THE ENVIRONMENTAL BIAS OF TRADE POLICY*

JOSEPH S. SHAPIRO

November 2020

Abstract

This paper describes a new fact, then analyzes its causes and consequences: in most countries, import tariffs and non-tariff barriers are substantially lower on dirty than on clean industries, where an industry's "dirtiness" is defined as its carbon dioxide (CO_2) emissions per dollar of output. This difference in trade policy creates a global implicit subsidy to CO_2 emissions in internationally traded goods and so contributes to climate change. This global implicit subsidy to CO_2 emissions totals several hundred billion dollars annually. The greater protection of downstream industries, which are relatively clean, substantially accounts for this pattern. The downstream pattern can be explained by theories where industries lobby for low tariffs on their inputs but final consumers are poorly organized. A quantitative general equilibrium model suggests that if countries applied similar trade policies to clean and dirty goods, global CO_2 emissions would decrease and global real income would change little.

JEL: Q50, Q56, F6, F13, F18, H23

*I thank Costas Arkolakis, Tim Armstrong, Kyle Bagwell, Brian Copeland, Arnaud Costinot, Thibault Fally, Penny Goldberg, Larry Goulder, Edgar Hertwich, Sam Kortum, Arik Levinson, Giovanni Maggi, Guillermo Noguera, Bill Nordhaus, Michael Peters, Steve Puller, Andrés Rodríguez-Clare, Katheryn Russ, Nick Ryan, Bob Staiger, Reed Walker, Marty Weitzman, and seminar participants at AEA, AERE, Columbia, Dartmouth, the Environmental Defense Fund, Georgetown, Harvard/MIT, Harvard Kennedy School, IDB, the Kiel Institute, NBER EEE, NBER ITI, NTA, Princeton, San Francisco Fed, Stanford, Sweden's Ministry for Foreign Affairs, UBC, UC Berkeley, U Hawaii, UC Santa Barbara, UC San Diego, U Mass Amherst, Vanderbilt, the World Bank, and Yale for useful discussions, Matt Fiedler and Ralph Ossa for sharing code, Elyse Adamic, Kenneth Lai, Eva Lyubich, Daisy Chen Sun, and Katherine Wagner for excellent research assistance, and the Alfred P. Sloan Foundation, NSF SES-1850790, and USDA NIFA #1021396 for generous support.
Contact: 714 University Hall #3310, Berkeley, CA 94720. (510) 642-3345, joseph.shapiro@berkeley.edu.

I Introduction

This paper describes a new fact, then analyzes its causes and consequences: in most countries, import tariffs and non-tariff barriers (NTBs) are lower on dirty than on clean industries, where an industry's "dirtiness" is measured by its carbon dioxide (CO_2) emissions per dollar of output. This difference between dirty and clean industries creates an implicit subsidy to CO_2 emissions in internationally traded goods and so contributes to climate change. I describe this pattern as trade policy's environmental bias.

This bias is widespread. I find it in most countries, in tariffs and NTBs, and in cooperative and non-cooperative tariffs. U.S. tariff data over the past 30 years suggest this bias has slowly diminished over time, though remains large. U.S. tariffs imposed in the 2018 trade war slightly attenuated but did not eliminate this bias. The global implicit subsidy in trade policy that I estimate, of \$85 to \$120 per ton of CO_2, is interesting because the global social cost of CO_2 emissions (and hence the optimal tax on CO_2 emissions) is usually estimated as around \$40 per ton of CO_2 (IWG 2016). The magnitude of the environmental bias of trade policy is therefore larger than what research suggests is an optimal tax on CO_2 emissions, and the sign is opposite—trade policy is imposing lower tax rates on dirtier goods, while an optimal carbon policy would impose higher tax rates on dirtier goods.

One way to interpret this fact is in terms of climate change policy. Optimal climate change policy would impose a uniform Pigouvian tax (or equivalent quantity mechanism like a cap-and-trade market) in all countries and industries, since CO_2 creates the same climate change externality regardless of where it is emitted. Researchers and policymakers often claim that imposing climate change policy in some countries but not others could harm domestic energy-intensive industries and lead to relocation or "leakage" of emissions, more than an absolute decrease in emissions. Climate change regulation is far from global and covers about 20 percent of global CO_2 emissions, including in the EU, California, and elsewhere (World Bank 2018). Carbon prices in these policies differ substantially across regulations and are generally below \$10/ton. Some countries have considered pairing such sub-global policy with an import tariff or border adjustment that is proportional to the CO_2 emitted from producing and transporting goods.[1]

Of course, most countries already impose tariffs and NTBs on traded goods. This paper asks whether dirty industries already face higher tariffs and NTBs, which would mean that countries already implicitly have carbon tariffs in their existing trade policies. Given media emphasis on dirty industries' political lobbying, one might expect dirtier industries to receive relatively greater trade protection. I show that this prediction is incorrect, and that dirtier industries face relatively low tariffs and NTBs.

I obtain these findings from regressions of tariff (or ad valorem NTB) rates on CO_2 intensity. I

[1] Some versions of this proposal would include rebates for exports. Several proposed U.S. climate change regulations would implement carbon tariffs, including the Waxman-Markey Bill (the American Clean Energy and Security Act), which passed the House but not the Senate in 2009; the American Opportunity Carbon Fee Act of 2014; and a current "carbon dividends" proposal by the U.S. Climate Leadership Council led by James Baker, Martin Feldstein, Greg Mankiw, and publicly endorsed by 27 economics Nobel laureates and 3500 economists. One common perception is that a carbon tariff is politically necessary (though so far not politically sufficient) to ensure support for any U.S. climate change regulation. Legal analyses suggest that regulations of the World Trade Organization (WTO) could allow such carbon tariffs, though disagree on exactly which type of carbon tariff WTO rules would allow (Hillman 2013; Pauwelyn 2013; Cosbey et al. 2017).

measure CO_2 intensity by inverting a global multi-region input-output table, which accounts for emissions embodied in intermediate goods. For example, the CO_2 emissions rate for U.S. kitchenware accounts for the Australian coal used to produce the Chinese steel used to produce a U.S. frying pan, and the bunker and diesel fuels used to transport each. The global input-output data this paper uses, from Exiobase, describe 48 countries and 163 industries, and so generate measures of CO_2 intensity for each international and intra-national trade flow in the global economy. The tariff data are even more detailed, with 200 million tariff measures that uniquely describe each origin×destination×industry. I obtain qualitatively similar results from several other datasets and sensitivity analyses.

Why have countries imposed more protection on clean than dirty industries? Theory and evidence suggest that countries do not explicitly consider CO_2 or intend to subsidize it in choosing trade policy; indeed, I believe that countries are not even aware of the implicit subsidy in trade policy this paper highlights, since previous literature has not tested for or identified it. Instead, this paper proposes that some forces which determine trade policy are correlated with CO_2 intensity.

To determine which forces account for the association between trade policy and CO_2 intensity, the analysis considers explanations based on 20 variables suggested by theoretical and empirical research. These explanations include optimal tariffs (inverse export supply elasticities), lobbying expenditures, unionization, labor and capital shares, declining or "sunset" industries, worker wages and education, firm size, industry concentration rates, intra-industry trade, levels and trend in trade exposure, dispersion in firm sizes and in firm locations, shipping costs, unemployment, "local" pollutants like sulfur dioxide, production efficiency, and an industry's upstream location. These variables are available for the U.S.; a subset is available for all countries. To address potential endogeneity, some specifications instrument a particular political economy explanation (e.g., mean wages in a specific industry) with its value from the ten other smallest countries in the data. I focus on the ten smallest other countries since they are more likely to take conditions in the rest of the world as given.

Among these potential explanations, linear regressions and a machine learning algorithm highlight an industry's location or "upstreamness" in global value chains as accounting for a large share of the association between CO_2 intensity and trade policy. The analysis measures upstreamness as the economic distance of each industry from final consumers (Antràs et al. 2012). More upstream industries have both lower protection and greater emissions.

I investigate one political economy explanation for the covariance of upstreamness and trade policy involving lobbying competition. Firms may lobby for high tariffs and NTBs on their own outputs, but also lobby for low tariffs on the goods they use directly and indirectly as inputs, so as to decrease production costs.[2] Because final consumers are poorly organized (Olson 1965), politicians give the least protection to

[2]Firms publicly emphasize this rationale. When President Trump initially proposed tariffs on steel, the American Automotive Policy Council announced, "The auto industry and the U.S. workers that the industry employs would be adversely affected and that [*sic*] this unintended negative impact would exceed the benefit provided to the steel industry" (Gibson 2017). The Consuming Industries Trade Action Coalition (CITAC), a twenty-year old U.S. lobby group focused on decreasing tariffs on upstream industries, experienced a doubling of membership during a "Stand up to Steel" campaign, and supported a bill in the U.S. House of Representatives (HR 2770) to give steel consumers greater standing in trade cases. When President Obama imposed tariffs on Chinese tires, CITAC responded, "[W]e believe that this case will undermine the jobs of many more US

the upstream industries (which are also the dirtiest) and the greatest protection to the most downstream industries (which are also the cleanest).

Figure I shows nonparametric local linear regressions that illustrate several key ideas in the paper. Each graph in this figure shows two lines. The downward-sloping dashed blue line shows a nonparametric regression of total CO_2 intensity on upstreamness. This line shows that the most upstream industries are dirtier. The upward-sloping solid red line shows a local linear regression of tariffs plus NTBs on upstreamness, which shows that the most upstream industries have the least protection. The patterns are similar for global and U.S. data. Previous research has not documented this systematic nonparametric relationship between trade policy and upstreamness, or between pollution and upstreamness. In these graphs, the relationships between each of these outcomes (CO_2 intensity, trade policy) and upstreamness are somewhat linear.

Appendix Figure I finds similar patterns in essentially each of the roughly 50 countries with data. This figure plots nonparametric relationships between CO_2 intensity and upstreamness, and between trade policy (tariffs+NTBs) and upstreamness, separately for each country in Exiobase. While this figure provides almost 50 separate small graphs, casual inspection shows the "X"-shaped pattern that in most countries, CO_2 intensity increases somewhat steadily with upstreamness, while tariffs and NTBs decrease.

A partial equilibrium back-of-the-envelope calculation suggests that this global implicit subsidy in trade policy to CO_2 emissions totals \$550 to \$800 billion dollars per year. This can be interpreted as revenue that a carbon tariff would collect if it had the same pattern as trade policy's environmental bias (i.e., -\$85/ton to -\$120/ton).

I then use a quantitative general equilibrium model to assess how counterfactual trade policies would affect CO_2 emissions and social welfare. This analysis uses strong assumptions that provide an imperfect approximation to reality.

The model incorporates several common features—input-output links, trade imbalances, CO_2 emissions from fossil fuel, tariffs that are lump-sum rebated, and NTBs (Costinot and Rodriguez-Clare 2014; Caliendo and Parro 2015; Eaton et al. 2016; Shapiro 2016). I study six sets of counterfactual policies. In the first, each country sets a single tariff per trading partner which applies to all industries, and which equals the country's mean baseline bilateral tariff. Each country implements a similar reform for NTBs. This counterfactual decreases global CO_2 emissions while leaving global real income unchanged or slightly increased. It has similar magnitude effects on CO_2 as two of the world's largest actual or proposed climate change policies, the EU Emissions Trading System and the U.S. Waxman-Markey Bill. In the second counterfactual, only the EU adopts this policy. One could think of this as a way for the EU to address leakage from its CO_2 cap-and-trade market, the EU Emissions Trading System. This decreases global CO_2 emissions by half the amount of the global policy, and again leaves global real income unchanged or slightly higher.

The third and fourth counterfactuals find that changing tariffs and NTBs to equal either the base-

workers in downstream industries..." (Business Wire 2009).

line level of the cleanest third or dirtiest third of industries decreases global CO_2 emissions by several percentage points. Fifth, I consider a counterfactual in which every country adds a tariff proportional to goods' CO_2 intensity, i.e., a carbon tariff. This has modest environmental benefits. Finally, if countries completely eliminated tariffs and NTBs, both global CO_2 emissions and real income would rise. Although turning off trade policy by definition eliminates trade policy's environmental bias, the resulting increase in income dwarfs this environmental effect.

This paper has potentially important policy implications. In a first-best setting where every country implemented uniform carbon prices on all CO_2 emissions, trade policy would have no role in efficient climate policy. In a second-best setting where political economy constraints make optimal climate change policy infeasible, considering environmental concerns in designing trade policy could potentially increase welfare. But in either setting, a trade policy which subsidizes CO_2 may be inefficient, and hence limiting the greater protection of clean relative to dirty goods could increase welfare. I believe that a reform which considers the CO_2 intensity of an industry in negotiating bilateral or multilateral trade policy across industries but without a formal carbon tariff has not been discussed in government or academia.[3] Such reforms may appeal to groups that typically disagree – dirty industries and environmentalists – because they can maintain protection of dirty domestic industries (at least relative to clean industries) while decreasing global CO_2 emissions. More broadly, the World Trade Organization (WTO) has sought to decrease protection of downstream relative to upstream industries, since such trade policy reforms would let developing countries sell more advanced technologies to industrialized countries. This paper suggests that such WTO goals may also help address climate change.

