CEWP 22-03

International Trade and the Environment: Three Remaining Empirical Challenges*

Jevan Cherniwchan
Carleton University

M.Scott Taylor
University of Calgary and NBER

May 9, 2022

CARLETON ECONOMICS WORKING PAPERS

Department of Economics
1125 Colonel By Drive
Ottawa, Ontario, Canada
K1S 5B6

International Trade and the Environment: Three Remaining Empirical Challenges*

Jevan Cherniwchan
Carleton University

M. Scott Taylor
University of Calgary
and NBER

May 9, 2022

Summary

Considerable progress that has been made in our understanding of the relationship between international trade and the environment since Gene Grossman and Alan Krueger published their now seminal working paper examining the potential environmental effects of the North American Free Trade Agreement in 1991. This review uses their original paper as a guide to highlight key developments along three main branches of research that all stem from their analysis: (i) the interaction between international trade, economic growth, and environmental outcomes, (ii) the role of environmental regulation in determining trade and investment flows, and (iii) estimating the relative magnitudes of the scale, composition, and technique effects induced by trade. It discusses key developments along each branch, with a particular focus on the empirical challenges that have impeded progress. It also highlights an area along each branch that is ripe for further study. These areas are termed the *Three Remaining Challenges*.

Keywords: International Trade and the Environment; Environmental Kuznets Curve; Pollution Haven Hypothesis
JEL Codes: F18; F64; O44; Q56

*This article is prepared for the Oxford Research Encyclopedia of Economics and Finance (https://oxfordre.com/economics/). We would like to thank Brian Copeland, Nouri Najjar and two anonymous referees for helpful comments. The usual disclaimer applies.

Introduction

In November 1991, Gene Grossman and Alan Krueger issued their now seminal working paper examining the potential environmental effects of the North American Free Trade Agreement (Grossman and Krueger, 1991). This work has been extremely influential since its initial circulation because it largely, and in our view correctly, identified the key mechanisms by which international trade can affect the environment while also providing a first attempt at estimating their relative strengths. As such, the paper launched several different branches of research - both theory and empirics - focused on both trade and growth's potential effects on the environment.

In this review, we revisit Grossman and Krueger's (henceforth GK) original paper focusing on its contribution to our understanding of how trade affects the environment. We use it as a springboard to help us explain subsequent work in the area, and in doing so, we pay particular attention to the key empirical challenges that researchers have faced along the way. We discuss and highlight these contributions by developing in our first section a simple model of trade, growth and the environment that allows us to replicate their key finding of an inverted-U shaped relationship between a nation's income per capita and measures of its environmental quality (the so-called *Environmental Kuznets Curve*, or EKC), and explain the importance of subsequent work estimating important forces identified in the model. The following two sections discuss our progress in estimating these key magnitudes, starting from the GK analysis and moving forward. In doing so we identify three areas that are ripe for further research, and present them as the *Three Remaining Challenges* for researchers in this area.

GK's method was deceptively simple and straightforward, with an argument that proceeds in four steps. First, they provided a description of how trade liberalization could alter emissions via the scale, composition, and technique effects.[1] Although their discussion was verbal, a little formalism proves useful in fixing ideas.[2] To that end, consider an economy comprised of N industries. Each industry i emits pollution Z_i as a by-product of producing value-added output at scale S_i measured at base period prices. By definition, total pollution emissions are the sum of industry level emissions. We can alternatively write industry level emissions as the product of an industry's emission intensity $E_i = Z_i/S_i$ and the scale of its production, S_i. Finally, since industries differ in their contribution to value-added, let $\Phi_i = S_i/S$ denote the share of industry i value-

[1] This verbal description was later explicitly formalized and given theoretical foundations in a series of papers by Copeland and Taylor (e.g. Copeland and Taylor (1994, 1995a,b)).

[2] Here we follow the original Copeland and Taylor (1994) decomposition and focus on emissions created in production, but a similar decomposition can be constructed to include pollution create by consumption.

added in total value added in the economy, S. Then we have, by this series of definitions, that economy wide emissions from production can be written as follows:

$$Z = \sum_{i=1}^{N} Z_i = \sum_{i=1}^{N} E_i S_i = S \sum_{i=1}^{N} E_i \Phi_i. \tag{1}$$

Total emissions are equal to the product of the scale of national output, multiplied by the value-added weighted average of industry emission intensities. Since Equation (1) is a definition, it has to hold everywhere and we can totally differentiate to identify its implications in terms of variation over time. This yields:

$$\hat{Z} = \hat{S} + \sum_{i=1}^{N} \Theta_i \hat{\Phi}_i + \sum_{i=1}^{N} \Theta_i \hat{E}_i \tag{2}$$

where a hat indicates percentage change, such that $\hat{Z} = dZ/Z$ etc., and $\Theta_i = Z_i/Z$ represents the share of industry i's emissions in total emissions.

The first term in Equation (2) is what GK called the *scale effect*, and it captures the change in aggregate pollution emissions due to changes in the level of economic activity. Holding constant the composition of national output and the dirtiness of the techniques of production, increases in the scale of output must raise emissions one-for-one. This is the simplest and most direct influence of economic activity on pollution.

The second term is the *composition effect*; it captures the change in pollution resulting from changes in the composition of production across industries. To understand whether composition effects raise or lower pollution recall that Φ_i is a share that sums to 1 across the N emitting industries. As a consequence, the changes in these shares must sum to zero implying that some elements in the composition term are negative while others are positive. As well, the emission intensities differ across industries and weight these changes. Therefore, the composition effect could either raise or lower emissions depending on the circumstance.[3]

Finally, the third term is the *technique effect*; it is the pollution share weighted change in each industry's emission intensities. If we believe that emission intensities can only fall (as is commonly thought), the technique effect captures the fall in pollution emissions coming from changes in the emission intensity of each industry.

With this conceptual division in hand, GK turned to provide preliminary empirical evidence as to the relative magnitudes of each effect. They started by examining the

[3]Subtract $\sum_{i=1}^{N} \hat{\Phi}_i \Phi = 0$ from the second composition term, and rearrange to find pollution rises if the value share of relatively dirty industries rises with the shock, or vice versa.

relative size of the scale and technique effects. This evidence was captured by the most influential finding presented by GK: that of the EKC.[4]

While the finding of an EKC stimulated a large subsequent literature, we begin our review by focusing on its implications for our understanding of relationship between international trade and the environment. As we discuss further in the first section of our review, entitled "Economic Growth and the Environment," the original goal of GK's analysis of the EKC was to provide evidence as to the potential effects of trade-induced growth in income levels on pollution; the inverted-U shape was interpreted as evidence that this growth would lead to a technique effect that would offset the scale effect from increased economic activity following NAFTA. The validity of this interpretation, however, depends on the mechanism or mechanisms that are driving the EKC.

Substantially more work is needed to understand the relationship between trade, economic growth and the environment. At present, we have some evidence that economic growth and environmental outcomes are related, and several different theories that imply different roles for trade in understanding the cross-country pattern of emissions. Yet, existing work has not successfully established the empirical relevance of these different explanations. Moreover, the empirical literature has largely ignored the possibility that trade liberalization itself affects growth rates, meaning that the environmental implications of trade will differ in the short and long run. Consequently, we set out as our first challenge the estimation of trade's long run consequences for the environment taking into account its potential impact on economic growth. We do so because even a cursory glance at the data shows a strong correlation amongst China's accession to the WTO, its growth rate, trade flows, and pollution levels. The case of China should, at the very least, make us suspect of no link whatsoever.

The next step in GK's analysis was to evaluate the likely composition effects induced by trade. The core of this evaluation was an examination of whether environmental regulations affect international trade and investment flows, and was stimulated by popular concerns over what is now termed the *pollution haven hypothesis*, that is the possibility that trade liberalization would lead to domestic industrial flight as "dirty" polluting industries relocated abroad to take advantage of weak or poorly enforced environmental policy. GK recognized that the concern over this type of trade-induced composition effect was predicated on an assumption that the costs associated with complying with environmental policy were significant enough to affect the pattern of trade and investment across countries by altering comparative advantage. Given the dearth of evidence

[4]This finding was reinforced in GK's subsequent work (Grossman and Krueger, 1995) that examined additional measures of environmental quality.

as to this *pollution haven effect* at the time, GK set about testing it by examining the effects of environmental regulations on trade between the US and Mexico in a cross-section of industries.[5]

The evidence of how environmental regulations affect trade and investment has changed substantially since GK's initial foray into the area. We highlight this advancement in the second section of our review, entitled "Environmental Regulations and the Composition Effect." This progress reflects changes in our understanding of the challenges researchers face in credibly identifying this causal relationship and in the methods for addressing them. While much of the initial work in the area, such of that of GK, found little-to-no evidence that environmental regulations affect trade or investment, it is now widely recognized how their estimates may have been affected by omitted variables bias, reverse causality, and measurement error. Subsequent work has adopted more transparent research designs that address these issues and identify the causal effects of environmental regulations on international trade and investment under a more plausible set of identification assumptions. The results of these studies provide compelling evidence that environmental regulations have a significant impact on international trade, altering both exports and imports, as well as affecting both inward and outward flows of FDI. In sum, there is now strong evidence of the pollution haven effect.

Despite this progress, work still remains. While the results of the recent literature provide overwhelming support for the *sign* of the pollution haven effect – that is, that environmental regulations negatively impact international trade and investment – they are less clear on its *magnitude*. This is an important limitation as understanding the size of the pollution haven effect is key to understanding the likelihood of the pollution haven hypothesis, and the potential role of environmental regulations in determining the compositional changes induced by trade. If the magnitude of the pollution haven effect is relatively small, particularly when compared to other factors that determine trade, then the concerns about the pollution haven hypothesis that have persisted since GK are likely unwarranted. As such, our second challenge is to obtain credible estimates of the magnitude of the pollution haven effect and understanding if it is larger or smaller than other determinants of comparative advantage.

The fourth, and final step, in GK's analysis was to provide an overall assessment of the potential effects of NAFTA by adding up their estimates of the scale, composition, and technique effects. Doing so pointed to a stunning conclusion: bilateral trade liberalization between the US and Mexico would not necessarily lead to a deterioration

[5]See Copeland and Taylor (2004) for a simple theory model linking the pollution haven effect to the pollution haven hypothesis.

of environmental quality in Mexico, as the large technique effects implied by the EKC would offset any potential deleterious effects brought about by changes in the scale or composition of production. Somewhat remarkably, this finding has been confirmed several times. As we discuss in the last section of our review, entitled "What is the Technique Effect?", researchers have consistently found that reductions in aggregate pollution emissions are primarily due to the technique effect regardless of whether they estimate the scale, composition and technique effects econometrically, or construct estimates using decomposition methods.

Although there appears to be an empirical consensus that technique effect are large, there is yet little agreement as to why. There has been, however, a general conclusion that this finding implies international trade has, at most, a minor role in determining aggregate pollution levels globally. Instead, researchers have argued that changes in aggregate pollution levels primarily reflect the effects of an ongoing tightening of environmental regulations, or simply the effects of continued changes in "technology", as these channels have been hypothesized as the primary drivers of the technique effect. However both of these hypotheses overlook a growing body of research that shows that the decreases in industry emission intensity that are driving the technique effect are at least partly explained by the direct effects of trade on firms. Moreover, there is some evidence that trade can also indirectly impact the technique effect by altering the technologies used by firms or the policies that are used to regulate environmental quality. Understanding the full extent of trade's impact on these margins is important given the technique effect's central role in driving aggregate environmental quality; if trade also significantly contributes to the technique effect, it is more likely that GK's initial finding that free trade will be good for the environment will hold generally. This yields our third challenge: obtaining credible estimates of the magnitude of the trade-induced technique effect.

This review differs from others in several significant ways. First, we employ a simple dynamic model to show that the forces behind the EKC result, are also key to determining the environmental impact of trade. In this way we bring some unity to the growth and trade contributions of the original GK work, but neglect a theoretical discussion of firm level adjustments to trade and their environmental role. For a review highlighting the potential role of firm level adjustment see Cherniwchan et al. (2017). Second, we identify key gaps in the empirical literature and present them as challenges for future research. Since the empirical literature directly on point is voluminous, our review is narrow in scope. For a more inclusive review of trade's effect on resource use, transport emissions and the role of trade and environmental policy, see Copeland et al. (2021).

Economic Growth and the Environment

The first, and perhaps most provocative, piece of GK's analysis was the empirical examination of the relationship between income per capita, and the level of environmental quality, as measured by pollution emissions, across countries. Since national income is tied to both the scale of national output, and incomes per capita which may be tied to greater demands for tighter environmental protection, this analysis was motivated as a means to provide evidence on the relative strength of scale vs technique effects that would result from NAFTA. This approach, while simple, was revolutionary.

