pollution halo" hypothesis argues that multinational firms may use more advanced technologies, cleaner production methods, and possess more developed environmental management systems and organizational techniques, which yields substantial environmental benefits to developing countries.

[7] See Hassler et al. (2016) for a review.

Lastly, this paper joins a growing body of literature on the growth effect of international trade and multinational production (Eaton and Kortum, 1999, 2001; Alvarez, Buera, Lucas et al., 2013; Santacreu, 2015; Sampson, 2016; Buera and Oberfield, 2020; Somale, 2021; Cai, Li, and Santacreu, 2022a; Lind and Ramondo, 2022a,b; Cai, Caliendo, Parro, and Xiang, 2022b; Cai and Xiang, 2023; Bai, Jin, Lu, and Wang, 2023). The static part of the model builds on Eaton and Kortum (2002), Ramondo and Rodríguez-Clare (2013), Arkolakis, Ramondo, Rodríguez-Clare, and Yeaple (2018), and Lind and Ramondo (2023) but departs from their work through the introduction of dynamics. Furthermore, I introduce two technology types to highlight the race between clean and dirty technological progress and firms' technology choices, which contrasts with the aforementioned single-technology static models. Cai and Xiang (2023) develop a framework where multinationals increase the quality of the knowledge pool of the host country and facilitate economic growth through knowledge diffusion. This paper complements Cai and Xiang (2023) by highlighting innovation instead of diffusion as a key driver of growth. This paper also complements Cai, Li, and Santacreu (2022a), who study how knowledge diffusion through international trade and endogenous innovation contribute to growth, by incorporating multinational production and technology choice to describe their effects on endogenous innovation and clean growth. Lind and Ramondo (2022a,b) present a model of global innovation and diffusion, assuming that ideas are created exogenously; this paper characterizes the endogenous R&D decisions of innovators.

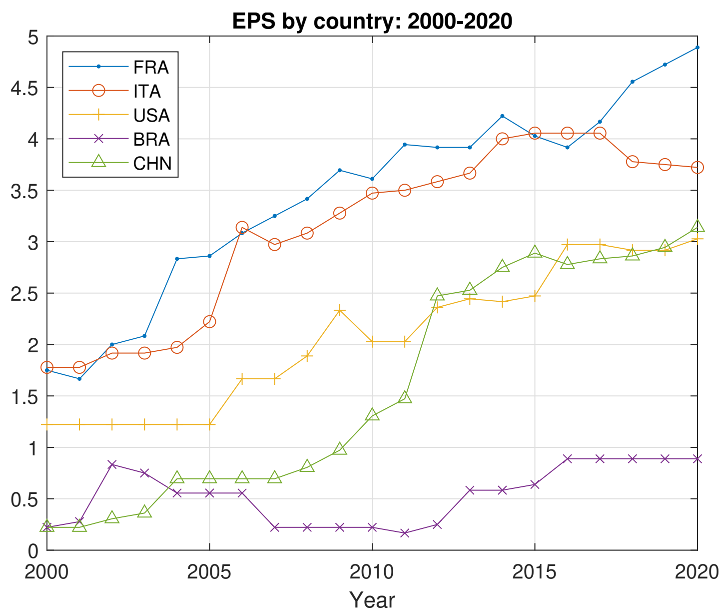
The remainder of the paper is organized as follows. Section 2 describes the data and documents motivating facts. Section 3 presents the theory. Section 4 describes model calibrations, and Section 5 reports the quantitative results. Section 6 concludes.

2 Motivating Facts

In this section, I show that more stringent local environmental policies are associated with an increase in the number of clean patents at the firm level. The association diminishes, however, as firms produce a larger share of their total output abroad.

To illustrate this association, I employ a triple-differences strategy that combines annual variation in environmental policy stringency across countries, sector-level variation in emissions intensity, and firm-level differences in multinational enterprise (MNE) status. The first component is $EPS_{c,t}$, the environmental policy stringency index compiled by the OECD, which varies across countries (c) and years (t). An important feature of this measure is that it encompasses both market-based and non-market-based practices and is comparable across countries and over time. Examples of market-based policies include carbon taxes, while non-market policies encompass practices like setting emissions limits. Altogether, there are thirteen subcategories, each representing a distinct policy. These policies are then mapped to an index ranging from 0 to 6. Figure 1 illustrates the EPS for France, Italy, the USA, Brazil, and China. While policy stringency levels differ, most countries tightened their environmental regulations between 2000-2020. Out of the five selected countries, France enforced the most stringent environmental policies, while Brazil adopted the most lenient approach. Notably, China is rapidly closing the gap, experiencing a nearly fifteen-fold increase in its EPS.

Figure 1: Environmental Policy Stringency 2000-2020 for Five Selected Countries



Notes: The figure shows the evolution of the Environmental Policy Stringency (EPS) between 2000-2020 for five selected countries: Brazil, China, France, Italy, and the United States. The EPS index is compiled by the OECD and varies across countries and years. It encompasses both market-based and non-market-based practices. These policies are converted into an index ranging from 0 to 6. The data is obtained from OECD.stat (https://stats.oecd.org/Index.aspx?DataSetCode=EPS).

The second component is the emissions intensity, $EM_{jc,t=2000}$, which is defined as CO_2 emissions per output of sector j in country c in the year 2000. This measure is reported by the World Input-Output Database (WIOD) from 2000 to 2014 and is used to gauge the pollutant intensity of a sector in a country (Corsatea et al., 2000).[8] I use the 2000 value for analysis to mitigate endo-

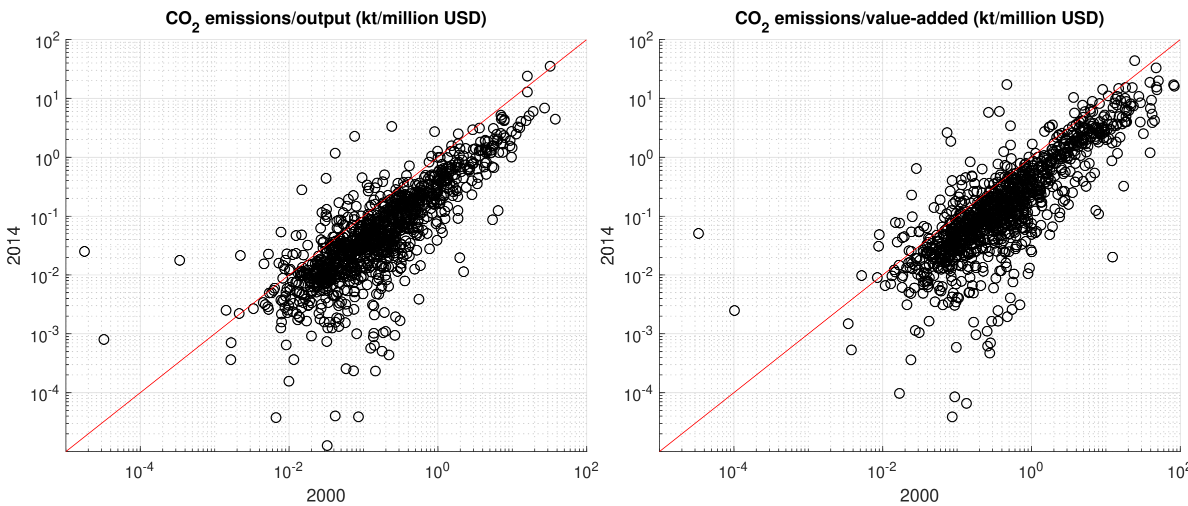
[8]The database covers 28 EU countries and 13 other major countries in the world for the period 2000 to 2014 for 64 sectors. The CO_2 emissions intensity is an imperfect measure, as CO_2 emissions intensities are not necessarily the same as pollution intensities. Here, we use CO_2 emissions intensity as a proxy for the pollution level of a sector.

geneity concerns, as the change in $EPS_{c,t}$ might affect $EM_{cj,t}$ over time. Figure 2a shows that the overall emissions intensity decreases over time, while the relative rank of emissions intensity remains stable. In Figure 2b, I show the same correlation but for emissions per value added. The two measures are highly correlated; hence, I use emissions per output for the subsequent regression analyses.

Figure 2: CO_2 Emissions Intensity

(a) CO_2 emissions per output: 2000 vs. 2014

(b) CO_2 emissions per value added: 2000 vs. 2014



Notes: The figure shows changes in CO_2 emissions intensities for forty-one countries and fifty-six (ISIC Rev. 4) sectors between 2000 and 2014. Each circle represents a sector in a country. The horizontal axis represents emissions intensities in 2000, while the vertical axis represents emissions intensities in 2014. Figure 2a shows the changes in emissions intensities defined as CO_2 emissions per output. Figure 2b shows the changes in the emissions intensities defined as CO_2 emissions per value added. The data is from the World Input-Output Database (WIOD) 2016.

The interaction $EPSEM_{jc,t} = EPS_{c,t} \times EM_{jc,t=2000}$ captures environmental policy exposure for each country sector pair (c, j). The relevance assumption required for the difference-in-differences shock $EPSEM_{jc,t} = EPS_{c,t} \times EM_{jc,t=2000}$ is that for a country c in a particular year t, the sector j with higher emission intensity will be more exposed to the same level of $EPS_{c,t}$.

I further compare the shock response of a given firm to that of a firm with more production activities abroad. To do so, I introduce another interaction term, $MNE_{i,t}$ and foreignshare$_{i,t}$. The variable $MNE_{i,t}$ is a dummy variable that characterizes the status of a firm i, with a value of 1 if the firm is an MNE and 0 otherwise. The variable foreignshare$_{i,t}$ represents the foreign revenue share of a firm i, calculated as its total generated revenues abroad divided by its total revenues. The variable foreignshare$_{i,t}$ is a continuous counterpart of $MNE_{i,t}$, with values ranging from 0 to

1. It equals 0 if the firm is entirely domestic and 1 if entirely foreign. With the addition of this third component, the environmental policy shock can be expressed as $EPSEM_{jc,t} \times MNE_{i,t}$ and $EPSEM_{jc,t} \times \text{foreignshare}_{i,t}$. The multinational production information is obtained from the Orbis Historical data, which reports ownership structure of a firm for each year from 2007 to 2019. It is significantly challenging to violate the exclusion restriction for a triple differences specification. A spurious result occurs if some confounding force covaries with nationwide environmental policy stringency, country-sector emissions intensity, and further differentially impacts firms with different multinational production status types.

I leverage the PATSTAT Global data maintained by the European Patent Office (EPO) for patent information. To identify clean technology, I use the cooperative patent classification (CPC) code. I follow the EPO document to identify climate change mitigation technologies (Angelucci et al., 2018).[9] I keep all patents that have been granted. Note that multi-filing is a major challenge when dealing with international patent data. It is common practice for innovators to file the same invention through various patent authorities in different countries. Furthermore, different aspects of a single invention could result in multiple patent filings. The patents that belong to a single invention form a patent family with a unique patent identifier. To address the double counting concern, my analysis uses the simple family ID (DOCDB family ID in PATSTAT database) to uniquely identify an invention. I merge EPO data with Orbis Historical data using the concordance provided by the Orbis Intellectual Database. I keep only manufacturing firms for the empirical analysis. The final data sample for regression analysis covers the years from 2007 to 2019.

2.1 Empirical Specification

To understand how the innovation activities of firms respond to environmental policies, I specify the regression equations as follows

$$(\text{clean patent \#})_{ijc,t} = \beta_1 EPSEM_{jc,t-k} \times m_{i,t-k} + \beta_2 EPSEM_{jc,t-k} + \beta_3 m_{i,t-k} + \varphi_t + \phi_i + u_{i,t}, \tag{1}$$

where i denotes a firm, j a sector, c a country, t a year, and k the years lagged. The variable (clean patent #) refers to the number of clean patent, while m is a measure of multinational production status which represents either $MNE_{i,t-k}$ or $\text{foreignshare}_{i,t-k}$. I control for year φ_t and firm ϕ_i fixed effects. The error term is denoted by $u_{i,t}$. The expected signs are $\beta_1 < 0$ and $\beta_2 > 0$.

[9]See https://www.epo.org/mobile/news-events/in-focus/classification/classification/updatesYO2andY04S.html.

2.2 Regression Results

I first present preliminary regression results, and then display the results from the specification in the previous subsection. Table 1 highlights the correlation between firm-level clean patent numbers and country-level EPS. Since I control for firm fixed effects, we don't need to include $EM_{jc,t=2000}$ in the regression. In Column (1), the correlation is positive among non-MNE firms, while Column (2) shows that the correlation is more pronounced if the non-MNE firm is in a sector with higher emissions intensity. Columns (3) and (4) reveal that the relationship vanishes for MNEs. Yet, in Columns (5) and (6), it is again positive among the largest 0.5% of non-MNEs, showing that the results in Columns (3) and (4) are unlikely to be driven by firm size. The results in Table 1 also suggests that $EPSEM_{jc,t} = EPS_{c,t} \times EM_{jc,t=2000}$ is a plausible measure for environmental policy stringency.

Table 1: Preliminary Results

Dept. Var. (clean patent #)$_{i,t}$	Non-MNE		MNE		Large non-MNE	
	(1)	(2)	(3)	(4)	(5)	(6)
$\text{EPS}_{c,t-1}$	0.2921***	0.2752***	-0.0537	-0.1058	0.3292***	0.3086***
	(0.091)	(0.086)	(0.218)	(0.209)	(0.096)	(0.083)
$\text{EPS}_{c,t-1} \times \text{EM}_{cj,2000}$		0.0350***		0.3530*		0.0397***
		(0.009)		(0.181)		(0.015)
N	2,000,396	2,000,396	12,359	12,359	9,333	9,333

Notes: The table reports the correlation between firm-level clean patent counts and country-level EPS using a Poisson Pseudo Maximum Likelihood (PPML) model. The dependent variable is the number of clean patents at the firm level. Column (1) and (2) report results for firms that are not multinationals (non-MNEs). Columns (3) and (4) report results for firms that are multinationals (MNEs). Column (5) and (6) show results for the largest 0.5% of non-MNEs. All columns control for year and firm fixed effects. Standard errors are clustered at the country-sector level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Next, I turn to the regression analysis with firm-level policy exposure. The results are reported in Table 2. Column (1) and (2) demonstrate that the number of clean patents at the firm level is positively associated with sector-level environmental policy stringency, and more importantly, this association is weakened if the firm has foreign production activities. These results are consistent with the idea that the incentive for clean innovation are attenuated if production location is substitutable. Columns (3) - (6) show that the results are robust and more pronounced if the policy occurred one or two years before, which is consistent with the fact that it takes time for the firm's innovation to take effect.

Table 2: Baseline Regression

Dept. Var. (clean patent #)$_{ijc,t}$	lag $k = 0$		lag $k = 1$		lag $k = 2$	
	(1)	(2)	(3)	(4)	(5)	(6)
$\text{EPSEM}_{jc,t-k}$	0.1571**	0.0987*	0.1844***	0.1219**	0.2110***	0.1342**
	(0.066)	(0.058)	(0.070)	(0.060)	(0.067)	(0.056)
$\text{EPSEM}_{jc,t-k} \times \text{MNE}_{i,t-k}$	-0.0743**		-0.0785**		-0.0912**	
	(0.034)		(0.039)		(0.041)	
$\text{EPSEM}_{jc,t-k} \times \text{foreignshare}_{i,t-k}$		-0.1522**		-0.1233*		-0.1309**
		(0.076)		(0.070)		(0.060)
$\text{MNE}_{i,t-k}$	0.5981***		0.5637***		0.6232***	
	(0.104)		(0.123)		(0.121)	
$\text{foreignshare}_{i,t-k}$		0.0653		0.0094		0.0757
		(0.165)		(0.140)		(0.118)
$\text{EPS}_{c,t-k}$	0.0528	0.1347	0.0511	0.1315	0.0430	0.1288
	(0.102)	(0.107)	(0.118)	(0.129)	(0.125)	(0.137)
N	2,941,259	2,941,259	2,012,755	2,012,755	1,483,191	1,483,191

Notes: The table reports the regression results for Equation 1 using a Poisson Pseudo Maximum Likelihood (PPML) model. The dependent variable is the number of clean patents at the firm level. In column (1) and (2), explanatory variables are contemporaneous. In column (3) and (4), explanatory variables are lagged by one period. In column (5) and (6), explanatory variables are lagged by two periods. All columns control for year and firm fixed effects. Standard errors are clustered at the firm level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

These results suggest that multinational production attenuates the effectiveness of local environmental regulation in fostering clean innovation. Motivated by these findings, I next develop a quantitative framework to better understand the interaction between clean innovation and multinational production.

3 Model

I begin with a description of innovation in a single country with exogenous R&D efforts. Second, I characterize the global deployment of clean (and dirty) technology by each country. Next, I demonstrate how equilibrium R&D efforts are pinned down by the global production pattern and local environmental policies. Finally, I characterize the key mechanisms through which local environmental policies affect clean innovation and production in the model. I relegate all derivations and proofs to Appendix B.

3.1 Innovation with Given R&D

To ease exposition, I consider innovation in one country for a particular technology type in this section. Later, in Section 3.2, I introduce two technology types (clean and dirty) and innovations in all countries.

The world economy consists of N production locations. In each location, there is a continuum of varieties ω produced in a unit interval $[0, 1]$. For each variety, there is a large set of potential firms who have different ideas for production. Each firm can potentially produce in any production location with its idea. An idea is a vector of productivity for a given variety; each element of the vector represents the productivity of this firm to produce a variety in a production location. Formally, I denote an idea to produce a variety ω by $\mathbf{Z}(\omega) = (Z_1(\omega), Z_2(\omega), ..., Z_N(\omega))$. Note that the dimension of the vector is equal to the number of potential production locations.

Between time t and $t + \Delta t$, researchers conduct R&D to generate new ideas on how to produce varieties in all production locations.[10] New ideas arrive stochastically at a Poisson intensity R_t that is equal to R&D effort at t.[11] That is, new ideas arrive faster with increased levels of research effort. Note that R&D efforts affect the arrival intensity of new ideas but not their quality. Each new idea corresponds to a new vector of productivity for global production. I assume the productivity of each new idea $\mathbf{Q} = (Q_1, Q_2, ..., Q_N)$ is randomly distributed. Researchers compare new ideas with existing ones and keep the best element for each production location. For example, consider a special case with two production locations. Suppose the productivity of the very first idea is represented by $(3, 1)$. The subsequent new idea has productivity levels$(1, 2)$, and researchers adopt it in the second but not in the first production location. Hence, the state-of-the-art idea becomes $(3, 2)$. We call the state-of-the-art productivity across production locations the technology frontier.

Characterizing the dynamics of ideas from a country necessitates keeping track of the productivity evolution in N production locations. Furthermore, once we start solving the full model with innovations in each country for two technology types, the number of state variables increases to $N \times N \times 2$ for each variety, even before considering the fact that varieties exist on a continuum. The large number of state variables presents a challenge in characterizing the dynamics of ideas

[10] This model abstracts away R&D by foreign affiliates of multinationals as Arkolakis et al. (2018) point out that this assumption can be justified by the fact that the parents of US multinationals account for 85% of its total R&D expenditure whereas 70% of its value-added according to 2009 BEA data. The assumption that innovation taking place in the headquarters country contributes to the productivity improvement of affiliates is motivated by Bilir and Morales (2020), who found that the median U.S. multinational firm realizes 20 percent of the return on its U.S. R&D investments abroad.

[11] To be clear, the entire idea vector, rather than each element of the vector, arrives with a Poisson intensity.

both analytically and computationally. In Assumption 1, I address this challenge by specifying the quality distribution and the applicability of new ideas. These assumptions allow me to characterize the technology frontier and its evolution over time in Proposition 1.

Assumption 1.

(i) The productivity of new ideas $\boldsymbol{Q}$ *follows a multivariate Pareto distribution*

$$H(\boldsymbol{z}) \equiv \mathbb{P}\left(\boldsymbol{Q} \leq \boldsymbol{z}\right) = 1 - \left(\sum_{j=1}^{N} z_j^{-\frac{\theta}{1-\eta}}\right)^{1-\eta}, \tag{2}$$

where the support of the distribution is $z_j \geq \underline{z} \equiv N^{(1-\eta)/\theta}$ *for all* j. [12]

(ii) The variety ω *to which a new idea is applied is uniformly distributed on* $[0,1]$.

The Pareto assumption highlights that ideas with higher productivity are less likely to be drawn.[13] The parameter $\theta > 0$ governs the shape of the right tail, while $\eta \in [0,1)$ captures the degree of correlation among elements z_j. If θ is lower, the tail is thicker, and new ideas are overall more productive. In the special case where $\eta = 0$, the only possible productivity vector is a $\boldsymbol{Q}$ with $Q_j > 0$ for some j and $Q^k_{j'h} > 0$ for all $j' \neq j$. When $\eta \to 1$, the elements of $\boldsymbol{Q}$ are pair-wise perfectly correlated.[14]

The assumption that the quality of new ideas follows a Pareto distribution implies that the technology frontier is Fréchet distributed, which in turn allows us to analytically characterize the production and technology choice of firms in Section 3.2.[15] In Proposition 1, I formally characterize the technology frontier, letting $T_t \equiv \int_0^t R_s ds$ denote the accumulated R&D efforts by date t.