Several caveats are worth noting. This paper refers to the higher tariff and NTB rates on clean relative to dirty goods as an implicit "subsidy" in trade policy to CO_2 emissions. This "subsidy" refers to a lower tax rate in a setting where most goods face positive taxes (tariffs and NTBs). This difference in trade policy may encourage countries to purchase more clean goods domestically and dirty goods from abroad; internationally traded goods within an industry are more CO_2-intensive both because they require long-distance transportation and because they tend to be outsourced to countries like China and India that rely heavily on coal for production and so are CO_2-intensive (Shapiro 2016). The difference in trade policy also encourages firms and final consumers to substitute from consuming cleaner to dirtier goods (e.g., substituting from aluminum to steel). For these reasons and since the quantitative analysis finds that the difference in trade policy between clean and dirty industries increases global CO_2 emissions, I refer to this difference in trade policy as an implicit "subsidy." This is a global subsidy—for example, if France imposes low import tariffs on dirty goods, this may increase CO_2 emissions from French trading partners and from the globe overall, though could decrease these emissions from within France.[4]

[3] Such reforms are likely feasible within WTO regulations. The WTO does not primarily regulate NTBs, so most changes in NTBs are permissible. WTO members negotiate maximum ("binding") tariffs on trading partners. The binding tariffs do constrain the maximum possible level, but WTO members have flexibility in choosing tariffs below those levels through bilateral or multilateral agreements.

[4] A carbon subsidy through trade policy can also produce a range of behavioral responses, not all modeled here. For example, increasing the price of energy-intensive goods might lead consumers to use existing goods for longer, which could make the net effect on carbon emissions ambiguous. An analogous pattern in vehicle fuel economy standards is sometimes called the

It is also worth discussing the implications of using a second-best tool like trade policy as an alternative or complement to traditional environmental taxes on production or consumption. Important debates have considered the merits of taxing pollution through trade policy (e.g., Kortum and Weisbach 2016). One point of this paper is that current trade policy is subsidizing pollution for political economy (not efficiency) reasons, which no theoretical or empirical arguments claim is efficient.

This paper builds on several literatures. I believe this paper is the first to report the association of tariffs or NTBs with the pollution emitted to produce different goods, and the first to quantify the environmental consequences of harmonizing trade policy between clean and dirty goods. Research on trade and the environment asks how hypothetical changes in aggregate trade flows affect pollution, studies hypothetical carbon tariffs, or investigates how environmental policies and attributes of industries affect trade flows though not trade policies (Antweiler 1996; Copeland and Taylor 2003; Frankel and Rose 2005; Fowlie et al. 2016; Shapiro and Walker 2018). A large literature studies the consequences of hypothetical carbon border tax adjustments, relying primarily on computable general equilibrium (CGE) models and largely or completely abstracting from existing patterns of tariffs or NTBs. An entire field of academia, industrial ecology, quantifies the pollution required to produce internationally traded goods. Research in industrial ecology and economics measures pollution embodied in traded goods (e.g., Antweiler 1996; Davis and Caldeira 2010; Aichele and Felbermayr 2012). None of this work compares its measures of pollution embodied in traded goods against actual current levels of tariffs or NTBs.

This paper also introduces tariffs and NTBs as a new setting to study political economy and the environment. Research on the political economy of environmental policy is limited. Some work does use Grossman and Helpman (1994)'s "Protection for Sale" model to study domestic environmental policy (Fredriksson 1997; Schleich and Orden 2000). Trade policy provides an appealing setting to study political economy and the environment because it governs the more than 20 percent of CO_2 that crosses international borders embodied in traded goods, substantially affects pollution, creates easily-observed tax rates that vary across industries and countries, and depends on political economy forces like lobbying.

This paper also builds on an older trade policy literature by providing the first nonparametric evidence of "tariff escalation" – the phenomenon that more processed goods face higher tariffs – using continuous measures of upstreamness; the first evidence of NTB escalation, which is important since NTBs create a larger trade barrier than tariffs in industrialized countries; and the first empirical link between tariff escalation and the environment. Corden (1966, p. 228) in the Journal of Political Economy described tariff escalation as "so well known that detailed substantiation is hardly needed." Research on tariff escalation has since become uncommon, despite renewed interest in global value chains. While the existing literature generally identifies tariff escalation by reporting three mean tariff rates, for "primary," "intermediate," and "consumer goods" (Balassa 1965; Golub and Finger 1979; Marvel and Ray 1983), I propose that upstreamness provides a natural and continuous measure to use for studying escalation. Upstreamness is also related to the explanation for tariff escalation that downstream industries may lobby for low tariffs on their intermediate inputs (Cadot et al. 2004; Gawande et al. 2012). I am not aware

"Gruenspecht effect" (Jacobsen and van Benthem 2015).

of prior work interpreting tariff escalation through the full measure of upstreamness, though Gawande et al. (2012) relate it to the simpler measure of the share of an industry's output sold as intermediates. I show that upstreamness is more strongly associated with tariffs or NTBs than are many other standard explanations for trade policy. The most relevant recent other parts of the trade policy literature link trade policy to global value chains (Antràs and Staiger 2012; Blanchard et al. 2016) and link trade policy to other domains like the environment (Copeland 2000; Maggi 2016).

The paper proceeds as follows. Section 2 describes the data and Section 3 the econometrics. Section 4 discusses the relationship between pollution intensity and trade policy. Section 5 evaluates political economy explanations. Section 6 evaluates consequences of counterfactual reforms. Section 7 concludes.

II Data

I combine data on three types of variables: trade policy, pollution emissions, and political economy. Unless otherwise noted, all data represent a cross-section for the year 2007 (which is the year Exiobase covers) or the closest available year. I show some estimates with multiple years of U.S. data. Appendix A provides additional detail on each set of files, including concordances that link different industry classifications, and Appendix A.6 compares the use of industry classifications versus input-output tables for measuring tariffs on intermediate versus final goods.

II.A Trade Policy

Tariffs are the most easily-quantified trade policy instrument, but NTBs are increasing in importance. I obtain data on tariff rates from the Market Access Map (Macmap) database. A 2-digit Harmonized System (HS) code version of these data is freely available online. I purchased the 6-digit HS code version from the French *Centre d'Etudes Prospectives et d'Informations* (CEPII) (Guimbard et al. 2012). The data provide the most comprehensive tariff records available. The data distinguish 5,000 different goods (6-digit Harmonized System codes) for 190 countries and account for most-favored nation tariffs, regional trade agreements, free trade agreements, customs unions, and tariff-rate quotas. The data cover all bilateral trading partners.

For tariffs on U.S. imports, I use records from the Census Bureau's Imports of Merchandise files. While Macmap lists statutory tariff rates (i.e., official policy), Census records list tariff duties actually paid, so permit calculation of effective tariff rates.

Non-tariff barriers (NTBs) include policy barriers to trade that are not tariffs, such as price regulations, product standards, quantity restrictions like quotas, or others.[5] I use data from Kee et al. (2009) on the

[5]A global social planner might set tariff rates to zero, since tariffs largely exist for political economy or terms-of-trade reasons. A global planner might set some NTBs to non-zero rates, since some NTBs could address market failures in health, safety, or the environment. I abstract from efficiency rationales for NTBs in part since I am not aware of data distinguishing the extent to which each country and industry's NTBs are efficient versus reflect rent-seeking and protectionism. It is generally believed that NTB rates have risen in recent decades partly in response to decreased tariff rates, which would suggest that NTBs primarily represent protection rather than correction of market failures.

dollar (i.e., ad valorem) equivalent of NTBs; they describe how they calculate these values from raw data in the World Bank's World Integrated Trade Solutions (WITS) system. These NTB values are calculated for each 6-digit HS code, for a year around 2000-2003 (the exact year varies across countries), and for about 100 countries.

II.B CO_2 Emissions

I first explain my approach to measuring CO_2 emissions informally for one closed economy, then explain it formally, then discuss multiple open countries, and finally describe data sources.

Consider two types of CO_2 emissions. First, an industry burns fossil fuels to produce output. Second, an industry purchases intermediate goods as inputs that themselves require CO_2 emissions to produce. I describe the first channel as "direct" CO_2 emissions and the second as "indirect." An input-output table for one country contains one row per industry and one column per industry. Each value in the table represents the dollars of output from an industry in a row required to produce a dollar of output of the industry indicated in a column. This permits calculation of direct CO_2 emissions, since it shows how many dollars of coal, oil, and natural gas are required to produce a dollar of output in each other industry. To calculate direct CO_2 emissions, I consider the rows for the coal extraction, oil extraction, and natural gas extraction industries. The analysis uses independent data on the national price per physical unit of each fossil fuel and on the physical emissions rate (i.e., the tons of CO_2 emitted per ton of coal, barrel of oil, or cubic foot of natural gas burned). Multiplying these coal, oil, and gas input expenditures by the tons of CO_2 emitted per dollar of fossil fuel burned gives the direct emissions rate. This approach to using an input-output matrix to account for pollution is standard (Miller and Blair 2009, p. 447) and resembles what the Intergovernmental Panel on Climate Change calls the "Tier 1" or "default" method of calculating CO_2 emissions. It is designed to measure emissions from producing goods, which is appropriate for an analysis of tariffs on internationally traded goods.[6] It obtains industry-level measures of CO_2-intensity, though abstracts from intra-industry heterogeneity (Lyubich et al. 2018).

This approach can calculate direct but not indirect emissions. For example, the emissions rate for cookware in this approach reflects fossil fuels burned to shape steel into a pan (which are listed in the cookware industry column) but not fossil fuels used to make the steel in the first place (which are listed in the steel industry column or its input industries like electricity). As shown formally below, inverting the input-output matrix permits calculation of total emissions, which equal the sum of direct and indirect emissions. This inverse indicates the dollars of coal, oil, and natural gas required to produce a dollar of output in each industry, including the coal, oil, and natural gas embodied in intermediate goods, and inputs to intermediates, and inputs to these inputs, etc. Environmental researchers call this a "life cycle" or "footprint" measure of emissions; international economists call it a "value chain" measure.

[6] One could also wonder how domestic or "behind-the-border" policies affect the choice among energy-consuming durable goods like cars or air conditioners. While Section IV.B discusses sensitivity analyses designed to account for energy used in consuming these goods, a detailed analysis of energy consumption for these goods and associated policies is the topic of an active body of research that uses models specialized to these sectors (e.g., Bento et al. 2009; Jacobsen et al. 2019).

Continuing this explanation for a single closed economy, let S denote the number of industries in the economy and let A be an $S \times S$ input-output table where each row lists the industry supplying inputs and each column lists the industry demanding outputs. Each entry in the matrix A describes the dollars of input from the industry in a given row required to produce a dollar of output for the industry in a given column. Let x be an $S \times 1$ column vector describing each industry's gross output and let d be an $S \times 1$ vector of final demand, including exports. An accounting identity states that each industry's gross output equals the value of its output used for intermediate goods in all industries plus the value of its output used for final demand: $x = Ax + d$. Simple algebra then reveals the total amount of intermediate inputs (including both direct and indirect inputs) required to produce a dollar of final demand: $x = (I - A)^{-1}d$. The matrix $(I - A)^{-1}$ is called the Leontief inverse or the matrix of total requirements. It describes the dollars of each input, including those required to produce intermediate inputs, and inputs to inputs, etc. required to produce an additional dollar of final demand. This approach does not account for changes in CO_2 emissions from goods that are complementary with or substitutes for the good of interest, which may be most relevant for energy-consuming durable goods like vehicles or housing.

Extending this approach to multiple open countries and industries is straightforward. Let N denote the number of countries. In a multi-region input-output table, A is an $NS \times NS$ matrix, where each row is a specific country×industry and each column is a specific country×industry. For example, one table entry might show the dollars of Chinese steel (one row) required to produce a dollar of U.S. cookware (one column). Then x and d are $NS \times 1$ column vectors describing gross output and final demand, respectively. Using a multi-region input-output table, the rest of the analysis proceeds as above.

Several data sources help measure CO_2 emissions. The main dataset is Exiobase, which combines trade data, input-output tables, and national accounts to construct a global multi-region input-output table. Exiobase reports the direct CO_2 emissions per million Euros of output for every country×industry. To construct data on CO_2 emissions per country×industry, Exiobase primarily uses emissions data from the International Energy Agency (IEA 2007a,b,c). I use Exiobase's calculated CO_2 emissions from fossil fuel combustion for each country×industry. Appendix A.2 provides additional details on Exiobase.

I then calculate total (direct+indirect) emissions rates from Exiobase as follows. Let L_{ijst} denote an entry of the Leontief inverse $L = (I - A)^{-1}$, i.e., the dollars of output from industry s in country i required to produce one dollar of output from industry t in country j, including the entire global value chain (inputs, inputs to inputs, etc.). Let E_{is}^{direct} be the direct emissions from producing a dollar of output from country i in industry s, i.e., the CO_2 emitted from the coal, oil, and natural gas used directly in this country×industry. Exiobase reports E_{is}^{direct}. Then the total emissions rate is $E_{jt} = \sum_{i,s} L_{ijst} E_{is}^{direct}$.

II.C Political Economy Explanations

Why do different industries face different trade policies? One explanation involves optimal tariffs and the terms of trade—a large country can privately benefit by imposing small import tariffs on its trading partners. In this classic explanation, a country's privately optimal tariff equals the inverse of the foreign export supply elasticity it faces (Bickerdike 1907). Optimal tariffs could correlate with CO_2 intensity, since optimal tariffs are higher on more differentiated industries, and clean industries may be more differentiated. The second set of theories involves political economy. The more influential of these theories focus on organized interest groups (Olson 1965; Grossman and Helpman 1994, 1995; Maggi and Rodríguez-Clare 1998, 2007). Organized industries can provide campaign contributions and use other means to obtain trade protection. Politicians may find it privately optimal to distort trade policy in response to this lobbying, but in most settings a social planner would not.

Some political economy variables are available separately for each country$\times$industry in Exiobase; I extract these variables and use them for the global analysis. A larger set of political economy variables are available for each industry in U.S. data; I use these data to analyze the U.S. only. The introduction lists each variable; Appendix A.3 describes measurement of each variable and their data sources. I choose variables to include following existing empirical trade policy research (Pincus 1975; Caves 1976; Anderson 1980; Ray 1981; Marvel and Ray 1987; Trefler 1993; Freund and Çaglar Özden 2008), especially Rodrik (1995).