GK investigated the relationship between per-capita income and pollution levels using a then novel data set that linked data from two main sources: income data from the Penn World Tables (Summers and Heston, 1991), and data on ambient pollution concentrations from the Global Environment Monitoring System (GEMS). The GEMS data were revolutionary as they included information on the emissions of three important air pollutants (sulphur dioxide, particulate matter, and smoke), collected via ground level air quality monitors using comparable methods in a set of major urban centres across the globe. The resulting dataset contained information on pollution in a large set of countries over the period 1977-1988.[6]

For each pollutant, GK estimated a series of specifications based on the following regression equation:

$$Z_{ijkt} = \beta_0 + \beta_1 Y_{kt} + \beta_2 Y_{kt}^2 + \beta_3 Y_{kt}^3 + \beta_4 C_{ijkt} + \beta_5 T_t + \mu_{ijkt} \tag{3}$$

where Z_{ijkt} is a measure of the ground level pollution concentrations at location i, in city j in country k, at time t,[7] Y_{kt} is per-capita income in country k at time t, T_t is a linear time trend, and C_{ijkt} is a vector of controls. These controls included an indicator reflecting the location's measurement device, the geographic location of a city (desert or coast location) and a measure of its population density, as well as a country specific control for political system (Communist or not) and a measure of a country's trade openness (exports plus imports divided by GDP). The remaining element, μ_{ijkt} is an error term which is then further divided into a common-to-city-component and an idiosyncratic

[6] The exact city and country composition varies over time owing to the changing city sample in the GEMS dataset.

[7] Since the GEMS system recorded ground level readings at high frequency (sometimes hourly), Z_{ijkt} was constructed as a summary statistic for one year's worth of observations. Given that pollution data tend to be highly skewed with annual measures often reflective of only a relatively few very bad pollution days, GK opted to measure measure Z_{ijkt} as either the median or 95th percentile reading, with the median being the focal choice.

site by time component. GK primarily treat the idiosyncratic component of the error as a mean zero random effect, although they also report a set of estimates where they treat site specific effects as constant over time.

While the precise estimates vary somewhat across specifications, GK's key empirical result was the finding that for two of the three pollutants (sulphur dioxide and particulates), pollution concentrations first rise and then fall with income per capita.[8] This inverted-U shaped relationship between pollution levels and income per capital would come to be known as the *Environmental Kuznets Curve* (or EKC).

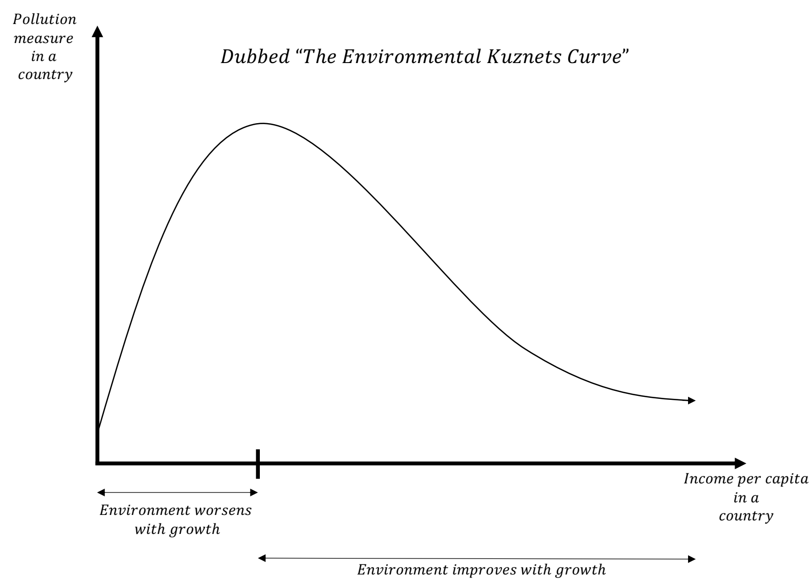


Figure 1: Environmental Kuznets Curve

An example of the EKC type relationship found by GK is illustrated if Figure 1, which plots the level of pollution associated with different levels of income per capita. As the figure shows, with an EKC, environmental quality initially deteriorates with growth but then begins to steadily improve once income per capita gets sufficiently high. GK interpreted this relationship as evidence that the scale effect dominates at lower levels of income per capita, but is eventually offset by the technique effect when incomes become sufficiently high. They suggested that this was due to changes in the demand for environmental quality as growth occurred; as incomes grow and pollution worsens, calls for tighter regulation create would eventually governments to enact more stringent policy, causing pollution to fall. Based on their analysis, GK suggested that the peak of the EKC would occur somewhere between $4,000 and $5,000 1985 USD per capita. Given that this corresponded to the then current income per capita of Mexico,

[8]For the last pollutant (smoke), pollution levels fall uniformly with increases in income per capita.

they concluded that income gains created by NAFTA could well lower - rather than raise - emissions. Not surprisingly, their methods, interpretations and conclusions stimulated a lively discussion.

Implications of the EKC Finding

One response to the EKC finding was a search for a fully articulated theory making sense of the empirical work. This search was led by theorists who, over the intervening years, provided several potential explanations for the hump shaped relationship. One explanation, most prominently tied to the work of Stokey (1998) is that income gains from growth drove pollution downward after an initial period where abatement was just not economic to undertake. If this was indeed the case, then a trade liberalization raising incomes may have a largely positive effect on subsequent emissions. Other explanations highlighted institutional rigidity and its impact on policy. Rigidities created barriers to the implementation of policy leading to threshold income levels beyond which policy was enacted, but below which policy was absent. If growth in income per capita created by a trade liberalization eased these constraints, the economy could shift regimes and become policy active as a consequence.[9] Others held that increasing returns to abatement were responsible for the eventual decline, or that growth itself naturally shifted from dirty capital deepening to cleaner human capital accumulation as development proceeded.[10] While there was soon no shortage of potential explanations for the EKC finding, very little effort was expended in evaluating these alternatives against other evidence.

Naturally, empirical researchers also responded. They worked very quickly to first replicate, extend, and finally, criticize the GK finding.[11] Although much of the empirical work following GK has been critical of their methods, we view most of this criticism as misplaced.[12] While it is clear that the cross-country evidence was sometimes in favor and other times not, the within country evidence for an EKC is quite strong.

For example, Brock and Taylor (2010) present raw data on 6 major pollutants in

[9] On thresholds in the EKC, see John and Pecchenino (1994) or Selden and Song (1994). Chapter 3 of Copeland and Taylor (2003) reviews several of the theories.

[10] For a discussion of the role of increasing returns to abatement, see Andreoni and Levinson (2001).

[11] This led to a voluminous empirical literature, which is too broad to discuss in much detail here. As such, we refer the interested reader to the discussion contained in the reviews by Copeland and Taylor (2004), Dinda (2004), Stern (2004) and Brock and Taylor (2005).

[12] Using methods very similar to GK, Harbaugh et al. (2002) found their results to be sensitive to data cleaning exercises by GEMS, as well as their choice of annual summary statistic, and econometric method. In contrast, Hilton and Levinson (1998) finds strong evidence for an EKC in an ingenious study of leaded gasoline levels.

the US from 1948-1998 showing that emissionsfirst rising and then fall over these fifty years. Given steady growth in income per capita over this period, this pattern produces a within-country demonstration of the EKC result. They also provide evidence of a similar pattern using a sample of 15 European countries over the period 1980 to 2000 showing most countries are now past their peak pollution levels.

More recent data paints a similar picture. This can be seen clearly in Table 1, which displays the year aggregate emissions of four common air pollutants (nitrogen oxides (NO_X), sulphur oxides (SO_X), volatile organic compounds (VOC) and carbon monoxide (CO)) peaked in thirty one countries over the period 1990-2019. These peaks are constructed using data from OECD (2022). As the table shows, in most cases, countries reached their peak emission level well before 2019 - despite their continuing economic growth - providing further within-country evidence of the EKC. Therefore, the question is not whether, for a broad range of countries, pollution problems first worsen and then improve with economic growth; the question is what this finding means for the relationship between international trade and the environment.

Its a matter of arithmetic that for pollution to fall, all theories need emissions per unit of national output to fall faster than the rate of growth of output. Not surprisingly then research soon focused on the conditions needed for abatement efforts to generate an EKC. To understand these conditions it is useful to employ the simple model of pollution abatement developed by Copeland and Taylor (1994). Versions of their approach to abatement appear in Stokey (1998), Forslid et al. (2018), Shapiro and Walker (2018), and many others.

To start, pollution emitted Z is the difference between that created and that abated, or in obvious notation, $Z = Z^C - Z^A$. Assume emissions created, Z^C, are proportional to economic activity F, and this factor of proportionality is $\Gamma > 0$. For the moment, treat Γ as a constant. Let output Y be produced with a constant returns to scale (CRS) neoclassical production function $Y = F(K, L)$ using capital and labor. Assume abatement is itself also a CRS economic activity that uses variable inputs to abate the pollution created by economic activity F. The variable factors are employed in the same proportion as in goods production, and hence abatement can be thought of an economic activity that uses the fraction θ of factors to abate the pollution created by F units of economic activity; i.e. $a(F, \theta F)$. If abatement at level a removes Γa units of pollution, then the amount of pollution abated is given by $Z^A = \Gamma a(F, \theta F)$. Since the fraction of inputs θ are taken from goods production, we must also have that final goods available for consumption

Table 1: Peak Emissions in 31 Countries: 1990-2019

Country	NO_X	SO_X	VOC	CO
Australia	2018	2002	1994	1997
Austria	2005	≤1990	≤1990	1991
Belgium	1992	1991	≤1990	≤1990
Canada	1999	≤1990	1994	≤1990
Czech Republic	≤1990	≤1990	≤1990	≤1990
Denmark	1991	1991	1991	1991
Estonia	≤1990	≤1990	≤1990	1996
Finland	≤1990	≤1990	≤1990	≤1990
France	1991	1991	1991	1991
Germany	≤1990	≤1990	≤1990	≤1990
Greece	2005	2005	2004	≤1990
Hungary	≤1990	1991	≤1990	≤1990
Iceland	1996	2012	1992	2014
Ireland	2000	1991	1991	≤1990
Italy	1992	≤1990	1992	1993
Japan	1997	≤1990	≤1990	≤1990
Latvia	≤1990	≤1990	≤1990	≤1990
Lithuania	1991	1991	1991	1991
Luxembourg	2005	1993	1991	≤1990
Netherlands	≤1990	≤1990	≤1990	≤1990
New Zealand	2019	2005	2019	2007
Norway	1998	≤1990	2001	≤1990
Poland	≤1990	≤1990	1996	1996
Portugal	1999	1992	1992	1992
Slovak Republic	≤1990	≤1990	≤1990	≤1990
Slovenia	1997	≤1990	1996	1996
Spain	1992	1991	1991	1991
Sweden	1991	≤1990	≤1990	1991
Switzerland	≤1990	≤1990	≤1990	≤1990
Turkey	1998	2012	2000	2000
United Kingdom	≤1990	≤1990	≤1990	1991
United States	≤1990	≤1990	≤1990	≤1990

Notes: Table reports the year of peak emissions for four key pollutants from 31 countries over the 1990-2019 period. Emissions peaks constructed using data on air emissions taken from OECD (2022). Dates listed under each heading indicate the year when emissions of the corresponding pollutant peaked, with ≤1990 implying that the peak occurred in 1990 or earlier.

(or investment) equal $Y = [1-\theta]F$. Together these assumptions imply:

$$Z = \Gamma[1 - a(1,\theta)]F \tag{4}$$

$$Z = \Gamma A(\theta) F \tag{5}$$

where $A(\theta) \equiv [1 - a(1,\theta)]$.[13]

Emissions per unit of economic activity Z/F are a function of only two things: $A(\theta)$ and Γ. $A(\theta)$ captures the impact of purposeful investments in abatement as measured by the scale of $\theta \leq 1$. Theories linking tighter environmental policies to income growth drive pollution downward via this channel. The second term Γ represents the state of existing abatement technology. And we might define anything that drives Γ down over time as *autonomous technological progress* in abatement. Naturally we think of these improvements coming from a myriad of sources: changes in facility design, improvement in energy efficiency, or changes in the pollution content of the energy mix created by fuel switching, etc. However, in our discussion of the sources of the technique effect, we will also show that within industry adjustments in firm numbers/types or firm processes can also masquerade as changes in Γ in aggregate data.[14]

The key problem with theories that make purposeful abatement carry all of the burden of explanation is an implication of constant returns. If a given level of pollution is to be abated, the application of increasingly large investments in abatement will necessarily run into diminishing returns: i.e. marginal abatement costs are upward sloping. But as growth proceeds and output expands, these investments must rise to keep pollution falling; i.e. θ has to approach 1 in the long run. But increasingly large and growing pollution abatement costs as a fraction of value-added output are not supported by the data, and hence income driven technique effects can only be part of the reason GK found the EKC.