Proposition 1. *(Technology frontier) Given Assumption 1, the productivity distribution of the best idea at date t follows a multivariate Fréchet distribution*

$$F_t\left(\boldsymbol{z}; T_t\right) \equiv \mathbb{P}\left(\boldsymbol{Z}_t\left(\omega\right) \leq \boldsymbol{z}\right) = \exp\left[-T_t \left(\sum_{j=1}^{N} \left(z_j\right)^{-\frac{\theta}{1-\eta}}\right)^{1-\eta}\right], \tag{3}$$

with support $z_j \geq N^{(1-\eta)/\theta}$ *for all* ω.

Proof. See Appendix B.1. $\square$

[12] $\mathbb{P}\left(\boldsymbol{Q} \leq \boldsymbol{z}\right) \equiv \mathbb{P}\left(Q_1 \leq z_1, Q_2 \leq z_2, ..., Q_N \leq z_N\right)$.

[13] This can easily be seen from $\partial H\left(z\right)/\partial z_j < 0, \forall j$.

[14] For other properties of this distribution, see Arkolakis et al. (2017) and Arkolakis et al. (2018).

[15] In addition, it implies a trend of constant productivity growth associated with constant growth in R&D, which closely aligns with the data (Kortum, 1997).

Proposition 1 shows that the technology frontier follows a multivariate Fréchet distribution with scale parameter T_t, which represents the knowledge stock by date t. Note that while equation (3) is for a given variety, the law of large numbers implies that it is also the fraction of varieties whose best production efficiencies are no greater than $\boldsymbol{z}$. This result addresses the aforementioned dimensionality problem because the state of productivity for N production location can be effectively reduced to one parameter T_t. In addition, we can use the result of Proposition 1 to characterize the evolution of technology frontier over time, which is simply $\dot{T}_t = R_t$.

Equilibrium R&D efforts R_t matter for the evolution of the technology frontier. But before we characterize the determinants of R_t, recall that clean technology is the central pillar of this paper. So I now introduce clean and dirty technologies into the model in Section 3.2. After that, I show how equilibrium R&D for each technology type hinges on the current state of technology and environmental regulations (on dirty technology), among other factors.

3.2 Preference, Technology, and Multinational Production

In this section, I describe preferences, technology choices, and the production location of firms.

Preference. In a country h, the representative household consists of a continuum of production workers and a continuum of researchers with measure $\bar{L}^P_{h,t}$ and $\bar{L}^R_{h,t}$, respectively. The representative household derives utility from a stream of consumption of a final good $C_{h,t}$

$$U_h = \mathbb{E}_0 \left[\int_0^\infty e^{-(\rho+n)t} \left(\log \left(C_{h,t} \right) + \log f_h \left(\boldsymbol{E}_t \right) \right) dt \right],$$

where ρ is the discount rate; n is the population growth rate; and $\boldsymbol{E}_t$ is a pollution externality associated with production of firms using dirty technology, which I discuss later.

The final good combines all varieties using a production function with constant elasticity of substitution σ. Let the $y_h(\omega)$ be the output of variety ω in h and $p_h(\omega)$ the corresponding price, then the output of final good and the price index are, respectively,

$$Y_{h,t} = C_{h,t} = \left[\int_0^1 y_{h,t}(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right]^{\frac{\sigma}{\sigma-1}}, \; P_{h,t} = \left[\int_0^1 p_{h,t}(\omega)^{1-\sigma} \right]^{\frac{1}{1-\sigma}}.$$

I assume that $\theta > \max(1, \sigma - 1)$ to guarantee that the price level is finite.

Pollution. $\boldsymbol{E}_t \equiv (\mathcal{E}_{1,t}, \mathcal{E}_{2,t}, ..., \mathcal{E}_{N,t})$ is a vector of pollution or emissions, where $\mathcal{E}_{j,t}$ denotes the total pollution or emissions originating in country j. I assume that environmental damage is associated with the production activity of firms adopting dirty technology. We can think of $\mathcal{E}_{j,t}$ as any environmental or climate damage generated by dirty technology, albeit in the quantitative analysis, we treat $\mathcal{E}_{j,t}$ as CO_2 emissions given data quality constraints. Following Shapiro (2016), I specify the functional form of environmental damages as

$$f_h(\boldsymbol{E}_t) = \frac{1}{1 + \left(\bar{u}_h \sum_j \mathcal{E}_{j,t}\right)^2},$$

where the parameter $\bar{u}_h$ captures the social cost of CO_2 emissions. I omit time subscript t throughout the rest of this subsection when characterizing static production choices.

Production. A potential producer can choose to adopt technology type $k \in \mathcal{K} = \{c, d\}$, where c denotes clean technology and d dirty technology. An idea of technology k with productivity Z^k uses production workers $L^{P,k}$ to produce the variety and exhibits constant returns to scale,

$$y^k(\omega) = Z^k(\omega) L^{P,k}(\omega), \tag{4}$$

where $y^k(\omega)$ denotes the quantity of variety ω produced by technology k. A firm in country h can seek to produce in any country j using technology k, with productivity Z^k_{jh}. An idea of a firm headquartered in h is thus characterized by a random productivity vector $\boldsymbol{Z}_h(\omega) = \left(\boldsymbol{Z}^c_h(\omega), \boldsymbol{Z}^d_h(\omega)\right)$, where $\boldsymbol{Z}^k_h(\omega) = (Z^k_{1h}, Z^k_{2h}, ..., Z^k_{Nh})$, $k \in \{c, d\}$.

Based on the results from the previous section, the frontier productivity of each technology type at any date follows a multivariate Fréchet distribution with dispersion θ, correlation η, and a country-technology-specific parameter T^k_h that governs the location of the distribution:

$$F^k_h(\boldsymbol{z}) = \exp\left[-T^k_h \left(\sum_{j=1}^{N} (z_j)^{-\frac{\theta}{1-\eta}}\right)^{1-\eta}\right]. \tag{5}$$

Firms with headquarters in country h that produce in country j incur iceberg-type MP costs $\gamma_{jh} \geq 1$, with $\gamma_{hh} = 1$. These costs can be due to differences in language, management practices, and judicial systems. Iceberg transportation costs τ_{ij} are also incurred when shipping goods from country j to country i, with $\tau_{jj} = 1$.

The government in each country j regulates emissions caused by dirty production in j. I assume a firm uses labor to pay for environmental regulation fees, which equals $\iota_j^k w_j L_j^{P,k}(\omega)$, with $\iota_j^d \geq 0$ and $\iota_j^c = 0$, so that ι_j^d captures the environmental regulation stringency, and w_j denotes the local wage of production workers in the production location. This allows us to define augmented cost bundles

$$\kappa_{ijh}^k \equiv \tau_{ij} w_j \gamma_{jh} \left(1 + \iota_j^k\right), \tag{6}$$

for a firm with its headquarter in h producing in j to serve market i, the marginal cost is $C_{ijh}^k(\omega) = \kappa_{ijh}^k / Z_{jh}^k(\omega)$, where $Z_{jh}^k(\omega)$ is drawn from $F_h^k(\boldsymbol{z})$.[16]

Firms engage in Bertrand competition for each variety. Consumers purchase each variety from the supplier that sells at the lowest cost. Formally, the cost of variety ω in market i is given by

$$C_i^*(\omega) = \min_{j,h,k} \left\{ C_{ijh}^k(\omega); k \in \mathcal{K}; j, h \in \mathcal{N} \right\}.$$

Note that here I assume clean and dirty technologies are perfect substitutes. Hence, in the equilibrium, a given variety for a given country will be produced either by clean or dirty technology.

The multivariate Fréchet distribution allows us to characterize the allocation of expenditures and production given knowledge stocks, T, and the augmented cost bundles, κ. Formally, I present the allocation results in Lemma 1.

Lemma 1. *The probability that consumers in market i buys from a firm with headquarters in h, producing in country j using technology k is*

$$\pi_{ijh}^k \equiv \mathbb{P}\left(C_i^*(\omega) = C_{ijh}^k\right) = \psi_{ijh}^k \cdot \pi_{i \cdot h}^k, \tag{7}$$

where

$$\psi_{ijh}^k = \frac{\left(\kappa_{ijh}^k\right)^{-\frac{\theta}{1-\eta}}}{\sum_{j'=1}^N \left(\kappa_{ij'h}^k\right)^{-\frac{\theta}{1-\eta}}},$$

is the probability that the variety is produced in j among those sold in market i using technology k from

[16] Note that this specification does not preclude the case where j equals h, i.e. a firm with its headquarter in h producing in its home country. It is possible that goods produced in country j are shipped back to the headquarter country h or sold locally in j, i.e. i can be equal to j or h.

headquarter h;

$$\pi^k_{i\cdot h} = \frac{T^k_h \left(\sum_{j=1}^N \left(\kappa^k_{ijh} \right)^{-\frac{\theta}{1-\eta}} \right)^{1-\eta}}{\sum_{k'=1}^K \sum_{h'=1}^N T^{k'}_{h'} \left(\sum_{j'=1}^N \left(\kappa^{k'}_{ij'h'} \right)^{-\frac{\theta}{1-\eta}} \right)^{1-\eta}}$$

is the probability that country i buys a variety produced by firms with their headquarters in h using technology k; $\kappa^k_{ijh} \equiv \tau_{ij} w_j \gamma_{jh} \left(1 + \iota^k_j\right)$ is from equation (6). The consumer price index for country i is

$$P_i = \mathcal{B} \Phi_i^{-\frac{1}{\theta}}, \tag{8}$$

where $\mathcal{B} = \left\{ \Gamma \left(\frac{2\theta - (\sigma-1)}{\theta} \right) \left[1 + \frac{\sigma-1}{\theta-(\sigma-1)} \bar{m}^{-\theta} \right] \right\}^{\frac{1}{1-\sigma}}$, $\Phi_i = \sum_{k=1}^K \sum_{h=1}^N T^k_h \left(\sum_{j=1}^N \left(\kappa^k_{ijh} \right)^{-\frac{\theta}{1-\eta}} \right)^{1-\eta}$, *and* $\bar{m} = \frac{\sigma}{\sigma-1}$.

Proof. See Appendix B.2. □

As there is a unit continuum of varieties, by the law of large numbers, π^k_{ijh} is also the expenditure share of country i on goods produced in j by firms from h using technology type k; ψ^k_{ijh} is the share of goods produced in j among those sold in market i using technology k from headquarter h.

Equation (7) illustrates that if the knowledge stock of clean technology in country h is higher, consumers in i are more likely to buy goods produced by clean technology from h, ceteris paribus. Take Switzerland and Indonesia as an example. Switzerland has accumulated a substantial knowledge stock in clean technology ($T^c_{Switzerland}$), whereas Indonesia has not accumulated much in comparison. It is likely that consumers opt for goods produced using Swiss clean technology as opposed to those using clean technology from Indonesia, i.e., $\frac{\pi^c_{i\cdot Switzerland}}{\pi^d_{i\cdot Switzerland}} > \frac{\pi^c_{i\cdot Indonesia}}{\pi^d_{i\cdot Indonesia}}$.

Trade and MP costs impact the selection of clean and dirty technologies for firms from the same home country symmetrically. A rise in shipping costs from the US to China ($\tau_{China,US}$) can prompt more US firms to produce in China and sell their products in China to circumvent such trade cost increases. Meanwhile, an increase in the cost imposed on US firms producing in China ($\gamma_{China,US}$) leads to reduced technology deployment by US firms in China, irrespective of the technology type. Labor costs influence the production location choices of firms. Countries with an abundant supply of labor, such as China and India, often have low labor costs, making them more likely to be selected as production locations.

How do environmental policies influence the choices firms make between clean and dirty technologies? How do they affect production locations choices of firms? I address these questions in Proposition 2 and 3.

Proposition 2. *(Production Location and Technology Choice) Suppose that $\iota_j^c = 0$ and $\iota_j^d = \iota_j$. The ratio of output, using clean technology to dirty technology, in production location j by firms from h follows*

$$\frac{Y_{jh}^c}{Y_{jh}^d} = \frac{T_h^c}{T_h^d} \left(1+\iota_j\right)^{\frac{\theta}{1-\eta}} \varpi_{jh},$$

where $\varpi_{jh} = \frac{\sum_i \left(\kappa_{ijh}^c\right)^{-\frac{\theta}{1-\eta}} \left(\sum_{j=1}^N \left(\kappa_{ijh}^c\right)^{-\frac{\theta}{1-\eta}}\right)^{-\eta} y_i / P_i^\theta}{\sum_i \left(\kappa_{ijh}^c\right)^{-\frac{\theta}{1-\eta}} \left(\sum_{j=1}^N \left(\kappa_{ijh}^d\right)^{-\frac{\theta}{1-\eta}}\right)^{-\eta} y_i / P_i^\theta}$, *and y_i denotes the income of country i. Holding wages and prices constant:*

(i) an increase in the environmental regulation in country h results in lower relative output using clean technology in country j by firms from h in the short run (keeping technology level constant),

$$\frac{\partial \ln Y_{jh}^c / Y_{jh}^d}{\partial \ln \left(1+\iota_h\right)} \leq 0,$$

where the weak inequality becomes an equality when $\eta = 0$;[17]

(ii) an increase in the environmental regulation in country j results in a higher relative output using clean technology in country j by firms from h,

$$\frac{\partial \ln Y_{jh}^c / Y_{jh}^d}{\partial \ln \left(1+\iota_j\right)} \geq \theta > 0,$$

where the weak inequality becomes an equality when $\eta = 0$.

Proof. See Appendix B.3. □

The expression for $\frac{Y_{jh}^c}{Y_{jh}^d}$ in Proposition 2 captures two key forces that determine the location and technology choice of a firm originating from country h. First, if country h has a higher comparative advantage in clean technology, i.e., a higher $\frac{T_h^c}{T_h^d}$, firms from that country are more likely to deploy clean technology in global production. Second, if the production location enforces more stringent environmental policies, firms tend to deploy cleaner technology for production in that location. This result generalizes the technology choice of a firm in Garcia-Lembergman et al. (2023), where

[17]When $\eta = 0$, $\frac{Y_{jh}^c}{Y_{jh}^d} = \frac{T_h^c}{T_h^d} \left(1+\iota_j\right)^{\frac{\theta}{1-\eta}}$. In this special case, $\frac{Y_{jh}^c}{Y_{jh}^d}$ does not depend on ι_h. Environmental regulations in location j, ι_j, affects Y_{jh}^d, while ι_h affects Y_{jh}^c and Y_{jh}^d symmetrically through general equilibrium.

they assume that multinational firms first choose a technology type and subsequently adhere to it for all their production worldwide.

Part (i) in Proposition 2 illustrates the relocation (or carbon leakage) effect caused by the imposition of new local environmental policies. It states that when country h tightens its environmental regulations, firms from country h relocate dirty production to a foreign country j. Thus production by firms in location j from country h becomes dirtier. Part (ii) in Proposition 2 demonstrates that the local environmental regulations of country j lead to cleaner local production by firms from all countries that produce there.

Now, I will describe how the composition of outputs, using both clean and dirty technology, is determined in country j.

Proposition 3. *(Clean and Dirty Technology Deployment) The relative output using clean technology in country j,*

$$\frac{Y_j^c}{Y_j^d} = \underbrace{\left(\sum_h \varpi_{jh}^c T_h^c\right)}_{\text{clean tech. deployment}} \underbrace{\left(\sum_h \varpi_{jh}^d T_h^d\right)}_{\text{dirty tech. deployment}}^{-1} (1+\iota_j)^{\frac{\theta}{1-\eta}},$$

where $\varpi_{jh}^c = \sum_i \frac{\left(\kappa_{ijh}^c/P_i\right)^{-\frac{\theta}{1-\eta}} y_i}{\left(\sum_{j=1}^N \left(\kappa_{ijh}^c/P_i\right)^{-\frac{\theta}{1-\eta}}\right)^{\eta}}$, *and* $\varpi_{jh}^d = \sum_i \frac{\left(\kappa_{ijh}^c/P_i\right)^{-\frac{\theta}{1-\eta}} y_i}{\left(\sum_{j=1}^N \left(\kappa_{ijh}^d/P_i\right)^{-\frac{\theta}{1-\eta}}\right)^{\eta}}$.

Proof. See Appendix B.4. □

Proposition 3 describes the global deployment of clean and dirty technologies. In production location j, firms from different headquarter countries deploy either clean technology or dirty technology. The relative output of goods using clean technology in a country j is determined by the relative force of the deployment of clean technology deployment ($\sum_h \varpi_{jh}^c T_h^c$) and dirty technology deployment ($\sum_h \varpi_{jh}^d T_h^d$), as well as the level of environmental policy stringency. The clean technology deployment term captures the composition of clean technology used by firms from all countries, with weights ϖ_{jh}^c. The dirty technology deployment term, on the other hand, depicts the composition of dirty technology. More stringent environmental policies result in a higher relative output produced by clean technology.

To see the role of MP, it is useful to compare my model to ones in the existing trade literature, which typically abstract away MP.[18] Allowing MP costs γ_{jh} to approach infinity for all $j \neq h$, we

[18] One exception is Garcia-Lembergman et al. (2023).

have

$$\lim_{\gamma_{jh}\to\infty,\forall j\neq h} \frac{Y_j^c}{Y_j^d} = \frac{T_j^c}{T_j^d}\left(1+\iota_j\right)^{\theta}.$$

Such a framework in the existing trade literature would not allow us to understand the relocation of dirty production in response to local environmental policies, i.e., the effect of ι_h on $\frac{Y_j^c}{Y_j^d}$. Furthermore, the output produced by clean technology relative to dirty ones depends only on the relative stock of knowledge at home and domestic environmental policies regardless of technology levels elsewhere in the world. As I will show later in the quantitative analysis, the global deployment of clean technology plays a crucial role in fortifying the long run efficacy of local environmental policies in reducing global carbon emissions.

Shifting focus from the immediate effects of environmental policies to a long-run perspective, I now describe the determinants of clean innovation.

3.3 Equilibrium R&D

In this section, I characterize the global equilibrium allocation of R&D, which is the focal point of the analysis. I proceed through the following steps: (i) I describe the global adoption of new ideas, (ii) I characterize the value of ideas, and (iii) I establish the determinants of investments in ideas for clean and dirty technology in each headquarter country.

The thrust of the argument is the following: the adoption of new ideas hinges on their productivity, with their local profitability determining whether they get adopted. Once an idea is embraced, it yields profits, but upon replacement in some locations, its overall profitability diminishes. The present discounted value of an idea is the discounted expected profit stream it generates over time. Crucially, the relative value of ideas for both clean and dirty technologies determines the relative allocation of R&D investments in these two technologies.

3.3.1 Adoption of New Ideas.

R&D efforts give rise to new ideas. Some new ideas might not be sufficiently productive to be adopted, in some or all markets, and even after being adopted they are subject to global Schumpeterian competition from even newer ideas. Nonetheless, some of these ideas do succeed and are adopted. In this section (Lemma 2), I formally characterize the probability that a new idea is adopted in different destinations.

Let $b_{ijh,t}^k$ denote the probability that an idea about technology k originated in h is competitive in the destination country i through production location j at date t.

Lemma 2. *The probability that the goods shipped to market i use ideas from h for technology k is*

$$b^k_{i\cdot h,t} = \sum_j b^k_{ijh,t} = \underline{q}\left(T^k_{h,t}\right)^{-1}\pi^k_{i\cdot h,t}\left[\left(\psi^k_{ihh,t}\right)^{1-\eta} + \sum_{j\neq h}\left(\psi^k_{ijh,t}\right)^{1-\eta}\right],$$

where $\underline{q} \equiv N^{-\eta}$.

Proof. See Appendix B.5. $\square$

Lemma 2 demonstrates that, to serve market i through production location j, a new idea from h has to overcome two hurdles. First, the new idea needs to beat the frontier of the existing knowledge from h. The larger the current stock of knowledge, $T^k_{h,t}$, the more difficult for a new idea to be adopted. Second, the new idea from h has to withstand global competition. Conditional on being the state-of-art idea in h, the idea has to be competitive in market i, which it is with probability $\pi^k_{i\cdot h,t}$. To be used for production in country j, the new idea has to hold up to competition from other multinationals and local firms, which it does with probability $\left(\psi^k_{ijh}\right)^{1-\eta}$.

The term $\left(\psi^k_{ijh}\right)^{1-\eta}$ also captures the role of environmental regulations. In the extreme case where there is a prohibitive regulation on dirty technology in h, ψ^d_{ihh} goes to zero. However, the domestic ban on the use of dirty technology does not prevent its use abroad, as indicated by the term $\sum_{j\neq h}\left(\psi^d_{ijh}\right)^{1-\eta}$. Firms may choose to produce their products in foreign locations that permit the use of dirty technology. Therefore, in this case, new ideas about dirty technology will still have a positive probability of being adopted.

To emphasize the impact of relocation possibilities on the adoption of new ideas, it is helpful to compare the equivalent probability in a closed economy where trade and production costs are arbitrarily high,

$$b^k_{h,t} = \left(T^k_{h,t}\right)^{-1}\frac{T^k_{h,t}\left(1+\iota^k_{h,t}\right)^{-\theta}}{\sum_{k'=1}^K T^{k'}_{h,t}\left(1+\iota^{k'}_{h,t}\right)^{-\theta}}.$$

In the closed-economy case, environmental regulations affect the usage of dirty technology. In the similar extreme case where there is a prohibitive regulation on dirty technology, $\left(1+\iota^d_{h,t}\right)^{-\theta}$ goes to zero. In this case, due to prohibitive regulations, dirty technology will not be used. Consequently, new ideas concerning dirty technology will not be adopted. When compared to a closed economy, the potential for multinational production dampens the impact of domestic environmental regulations on the probability of adopting new ideas for dirty technology.