I add a measure of "local" air pollution emissions and damages not discussed in the trade literature. Firms' emissions of air pollutants, in addition to emissions of pollution through water and land, create local external costs. These externalities could lead to policies like low tariffs and NTBs on dirty industries which seek to relocate polluting activity to other countries.[7]

I discuss the one variable here which turns out to be the most important. I measure each industry's "upstreamness" as the average economic distance of an industry from final use. One can also interpret upstreamness as the mean position of an industry's output in a vertical production chain (Antràs and Chor 2013) or as the share of an industry's output sold to relatively upstream industries (Fally 2012). If industry i is measured to be more upstream than industry j, this does not imply that industry i actually supplies industry j. Rather, this simply implies that industry i on average is further in economic distance from final consumers than industry j is. Appendix A.3 presents the equation used to measure upstreamness and discusses its measurement. Upstreamness is measured for each industry in the U.S. data and each country$\times$industry in the Exiobase global data.

III Econometrics

III.A Trade Policy and CO_2 Intensity

To measure differences in trade policy between clean and dirty industries, I estimate the following:a

$$t_{js} = \alpha E_{js} + \mu_j + \epsilon_{js} \tag{1}$$

[7]Another interpretation is that many regions impose domestic local zoning restrictions that relocate dirty production from richer to poorer areas. Similarly, imposing low tariffs and NTBs on dirty goods could reflect wealthy countries' efforts to relocate dirty production to poor countries.

The dependent variable t is the mean import tariff rate or ad valorem NTBs that destination country j imposes on goods in industry s. In the global data, s represents the foreign industry which produced the good, not the domestic industry which consumed it. For example, the emissions rate E for Mexican imports of steel reflects the mean emissions from steel production in all countries from which Mexico imports, while the tariffs t reflects Mexico's import tariffs on steel. Equation (1) has a j rather than both i and j subscript because the analyses averages across origin countries (weighted by the value of each trade flow) for three reasons: this enhances comparability between tariffs and NTBs, since the latter are defined only by destination country and industry; this helps address the presence of zero trade flows between some origin×destination×industry tuples; and this increases comparability of these regressions with political economy variables, which are observed at the country×industry level. I show some results with separate observations for each exporter×importer×industry ($i \times j \times s$) tuple. Appendix B provides some additional details on equation (1).

The main explanatory variable, E, represents the tons of CO_2 emitted per dollar of imported good. As discussed earlier, E is calculated from inverting an input-output table, so includes both direct CO_2 emissions, which are those emitted from industry s, and indirect emissions, which are those emitted from industries used as inputs to industry s, and inputs to inputs, etc.[8]

Equation (1) allows a useful interpretation: the parameter α represents the carbon tariff implicit in existing trade policy. The regression has this interpretation because t is measured in dollars of tariff duties (or NTB equivalent) per dollar of imports and E is measured in tons of CO_2 per dollar of imports. Therefore α represents duties collected per ton of CO_2 emitted.[9] For example, $\alpha = 40$ would imply that an additional \$40 of import duties (or NTB ad valorem equivalent) is collected for each additional ton of CO_2 embodied in a good. My finding of $\alpha \approx -85$ to -120 implies that current trade policy embodies a carbon subsidy in trade policy of 85 to 120 dollars per ton of CO_2. As mentioned in the introduction, I refer to this as a "subsidy" in part because it represents a lower tax rate for traded dirty goods, even though it occurs in a setting where most traded goods face positive taxes.

Standard errors in most regressions are clustered by industry. I also report some results with standard errors clustered by importer. As discussed in Appendix B, the main estimates in the paper, including those of equation (1), include only observations for manufacturing.

[8] Formally, $E_{js} = (\sum_{i \neq j,t} E_{ijst} X_{ijst}) / \sum_{i \neq j,t} X_{ijst}$, where E_{ijst} is the emissions rate from inverting the global input-output table, and X_{ijst} is the value of the trade flow from origin country i and origin industry s to destination country j and destination industry t. The summation excludes $i = j$ because the emissions rate relevant for carbon tariffs and international trade applies only to international imports, not to intra-national trade. The emissions rate E_{ijst} differs by importer×exporter×industry. For example, the emissions rate for U.S. steel imports from China differs from the emissions rate for U.S. steel imports from Canada. These emissions rates differ because China and Canada use different fossil fuel inputs, both directly and indirectly.

[9] Imports appear to be in the denominator of both the left- and right- hand sides of equation (1), which could produce spurious correlation. In practice in the global data, as Section II.A explains, t is measured as statutory tariffs (or NTB equivalents). Hence, t reflects published regulations about which tariff rate applies to different types of products, and t is not measured through dividing data on duties collected by data on imports. In the U.S. data, t equals duties collected divided by imports. But as Section II.B explains, in the U.S. data, E equals emissions from a U.S. industry's production of CO_2 (including life cycle CO_2) divided by the industry's gross output. Additionally, the U.S. value of E from the input-output table is instrumented with its value from the direct survey MECS. Hence, the measurement of these variables in data limits the scope for bias from spurious correlation.

Measuring CO_2 intensity from an input-output table may involve two types of measurement error. The first is potentially relevant to all analysis with input-output tables—the input-output table itself has errors-in-variables. Constructing an input-output table requires judgments of analysts from national statistical agencies and adjustment through linear programming (Horowitz and Planting 2006). Second, prices paid for each fossil fuel vary by industry, and input-output tables lack data on such industry-specific input prices. Both types of measurement error could attenuate OLS estimates of α.

To address potential measurement error in measures of CO_2 intensity, I use direct emissions as an instrumental variable for total emissions. The first-stage regression is

$$E_{js} = \beta E_{js}^{direct} + \mu_j + \eta_{js} \tag{2}$$

The second stage is equation (1). Here E_{js} measures total (direct+indirect) emissions from the input-output table and E_{js}^{direct} measures direct emissions. Direct emissions reflect fossil fuels used in industry s but not fossil fuels embodied in intermediate goods used in industry s. For example, the direct emissions for producing a car include the natural gas used to heat and weld the car parts together at the car factory but not the coal used to produce steel that is then shipped to the car factory.

For the U.S. data, the instrument is the direct emission rate, measured from the Manufacturing Energy Consumption Survey and the Census of Manufactures. For the global data, the instrument is the direct emissions rate in the 10 smallest other countries. The validity of such leave-out instruments can be less clear than the validity of some other types of instruments, in part due to concerns about the reflection problem (Manski 1993). Due to the possibility that measurement error persists in the global estimates, the true global subsidy in trade policy may be larger in absolute value than what I estimate.

III.B Political Economy Explanations

I then test the hypothesis that the association between trade policy and CO_2 intensity reflects variables that are omitted from equation (1) but that both determine trade policy and correlate with CO_2 intensity. I estimate linear regressions including potential variables F_{js} that are believed to explain trade policy, along with CO_2 intensity:

$$t_{js} = \beta E_{js} + \pi F_{js} + \mu_j + \epsilon_{js} \tag{3}$$

I estimate a separate regression for each political economy variable F_{js} then assess which of these political economy variables most attenuates the estimated covariance β between trade policy and carbon intensity.

In separate estimates, I control for all potential political economy explanations at once. I implement this regression using both linear regression and using the least absolute shrinkage and selection operator (Lasso), which is a common machine learning algorithm for automatic model selection (Tibshirani 1996). Identifying which variables Lasso includes in a model can be informative though also sensitive to specification (Mullainathan and Speiss 2017). These regressions test whether each variable, including CO_2 intensity, has additional explanatory power for trade policy beyond these other variables.

IV Results: Trade Policy and CO_2 Intensity

IV.A Summary Statistics

Table I describes the cleanest and dirtiest industries in the global data, ranked by total (direct+indirect) CO_2 emissions. Panel A shows the cleanest five industries while Panel B shows the dirtiest. Column (1) shows mean CO_2 rates across all countries, column (2) shows mean tariffs, and column (3) shows NTBs.

The cleanest five manufacturing industries primarily produce food products and have a mean global emissions rate of 0.37 tons CO_2 per thousand dollars of output. The dirtiest five manufacturing industries mostly produce heavy goods like bricks or steel and have a mean global emissions rate of 1.88 tons CO_2 per thousand dollars of output. Motor vehicles appear relatively clean in these data (also in U.S. input-output tables) because, as discussed earlier, most of the emissions due to vehicles come from a separate good that is complementary, refined petroleum, and later I explore estimates accounting for this complementarity.[10]

It may be informative to calculate the CO_2 externality these numbers imply. If each ton of CO_2 emitted creates a social cost of carbon of \$40 (IWG 2016), this comparison involves multiplying by 40/1000. This calculation implies that globally, pork products create a social cost from CO_2 emissions of about 1.5 percent of product value (=0.34*40/1000). Producing iron and steel creates a CO_2 externality equal to 7 percent of its product value (=1.74*40/1000).

In the ten industries of Table I, column (2) shows that the cleanest industries face over four times the mean tariff of the dirtiest industries, at 9 versus 2 percent. Column (3) shows a similar difference between the cleanest and dirtiest industries for NTBs (25 versus 5 percent). I now turn to regressions analyzing all industries.

IV.B Implicit Carbon Tariffs

Tariffs

Figure II, Panels A and B, plots hypothetical \$40/ton carbon tariffs. Each point in these graphs is a separate country×industry (Panel A, all countries) or industry (Panel B, U.S. only). The tariff rate is a constant multiple of the emissions rate, which makes both graphs linear. In this hypothetical policy, the mean carbon tariff for all countries in Panel B is three percent, which is slightly over half of current global mean tariff rates. The mean U.S. carbon tariff is about four percent, which is larger than prevailing mean U.S. tariffs (Table II).

Figure II, Panels C and D, shows actual tariff data. In these graphs, the pattern across industries is the opposite of hypothetical carbon tariffs. The hypothetical carbon tariffs in Panels A and B impose

[10] Some developing countries directly subsidize the consumption of raw fossil fuels; trade policy also to some extent reflects these patterns. For global trade in fossil fuel industries (coal, crude oil, natural gas, and refined petroleum), average global tariffs are 1.7 percent and NTBs are 3.6 percent; for all other industries, these averages are 3.9 percent and 9.6 percent, respectively. In developing countries, mean tariff and NTB rates for fossil fuels are 5.8 percent and 6.7 percent, and for all other industries they are 7.5 percent and 9.0 percent. These values are weighted and include all industries (not only manufacturing).

higher tariffs on dirtier industries (positively sloped line), but actual tariffs in Panels C and D impose lower tariffs on dirtier industries (negatively sloped line).

Table II reports regressions corresponding to these graphs. Panels A and B show estimates for the world and U.S., respectively. Odd-numbered columns are unweighted; even-numbered columns are weighted by the value of the trade flow. For the U.S., weighting provides an efficient response to heteroskedasticity, since U.S. effective tariff rates equal total duties divided by total trade value. Columns (1) and (2) show a first-stage regression of total CO_2 intensity on direct CO_2 intensity, corresponding to equation (2). Columns (3) and (4) show reduced-form regressions of tariffs on direct CO_2 intensity. Column (5) and (6) show OLS regressions of tariffs on total CO_2 intensity. Columns (7) and (8) report instrumental variables regressions of tariffs on total CO_2 intensity, instrumented by direct CO_2 intensity.

In Table II, Panel A, the negative signs in columns (3) through (8) imply that global tariffs have a subsidy in trade policy to CO_2 emissions, not a tax. Columns (7) and (8) show that the mean subsidy to CO_2 emissions in global tariffs is \$11 per ton of CO_2 weighted, or \$32/ton unweighted. The first-stage F-statistics show that most of the instruments are strong, though the unweighted U.S. estimates have marginally weak instruments (F-statistic of 9.8, versus a standard cutoff of 10), and hence are possibly biased towards OLS. The instrumental variables estimates are modestly larger than the corresponding OLS estimates, which is consistent with attenuation bias in OLS due to measurement error, though their qualitative results are similar.

Figure III shows the estimated association between CO_2 intensity and tariffs for the U.S., separately for each year of available data 1989-2017. The red circle shows the point estimate for each year and the vertical bar shows the 95 percent confidence interval. This graph shows statistically significant negative associations between U.S. tariffs and CO_2 intensity in every year. The estimated U.S. subsidy in trade policy was \$13/ton in 1989, then decreased gradually to \$6/ton around 1998, and remained near that value through 2017. One potential explanation is that the secular decline in mean tariffs overall decreased the difference between tariffs on clean and dirty goods.

Non-Tariff Barriers

Figure II, Panels E and F, plots NTBs against CO_2 emission rates. These graphs have similar structure to the tariff graphs. They show that dirtier industries face lower NTBs in both the global and U.S. data. Some of the cleanest industries have NTB ad valorem equivalent values close to 100 percent, while many of the dirtiest industries face little or no NTB protection.

Table III reports regressions corresponding to these graphs. The table structure is similar to the tariff regressions in Table II. Again the numbers in columns (3) through (8) are all negative, showing a carbon subsidy in trade policy rather than carbon tax. Columns (7) and (8) show that the implicit subsidy to CO_2 in global NTBs is \$76 in the unweighted regressions or \$90 in the weighted regressions. The instrumental variables estimates in columns (7) and (8) show a large subsidy to CO_2 emissions implicit in U.S. NTBs, of about \$37 to \$48/ton. Summing up subsidies in tariffs from Tables II and NTBs from Table III, columns (7) and (8), gives the global subsidy that I emphasize of about \$85 (weighted) to \$120 (unweighted).

These implicit subsidies appear in both tariffs and NTBs but have larger magnitude in absolute value in NTBs, perhaps in part since NTB mean values are greater. The mean U.S. ad valorem equivalent of NTBs is 8 to 11 percent, which is over four times the mean tariff rate (Tables II and III). This supports the common claim that U.S. NTBs are more restrictive than U.S. tariffs. Globally, NTBs create a larger barrier to trade than tariffs do, at 9 to 13 percent (NTBs) versus 3 to 5 percent (tariffs).