The key problem with explanations relying on autonomous technological progress is that they are opaque. Opaque theories are difficult to falsify, and as a result not very productive. As a partial remedy, Brock and Taylor (2010) develop several other observable implications of their theory of why GK found the EKC. These implications follow from the Solow model framework they adopt which generates the EKC prediction sought after, but also ancillary predictions such as conditional convergence of emissions

[13] A useful special case satisfying these assumptions makes $A(\theta) = [1-\theta]^{1/\beta}$ with $\beta < 1$, which allows for a direct production function representation where pollution appears to be a factor of production; i.e. $Y = \gamma z^{\beta} F^{1-\beta}$ with $\gamma > 0$.

[14] For an exhaustive, and somewhat exhausting, listing of potential within industry and within firm adjustments that are hidden behind changes in emissions per unit of output see the complete scale, composition and technique decomposition presented in Cherniwchan et al. (2017)

per person across countries and falling emission intensities despite a constant intensity of abatement over time. While these ancillary predictions are strongly supported by the data, it is still unclear whether the balance of the evidence lies in favor of the income effects or technological progress explanation. A recurring theme of our review is that while income driven policy changes clearly matter - observed reductions in emissions per unit of output (even for unregulated or lightly regulated pollutants) are very large - too large - to be credibly ascribed to the work of income effects alone. Therefore, technological progress or technology effects must also matter. Another recurrent theme of our review is that these aggregate technology effects probably reflect a wide range of within industry and firm level adjustments which may be tied to trading opportunities.

Unfortunately, this uncertainty over the cause of the EKC matters to our understanding of how trade affects the environment. Both theories provide an explanation for the EKC finding, but they provide quite different predictions for the implications of freer trade on pollution levels. To simplify drastically - the income effects explanation supports an essentially static and benign view of trade's environmental effects; policy is very active, successful and would be responsive to any income gains that a trade liberalization may well create. The technology explanation is essentially dynamic, and this forces us to contend with both the level and growth effects of international trade. Some of these effects, such as capital accumulation, are likely not environmentally friendly.[15]

To understand how these different explanations work, interact and compete, it its necessary to first revisit the Green Solow model of Brock and Taylor (2010) as a step towards an extension rich enough to entertain both the income, and technology effect explanations.

Making Solow Green

Suppose the intensity of abatement was constant over time so that $A(\theta)$ was a constant. Emissions per unit of output could still fall but only through changes in Γ which we assume falls at a constant rate g^* over time. We combine these abatement assumptions with the usual assumptions of the Solow (1956) model: there is one aggregate good Y produced using capital and labor; this aggregate output can now be consumed, invested or used in abatement; capital accumulates over time via investment; while technological progress makes inputs to both production and abatement more efficient over time. Following Solow, write the model in intensive - per unit effective labor (BL) - form and

[15] Even in models with endogenous and directed technological progress, the changes brought about by a trade liberalization may or may not cause environmentally friendly changes in the direction and pace of innovative activity. See for example the discussions in Di Maria and Smulders (2005) and Hémous (2016)

assume intensive production can be written $f(k) = k^{\alpha}$, then we have:

$$y = [1-\theta]k^{\alpha} \tag{6}$$

$$\dot{k} = s[1-\theta]k^{\alpha} - [n+g+\delta]k \tag{7}$$

$$Z = \Gamma A(\theta) k^{\alpha} BL \tag{8}$$

$$\frac{\dot{B}}{B} = g \quad \text{and} \quad \frac{\dot{\Gamma}}{\Gamma} = -g^{*} \tag{9}$$

where $k = K/BL$, $y = Y/BL$, and $K(0), \Gamma(0), L(0)$ are given.

Equations (6) and (7) are just the Solow model taking into account that a fraction of output, θ, is allocated to abatement, and a fraction s of the output that remains is allocated to capital accumulation. Equation (8) just repeats our pollution creation equation, while equation (9) details the assumptions over technological progress in both goods and abatement production.

Given it is a Solow model, its easy to see that the steady state value of capital per effective worker, k^{*}, is found by setting $\dot{k} = 0$. In the long run, growth in per-capita income is just g while national income grows at $g+n$. To understand what the model predicts for the growth rate of pollution over time, totally differentiate equation (8) to obtain:

$$\frac{\dot{Z}}{Z} = \alpha\frac{\dot{k}}{k} + [g+n-g^{*}] \tag{10}$$

In the short run there is a positive scale effect on pollution driven by capital accumulation given by the first term in $\dot{k}$. This term is positive, but when the economy's capital deepening is complete and we approach k^{*}, then whether pollution rises or falls depends only on the last three terms in brackets. These are the long run determinants. Pollution grows along the balanced growth path by virtue of a scale effect raising output at rate $g+n$, and a technique (or technology effect) lowering pollution at rate g^{*}.

Although its not apparent from the analysis here, pollution policy - in terms of a hypothetical pollution tax levied on firms in an appropriately decentralized solution - would not be constant during the transition or beyond. In fact, holding θ constant in the face of cost reducing improvements in technology requires the shadow or opportunity cost of pollution in terms of goods production to rise at the rate g^{*}. Therefore, while technological progress drives pollution down and incomes per capita up, pollution policy is active during this process. Policy changes are just not the cause of pollution declines - technological progress is.[16]

[16]To show this adopt the special form for $A(\theta)$ given earlier. Firms minimizing costs choose pollution

With this in mind consider the EKC. Let the growth rate of emissions along the balanced growth path be denoted $g_Z \equiv g + n - g^*$. Then define sustainable growth as one with an improving environment, $g_Z \leq 0$, and rising consumption per person along an economy's balanced growth path $g > 0$; an unsustainable path would have $g_Z > 0$ and $g > 0$. Then to see how the technology theory generates an EKC consider Figure 2 which is commonly used to discuss the Solow model's convergence properties.

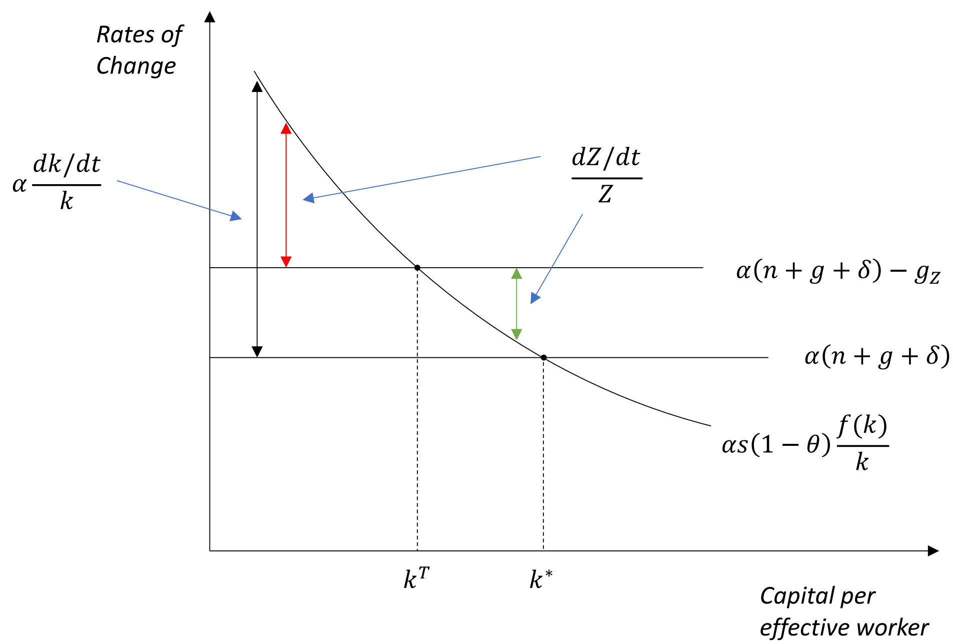


Figure 2: Convergence and the EKC

Figure 2 illustrates two relationships. First, to make the graphical analysis transparent we have multiplied equation (7) by α and then graphed its two component parts separately. One is the the downward sloping savings locus (proportional to the average product of capital $f(k)/k$); the other is the horizontal line reflecting depreciation. By construction these two components intersect at the steady state level of k^* as shown and have the property that the vertical distance between them is equal to α times the percentage rate of growth in capital (per effective labor unit). If an economy starts far to

so that $\tau z = \beta p$ where β is the share of pollution in the reformulated production function and p is the value of output. Using this result and the definition of $Y = [1-\theta]F$ shows $\Gamma[1-\theta]^{[1-\alpha]/[\alpha]} = \beta p/\tau$. With θ constant, and $p = 1$, τ - the opportunity cost of pollution - must rise at the rate g^*. Technological progress makes it easier to hit any z target, and in order for firms to be convinced to invest θ taxes must rise.

the left of k^* then growth in capital and output is at first very rapid but slows as the economy moves towards its steady state. This establishes the Solow prediction of rapid initial growth followed by slower growth along the BGP growth.

A second set of lines corresponds to the two components of the pollution accumulation equation (8). One is α times the savings locus, and the other is the horizontal depreciation line now shifted upwards by the positive quantity $-g_Z > 0$. When growth is sustainable, these two lines intersect at a k^T less than k^*. Also, by construction, the vertical distance between the lines is equal to the growth rate of emissions. As shown this growth is positive for $k < k^T$ and negative for $k > k^T$.

Putting all of these observations together we find an economy with little capital per person grows fast and has rising pollution levels. The rapid growth in output outstrips the decline in emissions being driven by technological progress. As growth slows, emissions growth also slows and eventually turns negative. Growth in national output continues but at a slowing pace as the economy approaches its balanced growth path. Therefore, in a sustainable economy with ongoing technological progress in goods and abatement production, the convergence properties of the Solow model alone, can very easily generate an EKC pattern in the data.[17]

The Green Solow model presents a strong case for why GK found their EKC results, but is silent on all questions about the role of trade. One way forward is to combine the static two sector model of Antweiler et al. (2001) with the simple dynamics of the Green Solow model much as done by Cherniwchan (2012). By doing so we can then sketch out both the short and long run implications of trade in a framework consistent with the original EKC finding and its two competing explanations.

A Model of Trade, Growth and the Environment

Any episode of trade liberalization has the potential to create a very complicated pollution response over time. In the short run we would expect changes in the scale and composition of output produced; the medium term may see environmental policy respond to income gains; and in the long run trade-inspired investments in capital and equipment could magnify or mollify these short and medium run responses. It should

[17]This very simple but powerful explanation for the EKC is also consistent with several other features of cross-country data. Brock and Taylor (2010), show that measures of pollution abatement costs are both small and without an upward trend over time consistent with the constant θ assumption; they show that emission intensities for several pollutants have been falling well before pollution peaked and indeed at almost constant rates; and they argue that small differences across countries in terms of initial conditions or production/abatement parameters generate heterogeneity in the EKC explaining the difficulty researchers have in estimating a common EKC.

be clear that any significant liberalization is likely to alter a country's environmental path for decades to come.

To discuss these responses we now employ a two sector extension of the Green Solow model very similar to that developed in Cherniwchan (2012). In this version of the GS model, one sector produces a final consumption good and the other an investment good. Both goods are produced using capital and effective labor in constant returns to scale production technologies. Final goods are assumed to be relatively capital intensive in production and create pollution as a joint product.[18] Pollution from final goods can be abated, and the technology for abatement uses factors in the same way as does the GS model. Since it is a Solow model the savings rate, s is fixed, and to match the GS model so too is θ. We choose the investment good as the numeraire, and start in a small open economy facing given relative prices.

With this formulation, at any point in time the model's supply side maximizes the value of gross national product, meaning it can be represented by a standard Gross National Product (GNP) function.[19] Not surprisingly, it proves useful to write the GNP function in intensive - per unit of effective labor - form and denote it as $G(p,k)$ where p is the relative price of the consumption good, and $k = K/BL$ as before. As $G(p,k)$ is also the income available to consumers, capital accumulation is governed by:

$$\dot{k} = sG(p,k) - [n+g+\delta]k \tag{11}$$

Given the properties of GNP functions, $G(p,k)$ is concave in k and hence dividing both sides of equation (11) by k generates an equation whose graphical representation is virtually identical to that shown in Figure 2. The monotonically declining average product of capital is now given by $G(p,k)/k$ rather than $f(k)/k$ as before and its intersection with the horizontal $n+g+\delta$ line defines the steady state k^*. For given prices, the evolution of k from any initial point is very similar to that in the GS model.

The key difference between the one and two sector models is the way pollution responds to capital accumulation during the transition to a steady state.[20] To investigate

[18]Assuming the final good is capital intensive ensures that the short and long run adjustments to the opening of trade are reinforcing - as they would be if we allowed for full optimization in a Ramsey style set-up. For a discussion of issues arising in two sector models with a fixed savings rate see Deardorff (2013).

[19]For further discussion see Cherniwchan (2012).