3.3.2 Profits and Values of Ideas

The profitability and churning rate of new ideas hinge on the probability that they are successfully adopted in different markets. In a scenario where firms engage in Bertrand competition within each variety market, the emergence of superior ideas that replace existing ones affects the evolution of equilibrium markups over time. At instances, newer ideas do not prove superior the best existing technology, but do pose new competitive threats as the second-best one, thereby reducing the markup charged by the incumbent producer. I characterize the equilibrium markup distribution in each country in Lemma 3.

Lemma 3. *The distribution of markups within each country follows Pareto distribution*

$$\mathbb{P}\left(\mathcal{M} \leq m\right) = \begin{cases} 1 - m^{-\theta} & \textit{if } 1 \leq m < \frac{\sigma}{\sigma-1}, \\ 1 & \textit{if } m = \frac{\sigma}{\sigma-1}. \end{cases}$$

Proof. See Appendix B.6. □

Lemma 3 establishes that the distribution of the markup remains constant over time, regardless of the destination, headquarters, or production location. This is primarily attributed to the dynamic interplay between the selection effect, favoring the dominance of superior ideas over time, and the "whittling down" effect, which tends to diminish the markups of existing ideas through the introduction of less competitive new ideas. These offsetting effects ultimately contribute to the stabilization of the markup distribution.

Patent Market and Competitive Auctions. Independent researchers search for new ideas. When new ideas are successful, researchers patent them and sell the patents in competitive auctions. Prospective producers, characterized as risk-neutral firms, engage in bidding wars, driving the patent's value upwards until it aligns with the expected present discounted value of the potential profits it can yield. A crucial aspect is that researchers appropriate all the profits linked to a new idea, thereby making innovation choices hinge on the discounted value of these ideas.[19] Further elaboration on this process will follow in the subsequent section.

[19] See Romer (1990) for similar market structure for ideas.

Values of Ideas. Using Lemma 3, we can show that profits generated from sales in country i, earned by either domestic or foreign firms selling there, are $\Pi_{i,t} = \frac{X_{i,t}}{1+\theta}$, where $X_{i,t}$ is the total expenditure of country i. With probability $b^k_{ijh,t}$, the value of a new idea about technology k from h to serve market i through production location j is

$$V^k_{ijh,t} = P_{h,t} \int_t^{\infty} \exp\left[-\int_t^s r_{h,u} du\right] \frac{\Pi_{i,s}}{P_{h,s}} \frac{b^k_{ijh,s}}{b^k_{ijh,t}} ds.$$

Knowing that $b^k_{ijh,t}$ represents the probability of an idea being successful at date t, the ratio of $\frac{b^k_{ijh,s}}{b^k_{ijh,t}}$ is the probability that the idea remains successful at date s given that it was successful at date t. Conditional on being successful at date s in market i, the profit generated in market i is $\Pi_{i,s}$. The integration of the stream of profits discounted by interest rate, $r_{h,u}$, gives the value of the idea. The expected value of an idea about technology k originated in country h is the weighted sum of its value across all production locations j and destination markets i, with the weight being $b^k_{ijh,t}$:

$$V^k_{h,t} = \sum_{i=1}^{N}\sum_{j=1}^{N} b^k_{ijh,t} V^k_{ijh,t} = \underline{q}\left(T^k_h\right)^{-1} P_{h,t} \sum_{i=1}^{N}\sum_{j=1}^{N} \pi^k_{i\cdot h}\left(\psi^k_{ijh}\right)^{1-\eta}\left\{\int_t^{\infty} \exp\left[-\int_t^s r_{h,u} du\right] \frac{\Pi_{i,s}}{P_{h,s}} \frac{b^k_{ijh,s}}{b^k_{ijh,t}} ds\right\}.$$

Several factors impact the value of new ideas in clean technology. First, the global stringency of environmental regulations plays a crucial role. When these regulations are stringent worldwide, it expands the market for clean technology, increasing the value of new ideas in clean technology. Second, when a production site j is a "pollution haven" of a country h, a rise in the production costs for firms from h choosing to operate in country j will decrease the market size for dirty technology among such firms. Recent instances, like the ongoing tensions between China and the US, have raised the production costs for US firms operating in China. Given that China used to be an important production location for US dirty production, this will lead to an increase in the value of clean technology for the US firms. The third factor involves scenarios where country j has less strict environmental regulations. If trade costs between j and h decrease, firms from country h may offshore dirty production to country j and then import goods produced by dirty technology from j. An example is China gaining Permanent Normal Trade Relations with the US, which led to the relocation of pollution-intensive activities from both the US and China (Choi et al., 2023), boosting the value of dirty technology originated in the US.

With the value of new ideas in hand, we are ready to characterize the equilibrium allocation of researchers between different technology types.

3.3.3 R&D Decisions

I assume that R&D involves researchers and exhibits decreasing returns to scale,

$$R^k_{h,t} = a^k_h \left(L^{k,R}_{h,t}\right)^{\nu},$$

where a^k_h is the exogenous research productivity of country h in technology k, and $L^{k,R}_{h,t}$ is the number of researchers.

From Proposition 1, we know that the scale parameter of the technology frontier $T^k_{h,t}$ equals the accumulated R&D efforts $R^k_{h,t}$ from time 0 to time t, i.e., $T^k_{h,t} = \int_0^t R^k_{h,s} ds$. Hence, from t to $t + dt$, the change of the level of technology k in country h follows

$$dT^k_{h,t} = a^k_h \left(L^{k,R}_{h,t}\right)^{\nu} dt. \tag{9}$$

The relative value of ideas for clean and dirty technologies is the key driver of the relative R&D investments in clean and dirty technologies. The optimality condition for researchers in country h is

$$L^{c,R}_{h,t} / L^{d,R}_{h,t} = \left(a^c_h V^c_{h,t} / a^d_h V^d_{h,t}\right)^{\frac{1}{1-\nu}}, \tag{10}$$

where the share of researchers for innovation in clean technology is proportional to the value and the research productivity of clean technology.

Next, I close the model with market clearing conditions and characterize the equilibrium.

3.4 General Equilibrium

I follow Eaton and Kortum (forthcoming) to define three type of profits for a location i: profits earned from sales there, denoted Π_i, profits earned from production there, denoted Π^P_i, and profits earned from innovation there, denoted Π^R_i. In country h, consumer expenditure equals income earned. Household income comes from wage payments to workers and profits generated by innovation, so that

$$P_h C_h = X_h = w_h L^P_h + \Pi^R_h. \tag{11}$$

The total profits earned from innovation in h are a sum of profits associated with production that uses ideas from h, which implies

$$\Pi_h^R = \sum_{j,h,k} \pi_{ijh}^k \Pi_i, \tag{12}$$

Since exporters earn profits from sales abroad, profits generated from production in j are

$$\Pi_j^P = \sum_i \pi_{ij} \Pi_i = \frac{1}{\theta} \left(w_j L_j^P \right), \tag{13}$$

where $\pi_{ij} = \sum_{h,k} \pi_{ijh}^k$ is the share of profits associated with sales in i but production in j, and profits earned from sales in i are

$$\Pi_i = \frac{X_i}{1+\theta} = \frac{1}{1+\theta} \left(w_i L_i^P + \Pi_i^R \right). \tag{14}$$

Market clearing conditions follow that total revenues from production in j equal total payments to goods produced in location j,

$$w_j L_j^P + \Pi_j^P = \sum_i \pi_{ij} X_i = \sum_i \pi_{ij} \left(w_i L_i^P + \Pi_i^R \right). \tag{15}$$

The adding-up constraints for labor markets are

$$\bar{L}_{j,t} = \bar{L}_{j,t}^P + \bar{L}_{j,t}^R, \tag{16}$$

$$\bar{L}_{j,t}^R = \sum_k L_{j,t}^{k,R}, \tag{17}$$

where $\bar{L}_{j,t}^P$ and $\bar{L}_{j,t}^R$ are supply of production workers and researchers.

The state of the economy is $\boldsymbol{L}_t = \left(L_t^P, L_t^R \right)$, where $\boldsymbol{L}_t^P = \{L_{h,t}^P\}_{h \in \mathcal{N}}$ and $\boldsymbol{L}_t^R = \{L_{h,t}^{k,R}\}_{h \in \mathcal{N}}^{k \in \mathcal{K}}$. I follow the quantitative trade literature and call bilateral trade cost $\boldsymbol{\tau}_t = \{\tau_{ij,t}^k\}_{i,j \in \mathcal{N}}^{k \in \mathcal{K}}$, MP cost $\gamma_t = \{\gamma_{jh,t}^k\}_{j,h \in \mathcal{N}}^{k \in \mathcal{K}}$, and environmental policies $\boldsymbol{\iota}_t = \{\iota_{j,t}^k\}_{j \in \mathcal{N}}$ time-variant fundamentals $\boldsymbol{\Theta}_t = (\tau_t, \gamma_t, \iota_t)$ and $\bar{\boldsymbol{\Theta}} = (\{a_h\}_{h \in \mathcal{N}})$ time-invariant fundamentals. I denote $\boldsymbol{T}_t = \{T_{h,t}^k\}_{h \in \mathcal{N}}^{k \in \mathcal{K}}$ the cross-section knowledge stock, $\boldsymbol{b}_t = \{b_{ijh,t}^k\}_{i,j,h \in \mathcal{N}}^{k \in \mathcal{K}}$, $\boldsymbol{V}_t^k = \{V_{h,t}^k\}_{h \in \mathcal{N}}$, $\boldsymbol{V}_t = \{V_t^k\}^{k \in \mathcal{K}}$, $\boldsymbol{w}_t = \{w_{j,t}\}_{j \in \mathcal{N}}$, $\boldsymbol{r}_t = \{r_{j,t}\}_{j \in \mathcal{N}}$, $\boldsymbol{\pi}_t = \{\pi_{ijh,t}^k\}_{i,j,h \in \mathcal{N}}^{k \in \mathcal{K}}$, $\boldsymbol{X}_t = \{X_{h,t}\}_{h \in \mathcal{N}}$, $\boldsymbol{P}_t = \{P_{h,t}\}_{h \in \mathcal{N}}$.

Sequential equilibrium. Next, I define a sequential equilibrium.

Definition 1. (*Sequential equilibrium*) A sequential competitive equilibrium is a sequence of wages, prices, idea values, knowledge stock, and labor allocation, $(\boldsymbol{w}_t, \boldsymbol{r}_t, \boldsymbol{V}_t, \boldsymbol{P}_t, \boldsymbol{T}_t, \boldsymbol{L}_t^P, \boldsymbol{L}_t^R)_{t=1}^{\infty}$, that solves the system of equations consisting of the trilateral shares (7), the consumer price index (8), the evolution of knowledge stocks (9), optimality conditions (10), budget balance conditions (11), profits balance conditions (12), (13), and (14), trade balance conditions (15), and labor market conditions (16) and (17), given an initial distribution of the knowledge stock $\boldsymbol{T}_0$, factor endowments $\boldsymbol{L}_0$, fundamentals $(\boldsymbol{\Theta}_t, \bar{\boldsymbol{\Theta}})$.

Balanced Growth Path. Let $g_y \equiv \frac{dy/dt}{y} = \frac{\dot{y}}{y}$ denote the growth rate of a variable y along the balanced growth path. Let $\bar{y}$ denote the variable y along the BGP if y is constant on it. I assume that the population growth rate is constant and equals n.

Definition 2. (Balanced Growth Equilibrium) A balanced growth equilibrium is a sequential equilibrium that satisfies the following balanced growth conditions. For all $i, j, h \in \mathcal{N}$, and all $k \in \mathcal{K}$

$$g_T = \frac{\dot{T}^k_{i,t}}{T^k_{i,t}} = n\nu,\ g_{b^k} = \frac{\dot{b}^k_{ijh,t}}{b^k_{ijh,t}} = -n\nu,\ g_{V^k_{ijh}} = \frac{\dot{V}^k_{ijh,t}}{V^k_{ijh,t}} = n,\ g_{V^k_h} = \frac{\dot{V}^k_{h,t}}{V^k_{h,t}} = (1-\nu)\,n,$$

$$g_{L^R} = \frac{\dot{L}^R_{i,t}}{L^R_{i,t}} = n,\ g_{L^P} = \frac{\dot{L}^P_{i,t}}{L^P_{i,t}} = n,\ g_P = \frac{\dot{P}_{i,t}}{P_{i,t}} = -n\nu/\theta,\ g_C = \frac{\dot{C}_{i,t}}{C_{i,t}} = n + n\nu/\theta.$$

and the gravity shares $\pi^k_{ijh,t} = \bar{\pi}^k_{ijh}$, $\psi^k_{ijh,t} = \bar{\psi}^k_{ijh}$, $\pi^k_{i\cdot h} = \bar{\pi}^k_{i\cdot h}$, wages $w_{i,t} = \bar{w}_i$ and the interest rate $r_{i,t} = \bar{r}_i$ are constant, with $\bar{r}_i = n + \rho$, and

$$\bar{\psi}^k_{ijh} = \frac{\left(\bar{\kappa}^k_{ijh}\right)^{-\frac{\theta}{1-\eta}}}{\sum_{j=1}^N \left(\bar{\kappa}^k_{ijh}\right)^{-\frac{\theta}{1-\eta}}},\ \bar{\pi}^k_{i\cdot h} = \frac{T^k_{h,t}\left(\sum_{j=1}^N \left(\bar{\kappa}^k_{ijh}\right)^{-\frac{\theta}{1-\eta}}\right)^{1-\eta}}{\sum_{k=1}^K \sum_{h=1}^N T^k_{h,t}\left(\sum_{j=1}^N \left(\bar{\kappa}^k_{ijh}\right)^{-\frac{\theta}{1-\eta}}\right)^{1-\eta}},\ \bar{\pi}^k_{ijh} = \bar{\psi}^k_{ijh}\bar{\pi}^k_{i\cdot h}.$$

The share of researchers for technology k in any country h is constant, which follows

$$L^{k,R}_{h,t}/L^R_{h,t} = \frac{\sum_{i=1}^N \sum_{j=1}^N \bar{\pi}^k_{i\cdot h}\left(\bar{\psi}^k_{ijh}\right)^{1-\eta}\left(\bar{w}_{i,t}L^P_{i,t} + \bar{\Pi}^R_{i,t}\right)}{\sum_{k=1}^K \sum_{i=1}^N \sum_{j=1}^N \bar{\pi}^k_{i\cdot h}\left(\bar{\psi}^k_{ijh}\right)^{1-\eta}\left(\bar{w}_{i,t}L^P_{i,t} + \bar{\Pi}^R_{i,t}\right)}.$$

The growth of the value of an idea for a given technology type has two components: n and $(1-\nu)$. The population growth rate n captures the market size effect. The return to researchers is captured by ν. It is harder to get an even better idea when ideas are good enough, so the value is discounted by $(1-\nu)$.

3.5 Clean Innovation

In this subsection, I determine how policies and underlying factors, such as MP/trade costs, collectively influence clean innovation in the long run. Let $s^I_{h,t} \equiv R^c_{h,t}/R^d_{h,t}$ denote the ratio of R&D allocated to clean technology to dirty technology in country h at date t. Demonstrating the mechanics of clean innovation analytically presents a challenge. To emphasize the core mechanisms involved, I investigate clean innovation along the balanced growth path.

Proposition 4. *(Clean innovation) Along the BGP, the relative clean R&D in country h follows*

$$\ln R^c_{h,t}/R^d_{h,t} = \ln a^c_{h,t}/a^d_{h,t} + \nu \ln T^c_{h,t}/T^d_{h,t} - \nu \ln \sum_j \bar{\chi}_{jh,t} \left(1 + \iota_{j,t}\right)^{-\theta},$$

where $\bar{\chi}_{jh,t} = \frac{\sum_{i=1}^N \left(\tau_{ij,t}\bar{w}_{j,t}\gamma_{jh,t}/P_{i,t}\right)^{-\theta}\left(\bar{w}_{i,t}L^P_{i,t}+\bar{\Pi}^R_{i,t}\right)}{\sum_{i=1}^N \sum_{j=1}^N \left(\tau_{ij,t}\bar{w}_{j,t}\gamma_{jh}/P_{i,t}\right)^{-\theta}\left(\bar{w}_{i,t}L^P_{i,t}+\bar{\Pi}^R_{i,t}\right)}$.

Proof. See Appendix B.9. □

This expression summarizes all relevant determinants for clean innovation. Specifically, the clean R&D relative to the dirty R&D depends on the relative research productivity a^c_h/a^d_h. Obviously, a higher return to researchers (a higher ν) implies more R&D. Notably, innovation features path dependence, as $s^I_{h,t}$ is proportional to the relative knowledge stock of clean technology, $T^c_{h,t}/T^d_{h,t}$, which is consistent with the empirical findings in the automobile industry in the US (Aghion et al., 2016). The term $\bar{\chi}_{jh,t}$ reflects the global production potential shaped by geographic barriers and labor cost; it also captures the exposure to the environmental policy in j (i.e., $\iota_{j,t}$). Increasing $\iota_{j,t}$ will affect the relative clean innovation $s^I_{h,t}$ in country h, and the strength of the effect is determined by the exposure to $\iota_{j,t}$:

$$\frac{\partial \ln s^I_{h,t}}{\partial \ln\left(1+\iota_{j,t}\right)} = \frac{\theta \bar{\chi}_{jh,t}\left(1+\iota_{j,t}\right)^{-\theta}}{\sum_j \bar{\chi}_{jh,t}\left(1+\iota_{j,t}\right)^{-\theta}} \leq \theta, \tag{18}$$

holding wages and prices constant. In the limit case where the MP cost is infinity,

$$\lim_{\gamma_{jh}\to\infty,\forall j\neq h} \frac{\partial \ln s^I_{h,t}}{\partial \ln\left(1+\iota_{j,t}\right)} = \begin{cases} \theta & j = h, \\ 0 & j \neq h. \end{cases} \tag{19}$$

Comparing equation (18) and (19), the effect of domestic environmental policies on home clean innovation is dampened ($\leq \theta$) in the presence of multinational production, compared to the case where no MP is allowed (θ). A change in the MP or trade costs will affect the environmental policy exposure $\bar{\chi}_{jh,t}$ and thus shapes the innovation response to the change in environmental policy stringency.

3.6 Connection to the Existing Models

Before proceeding to calibrate the model for quantitative analysis, I briefly summarize how my model relates to models in the existing trade literature. The model in this paper nests MP and trade models in the existing literature. Table 3 provides a summary of special cases and this paper's connection to the existing literature.

Table 3: Connection to the Existing Literature

	Trade	MP	Tech. choice	Innovation	Dynamics
Eaton and Kortum (2002)	✓				
Ramondo and Rodríguez-Clare (2013)	✓	✓			
Arkolakis et al. (2018)	✓	✓		✓	
Eaton and Kortum (2001) and Somale (2021)	✓			✓	✓
Farrokhi and Pellegrina (forthcoming)	✓		✓		
Garcia-Lembergman et al. (2023)	✓	✓	✓		
This paper	✓	✓	✓	✓	✓

Notes: The table shows the connection of this paper to the existing literature.

If we mute the channel of technology selection and innovation by letting $K = 1$ and $\nu = 0$, the model collapses to the static block of Cai and Xiang (2023) and the special case of Ramondo and Rodríguez-Clare (2013) without intermediate inputs. If we further let $\gamma_{jh} \to \infty, \forall j \neq h$, the model reduces to Eaton and Kortum (2002). Letting $\gamma_{jh} \to \infty, \forall j \neq h$ and $K = 1$ but allowing for $\nu > 0$ or $a_h^k > 0$, the model goes back to Eaton and Kortum (2001) and the single-sector version of Somale (2021).

4 Calibration

The quantitative exercises exploit a sample of 52 countries ranging from OECD countries to major developing economies such as Brazil, Russia, India, China, and South Africa (BRICS).[20]I also

[20]The countries are Argentina, Australia, Austria, Belgium, Bulgaria, Brazil, Canada, Switzerland, Chile, China, Colombia, Costa Rica, Czech Republic, Germany, Denmark, Spain, Estonia, Finland, France, United Kingdom, Greece, Croatia, Hungary, Indonesia, India, Ireland, Iceland, Italy, Japan, South Korea, Lithuania, Latvia, Morocoo, Mexico, Malta, Malaysia, Netherlands, Norway, New Zealand, Philippines, Poland, Portugal, Romania, Russia, Slovak Republic, Slovenia, Sweden, Thailand, Turkey, United States, Vietnam, and South Africa.

include Mexico as it is an important host country of foreign production not only for the North American Free Trade Agreement (NAFTA) countries but also for Japan and Germany. I calibrate the model to the most recently available data and assume it is on the BGP.

4.1 Calibration Procedures

The parameters to be calibrated are as follows,

$$\left(\underbrace{\gamma, \tau, \theta, \eta}_{\text{static}}, \underbrace{T, \nu}_{\text{dynamic}}, \underbrace{\bar{u}, e, \iota}_{\text{environment}} \right),$$

where matrices of MP cost γ's, matrices of trade cost τ's, shape parameters θ and η of the Fréchet distribution govern equilibrium production and trade pattern; $T's$ and ν are parameters related to $R\&D$ and the knowledge stock; $\bar{u}'s$, $e's$, and ι's are environment-related parameters.