Implicit Subsidies in Trade Policy, by Country

To investigate how these patterns vary by country, I sum together tariffs and the ad valorem equivalent of NTBs as a more complete measure of protection. I then estimate equation (1) separately for each country (hence, these regressions exclude country fixed effects).

Figure IV plots the result. Each point in this graph describes an estimate of the implicit carbon subsidy in trade policy for one country. Each point represents the subsidy to global emissions implicit in the trade policy of one country. The point for each country is estimated separately.[11] Points on the graph are ordered by the estimated implicit subsidy, with names shown for several countries of interest.

Almost every country in Figure IV has a negative value, implying that most countries have a carbon subsidy rather than a carbon tariff implicit in trade policy. European countries like France, Germany, Norway, and the UK have among the largest such subsidies in trade policy, with subsidy values exceeding \$175/ton. Russia, India, and China have smaller subsidies in trade policy. The y-axis of Figure IV shows each country's implicit subsidy. A country like China which has high emissions due to its reliance on coal can still nonetheless have a small value in Figure IV since its trade policies are not strongly correlated with industries' CO_2 emissions. The two regions with positive values in Figure IV are Romania and the Rest of the Middle East. Figure V plots these data in a global map which classifies countries by their subsidies implicit in trade policy.

The cross-country comparisons in Figures III and IV do not follow predictable patterns. Large subsidies in trade policy appear in both rich regions like the EU and poor regions like Africa; small subsidies also appear in both rich countries like Canada and poorer countries like Vietnam. Oil-intensive countries like Saudi Arabia and Iran have small subsidies in trade policy, while countries with strict environmental policies like Norway have large subsidies. This lack of patterns is consistent with the interpretation of Section 5 that these subsidies in trade policy are due to political economy forces which are correlated with CO_2 intensity.[12]

[11] Stacking the regression to account for the covariance structure across countries might increase efficiency in these estimates. Most of the estimates are significantly different from zero, though not significantly different from each other. I view the most striking feature of Figure IV as the fact that even with completely separate regressions for each country, most of the country-specific subsidies in trade policy are negative and large in absolute value; stacking the regression makes the estimates no longer independent across countries.

[12] In unreported results, I took the estimated country-level subsidy to CO_2 in trade policy plotted in Figures III and IV, and regressed it on several country characteristics. This regression finds that a country's GDP per capita, its mean tariff rate, and its quality of environmental management are all significantly associated with larger subsidies in trade policy (more negative regression coefficients). The regression also controlled for mean NTB rates, mean CO_2 emissions rates, an index of perceived

Sensitivity Analyses and Extensions

Appendix C discusses numerous alternative estimates of these implicit carbon tariffs that are shown in Appendix Table I, including from a tobit, alternative approaches to inference, nonlinear specifications of CO_2 intensity, winsorizing the data, including non-manufactured goods, including intra-national trade, separating direct and indirect emissions, including all greenhouse gases, separately accounting for CO_2 emissions from consumption and not merely production (e.g., the gasoline used to power a vehicle), the reverse regression, using the World Input-Output Database (WIOD), excluding manufactured agricultural and food products, and specifically analyzing the recent trade war by focusing on recent changes in U.S. import tariffs. Most of these results are qualitatively similar to the main estimates, though some vary in their magnitudes.

One interesting pattern in Appendix Table I, rows 26-28, is that the implicit subsidy persisted in U.S. tariff data through the year 2017. Tariffs imposed in the 2018 trade war (Fajgelbaum et al. 2020) slightly attenuated but did not eliminate this implicit subsidy. Even in these more recent U.S. data, controlling for upstreamness eliminates the estimated association between trade policy and CO_2 intensity.

Appendix C describes an additional analysis which suggests these implicit subsidies appear in both cooperative tariffs, such as those negotiated through the WTO, and non-cooperative tariffs, such as those the U.S. lists for trade with North Korea.

V Explanations for the Relationship Between Trade Policy and Pollution

Why do countries impose higher tariffs and NTBs on clean than on dirty goods? The existence of these subsidies in trade policy is surprising, so the question of why they exist is interesting. Additionally, because no prior research has tested for or demonstrated the existence of these subsidies in trade policy, explaining why they exist enhances their plausibility. Finally, understanding why these patterns of trade policy occur may provide insight into the political feasibility of changing them.

V.A Explanations: Omitted Variables

Which are the most important omitted variables in regressions of trade policy on CO_2 intensity? Appendix Table III shows that an industry's upstream location is likely to play an important role. The table shows the difference in each political economy variable between "dirty" and "clean" industries (i.e., those above and below the median CO_2 intensity), separately for global and U.S. data. All political economy variables are expressed in z-scores (i.e., I subtract the mean and divide by the standard deviation). Relative to

country corruption, and the country's mean upstreamness; these other variables had marginally significant (upstreamness) or no (other controls) association with the level of a country's implicit subsidy in trade policy. I do not show this cross-country, cross-sectional regression, which has 7 explanatory variables and less than 50 observations, since it may be hard to interpret economically; I mention it because it provides another way to summarize the data in Figures IV and V.

clean industries, dirty industries are significantly more upstream, have a lower labor share, lower wages, higher unionization rates, higher shipping costs, and higher local pollution emissions; the global data give conflicting patterns for intra-industry trade and the import penetration ratio. While dirty industries anecdotally have outsize political influence, dirty industries make marginally lower PAC contributions, though PAC contributions are believed to be a very imperfect measure of lobbying influence.

Appendix Table III shows that the association between emissions and upstreamness is stronger than the association between emissions and other political economy variables in U.S. and global data. All variables are in z-scores so have the same units. In the global data, some of the other variables are correlated with dirtiness, but the correlations are weaker than for upstreamness. The association between emissions and upstreamness in the global data is over four times stronger than the association between emissions and other political economy variables. Additionally, the regressions below imply that these other variables have less strong direct relationships to trade policy than upstreamness.

Table IV asks which political economy explanation is the most important omitted variable in regressions of trade policy on CO_2 intensity. It shows regressions of trade protection (tariffs+NTBs) on total CO_2 intensity while controlling for one political economy variable at a time, with specification corresponding to equation (3). Total CO_2 intensity is instrumented with direct CO_2 intensity. Panels A and B show estimates for all global trade; Panel C shows U.S. estimates. Column (1) includes no controls. Columns (2) through (6) each control for one political economy variable, observed at the level of a country×industry. Column (2) controls for upstreamness, column (3) for intra-industry trade, column (4) for the import penetration ratio, column (5) for the labor share, and column (6) for the mean wage.

Table IV, Panel B, uses data from the ten smallest other countries to construct instrumental variables for the focal country×industry. The regressions in Panel B have two instruments (direct CO_2 and political economy variables, both averaged across the ten smallest other countries) and two endogenous variables (CO_2 and a political economy variable, both for the focal country). These help address the possibility that some political economy explanations are endogenous. One example would be if trade policy affects wages in a given industry and country but not in the same industry in other countries. Analyses of agglomeration and import competition similarly use somewhat similar instruments (Ellison et al. 2010; Autor et al. 2013; Antràs et al. 2017).

Table IV, Panel A, column (1) restates the earlier result that the total subsidy to global CO_2 emissions implicit in global trade policy is around \$120/ton. Column (2) shows that controlling for upstreamness attenuates this estimate to \$33/ton. Columns (3) through (6) show that controlling for other political economy variables one at a time only slightly changes the estimated implicit subsidy in trade policy.

Table IV, Panel B, obtains similar estimates from instrumenting each political economy variable with its mean in the ten smallest other countries. In column (2), controlling for upstreamness eliminates the estimated implicit subsidy—the estimated association between CO_2 emissions and trade policy is -\$120 (34) with no political economy controls, but \$34 (39) when controlling for upstreamness. Columns (3) through (6) show that instrumenting does not substantially change the other estimates. These estimates have strong instruments.

Panel C finds similar patterns using U.S. data. The estimated U.S. subsidy from tariffs and NTBs is \$50 (10) per ton. Controlling for upstreamness attenuates this estimate, to \$3 (10) per ton. Other political economy controls do not substantially change the estimated subsidy in trade policy.

Figure VI graphs the U.S. estimates from Table IV, along with estimates controlling for other political economy variables that are available for the U.S. but not all countries. Each blue circle in these graphs is the coefficient from a regression of tariffs+NTBs on total CO_2 intensity (instrumented by direct CO_2 intensity), controlling for one political economy variable, and corresponding to equation (3). Each red horizontal line shows a 95% confidence interval. The "Main Estimates" restates results from Table IV, Panel C, column (1). Each of the other numbers controls for one additional variable. The "Firm size: mean" entry, for example, comes from a regression that controls for the mean firm size in each industry.

Figure VI shows that controlling for most political economy variables one-by-one produces little change in the association of trade policy with CO_2 intensity. Only one variable, upstreamness, eliminates the estimated implicit subsidy in trade policy, and renders it statistically indistinguishable from zero.[13]

Appendix C.1 discusses a wide range of sensitivity analyses, which collectively suggest that upstreamness accounts for an important share of the association between CO_2 intensity and trade policy.

Figure VI and Appendix Table V suggest that local pollution does not statistically account for the association between CO_2 emissions and trade policy, since controlling for local pollution emissions or damages does not substantially change the coefficient on CO_2 intensity, though does decrease its precision. A few additional reasons suggest why concern for local pollution emissions is unlikely to be the primary explanation for why dirty industries face lower tariffs and NTBs. First, many policymakers seek to maintain dirty industries' domestic production and directly regulate local pollution emissions by requiring installation of scrubbers or other technology that abates local air pollution (though not CO_2). Many environmental policies contain explicit provisions to prevent relocation of dirty production. Hence, relocating dirty industries abroad may not be a primary policy goal. Additionally, I am not aware of evidence that concern for local pollution emissions has led dirty industries to have lower tariffs or NTBs. Many trade agreements like NAFTA and TPP have side agreements dealing with the environment, but these agreements typically describe domestic environmental regulations or monitoring investments, not patterns of tariffs and NTBs. Many actually seek to prevent the relocation of dirty industries, by barring the use of weak domestic environmental policies to lure dirty production across borders. Moreover, this implicit subsidy in trade policy appears in most countries. Efforts to outsource local pollution would thus to some extent neutralize each other.

[13]This section's comparison of the carbon content of goods against their upstreamness and trade policy has similarities to Blanchard et al. (2016)'s value-added content logic that a country may choose trade policy for a good to reflect the domestic content which is embodied in the value chain for that good. One important difference is that each country may have preference over policy for its own domestic content embodied in traded goods. For CO_2 externalities, however, it does not matter whether the CO_2 embodied in a good was originally emitted from domestic or foreign fossil fuels, since CO_2 has the same effect on global climate regardless of the location of its emission.

V.B Explanations: Empirical Reasons why Upstreamness Substantially Accounts for Subsidies in Trade Policy

Why is an industry's upstreamness strongly correlated with its CO_2 intensity? Using U.S. data from the Bureau of Economic Analysis, Appendix Figure III graphs the share of each industry's revenue accounted for on the production side by intermediate goods, labor expenditures, profits and taxes, and fossil fuels. Appendix Figure III shows that upstream industries use a larger share of fossil fuels than downstream industries do. For the upstream industries, nearly five percent of production costs are devoted to fossil fuels; for the most downstream industries, less than one percent of costs are. Relative to upstream industries, downstream industries spend relatively more on labor and intermediate goods. Previous research has not shown these patterns but they make intuitive sense—upstream industries are taking raw materials extracted from the ground and transforming them, while downstream industries depend more on labor and other inputs.

Appendix Figure III also helps answer an important question. If downstream goods are just combinations of upstream goods, why would different import tariff rates on upstream versus downstream goods affect CO_2 emissions? Imagine an economy in which upstream goods were made exclusively from coal and downstream goods were made from upstream goods. In this hypothetical economy, upstream and downstream goods would have the same CO_2 intensity, and tariff escalation could not affect global CO_2 emissions. Appendix Figure III shows that this hypothetical economy is misleading because downstream industries use as inputs both upstream goods and relatively clean factors like labor. Hence, imposing high tariffs on downstream but not upstream goods can encourage consumers to substitute from demanding relatively clean factors like labor to demanding relatively dirty factors like energy.

Buyers can respond to changes in trade policy in many ways, including substituting between goods, changing total demand for an industry's products, and changing trading partners. To what extent can firms and consumers substitute between industries with different levels of upstreamness? Certainly in examples, goods that are substitutes have different levels of upstreamness and CO_2 intensity. For example, steel and aluminum are likely substitutes, and in the U.S. data which have greater industry detail, steel is both more upstream and more CO_2 intensive than aluminum.[14]

Appendix D informally discusses how theories of trade policy might rationalize these findings, and describes a few reasons why the Diamond and Mirrlees (1971) production efficiency theorem does not well account for these patterns of trade policy.

[14] The example compares iron and steel mills (NAICS industry 331111) against aluminum sheet, plate, and foil manufacturing (NAICS industry 331315).

VI Consequences of Implicit CO_2 Subsidies in Trade Policy

VI.A Partial Equilibrium Approximation

I use a few approaches to investigate the aggregate consequences of these patterns of trade policy. The first is a partial equilibrium calculation:

$$\hat{\alpha} \sum_{j,s} E_{js} \sum_{i \neq j} X_{ijs} \tag{4}$$

where $\sum_{i \neq j} X_{ijs}$ represents the value of international imports by country j in sector s. This represents the revenue that a carbon tariff would collect if it had the same pattern as trade policy's environmental bias (i.e., -\$85 to -\$120/ton). The parameter $\hat{\alpha}$ is the implicit carbon subsidy in trade policy from equation (1).