[20]We assume throughout that the steady state ensures the economy is diversified in production. Specialized steady states are indeed possible in two-sector versions of the Solow model (see Deardorff (1974)), but we abstract from them here as they do not add anything interesting to our discussion.

this difference note that pollution emissions are now given by:

$$Z = [A(\theta)\Gamma BL]\Phi(p,k)G(p,k)/p \tag{12}$$

where $\Phi(p,k) = pC/G(p,k)$ is the value share of the consumption good sector in GDP. Taking logs and differentiating with respect to time, we find:

$$\frac{\dot{Z}}{Z} = [\epsilon_{\Phi,k} + \epsilon_{G,k}]\frac{\dot{k}}{k} + [g + n - g^*] \tag{13}$$

As in GS, changes in pollution can be divided into forces that operate in the short and long run. To see this, assume the economy starts from an initial capital per effective worker below its steady state value of k^*. Then in the short run k rises and the economy grows rapidly due to capital deepening. This alters the economy's capital to effective labor ratio and at fixed prices shifts its production towards the capital intensive, polluting, consumption good. The magnitude of this response is proportional to the positive elasticity $\epsilon_{\Phi,k}$. As a result, the economy becomes dirtier along the transition path because of this dynamic composition effect. At the same time, holding the composition of output Φ constant, capital deepening also creates a scale effect proportional to the elasticity of GNP to k, $\epsilon_{G,k}$, which is positive.

As the economy approaches k^*, these short run impacts from capital deepening vanish. What remains is the difference between the long run scale effects given by $g + n$ and the long run technique effect given by g^*. For simplicity, we again take this difference to be negative implying that growth is sustainable in the long run. Not surprisingly, as long as the initial stock of capital per effective worker is far enough below k^*, this two sector model will also generate an EKC pattern in the data for much the same reasons as in GS. One important difference is that emissions per unit of output need not fall monotonically at a constant rate over time because short run composition effects can make aggregate output dirtier.[21]

The Movement to Free Trade

To understand the short and long run impacts of trade, consider a once for all movement to liberalized trade at different from autarky relative prices. For clarity, suppose we start with an economy already in its sustainable balanced growth path, with relative prices determined in autarky, and rates of technological progress in both goods and abatement given.

[21] See Cherniwchan (2012) for details.

Since autarky will be the counterfactual to the trading situation, we need to characterize the autarky BGP where (p, k^*) are constant in the long run. This solution is aided by some of the simple decision rules in the model. The demand for the investment good, in intensive form, is simply $I = sG(p, k)$ leaving the remainder of income spent on the consumption good implying $pC = (1 - s)G(p, k)$. Therefore, the relative demand (RD) for our two goods is, at any point in time,

$$[C/I]^D = [[1 - s]/sp] \tag{14}$$

As the relative price of consumption p rises, its relative demand falls.

The supply of the consumption good is equal to its value share in GNP divided by its price, or $\Phi G(p, k)/p$; the supply of the investment good $[1 - \Phi]G(p, k)$. This implies relative supply (RS) is, for given k,

$$[C/I]^S = \Phi/[1 - \Phi]p. \tag{15}$$

Equating demand and supply, shows that, for a given k, we must have that $s = [1 - \Phi]$. Since s is a constant, while the investment good's share of national income falls with p, this equilibrium is unique for given k. Moreover, since the investment good share of national income also falls with k, this equilibrium condition provides us with one relationship between $p = g(k)$ with $g' < 0$. When the economy turns to produce relatively more consumption because it has become more capital abundant, its relative price must fall.

To find our second relationship, set equation (11) to zero, to find steady state k and p are positively related. Intuitively, a higher relative price for the consumption good raises national income in terms of the (numeraire) investment good. This shifts the savings locus out leading to a higher k in steady state. A higher k would of course in turn raise the output of the consumption good. Therefore, the dynamics of capital accumulation imply a positive and magnifying relationship between p and k given by $p = f(k)$ with $f' > 0$. Putting these two relationships together generates the autarkic steady state pair (p^*, k^*).[22]

Short and Long Run Impacts

We employ three diagrams to identify the short run, medium run, and long run implications of a trade liberalization. These diagrams are shown in Figure 3. We start from an

[22]Deardorff (2013) contains a very clever graphical method for finding the autarky steady state.

existing (p^*, k^*) steady state for a country along its BGP. This existing situation is shown in the three panels of Figure 3 by the points labelled A.

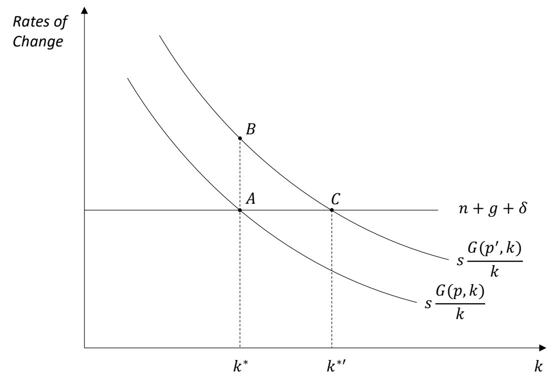
In panel (a), A identifies the initial steady state k^* conditional on p^*. In panel (b), A identifies the intersection of relative supply and demand which determines p^* when relative supply is evaluated at k^*. In panel (c), A identifies the log level of pollution in this economy at time time t_A. If nothing disturbs our existing autarkic steady state, then panels (a) and (b) repeat themselves ad infinitum. Panel (c) shows how autarky log pollution levels fall along the straight line through A and A' and beyond. This straight line through A and A' has a negative slope of g_Z because growth is sustainable. Since θ is constant throughout the opportunity cost of pollution to firms is rising at rate g^* which is faster than the long run rate of growth in per capita income g. The constellation of A points in the three panels represents the trajectory of our Autarky Counterfactual to any future trade opening.

Now consider an opening to trade at time t_B with given world prices $p' > p$. This implies our small open economy has a comparative advantage in the capital intensive, dirty, consumption good.[23] In the short run before capital can adjust, but relative prices, production and consumption can, the economy moves to the points labelled B in the three panels. In panel (a), the value of national income rises shifting the savings locus upwards. This creates a positive, but incipient, change in the economy's rate of capital accumulation. In panel (b), this increase in national income comes from a reallocation of production across sectors represented by the movement up its relative supply curve to B. The gap created between RS and RD at price p' is met by imports of the investment good and balanced by the value of exports of the consumption good. Since production of the consumption good is dirty, this composition change would drive the level of Z upwards. Absent any policy response, pollution may rise to some level above B in the third panel.[24] However, the movement to trade raised the value of GNP and real income. It is therefore possible that rising demands for better environmental protection may create a technique effect driving the intensity of abatement higher than our assumed constant θ. For example, if pollution taxes rose sufficiently, firms' would abate more intensively and the level of $A(\theta)$ in equation (12) falls. The pollution consequences of this change in production is captured in panel (c) by the now smaller movement

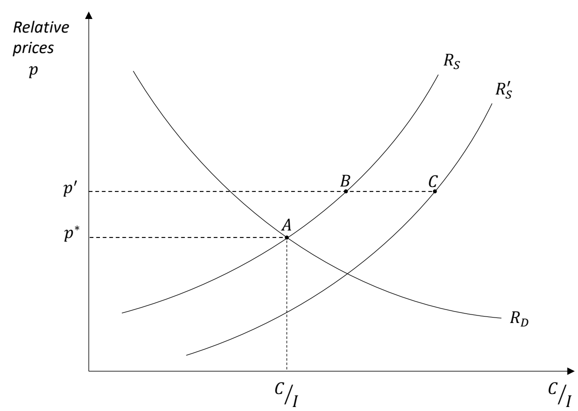
[23] This is of course not a necessity. The stringency of policy indexed by θ plus the remaining primitives set k^*. Together they determine the position of the RS in the second panel and hence comparative advantage.

[24] See Copeland and Taylor (2003) for a decomposition of trade inspired changes in pollution into scale and composition effects.

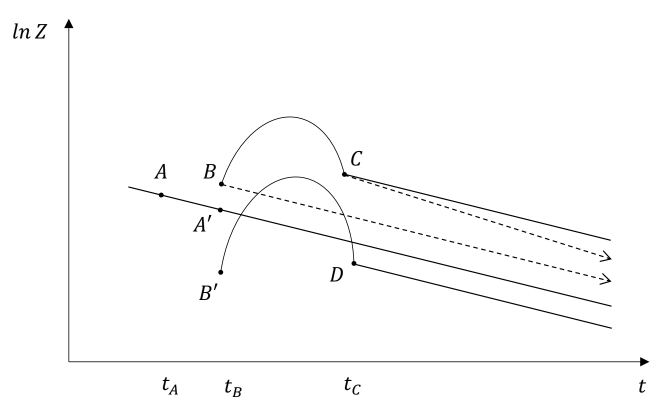
Figure 3: The Impacts of Trade Liberalization



(a) Saving and Depreciation



(b) Relative Supply and Relative Demand



(c) The Dynamics of Pollution

from A' to B at time t_B.[25] If the technique effect was extremely strong, it is possible for pollution to fall despite the economy's short run increase in dirty good production. Such a possibility is shown by the hypothetical downward movement to B'.[26]

Therefore in the short run, the economy moves to specialize in the relatively dirty consumption good. The size of scale, composition and technique effects will determine the short run response of pollution to somewhere between B and B'. If the movement is to B, then the impact of tightening regulation created by trade-inspired income gains is relatively ineffective in offsetting the scale and composition effects of trade. It could be ineffective because the impact of tighter regulation on competitiveness is small; i.e. it is because the *Pollution Haven Effect* created by the regulatory change is weak. Or it could be that the income gains from trade had little impact on policy at all; i.e. technique effects from tighter policy were small because increases in real income are not strong drivers of policy. Conversely, a movement to B' implies just the opposite - either production patterns change tremendously due to tighter policy or policy itself responds very vigorously turning even the relatively dirty production of C clean.

It is of course these contemporaneous changes in incomes, production and pollution that empirical studies of scale, composition, and technique effects try to measure. For example, one interpretation of the work of Antweiler et al. (2001) is that they are trying to measure how ongoing reductions in transport costs, which raise the relative price of a country's export goods almost continuously over time, affect domestic pollution levels. These small exogenous "price shocks" are then linked to the short run adjustments just outlined, and this over time variation (plus within and across country variation) is used to estimate scale, composition and technique effects. But studies such as these take production possibilities as given and unaffected by a trade liberalization, and hence they implicitly ignore the dynamic adjustments we will soon discuss. Implicitly, they assume trade has the conventional impacts created by moving around given production possibilities, and any overall trend in emissions per unit output simply continues during and after the trade liberalization. Under this assumption, the future trajectory of the economy follows either the dashed line through B, or the not drawn dashed line through B', parallel to the Autarky counterfactual path. Trade has level, but no dynamic effects

[25]Tighter policy would also shift the economy's relative supply curve inward (not shown).

[26]This seems unlikely given what we know from theory. An increase in income generated by neutral technological progress will lead to more pollution in our economy unless the elasticity of marginal utility exceeds one in magnitude. A trade liberalization such as the one described in which world prices are $p' > p$ at time t_B creates much stronger substitution effects towards dirty good production and hence would require an even stronger response on the demand side for pollution to fall as shown by the movement to B'.

on pollution.[27].

In this case, the key to knowing the environmental impact of trade is knowing the relative strength of composition and technique effects. Not surprisingly, a key component of GK's work was an evaluation of composition effects and the impact that tighter regulation had on trade flows. This work stimulated a large, vibrant and still growing literature trying to estimate what we would call the pollution haven effect. We turn to a discussion of that work in the next section, focusing on both the methods of identification and the magnitude of the impacts estimated.

As to the technique effect? The EKC analysis of GK is surely consistent with a strong technique effect, although its now clear that other factors could be responsible for the decline in pollution levels. Accordingly in our last section we turn to evaluate the strength of measured technique effects. To preempt slightly, we find good reason to question whether the magnitude of estimated technique effects reflect policy changes alone; instead other within industry and within firm adjustments are likely at play. In our aggregate model here, these very micro-adjustments that lower emissions intensity are obviously absent but would be captured by our g^*, lending support to the technology explanation for the EKC. If this is correct then the very dynamic behind the EKC tells us the impact of trade on the environment must have both short and long run effects. Moreover, if the path of emissions intensities is not driven by policy changes, the way is then open for other trade-inspired mechanisms such as cleaner fuels, technology adoption, and improved facility design to matter much more than previously thought.[28]

Trade may have at least two other dynamic implications missed by the typical discussion. The first is the result of trade-induced capital accumulation that reinforces the short run composition effect on pollution. Return to panel (a) to focus on capital accumulation. We know that in the long run, the economy moves from B to C. During the transition period, the rate of growth of output will rise above its BGP rate of $g + n$ because capital per effective worker is rising. This higher transitional growth will drive positive scale and composition effects and may, for a time, lead to rising emissions as shown in the movement from B to C in the third panel. Higher than BGP growth rates diminish over time as the economy approaches its new steady state and pollution eventually declines at a constant rate - albeit on a path of heightened pollution. Because the

[27] Antweiler et al. (2001) include either unrestricted time dummies or estimate a linear time trend which they find to be highly significant showing a decline in pollution levels of 3-5% per year independent of the scale, composition and technique effects. In theory this captures the movement along the line through B in panel (c).