Ideally, if we were to observe clean and dirty trade and MP flows separately, given the shape parameters of the Fréchet distribution, we could recover matrices of trade and MP cost by combining the gravity structure of the model with bilateral trade and MP flow data, as demonstrated in Arkolakis et al. (2018). However, in our case, we only have access to aggregate trade and MP flow data. The primary challenge thus concerns how to split the aggregate flow into clean and dirty components. To address this challenge, I rely on a key insight from the model: the technical comparative advantage of a firm's home country and the environmental regulations in its production locations jointly influence a firm's choice of technologies and locations. This, in turn, necessitates information on knowledge stocks for both clean and dirty technologies, as well as data on environmental regulations. To obtain this information, I utilize micro-level patent data to construct knowledge stocks T's and employ environment-related tax data from the OECD Policy Instrumentals for the Environment Database (OECD-PINE) to derive environmental taxes (ι's). I will provide a detailed explanation of this procedure in the subsection next. The static block of production and trade in the model bears similarities to the approach taken by Arkolakis et al. (2018). As such, I externally calibrate the parameters η and θ following Arkolakis et al. (2018). With knowledge stocks, environment-related taxes, shape parameters following a Fréchet distribution, and bilateral trade and MP flow data, I am able to recover matrices of MP cost (γ's) and trade cost (τ's). I follow the literature on semi-endogenous growth to set $\nu = 0.76$ (Bloom, Jones, Van Reenen, and Webb, 2020; Somale, 2021). To further ground the model, I target CO_2 emissions data at the country level to recover the emission intensity of dirty technology for each country. Additionally, I calibrate the environmental damage parameter ($\bar{u}$'s) that takes into account both the social cost of CO_2 emissions as well as country-specific environmental damages resulting from climate change. I will provide a detailed explanation of this calibration procedure in the following sections.

4.2 Calibration Results

To obtain the empirical counterpart of bilateral trade in the model, I use the CEPII Gravity database on trade flows from any country j to i for all $i \neq j$. For home sales, I use the multi-region input-output table (MRIO) from the Eora global supply chain database on trade flows.[21] As a result, I construct an $N \times N$ matrix of trade shares $\bar{\pi}_{ij}^{T}$, as well as $N \times 1$ vectors of aggregate expenditures $\bar{X}_i$ and aggregate production $\bar{Y}_j$. Following Cai and Xiang (2023), I get the empirical counterpart of bilateral MP flows in the model by using the Analytical AMNE database compiled by the OECD. The AMNE database contains a full matrix of output of foreign affiliates in 59 countries from 2005 to 2016 (Cadestin et al., 2018). Using this information, I construct an $N \times N$ matrix of MP shares $\bar{\pi}_{ij}^{M}$.

To construct the vector of labor supply, I proceed as follows. I combine "Researchers in R&D (per million people)" from the World Development Indicator as well as population data from the Penn World Table (PWT) 10.0 (Feenstra et al., 2015) to compute the total researchers $\bar{L}_i^R$ in each country.[22] The total employment is obtained from PWT 10.0.

Following Arkolakis et al. (2018), I set the shape parameter of the Fréchet distribution to be $\theta = 4.5$ and the correlation parameter to $\eta = 0.55$. The elasticity of substitution is set to be $\sigma = 3.79$, following Bernard, Eaton, Jensen, and Kortum (2003). The annual discount factor is set to be $\rho = 0.95$.

The calibration of the rest of the parameters proceeds in two steps. First, I estimate the knowledge stock of clean and dirty technology, respectively, by using patent data. Second, I recover the matrices of τ's and $\gamma's$, as well as the vector $\iota's$. To do so, I combine the gravity structure of the model and bilateral trade and MP flows data to inform the calibration of τ's and $\gamma's$. Given θ and η and setting $\gamma_{jj} = \tau_{jj} = 1$ for all j, I back out the trade costs τ_{ij} and MP costs γ_{jh} without impos-

[21] Conte et al. (2022) explains in detail the construction of the CEPII Gravity database. For more information about the MRIO from Eora, see Lenzen et al. (2013).

[22] For Australia, Brazil, Switzerland, India, New Zealand, Philippines, and Vietnam, the researcher information is missing in the year 2016. I imputed the data using available information in adjacent years assuming a constant growth rate. Linear inputation provides similar results.

Table 4: Calibrated Model Parameters and Data Targets

Parameters			Moments
Notation	Value	Description	Description
Internal			
T_h^c	0.26 (0.89)	Clean tech. stock in h ($\times 10^5$)	Accumulated citation-weighted clean patents in h
T_h^d	2.6 (10.8)	Non-clean tech. stock in h ($\times 10^5$)	Accumulated citation-weighted non-clean patents in h
τ_{ij}	4.4 (2.1)	Trade cost from j to i	Trade share from j to i
γ_{jh}	4.7 (2.3)	MP cost from h to j	MP share from h to j
ι_j^d	.0353 (.0208)	Environmental tax rate in j (%)	Environmental tax revenues in j
e_h	6.56 (8.35)	CO2 emissions intensities	CO2 emissions $\mathcal{E}_j$ in j
$\bar{u}_h$	8.12 (4.24)	Environment damage ($\times 10^{-10}$)	Social cost of one ton CO2 emissions (IPCC)
External			
θ	4.5	Fréchet shape parameter	Arkolakis et al. (2018)
η	0.55	Fréchet correlation parameter	Arkolakis et al. (2018)
ν	0.76	Returns to researchers	Bloom, Jones, Van Reenen, and Webb (2020) and Somale (2021)

Notes: The table summarizes the calibrated structural parameters. Parameter values for T_h^c, T_h^d, ι_j^d, e_h, and $\bar{u}_h$ are averages across N countries. Parameter values for τ_{ij} and γ_{jh} refer to averages across $N \times (N-1)$ country pairs. Standard deviations are in parenthesis.

ing symmetry. Next, I obtain the model-consistent environmental tax rate ι_j for all j by targeting the environment-related tax rates reported in the OECD Policy Instrumentals for the Environment Database (OECD-PINE). Given the matrices of τ's and $\gamma's$, the vector $\iota's$, and knowledge stocks T's, I recover all trilateral shares π_{ijh}^k for each technology type k. Table 4 summarizes the calibrated parameters and data targets.

Knowledge Stock $T_{h,t}^k$. I leverage the PATSTAT Global data maintained by the European Patent Office (EPO) to calculate the levels of clean and dirty technology, proceeding in four steps. First, I identify clean technology. Second, I classify patents into unique inventions. Third, using information on patent assignees, inventors, and patent authorities, I assign country ownership to each unique invention. Lastly, I construct a citation-weighted measure of the knowledge stock.

To identify clean technology, I use the cooperative patent classification (CPC) code. I follow the EPO document to identify climate change mitigation technologies (Angelucci et al., 2018).[23] I restrict my analysis to publications that have been formally granted. I define clean technology as a patent family that contains at least one patent application falling into one of the categories "Y02A", "Y02B", "Y02C", "Y02D", "Y02E", "Y02P", "Y02T","Y02W", and "Y04S".[24] I classify the remaining patents as non-clean (dirty) patents. Note that dirty patents do not particularly call out patents

[23]See https://www.epo.org/mobile/news-events/in-focus/classification/classification/updatesYO2andY04S.html.
[24]See Appendix A.1 for details.

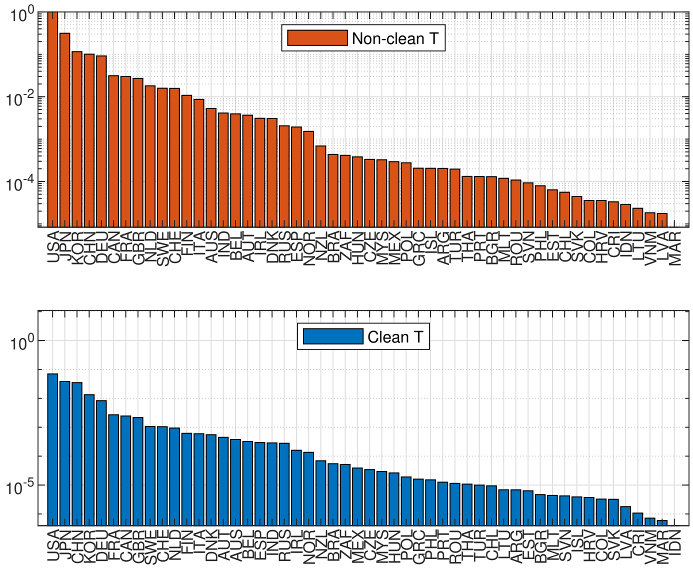
associated with dirty technology, but rather are defined as the complement of patents associated with clean technology. For the rest of this section, I use "dirty" and "non-clean" interchangeably.

As Liu and Ma (2021) point out, multi-filing is a major challenge when dealing with international patent data. It is common practice for innovators to file the same invention through various patent authorities in different countries. Furthermore, different aspects of a single invention could result in multiple patent filings. Patents that belong to a single invention form a patent family with a unique patent identifier. To address this double counting concern, my analysis uses the simple family ID (DOCDB family ID in PATSTAT database) to uniquely identify an invention.

Third, to assign country ownership to each patent (invention), this paper uses information about patent assignees, patent inventors, and patent authority offices. More specifically, the country of the patent assignee is used as the country of ownership for a patent. If there are multiple assignees, the patent is assigned evenly to the countries of each assignee. If information regarding assignees is missing, the inventor's country is used to determine the patent's ownership. Lastly, if information about assignees and inventors are both missing, the patent's country is determined by its patent authority.

After these previous steps, I am able to construct knowledge stock measures guided by my model. In my theory, each idea can be viewed as a potential patent. Each idea arrives at a Poisson arrival intensity specific to each country and technology type, the time integration of which is the mean of the Poisson distribution of idea counts by date t. The empirical counterpart of these idea counts is the accumulated number of unique inventions in each country for each technology type by date t. In other words, the knowledge stock $T^k_{h,t}$ is equal to the total number of unique inventions by date t for each technology type k. In the theory, the quality of a new idea is drawn from a distribution that is invariant across countries. To account for this theoretical assumption, I normalize the patent quality using citation weights. Following Akcigit, Caicedo, Miguelez, Stantcheva, and Sterzi (2018) and Prato (2021), I use the sum of citations in a three-year window after application for all patents to account for the issue of truncated citations. Using the components calculated in these four steps, I aggregate the total number of citation-weighted unique inventions by each year for clean and dirty technology to construct measures for $T^k_{h,t}$ (see Figure 3).

Figure 3: Clean and Non-clean Knowledge Stocks



Notes: The figure displays the accumulated citation-weighted patents from 1976 to 2017. The red bars represent regular (non-clean) patent stock, while the blue bars represent clean patent stocks. The non-clean patent stock in the US is normalized to 1. All bars represent values relative to US non-clean patent stocks. Patent application and citation data are sourced from the Global PATSTAT database maintained by the European Patent Office.

Trade and MP Costs. I now turn to the estimation of trade and MP costs. Let Y_i denote the value of production in country i, and X_{ijh} the total sales of firms originated in h that serve market i from location j. The aggregate trilateral flow can be expressed as

$$X_{ijh} = \sum_k \pi_{ijh}^k X_i = \sum_k \psi_{ijh}^k \pi_{i \cdot h}^k X_i.$$

This expression is useful for the construction of MP and trade shares. In particular, MP shares are defined as production shares across firms from different origins, $\pi_{jh}^M \equiv \frac{\sum_i X_{ijh}}{\sum_{i,h} X_{ijh}}$, while trade shares are given by expenditure shares across production locations, $\pi_{ij}^T \equiv \frac{\sum_h X_{ijh}}{\sum_{j,h} X_{ijh}}$. Therefore, MP and trade shares are, respectively,

$$\pi_{jh}^M = \sum_i X_{ijh}/Y_j = \sum_{k,i} \psi_{ijh}^k \pi_{i \cdot h}^k X_i / Y_j, \ \pi_{ij}^T = \sum_{k,h} \psi_{ijh}^k \pi_{i \cdot h}^k = \sum_{k,h} \pi_{ijh}^k.$$

Define $\xi_{jh,t}^M = \left(\gamma_{jh,t}\right)^{-\theta}$ and $\xi_{ij,t}^T = \left(\tau_{ij,t} w_{j,t}^P\right)^{-\theta}$, $\xi_j^k = \left(1 + \iota_j^k\right)^{-\theta}$. I assume that there is no regulation on clean production, i.e., $\iota_j^c = 0$. The bilateral trade and MP flows that are observed in the data can be expressed as

$$\pi^T_{ij,t} \equiv \sum_h \sum_k \frac{X^k_{ijh}}{X_i} = \sum_{h=1}^N \sum_{k=1}^N \frac{T^k_{h,t}\left(\xi^M_{jh,t}\xi^k_j\xi^T_{ij,t}\right)^{\frac{1}{1-\eta}}}{\sum_{j=1}^N T^k_{h,t}\left(\xi^T_{ij,t}\xi^M_{jh,t}\xi^k_j\right)^{\frac{1}{1-\eta}}} \frac{\left[\sum_{j=1}^N T^k_{h,t}\left(\xi^T_{ij,t}\xi^M_{jh,t}\xi^k_j\right)^{\frac{1}{1-\eta}}\right]^{1-\eta}}{\sum_{k=1}^K \sum_{h=1}^N \left[\sum_{j=1}^N T^k_{h,t}\left(\xi^T_{ij,t}\xi^M_{jh,t}\xi^k_j\right)^{\frac{1}{1-\eta}}\right]^{1-\eta}}, \tag{20}$$

$$\pi^M_{jh,t} \equiv \sum_i \sum_k \frac{X^k_{ijh}}{Y_j} = \sum_{i=1}^N \sum_{k=1}^N \frac{X_{i,t}}{Y_{j,t}} \frac{T^k_{h,t}\left(\xi^T_{ij,t}\xi^M_{jh,t}\xi^k_j\right)^{\frac{1}{1-\eta}}}{\sum_{j=1}^N T^k_{h,t}\left(\xi^T_{ij,t}\xi^M_{jh,t}\xi^k_j\right)^{\frac{1}{1-\eta}}} \frac{\left[\sum_{j=1}^N T^k_{h,t}\left(\xi^T_{ij,t}\xi^M_{jh,t}\xi^k_j\right)^{\frac{1}{1-\eta}}\right]^{1-\eta}}{\sum_{k=1}^K \sum_{h=1}^N \left[\sum_{j=1}^N T^k_{h,t}\left(\xi^T_{ij,t}\xi^M_{jh,t}\xi^k_j\right)^{\frac{1}{1-\eta}}\right]^{1-\eta}}. \tag{21}$$

The environment-related tax rate in each country is thus

$$\Omega_{j,t} = \iota^d_{j,t} \frac{\theta}{1+\theta} \sum_i \sum_h \tau_{ij} \pi^d_{ijh,t} X_{i,t}. \tag{22}$$

With data of $\pi^T_{ij,t}$, $\pi^M_{jh,t}$, $X_{i,t}$, and $Y_{j,t}$ constructed from the CEPII and AMNE data, $\Omega_{j,t}/Y^d_{j,t}$ from the OECD-PINE data base, and the $T^k_{h,t}$ computed above, I solve for $\xi^T_{ij,t}$, $\xi^M_{jh,t}$, and ξ^k_j using (20), (21), and (22).

The MP costs $\gamma_{jh,t}$ can be exactly backed out from the system of equations (20), (21), and (22), given the value of θ. In particular $\gamma_{jh,t} = \left(\xi^M_{jh,t}\right)^{-\frac{1}{\theta}}$. Imposing that $\tau_{jj} = 1$ for all j, then $\xi^T_{jj,t} = \left(x_{j,t}\right)^{-\theta}$, which implies that normalizing to one the diagonal of the matrix of $\xi^T_{ij,t}$, we could obtain a matrix of trade costs which do not necessarily have to be symmetric between i and j. With γ_{jh} and τ_{ij}, we can construct trilateral shares π^k_{ijh} using the gravity structure for all i, j, h, k.

Carbon Emissions Intensity e_h. Total emissions generated in each production location j are from dirty production conducted by local and foreign firms. It follows that

$$\mathcal{E}_{j,t} = \mathcal{E}^d_{j,t} = \sum_h e_h \int_0^1 y^d_{jh,t}(\omega) d\omega = \tilde{\mathcal{B}} \sum_h e_h \sum_i \frac{\tau_{ij,t} X_{i,t} \pi^d_{ijh,t}}{P_{i,t}},$$

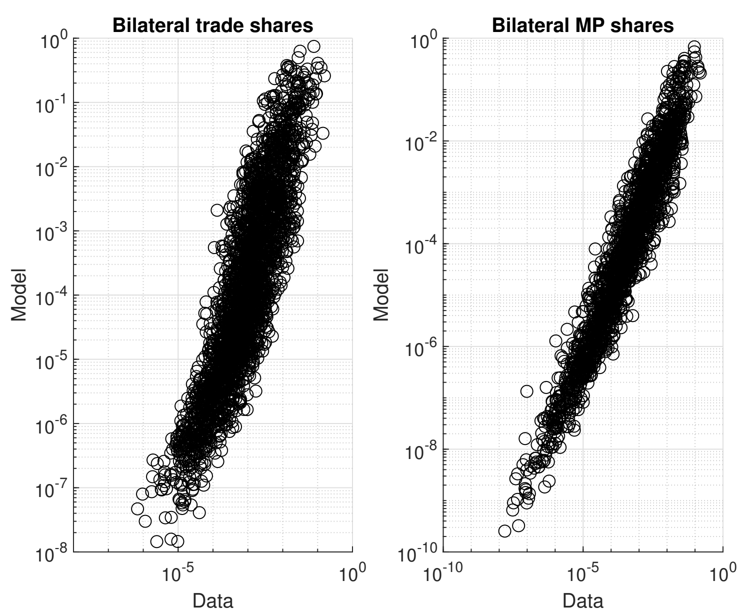
where $\tilde{\mathcal{B}} = \frac{2\theta+2+\sigma-\theta\sigma}{(\sigma-\theta-2)(\sigma-1)} \Gamma\left(\frac{2\theta-\sigma}{\theta}\right) \mathcal{B}^\sigma$. With π^d_{ijh}, P_i, and τ_{ij} computed above, as well as data on $\bar{X}_i$ and $\mathcal{E}_j$, we obtain e_h for all h.

Environmental Damages u_h. I follow Shapiro (2016) to calibrate $\bar{u}_h$ in the environmental damage function $f(\boldsymbol{E}_t) = \frac{1}{1+\left(\bar{u}_h \sum_j \mathcal{E}_{j,t}\right)^2}$. First, I differentiate the indirect utility with respect to $\sum_j \mathcal{E}_{j,t}$ to get the marginal impact of carbon emissions on indirect utility for each country. Second, following Shapiro (2016), I solve for relative values of u_h across countries using estimates about the country-specific damage due to a 2.5 °C warming by Nordhaus and Boyer (2000). Then, I pin down the level of u_h by matching the marginal global social cost of carbon which is equal to \$38 (Sheet, 2013).[25]

4.3 Model Fit

I next assess the model fit. Figure 4 plots the model fit of bilateral trade and MP shares. As in (Arkolakis et al., 2018), the model does not exactly match all the elements of the bilateral flow matrices as the data could feature imbalanced trade.

Figure 4: Trade and MP shares: model vs. data



Notes: The figure shows the model fit of bilateral trade and MP shares. The left panel shows the model fit of trade shares; the right panel shows the model fit of MP shares. Trade shares are imports from j divided by total expenditures of i, while MP shares are foreign outputs by firms from h divided by total outputs in j.

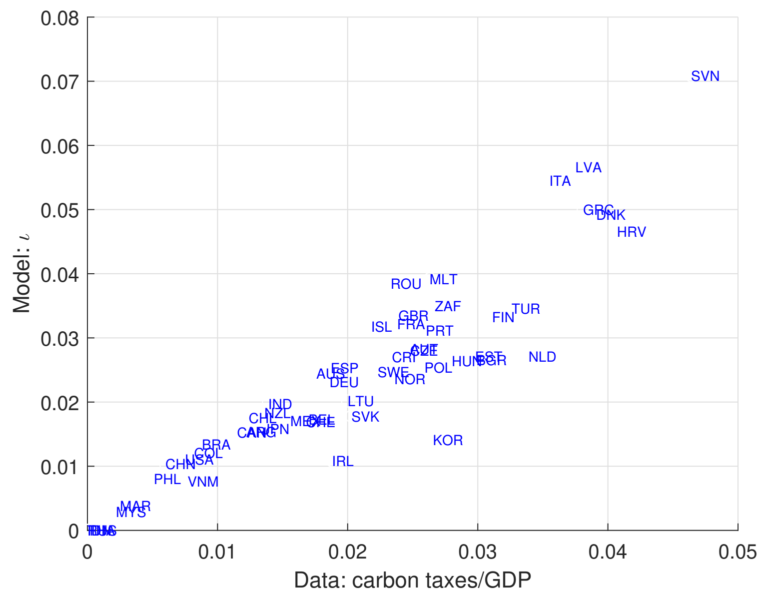
The figure demonstrates that the calibrated model fits well the bilateral trade and MP shares

[25]See https://eplanning.blm.gov/public_projects/nepa/66551/143342/176153/N15_WWP_Molvar_2014_Exhibit_1.pdf for "Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis" by Interagency Working Group on Social Cost of Carbon, United States Government. The social cost of carbon value \$38 for year 2016 linearly extrapolates the Interagency estimates of \$37 for 2015 and \$43 for 2020.

pattern in the data: the model captures 78 percent and 92 percent of the variation observed in the data for bilateral trade and MP shares, respectively.