My primary estimate of Equation (4) uses regression results from Tables II and III, columns (7) and (8). This implies that global trade policy provided an implicit subsidy of \$550 to \$800 billion in the year 2007 (measured in 2016 dollars). This can be calculated simply: 6.5 billion tons of CO_2 are embodied in international trade (including in intermediate goods), times \$85 to \$120 in subsidy per ton of CO_2 traded, gives $\$550 \approx 6.5 \times 85$ or $\$800 \approx 6.5 \times 125$.[15] Appendix Table I provides other estimates of the implicit subsidy α. These other regression estimates in turn lead to other estimates of the total global subsidy in trade policy. While the exact subsidy can vary with the regression specification, these magnitudes suggest this subsidy may have quantitatively important effects on trade and CO_2 emissions.

To put these estimates in perspective, global direct subsidies to fossil fuels were about \$530 billion in 2007 (IMF 2013). These direct subsidies are a focus of political debate. The CO_2 subsidies in trade policy, which have not been previously highlighted, have a similar magnitude. Of course, a direct subsidy to fossil fuel could have larger effects on CO_2 than an indirect subsidy through trade policy.

VI.B Analytical Model

The calculation of the previous subsection is simple but has limitations. I now turn to a model including several potentially important features—pollution can directly affect utility; pollution creates transboundary damages; countries may have pre-existing and sub-optimal trade policy on any goods; tariffs generate revenue which is lump-sum redistributed; and industries are connected through input-output links, so dirty industries can be upstream. This subsection describes a setting with two symmetric countries and two industries (one clean, one dirty). The next subsection describes a fuller quantitative model that

[15] In many regression settings, difference-in-difference analyses cannot measure the total effect of a policy, since such regressions are normalized against a comparison group and have fixed effects that remove any economy-wide effects. My summary calculations reflect a descriptive regression that is not differences-in-differences—the regression has no comparison group, and is not estimating a causal effect. Hence, this partial equilibrium calculation assumes that goods with zero tariff have zero subsidy.

impose fewer restrictive assumptions, but which thereby obtains numerical but not analytical results. These models make strong assumptions that are not literal descriptions of reality, like constant elasticity of substitution utility, but the benefit of these stylized descriptions is that they permit analysis of how specific counterfactual trade policy reforms affect CO_2 emissions and social welfare.

Preferences. The representative agent in country j maximizes national utility U_j.

$$U_j = \prod_s Q_{js}^{\beta_{js}} f(Z) \tag{5}$$

Here Q_{js} is a consumption aggregate given by $Q_{js} \equiv (\sum_i q_{ijs}^{(\sigma-1)/\sigma})^{\sigma/(\sigma-1)}$. The elasticity of substitution is $\sigma > 1$, which for simplicity here does not vary by sector. Utility depends on international trade (q_{ijs}, $i \neq j$) and intra-national trade (q_{jjs}) in each sector $s \in (1, 2)$. Global pollution emissions $Z = \sum_j Z_j$ create multiplicative damages $f(Z)$. The representative agent treats emissions as a pure externality, so ignores them in choosing expenditure. This analytical model does not need to specify the functional form of $f(Z)$, as I discuss below. Sector 1 represents dirty goods that emit pollution; sector 2 represents clean goods that do not. Preferences are Cobb-Douglas across sectors, with expenditure shares β_{js}. I describe an Armington model, in which countries have a taste for variety, and each country produces one variety per sector. Standard versions of Ricardian models with richer microfoundations, like Eaton and Kortum (2002) or Dornbusch et al. (1977), would produce the same aggregate equilibrium equations describing production, consumption, and trade (Arkolakis et al. 2012), and thus the same conclusions for emissions and social welfare.

The associated price index is

$$P_{js} = \left[\sum_i (c_{is}\phi_{ijs})^{-\epsilon} \right]^{-1/\epsilon} \tag{6}$$

Equivalently, one can write the trade elasticity $\epsilon > 0$ as $\epsilon = \sigma - 1$, where σ is the elasticity of substitution between goods from each country. The variable ϵ can be interpreted as the elasticity of trade flows with respect to trade costs. Goods face multiplicative trade costs $\phi_{ijs} \geq 1$. Trade costs may include a general iceberg friction, tariffs, and non-tariff barriers: $\phi_{ijs} = \tau_{ijs}(1 + t_{ijs})(1 + n_{ijs})$. Here $\tau_{ijs} \geq 1$ are iceberg trade costs, so τ goods must be shipped for one to arrive. Buyers pay bilateral import tariffs $t_{ijs} \geq 0$. Tariff revenues are lump-sum rebated to domestic consumers. I treat NTBs $n_{ijs} \geq 0$ as comparable to an iceberg trade cost. Intra-national trade costs equal one: $\phi_{jjs} = 1$.

Technology. The unit cost of producing goods is Cobb-Douglas in labor, which is sold at price w_i, and in intermediate goods from other sectors:

$$c_{is} = w_i^{1-\alpha_{is}} \prod_k P_{ik}^{\alpha_{iks}} \tag{7}$$

Here α_{is} is the labor share and α_{iks} is the cost share of industry k to produce output in country i and industry s.

Pollution. The pollution emitted to sell goods to a particular destination j is:[16]

$$Z_{ij1} = \frac{X_{ij1}}{c_{i1}(1+t_{ij,1})} \tag{8}$$

One could think of c_{is} as the factory-gate price. One country's pollution emissions are $Z_i = \sum_j Z_{ij1}$, and global pollution emissions are $Z = Z_1 + Z_2$. Here X_{ijs} is total expenditure on goods produced in origin country i, shipped to destination country j, in sector s.[17] Sales are deflated by the unit production cost c_{is}.

Trade flows. Consumer utility maximization implies the following international trade flows:

$$\begin{aligned} X_{ijs} &= \left(\frac{c_{is}\phi_{ijs}}{P_{js}}\right)^{-\epsilon} X_{js} \\ &= \lambda_{ijs} X_{js} \end{aligned} \tag{9}$$

where $\lambda_{ijs} \equiv X_{ijs}/X_{js} = (c_{is}\phi_{ijs}/P_{js})^{-\epsilon}$ denotes the share of country $j's$ expenditure on sector s varieties which is sourced from country i , and $X_{js} = \sum_i X_{ijs}$ is total expenditure on sector s goods in country j.

Equilibrium. In baseline data and counterfactuals, consumers maximize utility, firms maximize profits, and markets clear. Trade is balanced, so each country's revenues equal its expenditures. Global GDP is the numeraire, so $\sum_i Y_i^{'} = 1$, where Y_i is a country's factor payments. Market clearing for labor is $L_i = \sum_s L_{is}$, where L_{is} is labor supply, which I assume is inelastic. Total expenditure on goods in a country×sector equals the sum of expenditure on final and intermediate goods:

$$X_{js} = \beta_{js}(Y_j + T_j) + \sum_k \alpha_{jsk} R_{jk} \tag{10}$$

Here $X_j = Y_j + T_j$ is total expenditure, which comes from factor payments and from tariff revenues $T_j = \sum_{i,s} X_{ijs}t_{ijs}/(1+t_{ijs})$. Tariff revenues are lump-sum redistributed to the representative agent. The term R_{is} represents country×sector revenues, given by $R_{is} = \sum_j X_{ijs}/(1+t_{ijs})$.

Counterfactual Methodology To study counterfactuals, I express each variable in changes from baseline levels, sometimes called "exact hat algebra" (Dekle et al. 2008; Costinot and Rodriguez-Clare 2014). Let a denote a variable from the model in the baseline data and a' the value in a counterfactual. Define the proportional change in this variable due to counterfactual policy as

$$\hat{a} \equiv \frac{a'}{a} \tag{11}$$

Counterfactual Results and Interpretation: Pollution. Equation (9) implies that the change

[16] Pollution here is isomorphic to the consumption of fossil fuels. While analyzing air and water pollution is more complex and involves a range of abatement technologies firms can install, because end-of-pipe abatement technologies like carbon control and sequestration for fossil fuels are not economically viable, knowing the fossil fuels consumed is sufficient to measure the CO_2 emitted (Shapiro and Walker 2020).

[17] The right-hand side of equation (8) has 1 rather than s subscripts because only dirty goods (sector 1) emit pollution.

in expenditure shares due to a counterfactual trade policy is

$$\hat{\lambda}_{ijs} = \left(\frac{\hat{c}_{is}\hat{\phi}_{ijs}}{\hat{P}_{js}} \right)^{-\epsilon} \tag{12}$$

Similarly, equation (6) implies that the change in country×sector price index is

$$\hat{P}_{js} = \left[\sum_i \lambda_{ijs} (\hat{c}_{is}\hat{\phi}_{ijs})^{-\epsilon} \right]^{-1/\epsilon} \tag{13}$$

To derive the change in pollution emissions, I start from $\hat{Z} = Z^{'}/Z$, then substitute in the definitions of pollution, trade flows, and counterfactual changes, from equations (8), (9), and (11):

$$\hat{Z}_i = \frac{\hat{X}_{i1}}{\hat{c}_{i1}} \left[\frac{\hat{\lambda}_{ii1} Z_{ii1} + (\widehat{1+t_{ij1}})^{-1} \hat{\lambda}_{ij1} Z_{ij1}}{Z_i} \right] \tag{14}$$

To obtain another useful version of this result, I substitute in the equations for the changes in expenditure shares and the price index from equations (12) and (13):

$$\hat{Z}_i = \frac{\hat{X}_{i1}}{\hat{c}_{i1}} \left(\frac{\lambda_{ii1} + (\hat{\phi}_{ij1})^{-\epsilon} \frac{\lambda_{ij1}}{1+t^{'}_{ij1}}}{\lambda_{ii1} + (\hat{\phi}_{ij1})^{-\epsilon} \lambda_{ij1}} \right) \frac{1}{\lambda_{ii1} + \frac{\lambda_{ij1}}{1+t_{ij1}}} \tag{15}$$

The change in emissions due to a counterfactual policy equals the product of two terms: the change in real expenditure on dirty goods ($\hat{X}_{i,1}/\hat{c}_{i,1}$) and change in the pollution intensity of expenditure.

This interpretation is analogous to decomposing pollution into composition and technique effects. Scale is often part of such decompositions, but here is fixed by choice of the numeraire, at least in terms of nominal GDP. Equation (8) shows that pollution intensity in revenue terms is normalized to one, so that one dollar of dirty goods revenues always produces one unit of pollution. Because tariffs create a wedge between expenditures and revenues, pollution intensity in expenditure terms is less than one.

In equations (14) and (15), one can separate the change in pollution intensity into two channels that represent domestic and foreign spending. In equation (14), both appear in the numerator of the term in brackets—the first term reflects the counterfactual emissions from domestic sales and the second term reflects the counterfactual emissions from exports. In equation (15), the denominator of the term in parentheses equals the inverse of the change in the domestic expenditure share for dirty goods ($\hat{\lambda}_{111}^{-1}$). This denominator is less than one, so represents a channel by which a counterfactual policy increases emissions intensity. This reflects the idea that increasing tariffs on dirty goods increases pollution emitted to produce dirty goods sold domestically, which tends to increase the emissions intensity of expenditure. The magnitude of this effect grows as the trade elasticity (ϵ) increases, because this elasticity reflects how trade flows respond to trade costs, including tariffs.

In the numerator of the term in parentheses equation (15), the second term reflects emissions for exports. This term shows that increasing tariffs on dirty goods discourages trade in dirty goods. This result is intuitive and tends to decrease the emissions intensity of expenditure. This channel also becomes more important as the trade elasticity becomes larger.

The change in real spending on dirty goods appears in the first ratio of equations (14) and (15). In that ratio, the change in expenditure follows directly from the country×sector expenditure equation (10):

$$\hat{X}_{js} = \frac{\beta_{js}(Y_j + T_j') + \sum_k \alpha_{jsk} R_{jk}'}{\beta_{js}(Y_j + T_j) + \sum_k \alpha_{jsk} R_{jk}} \tag{16}$$

The change in unit costs follows from substituting the price index into the unit cost equations in changes for each industry, then solving this system of equations:

$$\hat{c}_{11} = [\lambda_{111} + \lambda_{211}(\hat{\phi}_{211})^{-\epsilon}]^{-\frac{1}{\epsilon}\frac{\alpha_{121}\alpha_{112}+(1-\alpha_{122})\alpha_{111}}{(1-\alpha_{111})(1-\alpha_{122})-\alpha_{112}\alpha_{121}}} [\lambda_{112} + \lambda_{212}(\hat{\phi}_{212})^{-\epsilon}]^{-\frac{1}{\epsilon}\frac{\alpha_{121}\alpha_{122}+(1-\alpha_{122})\alpha_{121}}{(1-\alpha_{111})(1-\alpha_{122})-\alpha_{112}\alpha_{121}}} \tag{17}$$

The exponents reflect the interdependence of the unit cost and price index in (6) and (7) due to input-output links.

Equations (16) and (17) show that the change in real expenditure on dirty goods reflects changes in spending on dirty final goods, dirty intermediate goods, and the production cost. The first term in the numerator and denominator of equation (16), $\beta_{js}(Y_j + T_j')$, represents expenditure on dirty final goods. The change in this expenditure occurs only due to changing tariff revenues. The assumption that expenditure on final goods is Cobb-Douglas implies that expenditure shares for final goods are fixed. Tariff revenues are typically a small percent of national income, so quantitatively, this channel may have small effects on emissions relative to the other channels. The second term in the numerator and denominator of equation (16), $\sum_k \alpha_{jsk} R_{jk}$, represents expenditure on dirty intermediate goods. Input-output links drive this channel. If all goods are final, this term equals zero.

Equation (17) represents the change in unit costs. If all goods are final in this model with symmetric countries, unit costs do not change ($\hat{c}_{11} = 1$). The first bracketed term represents the effects of changing trade policy for dirty goods, and the second bracketed term represents the effects of changing trade policy for clean goods. If trade policy for clean goods does not change, the second bracketed term in equation (17) equals one.