[28] Some of these changes could arise from trading opportunities altering the incentives for innovation and imitation thereby altering both the pace and direction of technological progress. For original theory developing these ideas see Acemoglu et al. (2014) and Di Maria and Smulders (2005)

economy becomes more capital intensive in the long run, the relative supply curve in panel (b) shifts to RS' which increases specialization of the economy in the relatively dirty consumption good. Therefore, at C the economy would have a dirtier composition of output than at B. The same is true for an alternative trajectory from B' to D.

Researchers have long known about these consequences, and some attempts to quantify their importance have been made. Despite the obvious difficulties in establishing a causal link between trade and capital accumulation, both Grossman and Krueger (1991) and Antweiler et al. (2001) present evidence on the magnitude of hypothetical responses of capital accumulation to trade liberalization. For example, GK report results from the computable general equilibrium model of Brown et al. (1991) showing a hypothetical 10% increase in Mexico's capital stock, together with NAFTA liberalization, would raise manufacturing output, and presumably pollution, by a similar amount. Following on this idea, Antweiler et al. (2001) use their estimates of scale, composition and technique elasticities to ask what effect a 1% increase in a nation's capital stock has on pollution emissions. They find a 1% rise in the capital stock would raise pollution by perhaps .7%. Both exercises, while not establishing a causal link between accumulation and trade liberalization demonstrate the potential importance of understanding the capital accumulation/trade liberalization link.

A second dynamic effect comes when trade accelerates those micro-processes already present in g^*. The availability of new abatement technologies, cleaner fuels, or within industry rationalization driving less productive dirty firms out of business may drive down the pollution intensity of output faster than previously. This possibility is shown in panel (c) by the dashed line starting at C with a steeper slope than g_Z.

The First Challenge: Estimating the Long-Run Impacts of Trade

There has been significant progress in our understanding of the links between trade, growth and environmental outcomes. GK's initial work on the EKC led to an explosion of research both empirical and theoretical that would not have existed otherwise. The end result is a much better understanding of the growth and environment nexus, and newly raised questions about the appropriate way to think about trade's environmental impact. The convention is to divide trade and growth's impact along taxonomic lines: the impact of trade comes from changes in prices or market access but holds production possibilities constant. Investigations of growth's impact holds the trade regime (often autarky) constant. While this division may serve a useful pedagogical purpose in teaching, it is a misleading view of the real world. Openness to international markets has been key to the development of many, if not most, nations worldwide. The opening of China is just the latest and greatest example of a path followed by many countries over many centuries. Therefore, openness to international markets will have both short and long run environmental consequences. And we have spelt out some of these likely consequences in detail because they followed logically from a leading explanation for why GK found the EKC. The remaining challenge is for researchers to take these possibilities seriously and devise empirical strategies that allow them to credibly identify, and estimate, both the short and long run, impacts of trade liberalization on the environment.

Environmental Regulations and the Composition Effect

The second major component of GK's analysis stems from their examination of the composition effect. While they also examined the potential roles of "traditional" determinants of comparative advantage (such as capital and labor stocks) in shifting production across countries[29], the core of this analysis was an examination of the effects of environmental regulations on international trade and investment flows between the US and Mexico. This investigation was motivated by popular concerns over what has come to be termed the *pollution haven hypothesis*: namely, the hypothesis that trade liberalization would lead to a change in the composition of domestic production, as polluting industries relocated abroad to take advantage of differences in the stringency and/or the enforcement of environmental policy.[30] Given the dearth of empirical evidence examining the effects of environmental regulation on trade at the time, it was unclear if the costs associated with complying with environmental policy were significant enough to impact international trade and investment, making it difficult to determine if this type of trade-induced composition effect was possible ex ante. As GK themselves noted:

> *...the question remains open as to whether the overall sectoral pattern of US economic relations with Mexico has been meaningfully affected by the higher costs of pollution abatement in the United States. If the pattern of specialization has been so influenced, then the composition effect of a further liberalization of trade and investment may be damaging to the environment. (Grossman and Krueger, 1991, pg. 22)*

Put differently, GK were unsure as to the existence of what is now referred to as the *pollution haven effect*, that is, the hypothesis that environmental regulation reduces ex-

[29] See section 3 of Grossman and Krueger (1991).

[30] These concerns are epitomized in the context of NAFTA by Ross Perot's famous claim that the agreement would lead to a "giant sucking sound going south" as US businesses moved production to Mexico to take advantage of regulatory differences that impacted the cost of production.

ports (or increases imports) of polluting goods. However, they keenly recognized that the pollution haven effect is a necessary condition for the pollution haven hypothesis to hold, and set about testing it in the context of trade between the US and Mexico.

To do so, GK collected data on the trade and investment flows that occurred between the US and Mexico in 1987, data from the US Census's 1988 survey of pollution abatement costs, as well as data on other industry characteristics. They then estimated three sets of specifications based on the following regression equation:

$$m_i = \alpha + \beta\tau_i + X_i'\gamma + \epsilon_i \tag{16}$$

where m_i denotes a measure of trade or investment in industry i, τ_i is a measure of environmental policy, X_i are other determinants of trade and ϵ_i is an idiosyncratic error term. In the first set of specifications, m_i is measured as the ratio US imports from Mexico to total US shipments in 1987 at the 3-digit SIC level. In the second set of specifications, m_i is again measured as the ratio of US imports from Mexico to total US shipments in 1987 by 3-digit SIC industry, but in this case imports are restricted to those entering the US under offshore assembly provisions (specifically, those that fall under import category 807.00 under the TSUSA tariff schedule). Finally, in the third set of specifications, m_i is measured as the ratio of value added at Mexican maquiladoras to US value added in 1987 by 2-digit industry. In all specifications, τ_i is measured as the ratio of pollution abatement costs to industry value added (our θ in the earlier theory), and X_i is a vector of controls that includes industry factor shares to account for traditional determinants of comparative advantage, the effective US import tariff rate by industry, as well as the industry injury rate to account for differences in labor regulations.

Based on the results of this exercise, GK concluded that NAFTA was unlikely to result in a trade-induced composition effect that was consistent with the predictions of the pollution haven hypothesis. They arrived at this conclusion because they found no evidence that would support the existence of a pollution haven effect between the US and Mexico; their estimates of $\hat{\beta}$ were either qualitatively small and statistically insignificant (such as when m_i was measured as the ratio US imports from Mexico to total US shipments) or the wrong sign given they hypothesized relationship between environmental regulations and trade and investment (as is the case when m_i was measured as the ratio of US category 807.00 imports from Mexico to total US shipments or the ratio of value added at Mexican maquiladoras to US value added). While GK suggested that this conclusion was not entirely unsurprising in light of the fact that, in their data, pollution abatement costs amounted to only 1.38 percent of value added across all manufacturing industries, it was incredibly provocative as it meant that popular concerns that NAFTA would enable US firms to circumvent domestic environmental policy by relocating to Mexico were likely overstated.

Three Reasons for Caution

While GK's results were surprising, there are at least three reasons to be cautious in interpreting them as evidence that the imposition of new or stricter environmental regulations do not *cause* changes in international trade and investment flows.

To see why, it is convenient to imagine the randomized experiment that GK may have instead run to understand the effects of US environmental regulations on trade with Mexico if the US government had afforded them the opportunity to do so. In such an idealized setting, GK would have been able to choose a particular form of environmental policy with a specified stringency, such as a uniform pollution tax or a technology standard, and randomly assign the firms in some industries to comply with the standard. The effect of the chosen environmental policy on trade could then be obtained by observing trade flows over some subsequent period and comparing average trade flows from industries that are "treated" or affected by the policy, with the average trade flows from "control" industries that are unaffected by the policy.[31] Moreover, this estimate could be assigned a causal interpretation as there is a clear source of exogenous variation in treatment: the randomization of policy across industries.

The first reason GK's estimates fall short of that produced by the experimental ideal stems from the fact that it is not clear what is driving the observed variation in pollution abatement costs across industries, creating the possibility that cross-industry differences in τ_i simply reflect differences in other industry characteristics that determine the cost of abating pollution. Indeed, GK acknowledge the possibility of such omitted variables bias explicitly, and attempt to mitigate it by controlling for other potentially important industry characteristics. However, there is strong reason to believe that these variables – the industry factor shares for human and physical capital, the industry tariff rate and the industry injury rate – are not the only factors that affect the cost of abating pollution. Indeed, more recent work, such as that of Bloom et al. (2010), suggests that other industry characteristics, such as managerial quality, affect the pollution abatement

[31] Of course, this comparison implicitly assumes that there are no general equilibrium interactions, such as re-allocations of factors or changes in input-output relationships, across industries as a result of environmental regulation. While this is a maintained assumption in the majority of the existing literature, research by Cherniwchan and Najjar (2022) suggests that these general equilibrium interactions may be small.

decisions of industries. This means that GK's estimates of β are likely capturing the effects of factors aside from environmental policy.

A second, albeit related, concern is the possibility that the observed variation in τ_i is, itself, a product of the US government altering environmental policy in response to the effects of international trade. There is strong reason to believe that this might be the case in GK's setting. As they themselves note, governments may alter environmental policy in response to trade-induced changes in the demand for environmental quality on the part of the electorate.[32] While the type of response articulated by GK would lead to a level change in environmental policy common to all industries, subsequent work has highlighted that trade can also affect the environmental policies faced by specific industries through mechanisms such as lobbying (e.g. Fredriksson (1997), Conconi (2003)) or regulatory capture (e.g. McAusland (2008)). These types of industry-specific changes to environmental regulation mean that the τ_i observed by GK most likely directly reflects the effects of trade.[33] This is problematic because it means that τ_i is endogenously determined, so that the resulting estimate of $\hat{\beta}$ is biased.

The third issue stems from the measurement of τ_i. Recall that, in an idealized setting, GK would have been able to evaluate the effects of a particular environmental policy on trade. However, they do not observe the actual environmental policies that are applied to US industries, but instead use a measure based on industry pollution abatement costs as a proxy variable. This pollution abatement cost data is constructed from surveys that examine on the environmental expenditures made by firms, meaning they capture both those costs that are directly determined by regulatory stringency, as well as environmental spending that is unrelated to environmental policy, such as cross-industry differences in the cost of the inputs used in pollution abatement (Brunel and Levinson, 2016). This is problematic because it means that τ_i is potentially measured with error; the observed variation across industries could arise for reasons other than differences in the stringency of environmental regulation. Indeed, attenuation bias created by such measurement error could potentially explain GK's finding that environmental regulations have little effect on the pattern of international trade and investment.

In sum, assigning a causal interpretation to GK's estimates of the effects of environmental policy on trade requires three very strong assumptions: (i) that there are no

[32]The underlying intuition is relatively simple: if trade alters real incomes and environmental quality is a normal good, then trade will affect how individuals value environmental quality. This intuition features prominently in the subsequent theoretical literature (see, for example, the work of Copeland and Taylor (1994, 1995b), Antweiler et al. (2001) or McAusland and Millimet (2013)).

[33]While empirical evidence of international trade's effect on how individual governments set and enact environmental policy is quite limited, research by Cherniwchan and Najjar (2021) suggests that trade liberalization can affect environmental regulation, at least in the case of the US.

omitted factors that are correlated with environmental policy that also affect trade, (ii) that trade does not cause governments to change environmental policy, and (iii) that there is no measurement error arising from the use of pollution abatement costs as a proxy for the stringency of environmental policy. As discussed, there are reasons to believe that each of these assumptions fail to hold in practice, making it difficult to conclude that the imposition of stringent environmental regulations would not lead to a pollution haven effect.

Three Potential Solutions

Subsequent work has sought to address these issues and estimate the causal effect of environmental regulations on international trade through three main empirical approaches.

The first of these is quite simple: researchers have collected data consisting of proxies for the environmental policies faced by industries, as well information their trade and investment flows and other characteristics over a number of years. Doing so allows for the possibility of estimating a richer, panel version of equation (16):

$$m_{it} = \alpha + \beta\tau_{it} + X_{it}'\gamma + \mu_i + \delta_t + \epsilon_{it}. \tag{17}$$

As can be seen by inspection, aside from the unit of observation now being the industry-year (*it*), equation (17) differs from (16) due to the inclusion of μ_i, an industry fixed effect, and δ_t, a year fixed effect. These fixed effects are useful as they provides a simple means for accounting for any unobserved industry-specific time-invariant factors or common aggregate shocks that may have been biasing GK's estimates of the effects of environmental regulations on trade. As such, the use of panel data provides one approach for investigating the possibility that their findings are due to omitted variables bias.