Figure 5 depicts the relationship between the calibrated environmental policy stringency (ι) and carbon taxes per capita in the data. The calibration effectively captures the overall trends across countries. In countries such as Denmark, Greece, Italy, and Slovenia, the calibrated ι exceeds the carbon taxes per GDP. This is due to the fact that these countries have a significant share of clean GDP, while ι is essentially equal to carbon taxes divided by total dirty output.

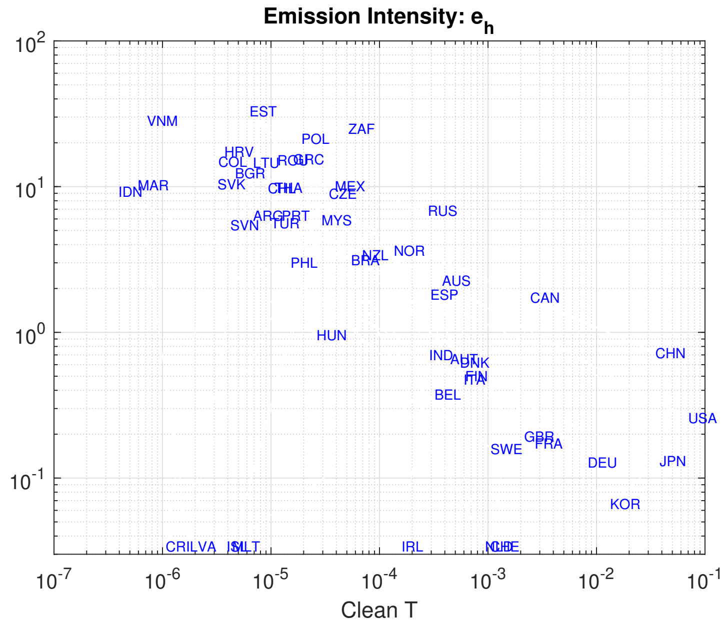
Figure 5: Environmental policy stringency and carbon tax rate



Notes: The figure shows a scatter plot of the calibrated environmental policy stringency, ι's, against per capita carbon taxes in the data. The carbon taxes data is from the OECD-PINE database, while the GDP data is from Penn World Table 10.0.

Figure 6 illustrates the relationship between the calibrated emission intensity of dirty technology, e_h, and the stock of clean technology, T_h^c. It demonstrates that countries with a higher stock of clean technology exhibit lower emission intensities associated with their dirty technologies.

Figure 6: Environmental policy stringency and carbon tax rate



Notes: The figure shows a scatter plot of the calibrated emission intensity of dirty technology, $e's$, against clean technology stock.

5 Counterfactual Experiments

Equipped with the calibrated model, I conduct a series of counterfactual exercises to learn the effects of environmental policies on innovation and production, as well as how multinational production shapes these effects. I am particularly interested in the following questions. First, how crucial is innovation in determining long-term progress in clean technological development? Second, what will be the environmental consequences of tightening environmental policies in the EU and the US? Lastly, how can more stringent environmental policies in major pollution havens boost clean innovation in developed countries?

5.1 Welfare Implications

I apply the exact hat algebra method for the counterfactual analysis. The real consumption in country i is

$$C_{h,t} = \frac{1+\theta}{\theta}\mathcal{B}^{-1}\left[\pi_{iii,t}^{-1}\sum_k T_i^k\left(\psi_{iii}^k\right)^{\eta}\left(1+\iota_j^k\right)^{-\theta}\right]^{\frac{1}{\theta}} L_{i,t}^P = \frac{1+\theta}{\theta}\mathcal{B}^{-1}\left[\frac{T_{i,t}^c}{\pi_{iii,t}^c}\left(\psi_{iii,t}^c\right)^{\eta}\right]^{\frac{1}{\theta}} L_{i,t}^P.$$

Define $\hat{x} \equiv \frac{x'}{x}$. The change in the real consumption at each date t due to a policy shock is

$$\hat{C}^P_{i,t} = \left[\frac{\hat{T}^c_{i,t}}{\hat{\pi}^c_{iii,t}} \left(\hat{\psi}^c_{iii,t}\right)^{\eta} \right]^{\frac{1}{\theta}},$$

and the change in emissions is

$$\hat{\mathcal{E}}_{j,t} = \frac{\mathcal{E}'_{j,t}}{\mathcal{E}'_{j,t}} = \frac{\sum_k \left(\mathcal{E}^k_{j,t}\right)'}{\sum_k \left(\mathcal{E}^k_{j,t}\right)} = \frac{\sum_h e_{h,t} \sum_i \tau'_{ij,t} X'_{i,t} \left(\pi^d_{ijh,t}\right)' / P'_{i,t}}{\sum_h e_{h,t} \sum_i \tau_{ij,t} X_{i,t} \pi^d_{ijh,t} / P_{i,t}}.$$

The change in welfare from a change in fundamental parameters, $\hat{W}_{i,t}$, can be measured as

$$\hat{W}_{i,t} \equiv \int_t^{\infty} e^{-(\rho+n)s} \frac{1 + \left(\bar{u}_h \sum_j \mathcal{E}_{j,t}\right)^2}{1 + \left(\bar{u}_h \sum_j \mathcal{E}'_{j,t}\right)^2} \log\left(\hat{C}_{i,s}\right) ds.$$

To perform counterfactual analysis, we define the following counterfactual equilibrium using the exact hat algebra as Dekle et al. (2008)(See Appendix C). I next turn to the the quantitative results.

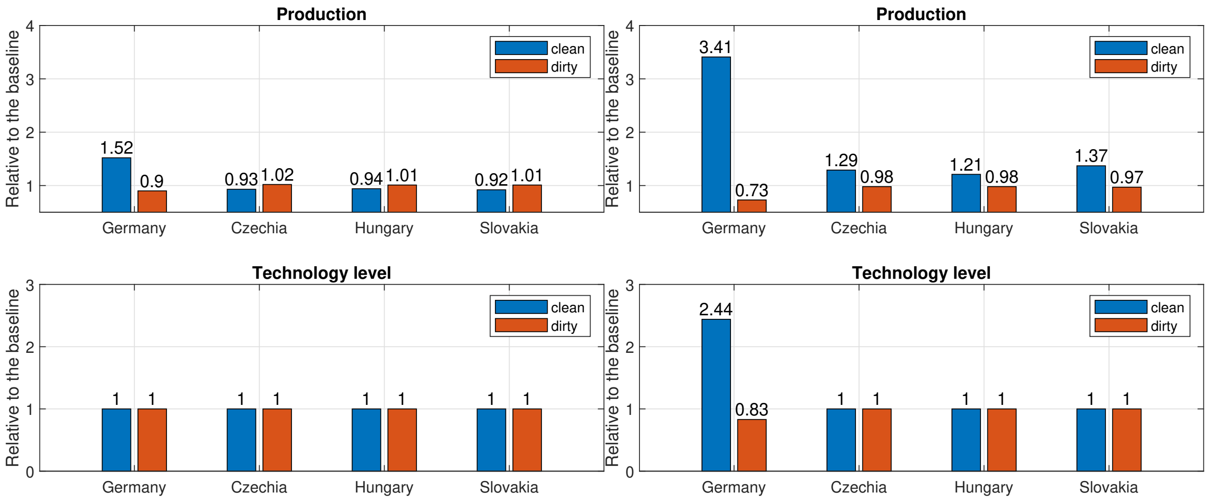
5.2 Innovation, Clean Technology Deployment, and Carbon Leakage

What is the role of innovation in determining the effect of environmental policies in the short and long run? I demonstrate that in the short run, when technology levels are kept constant, local environmental regulation leads to carbon leakage through the relocation of dirty production. However, in the long run, stimulating innovation will enhance a country's comparative advantage in clean technology, resulting in greater deployment of clean technology by firms from this country in foreign production locations. I focus on Germany, Czechia, Hungary, and Slovakia to illustrate this insight for two reasons. First, Germany is both prominent in the automobile industry, which is associated with pollution-intensive production, and home to large multinational firms that operate in various countries. Second, countries like Czechia, Hungary, and Slovakia are noteworthy because German firms account for more than 10% of production within their respective borders.

Figure 7: Carbon Leakage and Clean Technology Deployment

(a) Short run (no innovation)

(b) Long run (innovation only in Germany)



Notes: The figures display changes in production and technology levels for Germany, Czechia, Hungary, and Slovakia. Production is defined as the production within each country for each technology type, normalized by total global production, regardless of technology type. In terms of production, the blue bar represents production by clean technology, while the red bar represents production by dirty technology. Concerning technology levels, the blue bar signifies clean technology, while the red bar signifies dirty technology. The value on each bar indicates the change relative to the benchmark case. Figure 7a depicts a scenario without innovation, whereas figure 7b represents a situation in which only Germany conducts innovation.

Specifically, I increase $1 + \iota^d_{DEU}$ in Germany by 10%. Figure 7a indicates the carbon leakage effect when a country imposes a more stringent environmental policy. In this case, although production in Germany becomes cleaner, the dirty output share in the other three countries rises. Figure 7b illustrates a scenario in which only Germany conducts innovation while others do not. In this scenario, German firms have the incentive to enhance their clean technology levels in response to local environmental regulations. This leads to a higher comparative advantage for Germany in clean technology, shifting foreign production by German multinational firms toward the use of clean technology. As a result, the shares of clean output in Czechia, Hungary, and Slovakia will also increase. When we combine Figure 7a and 7b, we notice that clean technology deployment ultimately prevails over carbon leakage in the long term. In the short run, Germany may relocate dirty production to its major foreign production locations. However, in the long run, as Germany improves its clean technology and strengthens its comparative advantage in it, it also deploys more advanced clean technology to foreign production sites such as Czechia, Hungary, and Slovakia.

Table 5: Changes in CO_2 Emissions: $1 + \iota^{d}_{DEU}$ in Germany Increases by 10%

ΔCO_2	Germany	Czechia	Hungary	Slovakia
w/o innovation	-10%	0.55%	0.58%	0.24%
w/ innovation	-26%	0.44%	-0.56%	-0.30%

Notes: The table shows changes in CO_2 emissions if increasing $1 + \iota^{d}_{DEU}$ in Germany by 10%. The first row shows the changes in CO_2 emissions in a world without innovation, while the second row shows the changes in CO_2 emissions in a world where Germany can innovate to increase its technology level.

Table 5 shows changes in CO_2 emissions in the four countries. In the short run, when we keep the technology level fixed, tightening environmental regulations in Germany reduces its CO_2 emissions by 10%, while increasing CO_2 emissions in Czechia, Hungary, and Slovakia. The endogenous response of innovation in Germany enhances its comparative advantage in clean technology, eventually leading to the deployment of clean technology in Hungary and Slovakia. As a result, CO_2 emissions in these two countries decline in the long run. On the other hand, Czechia experiences an increase in CO_2 emissions even in the long run, suggesting a specialization effect. How can we reconcile the fact that the production share using dirty technology in Czechia decreases with the fact that CO_2 emissions in Czechia go up? This is because the output mix changes. As Germany shifts its technological comparative advantage toward clean technology, firms originating in Czechia specialize more in dirty production. Since the emissions intensities of the dirty technology of Czech firms are relatively higher than those of German firms, CO_2 emissions in Czechia increase.

5.3 Tightening Regulations in the US and EU

In this exercise, I investigate the environmental and welfare implications of tightening environmental policies in the EU and the US. Table 6 illustrates that when EU countries increase the stringency of their environmental policies to the level of the most stringent EU country (Slovenia), CO_2 emissions decrease in EU countries and increase in non-EU countries due to carbon leakages, holding technology levels fixed in the short run. In the long run, with endogenous innovation, both EU and non-EU countries see a decline in CO_2 emissions.

Table 6: Changes in CO_2 Emissions: Tightening Regulations in the US and EU

ΔCO_2	w/o innovation			w/ innovation		
	Global	EU	non-EU	Global	EU	non-EU
	(1)	(2)	(3)	(4)	(5)	(6)
EU uniform $\iota^* = \max_{j \in EU} \iota_j$	-0.4%	-6.4%	0.02%	-1.1%	-9.7%	-0.4%
EU + US uniform $\iota^* = \max_{j \in EU} \iota_j$	-1.3%	-6.4%	-1.0%	-4.7%	-11.9%	-4.1%

Notes: The table shows changes in CO_2 emissions. The first row shows changes in CO_2 emissions in a world where EU countries increase the stringency of their environmental policies to the level of Slovenia, while the second row shows changes in CO_2 emissions in a world where both EU countries and the US increase the stringency of their environmental policies to the level of Slovenia. Columns (1) - (3) report results when technology level is kept fixed, while columns (4) - (6) report results when technology level is allowed to adjust. Columns (1) and (4) document changes in global CO_2 emissions, columns (2) and (5) changes in CO_2 emissions in EU countries, and columns (3) and (6) changes in CO_2 emissions in non-EU countries.

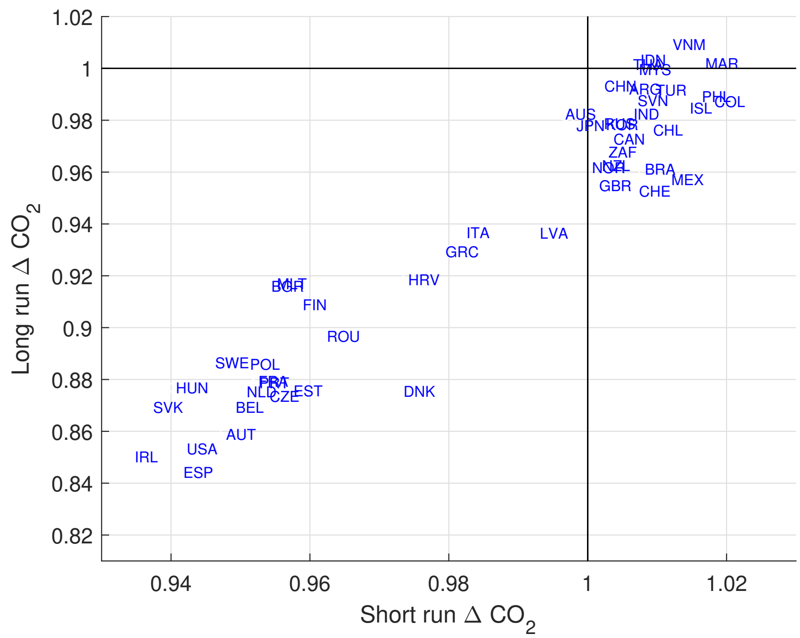
Further more, if both the US and EU countries increase the stringency of their environmental policies to the level of the most stringent EU country, global CO_2 emissions will decrease by 4.7%. What happens to each individual country? I contrast the change in CO_2 emissions in the short run, where the technology levels are fixed, with that in the long run, where endogenous innovation takes effect. Several interesting results emerge in Figure 8.

First, on the bottom left of the figure, countries reduce emissions in both the short term and the long term. These are mostly EU countries and the US, which face direct, tightened environmental regulations.

Second, on the top right of the figure, countries like Vietnam, Indonesia, Malaysia, and Thailand see an increase in emissions in both the short term and long term. In the short term, two effects take place in these countries. The first is that the relocation of dirty production from the EU and the US contributes to the emission increase in these countries. In addition, as the EU and the US impose higher production costs for dirty technology, their comparative advantage in dirty production weakens. As a consequence, countries like Vietnam, Indonesia, Malaysia, and Thailand specialize more in dirty production as they gain a comparative advantage. In the long run, when the aforementioned relocation and specialization effects dominate the clean technology deployment effect, these countries will see an increase in emissions, however modest, as the quantitative results seem to suggest.

For countries on the bottom right, emissions increase in the short run but decrease in the long run. The short-run emission increase is mainly driven by the relocation effect. In the long run, two effects drive the decrease in emissions. First, the clean technology deployed by EU and US firms in these countries becomes more productive. Second, the dirty technologies that some of these countries may deploy in the EU and US are subject to the same kinds of tightened regulations, thereby incentivizing these countries to improve their clean technology in the same way. Figure 8 highlights the heterogeneous effects of the same policy change on countries in different positions within the global production and trade network.

Figure 8: Carbon Leakage, Clean Technology Deployment, and Specialization



Notes: The figure shows a scatter plot of the changes in CO_2 emissions in the long run (with endogenous innovation) against the changes in CO_2 emissions in the short run (where the technology level is fixed), if both the US and EU countries increase the stringency of their environmental policies to the level of most stringent EU country (Slovenia).

While environmental regulation reduces real consumption, the reduction in global carbon emissions increases welfare. The welfare benefits are asymmetric: in steady state, the consumption-equivalent welfare gains are almost twice as large for EU countries as for the US. I summarize the welfare implications in Table 7.

Table 7: Welfare Implications

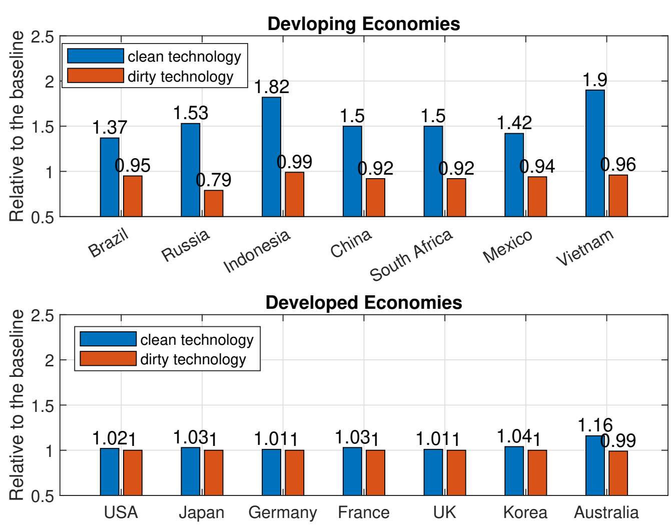
Average	$\Delta f\left(\boldsymbol{E}\right)$	$\Delta(w/P)$	ΔCE welfare
US	-9.2%	-5.0%	3.8%
EU	-9.2%	-3.1%	5.7%
World	-8.4%	-1.7%	6.4%

Notes: The table show welfare implications when both the US and EU countries increase the stringency of their environmental policies to the level of most stringent EU country (Slovenia). The first column shows the change in environmental damage $\Delta f\left(\boldsymbol{E}\right)$. The second column reports the change in real wages $\Delta\left(w/P\right)$, while the third column documents the change in consumption-equivalent welfare (CE welfare). The first row represents the US, the second row EU countries, and the third row the whole world. All results for EU countries and the whole world are a simple average across relevant countries.

5.4 Regulating Pollution Havens

If a country h has pollution activities in another country j, an increase in the environmental policy stringency in j will have an impact on country h's incentives to innovate in clean technology. Given this cross-country spillover, I investigate how regulating pollution haven countries can incentivize pollution dumping countries to conduct more R&D in clean technologies.

Figure 9: Regulating Pollution Havens



Notes: The figure shows the change in technology level associated with a 5% increase in environmental policy stringency, $1+\iota_j$, for $j=$ Brazil, Russia, Indonesia, China, South Africa, Mexico, and Vietnam. The blue bar represents clean technology, while the red bar represents dirty technology. The value on each bar indicates the change relative to the baseline case.

I increase the environmental policy stringency, $1+\iota_j$, by 5% for $j=$ Brazil, Russia, Indonesia, China, South Africa, Mexico, and Vietnam, to model increased regulation in pollution havens.

Figure 9 reveals the impact on innovation behaviors in pollution havens as well as in pollution dumping countries. The upper half shows that environmental policies have a significant impact on domestic clean innovation, while the lower half demonstrates that regulating pollution havens also stimulates more clean innovation in developed economies.

6 Concluding Remarks

Local environmental regulations can stimulate homegrown clean technology innovation, thereby contributing to the reduction of global emissions over the long term. However, in the short term, these regulations may lead to the displacement of polluting activities to other regions, thereby undermining the initial efficacy of the policy. In this paper, I construct a structural model incorporating both clean innovation and production relocation that can be used to assess the short- and long-term effects of environmental policies, particularly carbon policies.

The key qualitative insight of the model is that environmental policies can have opposing effects in the short and long run. Local environmental policies lead to the relocation of dirty production in the short run and thereby increase CO_2 emissions in foreign production locations. Endogenous innovation in clean technology that results from such policies enhances the country's technological comparative advantage in clean technology, ultimately leading to the deployment of clean technology in foreign production locations in the long run. As a result, CO_2 emissions in foreign production locations decline in the long run.

Quantitatively, if both the US and EU countries increase the stringency of their environmental policies to the level of the most stringent EU country, global CO_2 emissions will decrease by 4.7 percent. While environmental regulation reduces real consumption, the reduction in global carbon emissions increases welfare. The welfare benefits are asymmetric: in the steady state, the consumption-equivalent welfare gains are almost twice as large for EU countries compared to the US.

This framework has some limitations that future research could address. First, this paper does not study the energy market. Future studies could extend this framework by incorporating energy input for production. Second, this paper only analyzes the steady state of the model. Exploring the transition dynamics of clean innovation can further deepen our understanding of environmental policies. Despite these limitations, this framework remains useful for studying the long-run local and global impact of many real-world environmental policies.

References

Acemoglu, D., P. Aghion, L. Barrage, and D. Hémous (2023): "Climate change, directed innovation, and energy transition: The long-run consequences of the shale gas revolution," Tech. rep., National Bureau of Economic Research.

Acemoglu, D., P. Aghion, L. Bursztyn, and D. Hemous (2012): "The environment and directed technical change," *American economic review*, 102, 131–166.

Aghion, P., A. Dechezleprêtre, D. Hemous, R. Martin, and J. Van Reenen (2016): "Carbon taxes, path dependency, and directed technical change: Evidence from the auto industry," *Journal of Political Economy*, 124, 1–51.