This model allows for pre-existing tariffs and for changes in tariffs on clean and dirty goods. Neither baseline nor counterfactual tariffs on clean goods change the pollution intensity of expenditure, because clean good tariffs (t_{212}) do not appear directly in equation (15). Baseline and counterfactual tariffs on clean goods do affect real expenditure on dirty goods by changing tariff revenues and intermediate good flows in equation (16), and also through changing the prices of clean goods and thus the second bracketed term in equation (17).

This model assumes both sectors are traded. Adding a third, non-tradable sector provides similar results. If the non-traded sector is clean, the main change is in the cost of production (17), which then accounts for the contribution of non-traded prices to the production cost of dirty goods. If the non-traded sector is also dirty, then expression for emissions include an extra term reflecting emissions from non-tradables.

Counterfactual Results and Interpretation: Social Welfare. The change in social welfare due to a counterfactual equals the product of the change in real income and pollution damages: $\hat{W}_j = (\hat{X}_j/\hat{P}_j)\widehat{f(Z)}$. Substituting in the definition of expenditure and the price index from (13) gives

$$\hat{W}_j = \frac{Y_j + T_j^{'}}{Y_j + T_j} \prod_s [\sum_o \lambda_{ojs}(\hat{c}_{os}\hat{\phi}_{ojs})^{-\epsilon}]^{\frac{\beta_{js}}{\epsilon}} \widehat{f(Z)} \tag{18}$$

If all goods are final, given symmetry and choice of the numeraire, this simplifies to[18]

$$\hat{W}_j = \frac{Y_j + T_j^{'}}{Y_j + T_j} \left[\prod_s \left(\lambda_{11s} + \lambda_{21s}(\hat{\phi}_{21s})^{-\epsilon_s} \right)^{\frac{\beta_{js}}{\epsilon_s}} \right] \widehat{f(Z)} \tag{19}$$

Equations (18) and (19) show that the change in welfare equals the product of three terms. The first ratio represents the increase in nominal expenditure, which comes from tariff revenues ($T_j^{'}$). The second term represents the (inverse of the) price index. Increasing tariffs on dirty goods increases the price index, decreases this term, and therefore decreases real income, though decreasing tariffs on clean goods increases real income.

The third term represents the change in pollution damages, which can be written independently of the functional form of the damages $f(\cdot)$. Knowing the change in global emissions and the functional form of pollution damages would be sufficient to calculate the change in pollution damages. One general finding is that the magnitude of the damages from climate change tends to be much smaller than the change in real income (Shapiro 2016). Hence, for calculating social welfare changes, the first two channels (nominal income and the price index) may matter more than the change in emissions.

I consider counterfactual policies which increase protection for dirty goods but decrease protection for clean goods. How do these counterfactuals affect social welfare? Equations (18) and (19) show that the sign of how these policy changes affect social welfare is theoretically ambiguous. Even if they decrease CO_2 emissions, they involve both increases and decreases in protection, so the net effect of these changes on the price index and on tariff revenues depends on specific parameters and data.

VI.C Quantitative Model

I now turn to quantify effects of these counterfactuals in a richer model. The quantitative model is stylized but incorporates some additional features—many asymmetric countries; many industries; input-output links; trade imbalances; multiple dirty (fossil fuel) industries; distinctions between iceberg trade costs, non-tariff barriers, and tariffs; and others. Because the model resembles the"structural gravity" literature in trade (Costinot and Rodriguez-Clare 2014) and the simpler model of the previous subsection,

[18]The price index term in brackets from this equation can be rewritten in terms of the change in the share of a country's expenditure purchased from domestic producers (Costinot and Rodriguez-Clare 2014). I show the formulation in equation (18) since it is more similar to versions of the price index from the rest of this paper and allows simple interpretation.

I describe the model's formal assumptions and counterfactual methodology in Appendix E.[19]

I apply the model using data from Exiobase. For computation, I aggregate the data to 10 regions and 21 industries, shown in Appendix Tables VI and VII. I assume intra-regional tariffs are zero. Three regions comprise the EU: Western, Southern, and Northern Europe.

I use sector-specific trade elasticities from aggregating studies that estimate these parameters: Caliendo and Parro (2015), Shapiro (2016), Bagwell et al. (2018), and Giri et al. (2018). Within a study, I aggregate multiple estimates for a sector using inverse variance weighting, which minimizes variance (Hartung et al. 2008).[20] I calibrate the damages from CO_2 emissions so that a one-ton increase in CO_2 emissions decreases global welfare by \$40, which corresponds with prevailing estimates of the social costs of CO_2 emissions in 2007 (IWG 2016).

Choice of Counterfactuals

I use this model to analyze several counterfactual policies. The main counterfactual changes each country's bilateral import tariffs to the country's weighted mean baseline bilateral tariff, and similarly for NTBs, with weights equal to baseline trade: $t_{ijs}^{'} = [\sum_s t_{ijs} X_{ijs}/(1+t_{ijs})]/[\sum_s X_{ijs}/(1+t_{ijs})]\ \forall i \neq j$. Here t_{ijs} denotes the baseline tariff rate on goods from origin country i to destination country j and sector s, X_{ijs} denotes the tariff-inclusive baseline value of bilateral trade, and $t_{ijs}^{'}$ denotes the counterfactual tariff. The counterfactual makes a similar change for NTBs. Policies resembling this counterfactual could result from WTO multilateral negotiations focused on eliminating tariff escalation or from environmentalists lobbying for tariff harmonization between clean and dirty industries. In regions like the EU which already have a climate change policy, politicians could argue that this kind of reform decreases leakage. Such policies might even attract support from dirty industries.

Appendix E.4 describes several other counterfactuals, including one where only the EU imposes this policy change, where protection changes to the level of clean goods, where protection changes to the level of dirty goods, where I add a carbon tariff, and where I turn off all trade policy.

One important question is what channels account for any change in CO_2 emissions. I follow the

[19] One choice not standard in the "structural gravity" literature is the functional form of pollution damages. This does not matter in the analytical model, as discussed above, but does matter for exact numbers in the quantitative model. I specify pollution damages as multiplicative with utility from consuming goods:

$$U_j = \prod_s \left(\sum_i q_{ijs}{}^{\frac{\sigma_s - 1}{\sigma_s}} \right)^{\frac{\sigma_s}{\sigma_s - 1}\beta_{js}} [1 + \delta(Z - Z_0)]^{-1}$$

Here Z are global CO_2 emissions, Z_0 is a reference or baseline level of emissions, and δ is a damage parameter chosen so a one-ton increase in emissions from baseline levels creates a global social welfare cost equal to the social cost of carbon, or \$40. The other terms here are more standard: U_j is the utility of the representative agent in country j, s indexes sectors, q_{ijs} is the quantity of goods produced in country i and consumed in country j from sector s, σ_s is the elasticity of substitution, and β_{js} are Cobb-Douglas expenditure shares. I choose this functional form since it makes it so that a one-ton increase in emissions decreases global social welfare by \$40, which is a central estimate from the climate change literature.

[20] I take the median estimate across studies since confidence intervals for Giri et al. (2018) are small enough relative to the other papers that inverse variance weighting across studies implicitly puts disproportionately high weight on that study. Bartelme et al. (2018) take the median estimates across these studies to estimate trade elasticities.

environmental economics literature in decomposing the change in CO_2 emissions due to a counterfactual into three terms: the change in real output ("scale"), the change in the share of global output from each industry ("composition"), and the emissions intensity of each industry ("technique") (Grossman and Krueger 1993; Copeland and Taylor 2003; Levinson 2009; Shapiro and Walker 2018). Appendix E.3 describes this methodology.

It is not straightforward to assess the extent to which certain endogenous changes in the model, such as reallocation of production and transportation, account for the full effect of any counterfactual. I can, however, provide indirect evidence on the importance of transportation. Most energy used in transportation comes from petroleum, and most petroleum is used for transportation. Coal is disproportionately used for the heaviest industries, like electricity generation, cement manufacturing, and steel blast furnaces, while natural gas is used for other purposes. Hence, examining the change in emissions from each fossil fuel provides some insight as to the importance of these channels.

Counterfactuals: Results for Main Counterfactual

Table V, Panel A, analyzes the first counterfactual, in which each country sets the same tariffs and NTBs on clean and dirty industries. Column (1) shows the percentage change in global CO_2 emissions. Column (2) shows the percentage change in global real income, defined as the weighted sum of country-specific changes in real income, where the weights are each country's baseline real income. Column (3) shows the change in CO_2 intensity, which equals the change in CO_2 minus the change in real income (equal to column (1) minus column (2)). Column (4) shows the change in social welfare due to climate damages. Column (5) shows the change in social welfare due to both the gains from trade and climate damages. Differences in trade elasticities and values mean these counterfactuals can change trade's volume and benefits even if they don't change mean tariffs or NTBs.

I find that this counterfactual of harmonizing trade policy between clean and dirty industries would decrease global CO_2 emissions by about 3.6 percentage points but increase global real income by 0.7 percentage points (Table V, row 1). This counterfactual decreases CO_2 intensity by 4.2 percentage points.

Table V shows that the increase in social welfare due to the decreased CO_2 emissions is much smaller than the increase in social welfare due to the increased real income. The social welfare impact of changing CO_2 emissions in a counterfactual here scales roughly proportionally with the assumed social cost of CO_2. While I have assumed a social cost of $40 per ton, in line with central values from the prevailing literature, some studies provide estimates have ranged up to $100 per ton or greater (e.g., from an expert elicitation in Pindyck (2019), or from Stern (2006)). In general, the gains from trade are orders of magnitude larger than trade's climate change externality, whether the damages per ton are $40 or $100 (Shapiro 2016). In part, this finding may reflect the fact that prevailing estimates of climate damages assume a quadratic damage function that is parameterized from the historical experience of modest changes in climate, and may poorly reflect the costs of large future climate change. A large potential cost is the uncertain possibility that the climate could increase by more than 5 or even more than 10 degrees Celsius, which could create catastrophic damages not well measured in prevailing estimates (Weitzman 2009). In part for these reasons, I emphasize the physical decrease in CO_2 emissions more than the monetization of that decrease, though I show both.

Table V, Panel B, separates these changes by region, though interpreting it requires care. Because CO_2 mixes uniformly in the atmosphere, climate damages are the same regardless of where CO_2 emissions originate. Additionally, this regional allocation identifies where fossil fuels are extracted. Low protection on dirty industries in baseline data, as the EU has (Figure V), accelerates fossil fuel production and consumption in other regions like China and India but decreases it in the EU. Thus, changing prevailing patterns of trade policy tends to increase emissions in Europe and decrease them elsewhere. The largest driver of the differences in results across regions in Table V Panel B is the rate of baseline protection on dirty goods, shown in Figure V.

Accordingly, the regional allocation in Table V, Panel B, shows that this counterfactual causes the largest increases in emissions from Europe. The counterfactual causes the largest decreases in emissions from the Americas and Rest of the World. This counterfactual modestly increases real income in all regions; that is not predetermined but is driven by differences in trade elasticities and flows across regions. Some of the region-specific change are large, though within the range of historical experience.

Panel C separates these effects into scale, composition, and technique, using the methodology described in Appendix E.3. The scale effect shows that this counterfactual increases real output by 0.8 percentage points. The composition effect shows that this counterfactual reallocates production across industries to decrease emissions by about 1.3 percentage points. The technique effect shows that even holding composition and scale fixed, the (weighted) mean industry carbon intensity falls by about 2.5 percentage points.

Panel D reports the change by fossil fuel, which helps interpret the composition effect (since only fossil fuel directly emits CO_2 in this analysis). Coal production, which is primarily used for heavy industry, slightly increases. Oil and gas production each decrease by slightly more. A majority (though not all) of oil is used for transportation; this suggests that an important channel here for decreasing emissions is that dirtier goods are produced domestically and require less shipping. The decrease in gas suggests that decreased production of goods that rely heavily on gas is another important channel.

The introduction highlighted that using trade policy for environmental goals can produce a range of responses through changing sourcing countries, transportation, and input choices. This quantitative analysis suggests that each of these changes accounts for some of the results in this model. Panel B of Table V shows that this policy increases fossil fuel production from regions that are currently encouraging trade in dirty goods but decreases it in other regions; Panel D shows that this policy decreases in emissions from oil (one proxy for the change in emissions from transportation), which accounts for about half the policy's environmental benefits; and the policy reallocates expenditure across sectors (the composition effect), which also accounts for an important share of the decrease in emissions.

One way to benchmark these numbers is to compare against other climate change policies. The Waxman-Markey bill, which passed the House but not the Senate in 2009, would have created a U.S. cap-and-trade market for CO_2. The European Union Emissions Trading System (ETS), a large cap-and-trade market for CO_2, is the world's largest climate change policy (excluding China's incipient cap-and-trade market). These policies decrease global CO_2 emissions by 2.6 percent and 1.1 percent, respectively.[21] By comparison, I calculate that this trade policy counterfactual would decrease global greenhouse gas emissions by 3.6 percent, which is moderately more than the ETS. These calculations do compare a global trade policy reform against actual unilateral climate change policies, though most climate change policy to date has involved individual countries.

Counterfactuals: Decomposition

This subsection distinguishes several components of how changing trade policy affects emissions, and thereby seeks to provide some insight on why the model generates these results.

Appendix E.2 describes a decomposition of a counterfactual's effects on emissions through several channels—changes in the price of fossil fuels, changes in spending on domestic versus foreign fossil fuels, expenditure on fossil fuels as final versus intermediate goods, and other channels. Appendix Figure IV shows this decomposition visually, Appendix E.2 shows it mathematically, and Appendix Table IX shows it numerically.

Here I summarize the general patterns. The main counterfactual requires each country to impose the same trade policies on different goods, and thus it turns off differences in trade policy between clean and dirty goods. While this policy increases relative rates of protection for all energy-intensive goods, it also directly decreases international trade in fossil fuels. Domestic supply of fossil fuels does increase, but the domestic supply is not enough to offset the decrease in imported fuels.