This panel approach is typified by the work of Eskeland and Harrison (2003) and Ederington et al. (2005). Eskeland and Harrison examine the effects of environmental regulation (measured as pollution abatement costs as a fraction of value-added) on industry-level foreign direct investment flows into Côte d'Ivoire from 1977-1987, Morroco from 1985-1990, Mexico in 1990 and Venezuela from 1983-1988, as well as outbound foreign direct investment from US industries over the period 1982-1993. Ederington et al. (2005) examine the effects of environmental regulation (again measured as pollution abatement costs as a fraction of value-added) on net exports from US industries over the period 1978-1992. Both studies find some evidence consistent with a pollution haven effect, but this evidence is mixed. Eskeland and Harrison find that pollution abatement costs significantly affect FDI flows in some of their empirical specifications, but these results are not statistically robust. In contrast Ederington et al. (2005) report that the effects of environmental regulations are largely confined to the case of trade with developing countries, and they note their effects are largest for geographically mobile, or "footloose", industries.

One interpretation of the results from this approach is that they provide further support for GK's finding that environmental regulations have a limited effect on trade and investment flows, as they account for two potential sources of unobserved heterogeneity that may have been contaminating previous estimates. However, it is worth noting that assigning these estimates a causal interpretation still requires strong assumptions (albeit, weaker than in the case of GK). For one, the use of industry and time fixed effects does not address the possibility that the estimates are capturing the effects of omitted time-varying industry-specific factors, such as governments manipulating the stringency or enforcement of environmental policies faced by certain industries in response to political concerns over international competitiveness. Moreover, the relatively small estimates could again be potentially explained by the fact that environmental regulations are endogenously determined in response to trade, or by attenuation bias created by measurement error from the use of a proxy for the stringency of environmental policy.[34]

The second strategy that researchers have adopted to identify the effects of environmental regulations on trade and investment is the use of instrumental variables. This "IV approach" is appealing because it requires the articulation of a source of exogenous variation in τ, making it clear as to the conditions under which the resulting estimates can be assigned a causal interpretation.

The power of the IV approach is exemplified by the work of Levinson and Taylor (2008) and Kellenberg (2009). Levinson and Taylor use US panel data on industry pollution abatement costs, net exports and other characteristics over the period 1977-1989 to study the effects of environmental regulations on international trade between the US and Mexico. They construct instruments using variation in the factors that affect the demand for and supply of pollution created by geographic differences in where industries are located across the US. They find evidence of a modest pollution haven effect when using the panel-approach, and show that it becomes much stronger when they use a panel-IV approach, suggesting that Grossman and Krueger (1991)'s estimates were potentially capturing the effects of the endogenous relationship between international trade and environmental regulation as well as measurement error. Kellenberg (2009) uses panel data on the activities of US multinationals in 50 countries over the period 1999-2003 to

[34] Indeed, the attenuation bias arising from measurement error could be exacerbated due to the use of fixed effects.

study how environmental regulations affect multinational activity. He instruments for a country's environmental policy using neighboring country characteristics; this choice of instruments is motivated as the product of strategic policy interactions by competing governments. His results suggest that weak environmental regulations increases value added US affiliates, suggesting that multinationals make production decisions in a manner consistent with the intuition underlying the pollution haven effect.

While the results from the IV approach provide stronger support for the existence of the pollution haven effect, it is important to note that they are not without limitations. Indeed, the leading work in this area (such as that of Levinson and Taylor (2008) and Kellenberg (2009)) typically relies on model based arguments as to why the exclusion restriction holds and the resulting estimates can be interpreted causally. While these arguments are strongly grounded in economic theory, they are valid insofar as the underlying economic models are correctly specified. A potentially more significant issue is *what* causal effect the IV approach actually identifies. If, for example, some industries are unaffected by regulation because they already produce using the leading abatement technologies required by regulation, then the IV approach will only identify average effect of said regulation for the subset of industries for which it binds. This is potentially problematic given that the pollution haven effect is a statement about the average effect of regulation across all industries; if the IV approach is only identifying a *local* average effect for a subset of industries, then it is unclear how informative the estimates are for understanding the overall relationship between environmental regulation, and the pattern international trade and investment across industries.

The third approach researchers have used has been to study the effects of a single environmental policy on trade and investment flows. This "case-study" approach has two immediate benefits relative to relying on an indirect measure of environmental stringency such as pollution abatement costs. First, focusing on a single policy plausibly reduces measurement error, as the researcher is studying a specific regulation as it is applied in practice.[35] Second, a focus on a single policy has allowed researchers to highlight the particular institutional details that drive variation in regulatory stringency across firms or industries. As a result, it is clear what type of assumptions are required for identification and when they might fail, strengthening claims that the resulting estimates can be credibly interpreted as causal.

The usefulness of this approach can be seen clearly from studies that have examined

[35] It is important to note, however, that studying a single environmental policy need not eliminate measurement error completely, as the relevant statute or legislation observed by the researcher may not be enforced effectively.

the effects of air quality standards. Air quality standards are a popular form of environmental regulation around the world; for example, they have been used in the United States (the National Ambient Air Quality Standards (NAAQS) enacted under the Clean Air Act (CAA)), Canada (the Canada Wide Standards for Particulate Matter and Ozone (CWS)), and the European Union (the EU Clean Air Directives). These standards are designed to ensure that a country's minimum level of air quality meets a certain threshold; if ambient pollution concentrations in a region exceed the relevant threshold, then it is subject to more stringent environmental regulation.[36] This structure creates plausibly exogenous variation in the stringency of policy as ambient pollution concentrations are due, in part, to changes in weather conditions, and the standards are set federally, meaning they are unrelated to local tastes, characteristics or economic conditions. This facilitates the use of difference-in-difference type research designs in which the average outcomes of firms in regions that are subject to more stringent environmental regulation due to higher ambient pollution concentrations are compared with the average outcomes of firms in regions where ambient pollution concentrations are lower and environmental regulations are less stringent before and after regulation occurs.

This basic difference-in-difference design has been successfully exploited to study the effects of air quality standards on international trade and investment flows.[37]. One such example is the work of Hanna (2010), who uses a variant of this design to study the effects of the NAAQS enacted under the CAA on outbound FDI from affected US-based multinational firms over the period 1966-1999. Hanna's design exploits the fact that different multinational firms have different levels of exposure to regulation under the NAAQS due to the location of the manufacturing plants. As such, her modified difference-in-difference design effectively compares the average outbound FDI from multinationals that have a large fraction of their plants regulated with the average outbound FDI from those multinationals that have a small share of their plants regulated, before and after regulation. Using this approach Hanna finds evidence that US air quality standards caused US based multinationals to increase their foreign assets by 5.3% and their foreign output by 9%.

The work of Cherniwchan and Najjar (2022) is another example of study in which the identification strategy stems from the design of air quality standards. They study the

[36]Typically, these regulations have a two-part structure in which firms must adopt a "leading" technology that reduces pollution emissions, or firms have to pay a fine or face a production limit (Najjar and Cherniwchan, 2021).

[37]Variations of this design have also been used extensively elsewhere in environmental economics to study outcomes such as air quality (Henderson, 1996), plant location decisions (Becker and Henderson, 2000), industrial activity (Greenstone, 2002), labor markets (Walker, 2013), and firm emission intensity (Najjar and Cherniwchan, 2021).

effects of the CWS on the exports from affected Canadian manufacturing plants. While they employ the cross-region and temporal variation in regulation created by the design of the CWS, they also exploit an additional feature that is common to many air quality standards: that it designated a set of "targeted" industries that were to be subject to be the focus of more stringent regulation in areas with high pollution concentrations due to the fact that they were viewed as being responsible for the majority of the emissions of the targeted pollutants. This additional feature admits a triple difference research design in which the average outcomes of manufacturing plants that were both operating in targeted industries and located in regions that were subject to more stringent environmental regulations are compared with: (i) the average outcomes of manufacturing plants from the same region that operate in non-targeted industries, (ii) the average outcomes of plants from the same industry that operate in non-regulated regions, and (iii) the average outcomes of plants that operate in non-regulated regions and in non-targeted industries. This additional difference is useful because it provides a simple means to account for both time-varying industry and region specific factors such as industry demand shocks or localized recessions that may have otherwise confounded the effects of regulation. Using this approach, Cherniwchan and Najjar find evidence to suggest that the CWS air quality standards significantly affected exports from affected plants. Their results indicate that these regulations caused a 32% reduction in the volume of exports from the most affected continuing exporters and a 5 percentage point increase in the likelihood that affected plants stop exporting.

While focusing on the effects of a single environmental policy has clear benefits from the perspective of transparent identification, it is not necessarily without cost. Indeed, although the results presented by Hanna (2010) and Cherniwchan and Najjar (2022) are compelling evidence that environmental policy can significantly impact international trade and investment, they are difficult to interpret directly in the context of the debate over the pollution haven effect. This is due, in part, to the fact that a focus on a single environmental policy typically means the unit of observation is the individual firm. As such, the resulting research design often identifies *relative* changes in outcomes across *firms*, as in the case in the work of Hanna (2010) and Cherniwchan and Najjar (2022). These relative changes are not directly informative of the pollution haven effect as it is a statement about *absolute* changes in the volume of trade and investment across *industries*.

This issue is partially addressed by the work of Tanaka et al. (2021), who study the effects of US air quality standards on the potential relocation of lead-acid battery recycling from the US to Mexico over the period 2002-2015. Using a series of difference-in-difference designs, they show that the 2009 tightening of US's ambient lead standard reduced ambient lead concentrations near affected lead-acid battery recycling facilities in the US, increased exports of lead-acid batteries to Mexico, increased the volume of lead-acid battery recycling that occurs in Mexico, and negatively impacted the birth weight of infants born to mothers who reside in localities close to lead-acid battery recycling facilities in Mexico. These results are strongly suggestive of a pollution haven effect in the battery recycling industry, but as the authors themselves note, it is unclear if the findings would generalize to other industries that have other technological characteristics.

The Second Challenge: How Big is the Pollution Haven Effect?

Clearly, there has been significant progress in our understanding of how environmental regulations affect trade and investment since GK's initial work on the topic. The use of more advanced research designs, perhaps mirroring advancements in empirical methods throughout economics, has allowed researchers to make more credible claims that environmental regulations cause changes in trade and investment flows, at least for some firms and industries. When taken as a whole, the results of this recent literature provide strong evidence of the existence of the pollution haven effect.

Despite this progress in our understanding of the relationship between environmental regulations and trade, further research is required. While there is now clear evidence of the sign of the pollution haven effect, there is much less consensus as to its magnitude. This is due, at least in part, to the inherent limitations imposed by the empirical approaches that have been used in existing work. This is a significant issue given continued concerns over trade and pollution havens; as GK noted, understanding the magnitude of the pollution haven effect is necessary for understanding the likelihood with which the pollution haven hypothesis holds. If the pollution haven effect is relatively small, at least when compared to other traditional determinants of industrial specialization, then trade liberalization is unlikely to lead to the formation of pollution havens. Moreover, understanding the magnitude of pollution haven effect is a necessary condition for understanding the dynamic effects of trade liberalization; recall from our discussion of economic growth and the environment that the dynamic path of pollution for an economy depends on the short run effects of international trade, which are determined, at least in part, by how environmental regulations alter trade and investment flows. Given this, our second challenge is clear: obtain credible estimates of the pollution haven effect that can be directly compared with estimates of the effects other determinants of comparative advantage have on trade.

What is the Technique Effect?

Taken together, the results from GK's analysis point to a stunning conclusion: bilateral trade liberalization between the US and Mexico as part of NAFTA would not necessarily harm the environment in Mexico. This conclusion follows directly from adding up their estimates of the relative magnitudes of the scale, composition and technique effects, and hinges primarily on the existence of a very strong technique effect as implied by the EKC and the small composition effects as implied by their study of the determinants of comparative advantage underlying US-Mexico trade. Of course one criticism of this conclusion stems from the fact that GK arrive at their estimates of the relative magnitudes of the scale, composition, and technique effects indirectly through a variety of methods that may or may not be directly comparable.

Subsequent work has addressed this criticism head on by simultaneously estimating the magnitude of each effect. One approach for doing so has been to estimate the effects econometrically. Perhaps the most influential example of this approach is the work of Antweiler et al. (2001), who study the effects of trade on concentrations of sulphur dioxide using data from 290 observations sites in 108 cities located in 43 countries over the period 1971-1996. To do so, Antweiler et al. develop a simple two-sector general equilibrium model of a small open economy and use it to derive the estimating equation they use in their empirical analysis. The benefit of such a tight link between theory and empirics comes from the fact it is clear how the resulting estimates can be used to construct estimates of the scale, composition and technique effects. Moreover, because they come from a unified framework, the magnitudes of the estimates of each effect are directly comparable. Despite this, they arrive at a similar conclusion to that implied by GK: free trade is good for the environment for the majority of countries in their sample. This finding is also driven by a large technique effect and a relatively small composition effect; Antweiler et al. estimate the elasticity of pollution concentrations to income per capita growth to be larger than -1 in magnitude, while the elasticity of pollution concentrations to trade intensity ranges between -0.4 and -0.9. Subsequent work by Cole and Elliott (2003) and Managi et al. (2009) is also suggestive of large technique effects and relatively small composition effects, at least for some pollutants.