Akcigit, U., S. Caicedo, E. Miguelez, S. Stantcheva, and V. Sterzi (2018): "Dancing with the stars: Innovation through interactions," .

Albornoz, F., M. A. Cole, R. J. Elliott, and M. G. Ercolani (2014): "The environmental actions of firms: Examining the role of spillovers, networks and absorptive capacity," *Journal of environmental management*, 146, 150–163.

Alvarez, F. E., F. J. Buera, R. E. Lucas, et al. (2013): "Idea flows, economic growth, and trade," *NBER Working Paper*.

Angelucci, S., F. J. Hurtado-Albir, and A. Volpe (2018): "Supporting global initiatives on climate change: The EPO's "Y02-Y04S" tagging scheme," *World Patent Information*, 54, S85–S92.

Arkolakis, C., N. Ramondo, A. Rodríguez-Clare, and S. Yeaple (2018): "Innovation and Production in the Global Economy," *American Economic Review*, 108, 2128–73.

Arkolakis, C., A. Rodriguez-Clare, and J.-H. Su (2017): "A Multivariate Distribution with Pareto Tails and Pareto Maxima," *Unplulished*.

Arkolakis, C. and C. Walsh (2023): "Clean Growth," *NBER Working Paper*.

Bai, Y., K. Jin, D. Lu, and H. Wang (2023): "Optimal Trade Policy with International Technology Diffusion," *Working Paper*.

Bernard, A. B., J. Eaton, J. B. Jensen, and S. Kortum (2003): "Plants and Productivity in International Trade," *American Economic Review*, 93, 1268–1290.

BILIR, L. K. AND E. MORALES (2020): "Innovation in the global firm," *Journal of Political Economy*, 128, 1566–1625.

BLOOM, N., C. I. JONES, J. VAN REENEN, AND M. WEBB (2020): "Are ideas getting harder to find?" *American Economic Review*, 110, 1104–1144.

BRONER, F., P. BUSTOS, AND V. M. CARVALHO (2012): "Sources of comparative advantage in polluting industries," *NBER Working Paper*.

BUERA, F. J. AND E. OBERFIELD (2020): "The Global Diffusion of Ideas," *Econometrica*, 88, 83–114.

CADESTIN, C., K. DE BACKER, I. DESNOYERS-JAMES, S. MIROUDOT, D. RIGO, AND M. YE (2018): "Multinational enterprises and global value chains: the OECD analytical AMNE database," .

CAI, J., N. LI, AND A. M. SANTACREU (2022a): "Knowledge Diffusion, Trade, and Innovation across Countries and Sectors," *American Economic Journal: Macroeconomics*, 14, 104–45.

CAI, S., L. CALIENDO, F. PARRO, AND W. XIANG (2022b): "Mechanics of Spatial Growth," *NBER Working Paper*.

CAI, S. AND W. XIANG (2023): "Multinational Production, Technology Diffusion, and Economic Growth," *Yale University Working Paper*.

CAI, X., Y. LU, M. WU, AND L. YU (2016): "Does environmental regulation drive away inbound foreign direct investment? Evidence from a quasi-natural experiment in China," *Journal of development economics*, 123, 73–85.

CALEL, R. AND A. DECHEZLEPRÊTRE (2016): "Environmental policy and directed technological change: evidence from the European carbon market," *Review of economics and statistics*, 98, 173–191.

CHERNIWCHAN, J., B. R. COPELAND, AND M. S. TAYLOR (2017): "Trade and the environment: New methods, measurements, and results," *Annual Review of Economics*, 9, 59–85.

CHOI, J., J. HYUN, G. KIM, AND Z. PARK (2023): "The Cleanup of US Manufacturing through Pollution Offshoring," *Working Paper*.

CONTE, B., K. DESMET, D. K. NAGY, AND E. ROSSI-HANSBERG (2021): "Local sectoral specialization in a warming world," *Journal of Economic Geography*, 21, 493–530.

CONTE, M., P. COTTERLAZ, T. MAYER, ET AL. (2022): *The CEPII gravity database*, CEPII.

COPELAND, B. R., J. S. SHAPIRO, AND M. SCOTT TAYLOR (2022): "Chapter 2 - Globalization and the environment," in *Handbook of International Economics: International Trade, Volume 5*, ed. by G. Gopinath, E. Helpman, and K. Rogoff, Elsevier, vol. 5 of *Handbook of International Economics*, 61–146.

COPELAND, B. R. AND M. S. TAYLOR (2004): "Trade, growth, and the environment," *Journal of Economic literature*, 42, 7–71.

CORSATEA, T. D., S. LINDNER, I. ARTO, M. V. ROMÁN, J. M. RUEDA-CANTUCHE, A. VELÁZQUEZ AFONSO, A. F. AMORES, AND F. NEUWAHL (2000): "World input-output database environmental accounts," *Update*, 2016, 54.

CRUZ, J. L. (2021): "Global warming and labor market reallocation," .

CRUZ, J.-L. AND E. ROSSI-HANSBERG (2022): "Local Carbon Policy," .

——— (forthcoming): "The Economic Geography of Global Warming," *Review of Economic Studies*.

DEKLE, R., J. EATON, AND S. KORTUM (2008): "Global rebalancing with gravity: Measuring the burden of adjustment," *IMF Staff Papers*, 55, 511–540.

DESMET, K., R. E. KOPP, S. A. KULP, D. K. NAGY, M. OPPENHEIMER, E. ROSSI-HANSBERG, AND B. H. STRAUSS (2021): "Evaluating the Economic Cost of Coastal Flooding," *American Economic Journal: Macroeconomics*, 13, 444–86.

DESMET, K. AND E. ROSSI-HANSBERG (2015): "On the Spatial Economic Impact of Global Warming," *Journal of Urban Economics*, 88, 16 – 37.

——— (2023): "Climate Change Economics over Time and Space," .

DONALD, E. (2025): "Spillovers and the Direction of Innovation: An Application to the Clean Energy Transition," Tech. rep.

EATON, J. AND S. KORTUM (1999): "International Technology Diffusion: Theory and Measurement," *International Economic Review*, 40, 537–570.

——— (2001): "Technology, trade, and growth: A unified framework," *European economic review*, 45, 742–755.

——— (2002): "Technology, Geography, and Trade," *Econometrica*, 70, 1741–1779.

——— (forthcoming): "Technology and the Global Economy," *Annual Review of Economics*, 16.

ESKELAND, G. S. AND A. E. HARRISON (2003): "Moving to greener pastures? Multinationals and the pollution haven hypothesis," *Journal of development economics*, 70, 1–23.

FARROKHI, F. AND A. LASHKARIPOUR (2021): "Can trade policy mitigate climate change," .

FARROKHI, F. AND H. S. PELLEGRINA (forthcoming): "Trade, technology, and agricultural productivity," *Journal of Political Economy*.

FEENSTRA, R. C., R. INKLAAR, AND M. P. TIMMER (2015): "The next generation of the Penn World Table," *American economic review*, 105, 3150–82.

GARCIA-LEMBERGMAN, E., N. RAMONDO, A. RODRIGUEZ-CLARE, AND J. SHAPIRO (2023): "The Carbon Footprint of Multinational Production," *working in progress*.

GOLOSOV, M., J. HASSLER, P. KRUSELL, AND A. TSYVINSKI (2014): "Optimal taxes on fossil fuel in general equilibrium," *Econometrica*, 82, 41–88.

HANNA, R. (2010): "US environmental regulation and FDI: evidence from a panel of US-based multinational firms," *American Economic Journal: Applied Economics*, 2, 158–189.

HASSLER, J., P. KRUSELL, AND C. OLOVSSON (2021): "Directed technical change as a response to natural resource scarcity," *Journal of Political Economy*, 129, 3039–3072.

HASSLER, J., P. KRUSELL, AND A. A. SMITH JR (2016): "Environmental macroeconomics," in *Handbook of macroeconomics*, Elsevier, vol. 2, 1893–2008.

HÉMOUS, D. AND M. OLSEN (2021): "Directed technical change in labor and environmental economics," *Annual Review of Economics*, 13, 571–597.

KORTUM, S. S. (1997): "Research, patenting, and technological change," *Econometrica: Journal of the Econometric Society*, 1389–1419.

KORTUM, S. S. AND D. A. WEISBACH (2021): "Optimal unilateral carbon policy," .

KRUSELL, P. AND A. A. SMITH JR (2022): "Climate change around the world," Tech. rep., National Bureau of Economic Research.

LENZEN, M., D. MORAN, K. KANEMOTO, AND A. GESCHKE (2013): "Building Eora: a global multi-region input–output database at high country and sector resolution," *Economic Systems Research*, 25, 20–49.

LEVINSON, A. (2009): "Technology, international trade, and pollution from US manufacturing," *American economic review*, 99, 2177–2192.

LEVINSON, A. AND M. S. TAYLOR (2008): "Unmasking the pollution haven effect," *International economic review*, 49, 223–254.

LIND, N. AND N. RAMONDO (2022a): "Global innovation and knowledge diffusion," *NBER Working Paper*.

——— (2022b): "Global Knowledge and Trade Flows: Theory and Measurement," *NBER Working Paper*.

——— (2023): "Trade with correlation," *American Economic Review*, 113, 317–353.

LIU, E. AND S. MA (2021): "Innovation networks and r&d allocation," *NBER Working Paper*.

NORDHAUS, W. D. AND J. BOYER (2000): "Warming the world," .

PRATO, M. (2021): "The Global Race for Talent: Brain Drain, Knowledge Transfer, and Economic Growth," .

RAMONDO, N. AND A. RODRÍGUEZ-CLARE (2013): "Trade, multinational production, and the gains from openness," *Journal of Political Economy*, 121, 273–322.

ROMER, P. M. (1990): "Endogenous technological change," *Journal of political Economy*, 98, S71–S102.

SAMPSON, T. (2016): "Dynamic selection: an idea flows theory of entry, trade, and growth," *The Quarterly Journal of Economics*, 131, 315–380.

SANTACREU, A. M. (2015): "Innovation, diffusion, and trade: Theory and measurement," *Journal of Monetary Economics*, 75, 1–20.

SHAPIRO, J. S. (2016): "Trade costs, CO2, and the environment," *American Economic Journal: Economic Policy*, 8, 220–254.

Shapiro, J. S. and R. Walker (2018): "Why is pollution from US manufacturing declining? The roles of environmental regulation, productivity, and trade," *American Economic Review*, 108, 3814–54.

Sheet, E. F. (2013): "Social cost of carbon," *United States Environmental Protection Agency (EPA): Washington, DC, USA*.

Somale, M. (2021): "Comparative Advantage in Innovation and Production," *American Economic Journal: Macroeconomics*, 13, 357–396.

Weisbach, D. A., S. Kortum, M. Wang, and Y. Yao (2023): "Trade, leakage, and the design of a carbon tax," *Environmental and Energy Policy and the Economy*, 4, 43–90.

Appendix

A Data

In this section, I describe in detail the data I use in Section 2.

A.1 The Orbis Database

I use two Orbis databases in this paper. The first database is the BvD's Orbis Historical database, which is a point-in-time snapshot of the regular Orbis database. The second database is the Orbis Intellectual Property Database.

Orbis Historical data. It reports the usual firm-level information such as revenues, industrial classifications, corporation, and office addresses. More importantly, it provides us with information regarding the ownership structure of a firm. I obtain the multinational production status of each firm from this data. To do so, I need to know the global ownership structure of a firm. I proceed in the following steps. First, if a firm has a global ultimate owner with a minimum share of 50.01%, then this owner is automatically considered the global ultimate owner. Second, if a firm does not have a global ultimate owner with a minimum share of 50.01% but has a global ultimate owner with a minimum share of 25.01%, I define this owner as its global ultimate owner. Third, if a firm does not have either of the global ultimate owners mentioned above, I designate its largest shareholder as its global ultimate owner. The ownership data is available from 2007 to 2019.

I classify a firm as a multinational enterprise (MNE) if it is located in a different country from its global ultimate owner. Otherwise, I classify the firm as a non-MNE. I also calculate the share of foreign revenues out of its total revenues for each global ultimate owner.

To avoid double accounting for firms belonging to the same corporations, I use the consolidation status information and keep only unconsolidated firms.

To merge the firm-level data with sector-level emissions intensities, I leverage the 4-digit NACE Rev. 2 industrial classification codes reported in the data. For each firm, the core, primary and secondary industry codes are reported. When a firm have multiple primary and secondary industry codes, I use its core industry code.

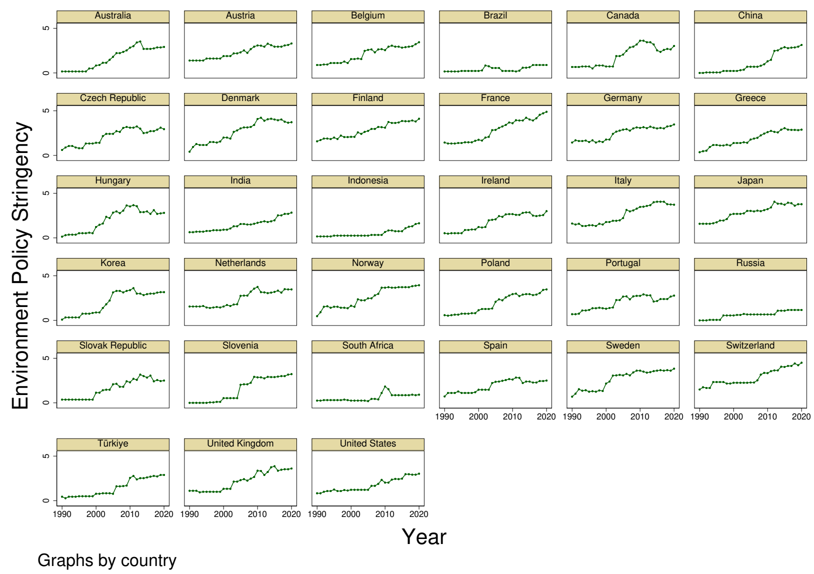
The Orbis Intellectual Property Data. The Orbis Intellectual Property Database merges firm-level data from Orbis Historical Database with global patent data from PATSTAT Global data maintained by the European Patent Office (EPO). This database allows me to obtain patent application and citation information at the firm level. To identify clean technology, I use the cooperative patent classification (CPC) code. I follow the EPO document to identify climate change mitigation technologies. I keep all publications that have been granted. I define clean technology as a patent family that contains at least one patent application falling into one of the categories "Y02A", "Y02B", "Y02C", "Y02D", "Y02E", "Y02P", "Y02T","Y02W", and "Y04S". Y02A denotes "technologies of applications for mitigation or adaptation against climate change." Y02B tags patents "relating to climate change mitigation technologies related to buildings, e.g. including housing and appliances or related end-user applications". Y02C covers all techniques for "capture, storage sequestration or disposal of greenhouse gases (GHG)". Y02D covers "climate change mitigation technologies in information and communication technologies (ICT), i.e., information and communication technologies aiming at the reduction of their won energy use." Y02E is assigned to patents devoted to "reduction of GHG emission, related to energy generation, transmission or distribution." Y02P includes "climate change mitigation technologies in the production or processing of goods." Y02T covers "climate change mitigation technologies related to transport". Y02W refers to "climate change mitigation technologies related to wastewater treatment or waste management." Y04S covers "Systems integrating technologies related to power network operation, communication or information technologies for improving the electrical power generation, transmission, distribution, management or usage, i.e. smart grids."

A.2 Environmental Policy Stringency

The OECD compile an index of environmental policy stringency (EPS), which is country-specific and internationally comparable. It encompasses both market-based and non-market-based policies. Market-based policies include CO_2 taxes, NO_x taxes, SO_x taxes, Diesel taxes, CO_2 certificate, and renewable energy certificates. Non-market-based policies include emission limit for CO_2, SO_x, particulate matter (PM), and Sulphur. These policies are converted into an index ranging from 0 (not stringent) to 6 (highest degree of stringency).

Figure 10 shows the evolution of EPS between 2000-2020 for countries covered in the OECD EPS data. Figure 10 shows the evolution of EPS between 2000-2020 for market-based and non-market-based policies.

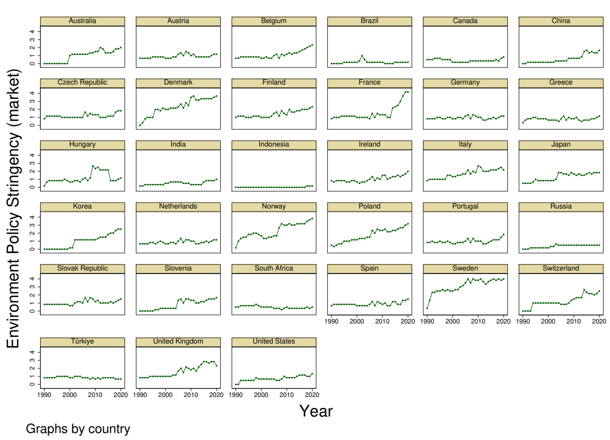
Figure 10: Environmental Policy Stringency Index: 2000-2020



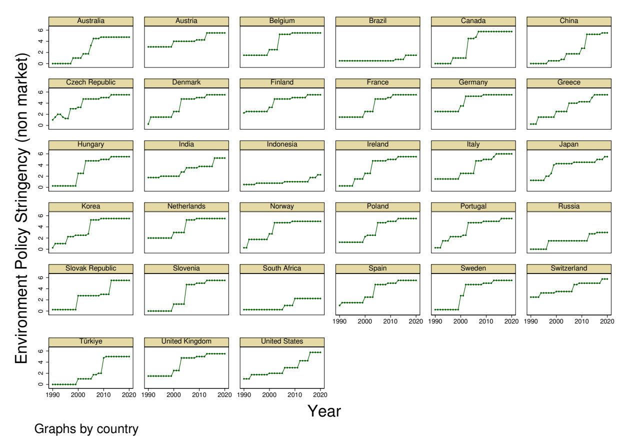
Notes: The figure shows the evolution of the Environmental Policy Stringency (EPS) between 2000-2020. The EPS index is compiled by the OECD and varies across countries and years. It encompasses both market-based and non-market-based policies. These policies are converted into an index ranging from 0 to 6.

Figure 11: Environmental Policy Stringency Index by Types: 2000-2020

(a) Market-based EPS: 2000-2020



(b) Non-market-based EPS: 2000-2020



Notes: The figure shows the evolution of the Environmental Policy Stringency (EPS) between 2000-2020. Figure 11a on the left shows the market-based EPS index. Figure 11b on the right shows the non-market-based EPS index. These market-based and non-market based policies are converted into an index ranging from 0 to 6.

B Model Derivations and Proofs

In this section, I show derivations of results in Section 3.

B.1 Proof of Proposition 1

I consider innovations for a particular technology type. Researchers draw ideas in the headquarter h about how to produce goods in country $j \in J$. The ideas arrive at a Poisson intensity $R_{h,t}$, where $R_{h,t}$ is the research effort of country h. An idea is the realization of two random variables. The first is the variety ω to which the idea is applied, where $\omega \in [0,1]$. The second is the efficiency $\boldsymbol{Q}(\omega) = (Q_1(\omega), Q_2(\omega), ..., Q_N(\omega))$ with which the variety ω can be produced. The random vector $\boldsymbol{Q}(\omega)$ follows a multivariate Pareto distribution

$$H(\boldsymbol{z}) \equiv \mathbb{P}(\boldsymbol{Q}(\omega) < \boldsymbol{z}) = 1 - \left(\sum_{j=1}^{N} (z_j)^{-\frac{\theta}{1-\eta}}\right)^{1-\eta},$$

and the support of the distribution is $z_j \geq N^{(1-\eta)/\theta}$ for all j. Define the total knowledge stock of country h by date t as

$$T_{h,t} = \int_0^t R_{h,s} ds.$$

The varieties fall on a unit interval, so the number of ideas, $\mathscr{N}$, for producing a variety by date t follows a Poisson distribution

$$\mathbb{P}(\mathscr{N} = n) = \frac{(T_{h,t})^n}{n!} \exp(-T_{h,t}).$$

Denote the efficient technology for producing ω by $\boldsymbol{Z}(\omega)$. The probability that the best idea has an efficiency below $\boldsymbol{z}$ when there have been n ideas is $[H(\boldsymbol{z})]^n$. Summing over all possible number of ideas, we have

$$
\begin{aligned}
\mathbb{P}_{h,t}\left[\mathbf{Z}(\omega)<\boldsymbol{z}\right] &= \sum_{n=0}^{\infty}\frac{\left(T_{h,t}\right)^{n}\mathrm{e}^{-T_{h,t}}}{n!}\left[H(\boldsymbol{z})\right]^{n} \\
&= \sum_{n=0}^{\infty}\frac{\left[T_{h,t}H(\boldsymbol{z})\right]^{n}\mathrm{e}^{-T_{h,t}H(\boldsymbol{z})}}{n!}\mathrm{e}^{-T_{h,t}\left[1-H(\boldsymbol{z})\right]} \\
&= \exp\left(-T_{h,t}\left[1-H(\boldsymbol{z})\right]\right) \\
&= \exp\left[-T_{h,t}\left(\sum_{j=1}^{N}\left(z_j\right)^{-\frac{\theta}{1-\eta}}\right)^{1-\eta}\right].
\end{aligned}
$$

The first equality holds by definition. The second equality is from multiplying and dividing the term $\mathrm{e}^{-T_{h,t}H(\boldsymbol{z})}$. The third equality uses the fact that $\sum_{n=0}^{\infty}\frac{\left[T_{h,t}H(\boldsymbol{z})\right]^{n}\mathrm{e}^{-T_{h,t}H(\boldsymbol{z})}}{n!}=1$, as it is the sum over a Poisson probability distribution with its mean equal to $T_{h,t}H(\boldsymbol{z})$. Substituting the expression for $H(\boldsymbol{z})$, we have the last equality. Technology frontier of country h thus follows

$$
F_{h,t}\left(\boldsymbol{z};T_{h,t}\right)=\exp\left[-T_{h,t}\left(\sum_{j=1}^{N}\left(z_j\right)^{-\frac{\theta}{1-\eta}}\right)^{1-\eta}\right],
$$

where $T_{h,t}=\int_0^t R_{h,s}ds$, $k=\{c,d\}$.