As to more detailed channels, increasing protection for energy-intensive goods increases the price of fossil fuels, in part because fossil fuel extraction itself uses other energy-intensive goods as an input, and the price of other energy-intensive goods rises as well. In addition, this counterfactual policy leads to decreases in the share of fossil fuels purchased from international sources. While it also increases the share purchased from domestic sources, the value of the fall in international purchasing exceeds the value of the rise in domestic production.[22] International trade is a small share of natural gas and coal, though is more common for oil. This counterfactual also tends to decrease spending on fossil fuels as an intermediate good, as it decreases demand for fossil fuels from other energy-intensive goods sectors. The change in spending on fossil fuels as a final good has mixed patterns—tariff revenue and factor incomes

[21] The Waxman-Markey bill would have decreased U.S. greenhouse gas emissions by 17 percent in the year 2020 relative to 2005 levels. The U.S. accounted for 15 percent of global CO_2 emissions in 2005. Although the Waxman-Markey bill did not pass, U.S. emissions were similar in 2014 as in 2005 (Climate Watch 2019). Assuming the Waxman-Markey bill would have decreased U.S. emissions by 17 percent, it would have decreased global emissions by 2.6 percent (=0.15*0.16). In 2005, the EU emitted 11 percent of global CO_2-equivalent (Climate Watch 2019). Some research estimates that the EU ETS decreased EU emissions relative to a counterfactual by about 10 percent (Dechezlepretre et al. 2018), which implies that the EU ETS decreased global emissions by about 1.1 percent.

[22] The main interesting exception discussed in Appendix E.2 is that India had exorbitant import tariffs on coal in the year 2007 (50 percent), and so this counterfactual actually decreases those tariffs and thus increases trade in coal, and emissions from coal trade.

both increase total spending (and thus spending on fossil fuels as a final good) in some cases but decrease it in others, and the net effect is small.

It may be useful to highlight similarities and differences between the full quantitative model, illustrated through this decomposition, and the analytical model of Section VI.B. Both models allow a natural separation of the change in emissions into a change in real expenditure on dirty goods, and a change in the pollution intensity of that expenditure. In both models, emissions reflect a tradeoff between emissions for domestic production and exports, where the trade elasticity magnifies that tradeoff. In addition, in both models, changes in tariff revenues drive changes in nominal expenditure on dirty goods, while input-output links play an important role in changing the price of dirty goods. One important difference is that even though the analytical model has input-output links, because it has only two sectors, one clean and one dirty, there is no distinction between energy-intensive goods like steel which are not fossil fuels versus relatively cleaner industries like food manufacturing; that distinction is more relevant in the full quantitative model. In addition, assumptions like symmetry and having two countries allow the analytical model to provide simpler and more intuitive equations for changes in pollution and other terms; the key ideas of those equations are still present in the fuller model, but is less direct.

Appendix E.4 discusses sensitivity analyses, which are shown in Appendix Table VIII, and give qualitatively similar results. Appendix E.4 and Appendix Table VIII also show results for the other counterfactuals, which are generally in line with the results of the main counterfactual.

The counterfactuals analyzed in the main text and Appendix suggest a few broader conclusions. Trade policy reforms can have quantitatively meaningful effects on CO_2 emissions; it is valuable to assess both the environmental and traditional costs and benefits of such trade policies; and policymakers concerned about the environment should consider decreasing protection on clean industries, not merely increasing protection on dirty industries.

VII Conclusions

This paper asks a simple but new question: how and why do tariffs and non-tariff barriers (NTBs) differ between clean and dirty industries? I define an industry's "dirtiness" by the total CO_2 emitted to produce a dollar of output. I find a simple answer: tariff and NTB rates are substantially higher on clean than on dirty goods. This relationship appears in most countries, in cooperative and non-cooperative trade policy, and in many years and ways of analyzing the data.

At a broad level, this paper suggests that trade policy can have important impacts on environmental outcomes. The implicit subsidy to CO_2 in trade policy this paper analyzes, which has not been previously identified, totals \$550 to \$800 billion per year. For comparison, all direct global subsidies to fossil fuel consumption, which are a major focus of political debates involving the U.S., EU, World Bank, and IMF, together total about \$530 billion per year. General equilibrium model-based analyses require strong assumptions but suggest that if countries imposed similar tariffs and NTBs on clean and dirty industries, global CO_2 emissions would fall, while global real income would largely not change or slightly increase.

The resulting change in global CO_2 emissions has similar magnitude to the estimated effects of some of the world's largest actual or proposed climate change policies.

I find that trade policy has this subsidy because political economy variables that determine trade policy are correlated with CO_2 emissions. The data show an important role for an industry's upstream location—the extent to which it sells to other firms versus final consumers. I describe theory and evidence consistent with the idea that firms lobby for high protection on their own outputs but low protection on their intermediate inputs. Because industries can be well organized but final consumers generally are not, countries end up with greater protection on downstream (and clean) goods, and less protection on upstream (and dirty) goods.

These conclusions are relevant to policy. Climate change is a classic externality that would be addressed efficiently with a Pigouvian tax on CO_2 emissions. Today, however, a fifth of global output faces carbon prices, and existing carbon prices are heterogeneous and below typical estimates of the social cost of carbon emissions. Countries that do implement carbon prices face concerns that they will decrease the competitiveness of domestic energy-intensive industries and cause "leakage" of dirty production from regulated to unregulated regions. A common proposal to address these concerns is a tariff that is proportional to the carbon embodied in imported goods, usually called a carbon tariff or carbon border adjustment. I show that countries are imposing greater protection on clean than on dirty goods, so instead of internationally adopting a carbon tariff, most countries have implicitly created a carbon subsidy in trade policy. Using trade policy negotiations to decrease this environmental bias of trade policy could help address climate change. This proposal is particularly relevant in regions like the EU which already have a domestic carbon price, but which currently have trade policies that may be encouraging leakage of dirty production to other regions rather than preventing it.

What is the political feasibility of harmonizing tariffs and NTBs between clean and dirty industries? The exact reform, of course, influences its political feasibility. For example, if one region implemented this reform and other regions did not respond with similar reforms, this could increase domestic emissions of "local" pollutants like particulate matter, even while decreasing global CO_2 emissions. Additionally, increasing tariffs and NTBs on upstream goods could disadvantage developing countries, which may have a comparative advantage in producing upstream goods, though environmental interest groups and dirty industries might support such reforms. Because dirty industries are disproportionately upstream, downstream industries lobbying for low tariffs on their inputs might oppose such reforms. Lobbying from energy-intensive industries is usually a problem for climate change policy, but for these reforms would actually increase their feasibility. More generally, climate change and the environment have never been part of the argument against tariff escalation, and are rarely part of the debate in choosing relative levels of tariffs and NTBs across industries. This paper suggests that making the environment part of these policy conversations could produce important benefits.

University of California, Berkeley, and National Bureau of Economic Research

References

Aichele, R. and G. Felbermayr (2012). Kyoto and the carbon footprint of nations. *Journal of Environmental Economics and Management 63(3)*, 336–354.

Anderson, K. (1980). The political market for government assistance to australian manufacturing industries. *Economic Record 56(153)*, 132–44.

Antràs, P. and D. Chor (2013). Organizing the global value chain. *Econometrica*, 2127–2204.

Antràs, P., D. Chor, T. Fally, and R. Hillberry (2012). Measuring the upstreamness of production and trade flows. *American Economic Review Papers and Proceedings 102(3)*, 412–416.

Antràs, P., T. C. Fort, and F. Tintelnot (2017). The margins of global sourcing: Theory and evidence from u.s. firms. *American Economic Review 107*(9), 2514–64.

Antràs, P. and R. W. Staiger (2012). Offshoring and the role of trade agreements. *American Economic Review 102*(7), 3140–3183.

Antweiler, W. (1996). The pollution terms of trade. *Economic Systems Research 8*(4), 361–366.

Arkolakis, C., A. Costinot, and A. Rodriguez-Clare (2012). New trade models, same old gains? *American Economic Review 102(1)*, 94–130.

Autor, D. H., D. Dorn, and G. H. Hanson (2013). The china syndrome: Local labor market effects of import competition in the united states. *American Economic Review 103*(6), 2121–68.

Bagwell, K., R. W. Staiger, and A. Yurukoglu (2018). Quantitative analysis of multi-party tariff negotiations.

Balassa, B. (1965). Tariff protection in industrial countries: An evaluation. *Journal of Political Economy 73*(6), 573–594.

Bartelme, D., A. Costinot, D. Donaldson, and A. Rodriguez-Clare (2018). Economies of scale and industrial policy: A view from trade.

Bento, A. M., L. H. Goulder, M. R. Jacobsen, and R. H. von Haefen (2009). Didistribution and efficiency impacts of iincrease us gasoline taxes. *American Economic Review 99(3)*, 667–699.

Bickerdike, C. F. (1907). Review of a. c. pigou's protective and preferential import duties. *Economic Journal 17*, 98–108.

Blanchard, E. J., C. P. Bown, and R. C. Johnson (2016). Global value chains and trade policy. NBER Working Paper 21883.

Business Wire (2009). Citac: Obama administration tariffs on chinese tires will hurt u.s. automakers and other downstream industries.

Cadot, O., J. de Melo, and M. Olarreaga (2004). Lobbying, counterlobbying, and the structure of tariff protection in poor and rich countries. *World Bank Economic Review*.

Caliendo, L. and F. Parro (2015). Estimates of the trade and welfare effects of nafta. *Review of Economic Studies 82*, 1–44.

Caves, R. E. (1976). Economic models of political choice: Canada's tariff structure. *Canadian Journal of Economics 9*(2), 278–300.

Climate Watch (2019). Historical ghg emissions.

Copeland, B. R. (2000). Trade and environment: policy linkages. *Environment and Development Economics 5*, 405–432.

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

Corden, W. M. (1966). The structure of a tariff system and the effective protective rate. *Journal of Political Economy 74(3)*, 221–237.

Cosbey, A., S. Droege, C. Fischer, and C. Munnings (2017). Developing guidance for implementing border carbon adjustments: Lessons, cautions, and research needs from the literature. Mimeo, RFF.

Costinot, A. and A. Rodriguez-Clare (2014). *Handbook of International Economics, Vol. 4*, Chapter Trade Theory with Numbers: Quantifying the Consequences of Globalization, pp. 197–262. North-Holland.

Davis, S. J. and K. Caldeira (2010). Consumption-based accounting of co2 emissions. *Proceedings of the National Academy of Sciences 207(12)*, 5687–5692.

Dechezlepretre, A., D. Nachtigall, and F. Venmans (2018). The joint impact of the european union emissions trading system on carbon emissions and economic performance.

Dekle, R., J. Eaton, and S. Kortum (2008). Global rebalancing with gravity: Measuring the burden of adjustment. *IMF Staff Papers 55(3)*, 511–539.

Diamond, P. A. and J. A. Mirrlees (1971). Optimal taxation and public production i: Production efficiency. *American Economic Review 61(1)*, 8–27.

Dornbusch, R., S. Fischer, and P. Samuelson (1977). Comparative advantage, trade, and payments in a ricardian model with a continuum of goods. *American Economic Review 67*(5), 823–839.

Eaton, J. and S. Kortum (2002). Technology, geography, and trade. *Econometrica 70(5)*, 1741–1779.

Eaton, J., S. Kortum, B. Neiman, and J. Romalis (2016). Trade and the global recession. *American Economic Review 106*(11), 3401–38.

Ellison, G., E. L. Glaeser, and W. R. Kerr (2010). What causes industry agglomeration? evidence from coagglomeration patterns. *American Economic Review 100*, 1195–1213.

Fajgelbaum, P., P. Goldberg, P. Kennedy, and A. Khandelwal (2020). The return to protectionism. *Quarterly Journal of Economics*.

Fally, T. (2012). Production staging: Measurement and facts. Mimeo, UC Berkeley.

Fowlie, M., M. Reguant, and S. P. Ryan (2016). Market-based emissions regulation and industry dynamics. *Journal of Political Economy 124(1)*, 249–302.

Frankel, J. A. and A. K. Rose (2005). Is trade good or bad for the environment? sorting out the causality. *Review of Economics and Statistics 87(1)*, 85–91.

Fredriksson, P. G. (1997). The political economy of pollution taxes in a small open economy. *Journal of Environmental Economics and Management 33*, 44–58.

Freund, C. and Çaglar Özden (2008). Trade policy and loss aversion. *American Economic Review 98(4)*, 1675–1691.

Gawande, K., P. Krishna, and M. Olarreaga (2012). Lobbying competition over trade policy. *International Economic Review 53(1)*, 115–132.

Gibson, G. (2017). Behind the scenes, companies fight trump on u.s. steel tariffs.

Giri, R., K.-M. Yi, and H. Yilmazkuday (2018). Gains from trade: Does sectoral heterogeneity matter?

Golub, S. S. and J. M. Finger (1979). The processing of primary commodities: Effects of developed-country tariff escalation and developing-country export taxes. *Journal of Political Economy 87*(3), 559–577.

Grossman, G. and E. Helpman (1995). Trade wars and trade talks. *Journal of Political Economy 103(4)*, 675–708.

Grossman, G. M. and E. Helpman (1994). Protection for sale. *American Economic Review 84*(4), 833–850.

Grossman, G. M. and A. B. Krueger (1993). *The Mexico-U.S. Free Trade Agreement*, Chapter Environmental Impacts of a North American Free Trade Agreement. M.I.T. Press.

Guimbard, H., S. Jean, M. Mimouni, and X. Pichot (2012). Macmap-hs6 2007, an exhaustive and consistent measure of applied protection in 2007. *International Economics 2(130)*, 99–121.

Hartung, J., G. Knapp, and B. K. Sinha (2008). *Statistical Meta-Analysis With Applications*. Wiley.