An alternative approach that has become quite popular in the literature is to generate estimates of the relative strength of the scale, composition, and technique effects by taking equation (2) directly to the data. This approach follows from the work of Levinson (2009), who builds on an earlier study by Selden et al. (1999) and employs a version of equation (2) to measure the relative importance of each effect in driving changes in the level of pollution emitted by US manufacturing over the period 1987-2001. Levinson reports two key results. First, he finds that composition effects have played a small role in determining the observed decline in the emission intensity of four common air pollutants from the US manufacturing. Second, he shows that the lion's share – that is, nearly 80% – of this "cleanup" of US manufacturing can be attributed to the technique effect. Subsequent research has produced similar findings in other settings, including for other time periods in the US (Levinson, 2015; Shapiro and Walker, 2018), in Europe (Brunel, 2016), China (Cole and Zhang, 2019), and Canada (Najjar and Cherniwchan, 2021), as well as for the world as whole (Grether et al., 2009), suggesting that the technique effect is the primary driver of aggregate environmental quality globally.

Somewhat surprisingly, the estimation and decomposition approaches produce two similar results: technique effects are relatively large and composition effects are relatively small. This begs the question: why are technique effects so large?

Unbundling the Technique Effect

One hypothesis for the large technique effect that has been advanced in the decomposition literature is that it is due to the effects of changes in "technology." At first glance, this is, perhaps, a natural interpretation given that the technique effect is driven by changes in the emission intensity of individual industries. Such changes necessarily reflect changes in the level of pollution emitted per unit of output in an industry, which is consistent with the idea that pollution is changing because of changes in the methods or "technologies" that are used to produce in the industry. However, for this interpretation to be correct, it must be the case that changes in $\hat{E}_i$ are purely driven by reductions in the emission intensities of the firms that comprise each industry i.

To see this more directly, it is useful to extend the accounting identity underlying the industry level decomposition to "unbundle" the technique effect. Following Cherniwchan et al. (2017), suppose industry i is comprised of a continuum of firms on the interval $[0, n_i]$, and let $z_i(n)$ and $x_i(n)$ denote pollution and value added from firm n. Then industry emission intensity can be expressed as:

$$E_i = \frac{Z_i}{X_i} = \int_0^{n_i} e_i(n)\varphi_i(n)dn \tag{18}$$

where $e_i(n) = z_i(n)/x_i(n)$ is firm n's emission intensity and $\varphi_i(n) = x_i(n)/X_i$ is the value share of firm n's production in industry i. Clearly, the emission intensity of industry i is simply a weighted average of the emission intensities of all firms that operate in that industry. Following the same approach as for the aggregate industry decomposition, taking logs and totally differentiating yields:

$$\hat{E}_i = \int_0^{n_i} \hat{e}_i(n)\theta_i(n)dn + \int_0^{n_i} \hat{\varphi}_i(n)\theta_i(n)dn + n_i[\theta_i(n_i) - \varphi_i(n_i)]\hat{n}_i \tag{19}$$

where hats again indicate percentage changes, and $\theta_i(n) = z_i(n)/Z_i$ is firm n's share of emissions in industry i.

Equation (19) indicates that the technique effect as measured by Levinson and others necessarily depends on three within-industry changes, that Najjar and Cherniwchan (2021) term the *process*, *reallocation* and *selection* effects, respectively. The process effect, given by the first term in equation (19), reflects the change in an industry's emission intensity driven by changes in individual firms' emission intensities.[38] The second term, or reallocation effect, captures the change in industry emission intensity that is owing to changes in the relative sizes of firms. The third term, or selection effect, is the change in industry emission intensity caused by the entry and/or exit of firms.

As equation (19) makes clear, interpreting the technique effect as the sole product of changes in "technology" requires that it be driven purely by the process effect. This interpretation is not supported by the evidence. Recent work by Holladay and LaPlue III (2021) that applies the Levinson (2009) approach to firm level data, shows that, for some common pollutants, the reallocation and selection effects account for a significant fraction (i.e. greater than 40%) of the aggregate reduction in the emission intensity of the US manufacturing sector, meaning that the technique effect cannot be determined by changes in "technology" alone.

An alternative to the technology hypothesis is that the changes in process, reallocation and selection effects driving the technique effect are simply a product of the effects of environmental regulations. There is reason to believe that environmental policies may be the likely culprit; as many of the countries that have exhibited a large technique effect also experienced large gains in real income over the corresponding period of study. As highlighted in our discussion of economic growth and the environment, real income gains can generate an increase in the demand for new or more stringent environmental regulation, meaning the observed technique effects could simply reflect shifts and changes in the economic activities of firms brought about by governments enacting new environmental policies in response to the demands of their constituents.

Indeed, the work of Shapiro and Walker (2018) suggests that technique effects are

[38] Of course, one could extend the logic of the exercise used to obtain equation (19) to ask what is driving the process effect. See the work of Cherniwchan et al. (2017) for such an extension.

driven almost entirely by environmental regulation. They develop a multi-sector quantitative general equilibrium model that features heterogeneous firms a la Melitz (2003), in which firms pollute and abate using a variant of the Copeland and Taylor (1994) joint production technology. Shapiro and Walker estimate the model's key parameters and then engage in a series of counterfactual experiments to examine whether the observed reduction in the emission intensity of the US manufacturing sector over the period 1990-2008 can be attributed to changes in the shadow cost of pollution, productivity shocks, or changes in trade costs. Their analysis suggests that the implicit shadow price of pollution nearly doubles over their period of study, and that the effects of this change are the primary determinant of reductions in industry emission intensity. Given that the shadow price reflects the implicit cost of polluting for firms in their model, Shapiro and Walker interpret this measure as reflective of the costs of complying with environmental regulations and conclude that environmental regulations have driven the cleanup of manufacturing. Given that the reduction in the emission intensity of US manufacturing is almost entirely due to the technique effect, their result suggest that firm level responses to environmental regulations are driving the technique effect.

While the analysis presented by Shapiro and Walker (2018) is suggestive of the role of environmental regulations, there are at least three reasons to believe that it is not the only cause of the technique effect. The first of these is the findings of a study by Najjar and Cherniwchan (2021), who estimate the effects of Canadian air quality standards – the Canada-Wide Standards for Particulate Matter and Ozone (CWS) – on the pollution emitted by affected Canadian manufacturing plants. Taken together, Najjar and Cherniwchan's estimates indicate that the process, selection and reallocation effects caused by these regulations account for just under 38% of the technique effect that occurred in Canada during their period of study. As the CWS was the primary form of environmental regulation targeting fine scale particulate matter – the pollutant examined by Najjar and Cherniwchan – in Canada over this period, this finding is suggestive of the possibility that other factors aside from environmental regulation have contributed to the technique effect.

The second piece of evidence that casts doubt on the conclusion that the technique effect is solely due to the effects of environmental regulation comes from studies that have used Levinson-type decompositions to examine carbon dioxide and greenhouse gasses (GHGs) in a number of countries. These studies have also identified large technique effects similar to what has been found in other studies. For example, this is true for India (Martin, 2012), the US (Brunel and Levinson, 2021), and for a panel of over 40 developed and less developed countries (Copeland et al., 2021). These findings are difficult to reconcile with the role of environmental policy as carbon dioxide and GHGs have been virtually unregulated in most countries over the last 40 years. This also suggests other factors must be driving emissions downward.

The third piece of evidence comes from a large body of work examining the relationship between trade and the environment at the firm level. Much of the work in this area has simply documented a correlation between exporting and the pollution produced by firms, whereby exporters are less pollution intensive than non-exporters (e.g. Holladay (2016), Richter and Schiersch (2017), Forslid et al. (2018)). On its own, this correlation is suggestive of a role for trade in determining the technique effect as it consistent with international trade inducing a reallocation of economic activity across firms according to their emission intensity. This interpretation is supported by a number of recent papers that examine the direct effects of trade integration on the pollution emitted by firms. For example, Cherniwchan (2017) shows that bilateral trade liberalization between the US and Mexico caused a reduction in the emission intensity of US manufacturing firms affected by the agreement. Similarly, Gutiérrez and Teshima (2018) show that import competition affected the abatement decisions of affected Mexican manufacturing firms, and Barrows and Ollivier (2021) show that foreign demand shocks significantly impact the emission intensity of affected firms in India. While these studies are largely suggestive of trade causing a process effect, recent work by Lim (2021) also provides evidence that trade can also impact the technique effect through reallocation and selection effects in addition to the process effect. She develops a quantitative model and shows that 8-10% of the observed technique effect for NO_X in the US over the period 1998-2014 can be attributed to reductions in the cost of importing intermediate inputs. Importantly, she shows that over 40% of this effect can be attributed to the reallocation and selection effects induced by reductions in the cost of importing intermediate inputs.

The Third Challenge: The Trade-Induced Technique Effect

From this discussion, it should be clear that neither technology, nor environmental regulations are the sole driver of the technique effect; international trade unquestionably plays a role in determining the emission intensity of individual industries. What is not yet clear, exactly, is the extent of trade's contribution. Determining trade's role is complicated by the fact that it can also indirectly impact the technique effect by altering the technologies used by firms, or the policies used to regulate environmental problems, two channels discussed explicitly by GK.[39] However, doing so is important because the cur-

[39]For examples of the first see Reppelin-Hill (1999) or Khanna and Zilberman (2001). Recent work by Cherniwchan and Najjar (2021) provides evidence of the latter channel. It is worth noting that the two

rent conclusion that trade is good for the environment hinges primarily on the strength of the technique effect. If trade is a leading cause of the technique effect, then it is more likely that this conclusion is, indeed, correct. This sets the stage for our third, and final challenge: obtaining estimates of the magnitude of the trade-induced technique effect.

Concluding Remarks

We used the original contribution of GK as a springboard for a discussion of the research that has been done into the relationship between international trade and the environment following the release of their seminal paper studying the potential environmental effects of NAFTA in 1991. We did so paying particular attention to their original contributions, and how this laid the foundation for subsequent work. We started by revisiting their somewhat controversial empirical finding of an EKC, and leveraged this discussion to identify the short and long run impacts of trade on the environment. The net result of these various impacts is apriori uncertain, and so we turned to review the empirical literature estimating the strength of composition effects and the role policy changes play in driving emission intensities downward. By reviewing very recent and past empirical work, we discovered three empirical challenges for future research. Our hope is that by identifying these challenges, and presenting a simple theory to frame their discussion, we have provided both the means and the motivation for others to follow.

channels can interact; as shown by Lovely and Popp (2011) trade-induced technology transfer that affects pollution can also impact a country's environmental policy.

Further Reading

Acemoglu, D., Aghion, P., and Hémous, D. (2014). The Environment and Directed Technical Change in a North–South Model. *Oxford Review of Economic Policy*, 30(3):513–530.

Brock, W. A. and Taylor, M. S. (2010). The Green Solow Model. *Journal of Economic Growth*, 15(2):127–153.

Cherniwchan, J. (2012). Economic Growth, Industrialization, and the Environment. *Resource and Energy Economics*, 34(4):442–467.

Cherniwchan, J. (2017). Trade Liberalization and the Environment: Evidence from NAFTA and US Manufacturing. *Journal of International Economics*, 105:130–149.

Cherniwchan, J., Copeland, B. R., and Taylor, M. S. (2017). Trade and the Environment: New Methods, Measurements, and Results. *Annual Review of Economics*, 9:59–85.

Cherniwchan, J. and Najjar, N. (2022). Do Environmental Regulations Affect the Decision to Export? *American Economic Journal: Economic Policy*, 14(2):125–160.

Copeland, B. R., Shapiro, J. S., and Taylor, M. S. (2021). Globalization and the Environment. Working Paper 28797, National Bureau of Economic Research.

Di Maria, C. and Smulders, S. A. (2005). Trade Pessimists vs Technology Optimists: Induced Technical Change and Pollution Havens. *Advances in Economic Analysis & Policy*, 4(2).

Grossman, G. and Krueger, A. B. (1991). Environmental Impacts of a North American Free Trade Agreement. Working Paper No. 3914, National Bureau of Economic Research.

Grossman, G. and Krueger, A. B. (1995). Economic Growth and the Environment. *Quarterly Journal of Economics*, 110:353–377.

Hanna, R. (2010). US Environmental Regulation and FDI: Evidence from a Panel of US-Based Multinational Firms. *American Economic Journal: Applied Economics*, 2(3):158–89.

Holladay, J. S. and LaPlue III, L. D. (2021). Decomposing Changes in Establishment-Level Emissions with Entry and Exit. *Canadian Journal of Economics*, 54(3):1046–1071.