B.2 Proof of Lemma 1

For a given variety ω, the probability that the lowest cost is c in destination country i follows

$$
\begin{aligned}
G_i(c) &= 1-\Pr\left(\frac{\kappa_{ijh}^{k}}{Z^{k}(\omega)}\geq c_{ijh}^{k};k\in\mathcal{K},j,h\in\mathcal{N}\right)|_{c_{ijh}^{k}=c} \\
&= 1-\exp\left[-\sum_{k=1}^{K}\sum_{h=1}^{N}T_h^k\left(\sum_{j=1}^{N}\left(\frac{\kappa_{ijh}^{k}}{c_{ijh}^{k}}\right)^{-\frac{\theta}{1-\eta}}\right)^{1-\eta}\right]|_{c_{ijh}^{k}=c} \\
&= 1-\exp\left[-\sum_{k=1}^{K}\sum_{h=1}^{N}T_h^k\left(\sum_{j=1}^{N}\left(\kappa_{ijh}^{k}\right)^{-\frac{\theta}{1-\eta}}\right)^{1-\eta}c^{\theta}\right].
\end{aligned}
$$

The marginal probability density for j, h and k, which is, the probability that the technology k from country h delivers the lowest cost at c through production location j $\Pr\left(\min C_i(\omega)=C_{ijh}^{k}\cap C_i(\omega)=c\right)=$

$\frac{\partial G(c)}{\partial c^k_{ijh}}|_c$ follows

$$\frac{\partial G(c)}{\partial c^k_{ijh}}|_c = \theta \exp\left[-\sum_{k=1}^{K}\sum_{h=1}^{N} T_h^k \left(\sum_{j=1}^{N}\left(\frac{\kappa^k_{ijh}}{c^k_{ijh}}\right)^{-\frac{\theta}{1-\eta}}\right)^{1-\eta}\right] T_h^k \left(\sum_{j=1}^{N}\left(\frac{\kappa^k_{ijh}}{c^k_{ijh}}\right)^{-\frac{\theta}{1-\eta}}\right)^{-\eta} \left(\kappa^k_{ijh}\right)^{-\frac{\theta}{1-\eta}} \left(c^k_{ijh}\right)^{\frac{\theta}{1-\eta}-1}|_c$$
$$= \exp\left[-\sum_{k=1}^{K}\sum_{h=1}^{N} T_h^k \left(\sum_{j=1}^{N}\left(\kappa^k_{ijh}\right)^{-\frac{\theta}{1-\eta}}\right)^{1-\eta} c^\theta\right] T_h^k \left(\sum_{j=1}^{N}\left(\kappa^k_{ijh}\right)^{-\frac{\theta}{1-\eta}}\right)^{-\eta} \left(\kappa^k_{ijh}\right)^{-\frac{\theta}{1-\eta}} \theta c^{\theta-1}.$$

The probability the technology k from country h delivers the lowest cost through production location j is $\Pr\left(\min C_i(\omega) = C^k_{ijh}\right) = \pi^k_{ijh}$, and

$$\pi^k_{ijh} = \int_0^\infty \exp\left[-\sum_{k=1}^{K}\sum_{h=1}^{N} T_h^k \left(\sum_{j=1}^{N}\left(\kappa^k_{ijh}\right)^{-\frac{\theta}{1-\eta}}\right)^{1-\eta} c^\theta\right] T_h^k \left(\sum_{j=1}^{N}\left(\kappa^k_{ijh}\right)^{-\frac{\theta}{1-\eta}}\right)^{-\eta} \left(\kappa^k_{ijh}\right)^{-\frac{\theta}{1-\eta}} \theta c^{\theta-1} dc$$
$$= \frac{T_h^k \left(\sum_{j=1}^{N}\left(\kappa^k_{ijh}\right)^{-\frac{\theta}{1-\eta}}\right)^{-\eta}\left(\kappa^k_{ijh}\right)^{-\frac{\theta}{1-\eta}}}{\sum_{k=1}^{K}\sum_{h=1}^{N} T_h^k \left(\sum_{j=1}^{N}\left(\kappa^k_{ijh}\right)^{-\frac{\theta}{1-\eta}}\right)^{1-\eta}} = \frac{\left(\kappa^k_{ijh}\right)^{-\frac{\theta}{1-\eta}}}{\sum_{j=1}^{N}\left(\kappa^k_{ijh}\right)^{-\frac{\theta}{1-\eta}}} \cdot \frac{T_h^k \left(\sum_{j=1}^{N}\left(\kappa^k_{ijh}\right)^{-\frac{\theta}{1-\eta}}\right)^{1-\eta}}{\sum_{k=1}^{K}\sum_{h=1}^{N} T_h^k \left(\sum_{j=1}^{N}\left(\kappa^k_{ijh}\right)^{-\frac{\theta}{1-\eta}}\right)^{1-\eta}}.$$

B.3 Proof of Proposition 2

By definition, we have,

$$\frac{\sum_i X^c_{ijh}}{\sum_i X^d_{ijh}} = \frac{\sum_i \pi^c_{ijh} y_i}{\sum_i \pi^d_{ijh} y_i} = \frac{\sum_i \left(\kappa^c_{ijh}\right)^{-\frac{\theta}{1-\eta}} T_h^c \left(\sum_{j=1}^{N}\left(\kappa^c_{ijh}\right)^{-\frac{\theta}{1-\eta}}\right)^{-\eta} y_i}{\sum_i \left(\kappa^d_{ijh}\right)^{-\frac{\theta}{1-\eta}} T_h^d \left(\sum_{j=1}^{N}\left(\kappa^d_{ijh}\right)^{-\frac{\theta}{1-\eta}}\right)^{-\eta} y_i}$$
$$= \frac{T_h^c}{T_h^d}\left(1+\iota_j\right)^{\frac{\theta}{1-\eta}} \frac{\sum_i \left(\kappa^c_{ijh}\right)^{-\frac{\theta}{1-\eta}} \left(\sum_{j=1}^{N}\left(\kappa^c_{ijh}\right)^{-\frac{\theta}{1-\eta}}\right)^{-\eta} y_i}{\sum_i \left(\kappa^c_{ijh}\right)^{-\frac{\theta}{1-\eta}} \left(\sum_{j=1}^{N}\left(\kappa^d_{ijh}\right)^{-\frac{\theta}{1-\eta}}\right)^{-\eta} y_i}.$$

Taking first order derivative holding all prices and technology levels constant, we have

$$\frac{\partial \ln \sum_i X_{ijh}^c / \sum_i X_{ijh}^d}{\partial \left(1+\iota_j\right)} = \frac{\theta}{1-\eta}\frac{1}{1+\iota_j} - \frac{\theta\eta}{1-\eta}\frac{\sum_i \left(\kappa_{ijh}^c\right)^{-\frac{\theta}{1-\eta}} \left(\sum_{j=1}^N \left(\kappa_{ijh}^c\right)^{-\frac{\theta}{1-\eta}} \left(1+\iota_j\right)^{-\frac{\theta}{1-\eta}}\right)^{-\eta-1} y_i \left(\left(\kappa_{ijh}^c\right)^{-\frac{\theta}{1-\eta}} \left(1+\iota_j\right)^{-\frac{\theta}{1-\eta}-1}\right)}{\sum_i \left(\kappa_{ijh}^c\right)^{-\frac{\theta}{1-\eta}} \left(\sum_{j=1}^N \left(\kappa_{ijh}^c\right)^{-\frac{\theta}{1-\eta}} \left(1+\iota_j\right)^{-\frac{\theta}{1-\eta}}\right)^{-\eta} y_i}$$

$$= \frac{\theta}{1-\eta}\frac{1}{1+\iota_j} - \frac{\theta\eta}{1-\eta}\frac{\sum_i \left(\kappa_{ijh}^c\right)^{-\frac{\theta}{1-\eta}} \left(\sum_{j=1}^N \left(\kappa_{ijh}^c\right)^{-\frac{\theta}{1-\eta}} \left(1+\iota_{j,t}\right)^{-\frac{\theta}{1-\eta}}\right)^{-\eta} y_i \frac{\left(\kappa_{ijh}^d\right)^{-\frac{\theta}{1-\eta}}}{\sum_{j=1}^N \left(\kappa_{ijh}^d\right)^{-\frac{\theta}{1-\eta}}}}{\sum_i \left(\kappa_{ijh}^c\right)^{-\frac{\theta}{1-\eta}} \left(\sum_{j=1}^N \left(\kappa_{ijh}^c\right)^{-\frac{\theta}{1-\eta}} \left(1+\iota_{j,t}\right)^{-\frac{\theta}{1-\eta}}\right)^{-\eta} y_i}\frac{1}{1+\iota_j}$$

$$> \frac{1}{1+\iota_j}\left[\frac{\theta}{1-\eta} - \frac{\theta\eta}{1-\eta}\right] = \frac{\theta}{1+\iota_j} > 0.$$

Similarly, we have

$$\frac{\partial \ln \sum_i X_{ijh}^c / \sum_i X_{ijh}^d}{\partial \ln \left(1+\iota_h\right)} = -\frac{\theta\eta}{1-\eta}\frac{\sum_i \left(\kappa_{ijh}^c\right)^{-\frac{\theta}{1-\eta}} \left(\sum_{j=1}^N \left(\kappa_{ijh}^c\right)^{-\frac{\theta}{1-\eta}} \left(1+\iota_j\right)^{-\frac{\theta}{1-\eta}}\right)^{-\eta-1} y_i \left(\kappa_{ihh}^c\right)^{-\frac{\theta}{1-\eta}} \left(1+\iota_h\right)^{-\frac{\theta}{1-\eta}}}{\sum_i \left(\kappa_{ijh}^c\right)^{-\frac{\theta}{1-\eta}} \left(\sum_{j=1}^N \left(\kappa_{ijh}^c\right)^{-\frac{\theta}{1-\eta}} \left(1+\iota_j\right)^{-\frac{\theta}{1-\eta}}\right)^{-\eta} y_i}$$

$$= -\frac{\theta\eta}{1-\eta}\left(1+\iota_{h,t}\right)^{-\frac{\theta}{1-\eta}}\frac{\sum_i \left(\kappa_{ijh}^c\right)^{-\frac{\theta}{1-\eta}} \left(\sum_{j=1}^N \left(\kappa_{ijh}^c\right)^{-\frac{\theta}{1-\eta}} \left(1+\iota_{j,t}\right)^{-\frac{\theta}{1-\eta}}\right)^{-\eta} y_i \frac{\left(\kappa_{ihh}^c\right)^{-\frac{\theta}{1-\eta}}}{\left(\sum_{j=1}^N \left(\kappa_{ijh}^d\right)^{-\frac{\theta}{1-\eta}}\right)^{-\eta}}}{\sum_i \left(\kappa_{ijh}^c\right)^{-\frac{\theta}{1-\eta}} \left(\sum_{j=1}^N \left(\kappa_{ijh}^c\right)^{-\frac{\theta}{1-\eta}} \left(1+\iota_j\right)^{-\frac{\theta}{1-\eta}}\right)^{-\eta} y_i}$$

$$= -\frac{\theta\eta}{1-\eta}\frac{\sum_i \left(\kappa_{ijh}^c\right)^{-\frac{\theta}{1-\eta}} \left(\sum_{j=1}^N \left(\kappa_{ijh}^d\right)^{-\frac{\theta}{1-\eta}}\right)^{-\eta} y_i \frac{\left(\kappa_{ihh}^d\right)^{-\frac{\theta}{1-\eta}}}{\sum_{j=1}^N \left(\kappa_{ijh}^d\right)^{-\frac{\theta}{1-\eta}}}}{\sum_i \left(\kappa_{ijh}^c\right)^{-\frac{\theta}{1-\eta}} \left(\sum_{j=1}^N \left(\kappa_{ijh}^d\right)^{-\frac{\theta}{1-\eta}}\right)^{-\eta} y_i}$$

$$= -\frac{\theta\eta}{1-\eta}\left(1+\iota_h\right)^{-\frac{\theta}{1-\eta}}\bar{W}_{jh} < 0.$$

B.4 Proof of Proposition 3

Recall that

$$\frac{Y_{j,t}^c}{Y_{j,t}^d} = \frac{\sum_{i,h} X_{ijh,t}^c}{\sum_{i,h} X_{ijh,t}^d} = \left(\sum_h \varpi_{jh,t}^c T_{h,t}^c\right)\left(\sum_h \varpi_{jh,t}^d T_{h,t}^d\right)^{-1}\left(1+\iota_{j,t}\right)^{\frac{\theta}{1-\eta}}.$$

Taking logs, we have

$$\ln\frac{Y_{j,t}^c}{Y_{j,t}^d} = \ln\frac{\sum_{i,h} X_{ijh,t}^c}{\sum_{i,h} X_{ijh,t}^d} = \ln\left(\sum_h \varpi_{jh,t}^c T_{h,t}^c\right) - \ln\left(\sum_h \varpi_{jh,t}^d T_{h,t}^d\right) + \frac{\theta}{1-\eta}\ln\left(1+\iota_{j,t}\right).$$

Local effects of environmental policies can be expressed as

$$\frac{\partial \ln Y^c_{j,t}/Y^d_{j,t}}{\partial \ln\left(1+\iota_{j,t}\right)} = \frac{\theta}{1-\eta} - \frac{\eta\frac{\theta}{1-\eta}T^d_{h,t}\sum_i \frac{\left(\tau_{ij,t}w^P_{j,t}\gamma_{jh,t}/P_{i,t}\right)^{-\frac{\theta}{1-\eta}}y_{i,t}}{\left(\sum_{j=1}^N\left(\tau_{ij,t}w^P_{j,t}\gamma_{jh,t}\left(1+\iota_{j,t}\right)/P_{i,t}\right)^{-\frac{\theta}{1-\eta}}\right)^{\eta}}\frac{\left(\tau_{ij,t}w^P_{j,t}\gamma_{jh,t}\left(1+\iota_{j,t}\right)/P_{i,t}\right)^{-\frac{\theta}{1-\eta}}}{\sum_{j=1}^N\left(\tau_{ij,t}w^P_{j,t}\gamma_{jh,t}\left(1+\iota_{j,t}\right)/P_{i,t}\right)^{-\frac{\theta}{1-\eta}}}}{\sum_h \varpi^d_{jh,t}T^d_{h,t}}$$

$$\geq \frac{\theta}{1-\eta} - \eta\frac{\theta}{1-\eta}\frac{\varpi^d_{jh,t}T^d_{h,t}}{\sum_h \varpi^d_{jh,t}T^d_{h,t}} \geq \frac{\theta}{1-\eta} - \eta\frac{\theta}{1-\eta} = \theta > 0.$$

Cross-country effects of environmental policies can be expressed as

$$\frac{\partial \ln Y^c_{j,t}/Y^d_{j,t}}{\partial \ln\left(1+\iota_{h,t}\right)} = -\eta\frac{\theta}{1-\eta}\frac{T^d_{h,t}\sum_i \frac{\left(\tau_{ih,t}w^P_{h,t}\gamma_{hh,t}/P_{i,t}\right)^{-\frac{\theta}{1-\eta}}w^P_{i,t}L^P_{i,t}}{\left(\sum_{j=1}^N\left(\tau_{ij,t}w^P_{j,t}\gamma_{jh,t}\left(1+\iota_{j,t}\right)/P_{i,t}\right)^{-\frac{\theta}{1-\eta}}\right)^{\eta}}\frac{\left(\tau_{ih,t}w^P_{h,t}\gamma_{hh,t}(1+\iota_{h,t})/P_{i,t}\right)^{-\frac{\theta}{1-\eta}}}{\sum_{j=1}^N\left(\tau_{ij,t}w^P_{j,t}\gamma_{jh,t}\left(1+\iota_{j,t}\right)/P_{i,t}\right)^{-\frac{\theta}{1-\eta}}}}{\sum_h \varpi^d_{jh,t}T^d_{h,t}} \leq 0.$$

where

$$\varpi^d_{jh,t} = \sum_i\left(\tau_{ij,t}w^P_{j,t}\gamma_{jh,t}/P_{i,t}\right)^{-\frac{\theta}{1-\eta}}w^P_{i,t}L^P_{i,t}\left(\sum_{j=1}^N\left(\tau_{ij,t}w^P_{j,t}\gamma_{jh,t}\left(1+\iota_{j,t}\right)/P_{i,t}\right)^{-\frac{\theta}{1-\eta}}\right)^{-\eta}.$$

B.5 Proof of Lemma 2

The marginal probability density for the Pareto distribution $H(\boldsymbol{q})$ follows

$$\frac{\partial H(\boldsymbol{q})}{\partial Q_j} = \theta\left[\sum_{j=1}^N\left(q_j\right)^{-\frac{\theta}{1-\eta}}\right]^{-\eta}\left(q_j\right)^{-\frac{\theta}{1-\eta}}q_j^{-1}|_{\boldsymbol{q}} = \theta N^{-\eta}q^{-\theta-1} \equiv h_j\left(q\right).$$

The lowest cost available in country i that is higher than value c has the following productivity distribution

$$\Pr\left(\frac{\kappa^1_{i11}}{Z^1_{11}\left(\omega\right)} \geq c, ..., \frac{\kappa^k_{i11}}{Z^k_{11}\left(\omega\right)} \geq c, ...\frac{\kappa^1_{ijh}}{Z^1_{jh}\left(\omega\right)} \geq c, ..., \frac{\kappa^k_{ijh}}{Z^k_{jh}\left(\omega\right)} \geq c, ..., \frac{\kappa^1_{iNN}}{Z^1_{NN}\left(\omega\right)} \geq c, ..., \frac{\kappa^k_{iNN}}{Z^k_{NN}\left(\omega\right)} \geq c\right)$$
$$= \exp\left[-\sum_{k=1}^K\sum_{h=1}^N T^k_h\left(\sum_{j=1}^N\left(\kappa^k_{ijh}\right)^{-\frac{\theta}{1-\eta}}\right)^{1-\eta}c^{\theta}\right]$$

To get the probability that $\frac{\kappa_{ijh}^k}{q}$ is the lowest cost in market i, we evaluate the above productivity distribution at $c = \frac{\kappa_{ijh}^k}{q}$,

$$\exp\left[-\sum_{k=1}^{K}\sum_{h=1}^{N} T_h^k \left(\sum_{j=1}^{N}\left(\kappa_{ijh}^k\right)^{-\frac{\theta}{1-\eta}}\right)^{1-\eta}\left(\kappa_{ijh,t}^k\right)^{\theta} q^{-\theta}\right].$$

The probability that j provides the lowest cost for market i using ideas from h follows

$$\begin{aligned}
b_{ijh}^k &= \int_{\underline{q}}^{\infty} \exp\left[-\sum_{k=1}^{K}\sum_{h=1}^{N} T_h^k \left(\sum_{j=1}^{N}\left(\kappa_{ijh}^k\right)^{-\frac{\theta}{1-\eta}}\right)^{1-\eta}\left(\kappa_{ijh,t}^k\right)^{\theta} q^{-\theta}\right] h_j\left(q\right) dq \\
&= \underline{q}\frac{\pi_{i\cdot h}^k}{T_h^k}\left[\frac{\left(\kappa_{ijh,t}^k\right)^{-\frac{\theta}{1-\eta}}}{\sum_{j=1}^{N}\left(\kappa_{ijh}^k\right)^{-\frac{\theta}{1-\eta}}}\right]^{1-\eta} = \underline{q}\left(T_h^k\right)^{-1}\pi_{i\cdot h}^k\left(\psi_{ijh}^k\right)^{1-\eta}, \text{ where } \underline{q} \equiv N^{-\eta}.
\end{aligned}$$

The probability that an idea of technology k in headquarter h succeeds in market i follows

$$b_{i\cdot h}^k = \sum_{j=1}^{N} b_{ijh,t}^k = \underline{q}\left(T_h^k\right)^{-1}\pi_{i\cdot h}^k\left[\left(\psi_{ihh}^k\right)^{1-\eta} + \sum_{j\neq h}\left(\psi_{ijh}^k\right)^{1-\eta}\right].$$

B.6 Proof of Lemma 3

First, I derive the price index in destination country i. Denote the lowest cost of selling ω in i by $C_i^{(1)}\left(\omega\right)$, and denote the second lowest such cost by $C_i^{(2)}\left(\omega\right)$. Let $m_i\left(\omega\right) = \frac{C_i^{(2)}\left(\omega\right)}{C_i^{(1)}\left(\omega\right)}$, $m_i'\left(\omega\right) = \min\left\{m_i\left(\omega\right), \bar{m}\right\}$, and $\bar{m} = \frac{\sigma}{\sigma-1}$. We have

$$\begin{aligned}
P_i^{1-\sigma} &= \int_0^1 p_i\left(\omega\right)^{1-\sigma} d\omega = \int_0^1 \left[C_i^{(1)}\left(\omega\right) m_i'\left(\omega\right)\right]^{1-\sigma} d\omega \\
&= \int_0^1 \left[C_i^{(2)}\left(\omega\right)\frac{m_i'\left(\omega\right)}{m_i\left(\omega\right)}\right]^{1-\sigma} d\omega = \mathbb{E}\left[\left(C_i^{(2)}\right)^{1-\sigma}\right]\mathbb{E}\left[\left(\frac{m_i'\left(\omega\right)}{m_i\left(\omega\right)}\right)^{1-\sigma}\right].
\end{aligned}$$

Let $\Phi_i = \sum_{k=1}^{K}\sum_{h=1}^{N} T_h^k\left(\sum_{j=1}^{N}\left(\kappa_{ijh}^k\right)^{-\frac{\theta}{1-\eta}}\right)^{1-\eta}$ and $G_i^{(2)}\left(c\right)$ denote the distribution of the second lowest cost serving market i for a given variety,

$$G_i^{(2)}\left(c\right) = \Pr\left[C_i^{(2)} \leq c\right] = 1 - \exp\left(-\Phi_i c^{\theta}\right) - \Phi_i c^{\theta}\exp\left(-\Phi_i c^{\theta}\right).$$

Hence, $dG_i^{(2)}(c) = \theta\Phi_i c^{\theta-1}\exp\left(-\Phi_i c^\theta\right)\left(1+\Phi_i c^\theta\right) - \theta\Phi_i c^{\theta-1}\exp\left(-\Phi_i c^\theta\right) = \theta\left(\Phi_i\right)^2 c^{2\theta-1}\exp\left(-\Phi_i c^\theta\right)$. The expectation $\mathbb{E}\left[\left(C_i^{(2)}\right)^{1-\sigma}\right]$ follows

$$\begin{aligned}\mathbb{E}\left[\left(C_i^{(2)}\right)^{1-\sigma}\right] &= \int c^{1-\sigma} dG_i^{(2)}(c) \\ &= \theta\left(\Phi_i\right)^2 \int c^{1-\sigma}\cdot c^{2\theta-1}\exp\left(-\Phi_i c^\theta\right) dc = \Gamma\left(\frac{2\theta-\sigma+1}{\theta}\right)\Phi_i^{-\frac{1-\sigma}{\theta}}.\end{aligned}$$

For $\mathbb{E}\left[\left(\frac{m_i'(\omega)}{m_i(\omega)}\right)^{1-\sigma}\right]$, we have

$$\begin{aligned}\mathbb{E}\left[\left(\frac{m_i'(\omega)}{m_i(\omega)}\right)^{1-\sigma}\right] &= \int_1^{\bar{m}} d\mathbb{P}(m) + \int_{\bar{m}}^{\infty}\left(\frac{\bar{m}}{m}\right)^{1-\sigma} d\mathbb{P}(m) \\ &= 1 - \bar{m}^{-\theta} + \bar{m}^{1-\sigma}\int_{\bar{m}}^{\infty} m^{\sigma-1}\theta m^{-\theta-1} dm \\ &= 1 + \frac{\sigma-1}{\theta-(\sigma-1)}\bar{m}^{-\theta}.\end{aligned}$$

Therefore, the price index follows

$$P_i^{1-\sigma} = \Gamma\left(\frac{2\theta-(\sigma-1)}{\theta}\right)\Phi_i^{-\frac{1-\sigma}{\theta}}\left[1+\frac{\sigma-1}{\theta-(\sigma-1)}\bar{m}^{-\theta}\right].$$

Then I derive the profits. Denote by $X_i(\omega)$ the expenditure of country i on variety ω, so the cost to produce a variety ω sold in country i,

$$\begin{aligned}C_i(\omega) &= \frac{X_i(\omega)}{m_i'(\omega)} = \frac{1}{m_i'(\omega)}\left[\frac{p_i(\omega)}{P_i}\right]^{1-\sigma} X_i = \frac{P_i^{\sigma-1}}{m_i'(\omega)} X_i p(\omega)^{1-\sigma} = \frac{P_i^{\sigma-1}}{m_i'(\omega)} X_i\left[C_i^{(1)}(\omega)\, m_i'(\omega)\right]^{1-\sigma} \\ &= \frac{P_i^{\sigma-1}}{m_i'(\omega)} X_i\left[C_i^{(2)}(\omega)\frac{m_i'(\omega)}{m_i(\omega)}\right]^{1-\sigma} = X_i P_i^{\sigma-1}\left[C_i^{(2)}(\omega)\right]^{1-\sigma}\left[\frac{m_i'(\omega)^{-\sigma}}{m_i(\omega)^{1-\sigma}}\right].\end{aligned}$$

The cost index is

$$C_i = X_i P_i^{\sigma-1}\mathbb{E}\left[C_i^{(2)}(\omega)^{1-\sigma}\right]\mathbb{E}\left[\frac{m_i'(\omega)^{-\sigma}}{m_i(\omega)^{1-\sigma}}\right] = X_i P_i^{\sigma-1}\Gamma\left(\frac{2\theta-\sigma+1}{\theta}\right)\Phi_i^{-\frac{1-\sigma}{\theta}}\mathbb{E}\left[\frac{m_i'(\omega)^{-\sigma}}{m_i(\omega)^{1-\sigma}}\right].$$

For $\mathbb{E}\left[\frac{m_i'(\omega)^{-\sigma}}{m_i(\omega)^{1-\sigma}}\right]$, we have

$$\begin{aligned}\mathbb{E}\left[\frac{m_i'(\omega)^{-\sigma}}{m_i(\omega)^{1-\sigma}}\right] &= \int_1^{\bar{m}} \frac{1}{m} d\mathbb{P}(m) + \bar{m}^{-\sigma} \int_{\bar{m}}^{\infty} \left(\frac{1}{m}\right)^{1-\sigma} d\mathbb{P}(m) \\ &= \int_1^{\bar{m}} \frac{1}{m} \theta m^{-\theta-1} dm + \bar{m}^{-\sigma} \int_{\bar{m}}^{\infty} \left(\frac{1}{m}\right)^{1-\sigma} \theta m^{-\theta-1} dm \\ &= \frac{\theta}{\theta+1}\left[1 + \frac{\sigma-1}{\sigma-\theta-1}\bar{m}^{-\theta}\right].\end{aligned}$$

The last equality uses the fact that $\bar{m} = \frac{\sigma}{\sigma-1}$. Then $C_i = \frac{\theta}{1+\theta}X_i$. This implies that $\Pi_i = \frac{X_i}{1+\theta}$.

B.7 Total Emissions

The total emissions in production location j is define as

$$E_j = \sum_h y_{jh}^d \times e_h,$$

which is

$$\sum_h \int y_{jh}^d(\omega) \times e_h d\omega = \sum_h e_h \int_0^1 y_{jh}^d(\omega) d\omega.$$

Solving for $\int_0^1 y_{jh}^d(\omega)d\omega$, we have

$$\begin{aligned}\int_0^1 y_{jh}^d(\omega)\, d\omega &= \sum_i \int_0^1 \frac{\tau_{ij} X_{ijh}^d(\omega)}{p_i(\omega)} d\omega = \sum_i \int_0^1 \frac{\tau_{ij} X_i(\omega)\, \pi_{ijh}^d}{p_i(\omega)} d\omega \\ &= \sum_i \tau_{ij} \pi_{ijh}^d \int_0^1 \frac{1}{p_i(\omega)} \left(\frac{p_i(\omega)}{P_i}\right)^{1-\sigma} X_i d\omega \\ &= \sum_i \frac{\tau_{ij} X_i \pi_{ijh}^d}{P_i^{1-\sigma}} \int_0^1 p_i(\omega)^{-\sigma} d\omega\end{aligned}$$

The first equality is the definition of real output in j from firms originating in h using non-clean technology. The second equality is from the fact that $\pi_{ijh}^d = \pi_{ijh}^d(\omega)$ for all ω. The third equality is from the fact that $X_i(\omega) = \left(\frac{p_i(\omega)}{P_i}\right)^{1-\sigma} X_i$. Now we deal with $\int_0^1 p_i(\omega)^{-\sigma} d\omega$.

$$\int_0^1 p_i(\omega)^{-\sigma} d\omega = \int_0^1 \left[C_i^{(1)}(\omega)\, m_i'(\omega)\right]^{-\sigma} d\omega = \mathbb{E}\left[\left(C_i^{(2)}\right)^{1-\sigma}\right] \mathbb{E}\left[\left(\frac{m_i'(\omega)}{m_i(\omega)}\right)^{1-\sigma}\right]$$
$$= \int_0^1 \left[C_i^{(2)}(\omega) \frac{m_i'(\omega)}{m_i(\omega)}\right]^{-\sigma} d\omega = \mathbb{E}\left[\left(C_i^{(2)}\right)^{-\sigma}\right] \mathbb{E}\left[\left(\frac{m_i'(\omega)}{m_i(\omega)}\right)^{-\sigma}\right].$$

For $\mathbb{E}\left[\left(C_i^{(2)}\right)^{-\sigma}\right]$, we have

$$\mathbb{E}\left[\left(C_i^{(2)}\right)^{-\sigma}\right] = \int c^{-\sigma} dG_i^{(2)}(c) = \theta (\Phi_i)^2 \int c^{-\sigma} \cdot c^{2\theta-1} \exp\left(-\Phi_i c^{\theta}\right) dc$$
$$= \theta (\Phi_i)^2 \int c^{2\theta-\sigma-1} \exp\left(-\Phi_i c^{\theta}\right) dc = \Gamma\left(\frac{2\theta-\sigma}{\theta}\right) \Phi_i^{\frac{\sigma}{\theta}}.$$

In addition,

$$\mathbb{E}\left[\left(\frac{m_i'(\omega)}{m_i(\omega)}\right)^{-\sigma}\right] = \int_1^{\bar{m}} d\mathbb{P}(m) + \int_{\bar{m}}^{\infty} \left(\frac{\bar{m}}{m}\right)^{-\sigma} d\mathbb{P}(m)$$
$$= 1 - \bar{m}^{-\theta} + \bar{m}^{-\sigma} \int_{\bar{m}}^{\infty} m^{\sigma} \theta m^{-\theta-1} dm$$
$$= \frac{2\theta + 2 + \sigma - \theta\sigma}{(\sigma-\theta-2)(\sigma-1)}.$$

Therefore,
$$\int_0^1 p_i(\omega)^{-\sigma} d\omega = \frac{2\theta + 2 + \sigma - \theta\sigma}{(\sigma-\theta-2)(\sigma-1)} \Gamma\left(\frac{2\theta-\sigma}{\theta}\right) \Phi_i^{\frac{\sigma}{\theta}}.$$

We then have

$$\int_0^1 y_{jh}^d(\omega) d\omega = \sum_i \frac{\tau_{ij} X_i \pi_{ijh}^d}{P_i^{1-\sigma}} \int_0^1 p_i(\omega)^{-\sigma} d\omega$$
$$= \frac{2\theta + 2 + \sigma - \theta\sigma}{(\sigma-\theta-2)(\sigma-1)} \Gamma\left(\frac{2\theta-\sigma}{\theta}\right) \sum_i \frac{\tau_{ij} X_i \pi_{ijh}^d}{P_i^{1-\sigma}} \Phi_i^{\frac{\sigma}{\theta}}$$
$$= \tilde{\mathcal{B}} \sum_i \frac{\tau_{ij} X_i \pi_{ijh}^d}{P_i}.$$

Total emissions in country j is thus

$$E_j = \sum_h e_h \int_0^1 y_{jh}^d(\omega) d\omega = \tilde{\mathcal{B}} \sum_h e_h \sum_i \frac{\tau_{ij} X_i \pi_{ijh}^d}{P_i}.$$

B.8 Balanced Growth Path

I characterize the balanced growth path of the economy in this section. Let $\hat{x} \equiv \frac{dx/dt}{x}$ for any variable x. We know that the law of motion of the knowledge stock follows

$$dT_{h,t}^k = a_h^k \left(L_{h,t}^{k,R}\right)^\nu dt.$$

Let $T_{h,t}^k$ grow at n_T, we have

$$n_T = a_h^k \left(L_{h,t}^{k,R}\right)^\nu / T_{h,t}^k.$$

Taking logs on both sides of the above equation, we have

$$\nu \log\left(L_{h,t}^{k,R}\right) = \log\left(T_{h,t}^k\right),$$

which implies that $\nu \hat{L}_{h,t}^{k,R} = \hat{T}_{h,t}^k$. On the BGP the population growth n implies that $\hat{L}_{h,t}^{k,R} = \hat{L}_{h,t}^P = n$, so that $\hat{T}_{h,t}^k = \nu n$. Using equation (15), it is evident that on the BGP if the wage $w_{i,t}$ is a solution to (15) then so is $w_{i,s}$ for $\forall s > t$. We set $w_{i,t}$ to be constant. So $\Phi_{i,t}$ grows at the same rate as $T_{h,t}^k$, so $\hat{\Phi}_i = \nu n$, $\hat{P}_h = -n\nu/\theta$. That is, the real wage grows at $n\nu/\theta$. With a constant wage, $\hat{X}_{h,t} = n$. Knowing that $X_{h,t} = P_{h,t}C_{h,t}$, $C_{h,t} = n + n\nu/\theta$.

$$\frac{dC_{h,t}/C_{h,t}}{dt} = r_{h,t} - \frac{dP_{h,t}/P_{h,t}}{dt} - \rho$$

implies $r_h = n + n\nu/\theta - n\nu/\theta + \rho = n + \rho$. Next, I analyze the growth rate of $V_{h,t}^k$. Along the BGP,

$$V_{ijh,t}^k = \frac{\Pi_{i,t}}{\rho + n\nu - n\nu/\theta} = \frac{X_{i,t}}{(1+\theta)(\rho + n\nu - n\nu/\theta)} = \frac{\left(\bar{w}_{i,t} L_{i,t}^P + \bar{\Pi}_{i,t}^R\right)}{(1+\theta)(\rho + n\nu - n\nu/\theta)}.$$

Hence,

$$V_{h,t}^k = \sum_{i=1}^N \sum_{j=1}^N b_{ijh,t}^k V_{ijh,t}^k = \frac{\underline{q}\left(T_{h,t}^k\right)^{-1}}{\theta(\rho + n\nu - n\nu/\theta)} \sum_{i=1}^N \left(\bar{w}_{i,t} L_{i,t}^P + \bar{\Pi}_{i,t}^R\right) \sum_{j=1}^N \pi_{i\cdot h,t}^k \left(\psi_{ijh,t}^k\right)^{1-\eta}.$$

Knowing that the law of motion of the knowledge stock

$$\frac{\dot{T}^k_{h,t}}{T^k_{h,t}} = a^k_h \left(L^{k,R}_{h,t}\right)^{\nu} / T^k_{h,t} = n\nu,$$

we have that along the BGP

$$\left(L^{k,R}_{h,t}\right)^{\nu-1} = \left(n\nu T^k_{h,t}/a^k_{h,t}\right)^{\frac{\nu-1}{\nu}}. \tag{1}$$

We know that $L^{c,R}_{h,t}/L^{d,R}_{h,t} = \left(a^c_h V^c_{h,t}/a^d_h V^d_{h,t}\right)^{\frac{1}{1-\nu}}$. Hence, $\left(L^{c,R}_{h,t}/L^{d,R}_{h,t}\right)^{1-\nu} = \left(\frac{T^d_{h,t}/a^d_{h,t}}{T^c_{h,t}/a^c_{h,t}}\right)^{\frac{\nu-1}{\nu}}$. We have

$$L^{k,R}_{h,t}/L^R_{h,t} = \frac{\sum_{i=1}^N \sum_{j=1}^N \bar{\pi}^k_{i\cdot h} \left(\bar{\psi}^k_{ijh}\right)^{1-\eta} \left(\bar{w}_{i,t} L^P_{i,t} + \bar{\Pi}^R_{i,t}\right)}{\sum_{k=1}^K \sum_{i=1}^N \sum_{j=1}^N \bar{\pi}^k_{i\cdot h} \left(\bar{\psi}^k_{ijh}\right)^{1-\eta} \left(\bar{w}_{i,t} L^P_{i,t} + \bar{\Pi}^R_{i,t}\right)}.$$

B.9 Proof of Proposition 4

The allocation of researchers follows,

$$L^{c,R}_{h,t}/L^{d,R}_{h,t} = \frac{\sum_{i=1}^N \sum_{j=1}^N \bar{\pi}^c_{i\cdot h} \left(\bar{\psi}^c_{ijh}\right)^{1-\eta} \left(\bar{w}_{i,t} L^P_{i,t} + \bar{\Pi}^R_{i,t}\right)}{\sum_{i=1}^N \sum_{j=1}^N \bar{\pi}^d_{i\cdot h} \left(\bar{\psi}^d_{ijh}\right)^{1-\eta} \left(\bar{w}_{i,t} L^P_{i,t} + \bar{\Pi}^R_{i,t}\right)} = \frac{T^c_{h,t}}{T^d_{h,t}} \left(\sum_{j=1}^N \frac{\sum_{i=1}^N \left(\tau_{ij}\bar{w}_{j,t}\gamma_{jh}/P_{i,t}\right)^{-\theta} \left(\bar{w}_{i,t} L^P_{i,t} + \bar{\Pi}^R_{i,t}\right)}{\sum_{i=1}^N \sum_{j=1}^N \left(\tau_{ij}\bar{w}_{j,t}\gamma_{jh}/P_{i,t}\right)^{-\theta} \left(\bar{w}_{i,t} L^P_{i,t} + \bar{\Pi}^R_{i,t}\right)} \right. (1$$

Hence,

$$\ln R^c_{h,t}/R^d_{h,t} = \ln a^c_h/a^d_h + \nu \ln T^c_{h,t}/T^d_{h,t} - \nu \ln \sum_j \bar{\chi}_{jh,t} \left(1+\iota_j\right)^{-\theta},$$

where $\bar{\chi}_{jh,t} = \frac{\sum_{i=1}^N \left(\tau_{ij,t}\bar{w}_{j,t}\gamma_{jh,t}/P_{i,t}\right)^{-\theta} \left(\bar{w}_{i,t} L^P_{i,t} + \bar{\Pi}^R_{i,t}\right)}{\sum_{i=1}^N \sum_{j=1}^N \left(\tau_{ij,t}\bar{w}_{j,t}\gamma_{jh}/P_{i,t}\right)^{-\theta} \left(\bar{w}_{i,t} L^P_{i,t} + \bar{\Pi}^R_{i,t}\right)} \left(1+\iota_{j,t}\right)^{-\theta}$.

C Quantification

To perform counterfactual analysis, we define the following counterfactual equilibrium using the exact hat algebra.

Definition 3. (*Counterfactual equilibrium*) Let x' denote a variable x after shock, the change in the endogenous variables before and after a shock satisfies

$$
\begin{aligned}
\hat{\hat{\kappa}}_{ijh}^{k} &= \hat{\hat{w}}_{j}\hat{\gamma}_{jh}^{k}\hat{\tau}_{ij}^{k}\widehat{\left(1+\iota_{j}^{k}\right)}, \\
\hat{\hat{P}}_{i} &= \left[\sum_{k=1}^{K}\sum_{h=1}^{N}\bar{\pi}_{i\cdot h}^{k}\hat{\bar{T}}_{h}^{k}\left(\sum_{j=1}^{N}\bar{\psi}_{ijh}^{k}\left(\hat{\hat{\kappa}}_{ijh}^{k}\right)^{-\frac{\theta}{1-\eta}}\right)^{1-\eta}\right]^{-\frac{1}{\theta}}, \\
\left(\bar{\psi}_{ijh}^{k}\right)' &= \frac{\bar{\psi}_{ijh}^{k}\left(\hat{\hat{\kappa}}_{ijh}^{k}\right)^{-\frac{\theta}{1-\eta}}}{\sum_{j=1}^{N}\bar{\psi}_{ijh}^{k}\left(\hat{\hat{\kappa}}_{ijh}^{k}\right)^{-\frac{\theta}{1-\eta}}}, \\
\left(\bar{\pi}_{i\cdot h}^{k}\right)' &= \bar{\pi}_{i\cdot h}^{k}\hat{\bar{T}}_{h}^{k}\left(\sum_{j=1}^{N}\bar{\psi}_{ijh}^{k}\left(\hat{\hat{\kappa}}_{ijh}^{k}\right)^{-\frac{\theta}{1-\eta}}\right)^{1-\eta}\left(\hat{\hat{P}}_{i,t}\right)^{\theta}, \\
\hat{\bar{T}}_{h}^{k} &= \left(\hat{\bar{L}}_{h}^{k,R}\right)^{\nu}, \\
\hat{\bar{w}}_{j}\bar{w}_{j}\bar{L}_{j}^{P}+\Pi_{j}^{P}\hat{\Pi}_{j}^{P} &= \sum_{k=1}^{N}\sum_{i=1}^{N}\sum_{h=1}^{N}\left(\bar{\pi}_{ijh}^{k}\right)'\left(\hat{\bar{w}}_{i}\bar{w}_{i}\bar{L}_{i}^{P}+\Pi_{i}^{R}\hat{\Pi}_{i}^{R}\right), \\
\sum_{k=1}^{K}\bar{L}_{h}^{k,R} &= \sum_{k=1}^{K}\hat{\bar{L}}_{h}^{k,R}\bar{L}_{h}^{k,R}.
\end{aligned}
$$