Levinson, A. (2009). Technology, International Trade and Pollution From US Manufacturing. *American Economic Review*, 99(5):2177–2192.

Levinson, A. and Taylor, M. S. (2008). Unmasking the Pollution Haven Effect. *International Economic Review*, 49(1):223–254.

Shapiro, J. S. and Walker, R. (2018). Why is Pollution from US Manufacturing Declining? The Roles of Environmental Regulation, Productivity, and Trade. *American Economic Review*, 108(12):3814–54.

Tanaka, S., Teshima, K., and Verhoogen, E. (2021). North-South Displacement Effects of Environmental Regulation: The Case of Battery Recycling. *American Economic Review: Insights*, page Forthcoming.

References

Acemoglu, D., Aghion, P., and Hémous, D. (2014). The Environment and Directed Technical Change in a North–South Model. *Oxford Review of Economic Policy*, 30(3):513–530.

Andreoni, J. and Levinson, A. (2001). The Simple Analytics of the Environmental Kuznets Curve. *Journal of Public Economics*, 80(2):269–286.

Antweiler, W., Copeland, B. R., and Taylor, M. S. (2001). Is Free Trade Good for the Environment? *American Economic Review*, 91:877–908.

Barrows, G. and Ollivier, H. (2021). Foreign Demand, Developing Country Exports, and CO2 Emissions: Firm-level Evidence from India. *Journal of Development Economics*, 149:102587.

Becker, R. and Henderson, V. (2000). Effects of Air Quality Regulations on Polluting Industries. *Journal of Political Economy*, 108:379–421.

Bloom, N., Genakos, C., Martin, R., and Sadun, R. (2010). Modern Management: Good for the Environment or Just Hot Air? *Economic Journal*, 120(544):551–572.

Brock, W. A. and Taylor, M. S. (2005). Economic Growth and the Environment: A Review of Theory and Empirics. *Handbook of Economic Growth*, 1:1749–1821.

Brock, W. A. and Taylor, M. S. (2010). The Green Solow Model. *Journal of Economic Growth*, 15(2):127–153.

Brown, D. K., Deardorff, A. V., and Stern, R. M. (1991). A North American Free Trade Agreement: Analytical Issues and a Computational Assessment. Working Paper 289, Institute of Public Policy Studies University of Michigan.

Brunel, C. (2016). Pollution Offshoring and Emissions Reductions in EU and US Manufacturing. *Environmental and Resource Economics*, 68(3):621–641.

Brunel, C. and Levinson, A. (2016). Measuring the Stringency of Environmental Regulations. *Review of Environmental Economics and Policy*, 10(1):47–67.

Brunel, C. and Levinson, A. (2021). Globalization and Greenhouse Gas Emissions: Evidence From the United States. Working Paper No. 28372, National Bureau of Economic Research.

Cherniwchan, J. (2012). Economic Growth, Industrialization, and the Environment. *Resource and Energy Economics*, 34(4):442–467.

Cherniwchan, J. (2017). Trade Liberalization and the Environment: Evidence from NAFTA and US Manufacturing. *Journal of International Economics*, 105:130–149.

Cherniwchan, J., Copeland, B. R., and Taylor, M. S. (2017). Trade and the Environment: New Methods, Measurements, and Results. *Annual Review of Economics*, 9:59–85.

Cherniwchan, J. and Najjar, N. (2021). Free Trade and the Formation of Environmental Policy: Evidence from US Legislative Votes. Working Paper 21-11, Carleton University.

Cherniwchan, J. and Najjar, N. (2022). Do Environmental Regulations Affect the Decision to Export? *American Economic Journal: Economic Policy*, 14(2):125–160.

Cole, M. A. and Elliott, R. J. (2003). Determining the Trade-Environment Composition Effect: the Role of Capital, Labor, and Environmental Regulations. *Journal of Environmental Economics and Management*, 46:363–383.

Cole, M. A. and Zhang, L. (2019). The Clean-up of Chinese Manufacturing: Examining the Role Played by Changing Techniques of Production. *Economics Letters*, 180:11–14.

Conconi, P. (2003). Green Lobbies and Transboundary Pollution in Large Open Economies. *Journal of International Economics*, 59(2):399–422.

Copeland, B. R., Shapiro, J. S., and Taylor, M. S. (2021). Globalization and the Environment. Working Paper 28797, National Bureau of Economic Research.

Copeland, B. R. and Taylor, M. S. (1994). North-South Trade and the Environment. *Quarterly Journal of Economics*, 109:755–787.

Copeland, B. R. and Taylor, M. S. (1995a). Trade and the Environment: A Partial Synthesis. *American Journal of Agricultural Economics*, 77:765–771.

Copeland, B. R. and Taylor, M. S. (1995b). Trade and Transboundary Pollution. *American Economic Review*, 85:716–737.

Copeland, B. R. and Taylor, M. S. (2003). *Trade and the Environment*. Princeton University Press, Princeton, NJ.

Copeland, B. R. and Taylor, M. S. (2004). Trade, Growth, and the Environment. *Journal of Economic Literature*, 42:7–71.

Deardorff, A. V. (1974). A Geometry of Growth and Trade. *Canadian Journal of Economics*, 7(2):295–306.

Deardorff, A. V. (2013). Growth or Decline of Comparative Advantage. *Journal of Macroeconomics*, 38:12–18.

Di Maria, C. and Smulders, S. A. (2005). Trade Pessimists vs Technology Optimists: Induced Technical Change and Pollution Havens. *Advances in Economic Analysis & Policy*, 4(2).

Dinda, S. (2004). Environmental Kuznets Curve Hypothesis: A Survey. *Ecological Economics*, 49(4):431–455.

Ederington, J., Levinson, A., and Minier, J. (2005). Footloose and Pollution-Free. *The Review of Economics and Statistics*, 87(1):92–99.

Eskeland, G. S. and Harrison, A. E. (2003). Moving to Greener Pastures? Multinationals and the Pollution Haven Hypothesis. *Journal of Development Economics*, 70(1):1–23.

Forslid, R., Okubo, T., and Ulltveit-Moe, K. H. (2018). Why Are Firms That Export Cleaner? International Trade, Abatement and Environmental Emissions. *Journal of Environmental Economics and Management*, 91:166–183.

Fredriksson, P. G. (1997). The Political Economy of Pollution Taxes in a Small Open Economy. *Journal of Environmental Economics and Management*, 33(1):44–58.

Greenstone, M. (2002). The Impacts of Environmental Regulations on Industrial Activity: Evidence from the 1970 and 1977 Clean Air Act Amendments and the Census of Manufactures. *Journal of Political Economy*, 110(6):1175–1219.

Grether, J.-M., Mathys, N. A., and de Melo, J. (2009). Scale, Technique and Composition Effects in Manufacturing SO_2 Emissions. *Environmental and Resource Economics*, 43:257–274.

Grossman, G. and Krueger, A. B. (1991). Environmental Impacts of a North American Free Trade Agreement. Working Paper No. 3914, National Bureau of Economic Research.

Grossman, G. and Krueger, A. B. (1995). Economic Growth and the Environment. *Quarterly Journal of Economics*, 110:353–377.

Gutiérrez, E. and Teshima, K. (2018). Abatement Expenditures, Technology Choice, and Environmental Performance: Evidence from Firm Responses to Import Competition in Mexico. *Journal of Development Economics*, 133:264–274.

Hanna, R. (2010). US Environmental Regulation and FDI: Evidence from a Panel of US-Based Multinational Firms. *American Economic Journal: Applied Economics*, 2(3):158–89.

Harbaugh, W. T., Levinson, A., and Wilson, D. M. (2002). Reexamining the Empirical Evidence for an Environmental Kuznets Curve. *Review of Economics and Statistics*, 84(3):541–551.

Hémous, D. (2016). The Dynamic Impact of Unilateral Environmental Policies. *Journal of International Economics*, 103:80–95.

Henderson, J. V. (1996). Effects of Air Quality Regulation. *American Economic Review*, pages 789–813.

Hilton, F. H. and Levinson, A. (1998). Factoring the Environmental Kuznets Curve: Evidence from Automotive Lead Emissions. *Journal of Environmental Economics and Management*, 35(2):126–141.

Holladay, J. S. (2016). Exporters and the Environment. *Canadian Journal of Economics*, 49(1):147–172.

Holladay, J. S. and LaPlue III, L. D. (2021). Decomposing Changes in Establishment-Level Emissions with Entry and Exit. *Canadian Journal of Economics*, 54(3):1046–1071.

John, A. and Pecchenino, R. (1994). An Overlapping Generations Model of Growth and the Environment. *Economic Journal*, 104(427):1393–1410.

Kellenberg, D. (2009). An Empirical Investigation of the Pollution Haven Effect with Strategic Environment and Trade Policy. *Journal of International Economics*, 78:242–255.

Khanna, M. and Zilberman, D. (2001). Adoption of Energy Efficient Technologies and Carbon Abatement: The Electricity Generating Sector in India. *Energy Economics*, 23:637–658.

Levinson, A. (2009). Technology, International Trade and Pollution From US Manufacturing. *American Economic Review*, 99(5):2177–2192.

Levinson, A. (2015). A Direct Estimate of the Technique Effect: Changes in the Pollution Intensity of US Manufacturing 1990-2008. *Journal of the Association of Environmental and Resource Economists*, 2(1):43–56.

Levinson, A. and Taylor, M. S. (2008). Unmasking the Pollution Haven Effect. *International Economic Review*, 49(1):223–254.

Lim, H. R. (2021). Trade in Intermediates and US Manufacturing Emissions. Unpublished manuscript.

Lovely, M. and Popp, D. (2011). Trade, Technology, and the Environment: Does Access to Technology Promote Environmental Regulation? *Journal of Environmental Economics and Management*, 61(1):16–35.

Managi, S., Hibiki, A., and Tsurumi, T. (2009). Does Trade Openness Improve Environmental Quality? *Journal of Environmental Economics and Management*, 58(3):346–363.

Martin, L. A. (2012). Energy Efficiency Gains from Trade: Greenhouse Gas Emissions and India's Manufacturing Sector. Unpublished manuscript.

McAusland, C. (2008). Trade, Politics, and the Environment: Tailpipe vs. Smokestack. *Journal of Environmental Economics and Management*, 55(1):52–71.

McAusland, C. and Millimet, D. L. (2013). Do National Border Matter? Intranational Trade, International Trade, and the Environment. *Journal of Environmental Economics and Management*, 65(3):411–437.

Melitz, M. (2003). The Impact of Trade on Aggregate Industry Productivity and Intra-Industry Reallocations. *Econometrica*, 71(6):1695–1725.

Najjar, N. and Cherniwchan, J. (2021). Environmental Regulations and the Cleanup of Manufacturing: Plant-Level Evidence. *Review of Economics and Statistics*, 103(3):476–491.

OECD (2022). OECD.Stat Database. https://stats.oecd.org/. Accessed: 2022-04-21.

Reppelin-Hill, V. (1999). Trade and Environment: An Empirical Analysis of the Technology Effect in the Steel Industry. *Journal of Environmental Economics and Management*, 38(3):283–301.

Richter, P. M. and Schiersch, A. (2017). CO2 Emission Intensity and Exporting: Evidence from Firm-Level Data. *European Economic Review*, 98:373–391.

Selden, T. M., Forrest, A. S., and Lockhart, J. E. (1999). Analyzing the Reductions in US Air Pollution Emissions: 1970 to 1990. *Land Economics*, 75(1):1–21.

Selden, T. M. and Song, D. (1994). Environmental Quality and Development: Is There a Kuznets Curve for Air Pollution Emissions? *Journal of Environmental Economics and Management*, 27(2):147–162.

Shapiro, J. S. and Walker, R. (2018). Why is Pollution from US Manufacturing Declining? The Roles of Environmental Regulation, Productivity, and Trade. *American Economic Review*, 108(12):3814–54.

Solow, R. M. (1956). A Contribution to the Theory of Economic Growth. *Quarterly Journal of Economics*, 70(1):65–94.

Stern, D. I. (2004). The Rise and Fall of the Environmental Kuznets Curve. *World Development*, 32(8):1419–1439.

Stokey, N. L. (1998). Are There Limits to Growth? *International Economic Review*, 39(1):1–31.

Summers, R. and Heston, A. (1991). The Penn World Table (Mark 5): an Expanded Set of International Comparisons, 1950–1988. *Quarterly Journal of Economics*, 106(2):327–368.

Tanaka, S., Teshima, K., and Verhoogen, E. (2021). North-South Displacement Effects of Environmental Regulation: The Case of Battery Recycling. *American Economic Review: Insights*, page Forthcoming.

Walker, W. R. (2013). The Transitional Costs of Sectoral Reallocation: Evidence from the Clean Air Act and the Workforce. *Quarterly Journal of Economics*, 128(4):1787–1835.