Hillman, J. (2013). Changing climate for carbon taxes: Who's afraid of the wto? Technical report, German Marshall Fund. http://www.gmfus.org/publications/changing-climate-carbon-taxes-whos-afraid-wto.

Horowitz, K. and M. Planting (2006). *Concepts and Methods of the U.S. Input-Output Accounts*. Bureau of Economic Analysis of the U.S. Department of Commerce.

IEA (2007a). Energy balances of non-oecd count2007 2007. Technical report, IEA.

IEA (2007b). Energy balances of oecd countries 2007. Technical report, IEA.

IEA (2007c). Key world energy statistics. Technical report, IEA. https://www.iea.org/publications/freepublications/publication/KeyWorld2017.pdf.

IMF (2013). Energy subsidy reform: Lessons and implications. Technical report, IMF. https://www.imf.org/en/Publications/Policy-Papers/Issues/2016/12/31/Energy-Subsidy-Reform-Lessons-and-Implications-PP4741.

IWG (2016). Technical support document: technical update of the social cost of carbon for regulatory impact analysis under executive order 12866. Technical report, Interagency Working Group on Social Cost of Greenhouse Gases, United States Government.

Jacobsen, M. R., C. R. Knittel, J. M. Sallee, and A. A. van Benthem (2019). The use of regression statistics to analyze imperfect pricing policies. *Journal of Political Economy*.

Jacobsen, M. R. and A. A. van Benthem (2015). Vehicle scrappage and gasoline policy. *American Economic Review 105(3)*, 1312–38.

Kee, H. L., A. Nicita, and M. Olarreaga (2009). Estimating trade restrictiveness indices. *Economic Journal 119*, 172–199.

Kortum, S. and D. Weisbach (2016). Border adjustments for carbon emissions: Basic concepts and design. RFF Discussion Paper 16-09.

Levinson, A. (2009). Technology, international trade, and pollution from us manufacturing. *American Economic Review 99(5)*, 2177–2192.

Lyubich, E., J. S. Shapiro, and R. Walker (2018). Regulating mismeasured pollution: Implications of firm heterogeneity for environmental policy. *AEA Papers and Proceedings 108*.

Maggi, G. (2016). *Handbook of Commercial Policy*, Chapter Issue Linkage, pp. 513–564. Elsevier.

Maggi, G. and A. Rodríguez-Clare (1998). The value of trade agreements in the presence of political pressures. *Journal of Political Economy 106(3)*, 574–601.

Maggi, G. and A. Rodríguez-Clare (2007). A political-economy theory of trade agreements. *American Economic Review 97(4)*, 1374–1406.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *Review of Economic Studies 60*, 531–542.

Marvel, H. P. and E. J. Ray (1983). The kennedy round: Evidence on the regulation of international trade in the united states. *American Economic Review 73(1)*, 190–197.

Marvel, H. P. and E. J. Ray (1987). Intraindustry trade: Sources and effects on protection. *Journal of Political Economy 95(6)*, 1278–1291.

Miller, R. E. and P. D. Blair (2009). *Input-Output Analysis: Foundation and Extensions*. Cambridge University Press.

Mullainathan, S. and J. Speiss (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives 31*(2), 87–106.

Olson, M. (1965). *The Logic of Collective Action*. Harvard University Press.

Pauwelyn, J. (2013). *Research Handbook on Environment, Health and the WTO*, Chapter Carbon leakage measures and border tax adjustments under WTO law, pp. 448–506. Edward Elgar.

Pincus, J. J. (1975). Pressure groups and the pattern of tariffs. *Journal of Political Economy 83(4)*, 757–778.

Pindyck, R. S. (2019). The social cost of carbon revisited. *Journal of Environmental Economics and Management 94*, 140–160.

Ray, E. J. (1981). The determinants of tariff and nontariff trade restrictions in the united states. *Journal of Political Economy 89(1)*, 105–121.

Rodrik, D. (1995). *Handbook of International Economics*, Chapter Political Economy of Trade Policy, pp. 1457–1494. Elsevier.

Schleich, J. and D. Orden (2000). Environmental quality and industry protection with noncooperative versus cooperative domestic and trade policies. *Review of International Economics 8(4)*, 681–697.

Shapiro, J. S. (2016). Trade costs, co_2, and the environment. *American Economic Journal: Economic Policy 8*(4), 220–54.

Shapiro, J. S. and R. Walker (2018). Why is pollution from u.s. manufacturing declining? the roles of environmental regulation, productivity, and trade. *American Economic Review 108*(12), 3814–54.

Shapiro, J. S. and R. Walker (2020). Is air pollution regulation too stringent?

Stern, N. (2006). *The Economics of Climate Change: The Stern Review*. Cambridge University Press.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B 58*(1), 267–288.

Trefler, D. (1993). International factor price differences: Leontief was right! *Journal of Political Economy 101(6)*, 961–987.

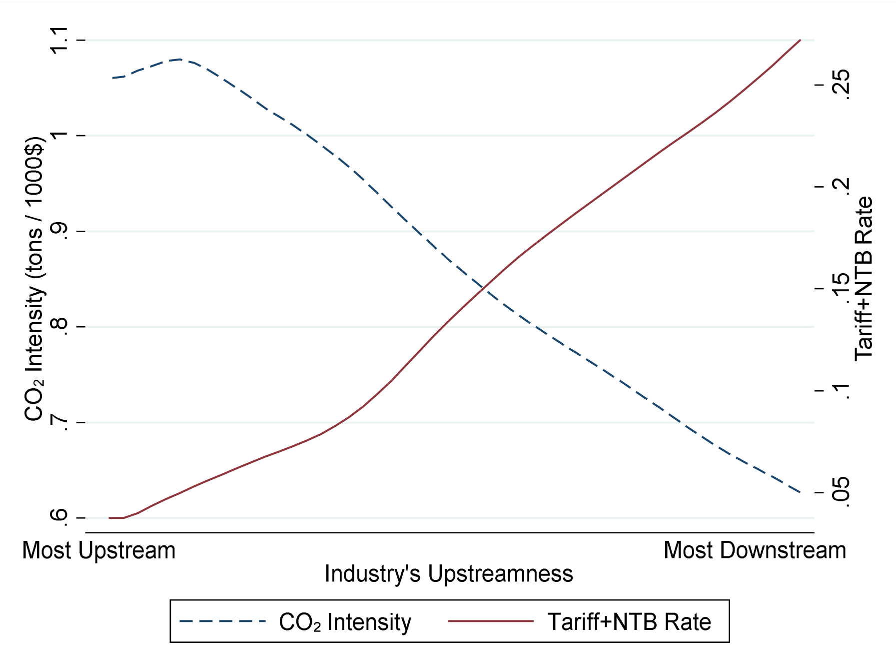
Weitzman, M. L. (2009). On modeling and interpreting the economics of catastrophic climate change. *Review of Economics and Statistics 91*(1), 1–19.

World Bank (2018). State and trends of carbon pricing 2018. Technical report.

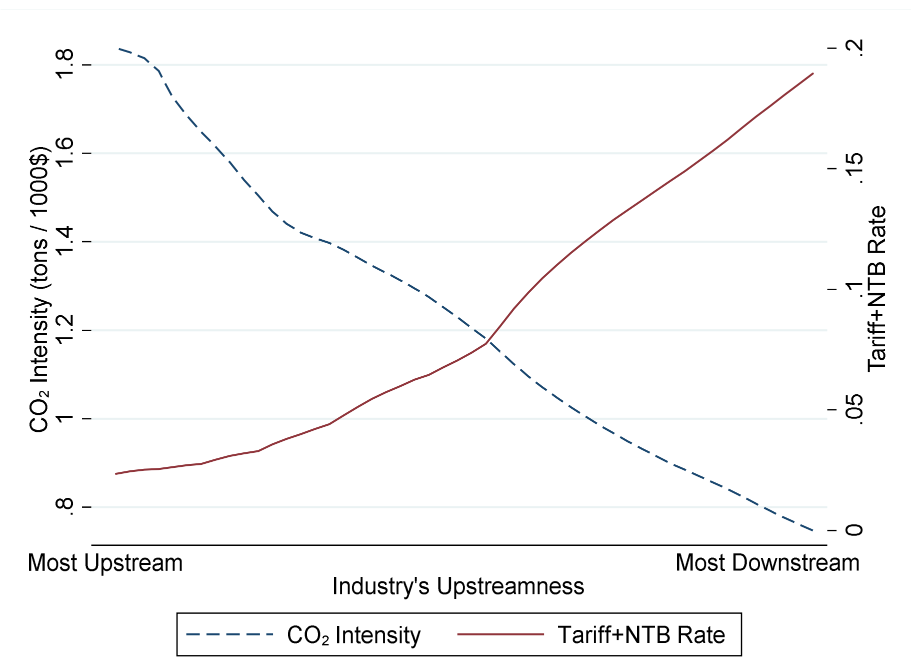
FIGURE I

Upstreamness, CO_2 Intensity, and Trade Policy

Panel A: All Global Trade



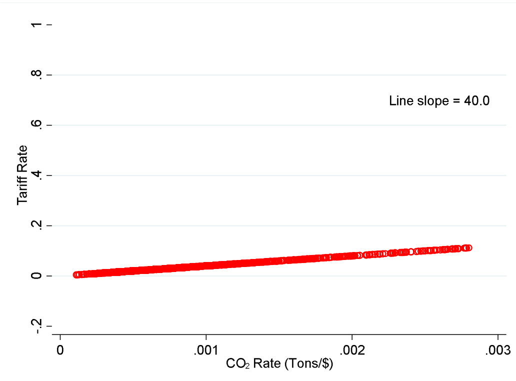
Panel B: U.S. Only



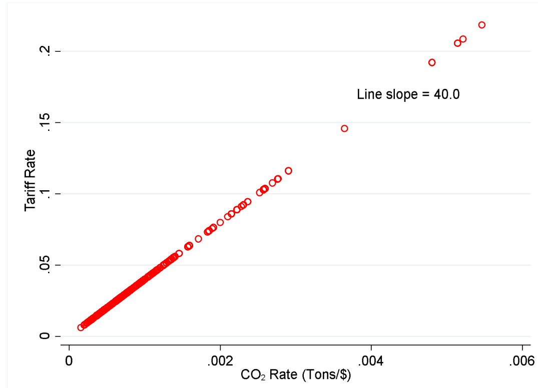
Notes: Solid line is local linear regression of tariffs plus NTBs on upstreamness. Dashed line is local linear regression of CO_2 intensity on upstreamness. Each observation is an importer×industry (Panel A) or an industry (Panel B). All lines use Epanechnikov kernel with bandwidth of 0.75.

FIGURE II
Trade Protection Versus CO_2 Emission Rates

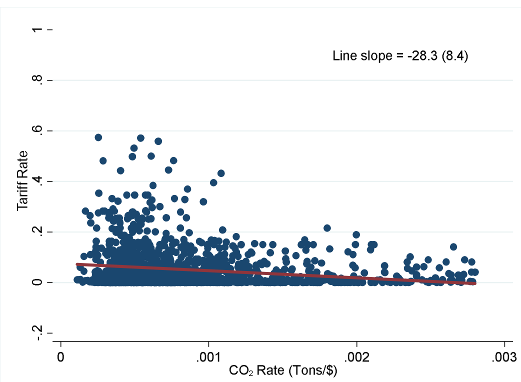
Panel A. Global hypothetical \$40/ton carbon tariff



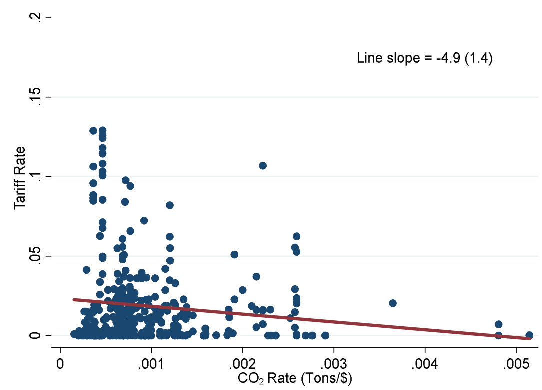
Panel B. U.S. hypothetical \$40/ton carbon tariff



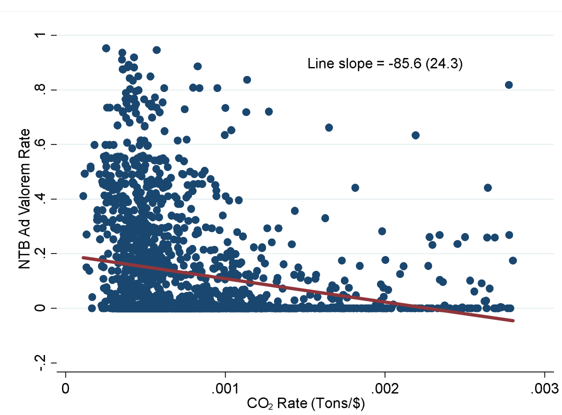
Panel C. Actual global tariffs



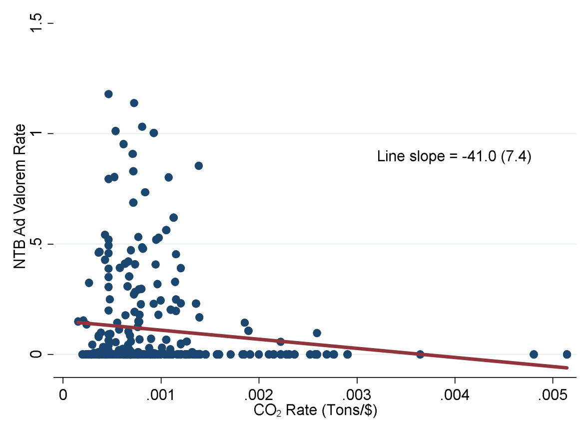
Panel D. Actual U.S. tariffs



Panel E. Actual global non-tariff barriers



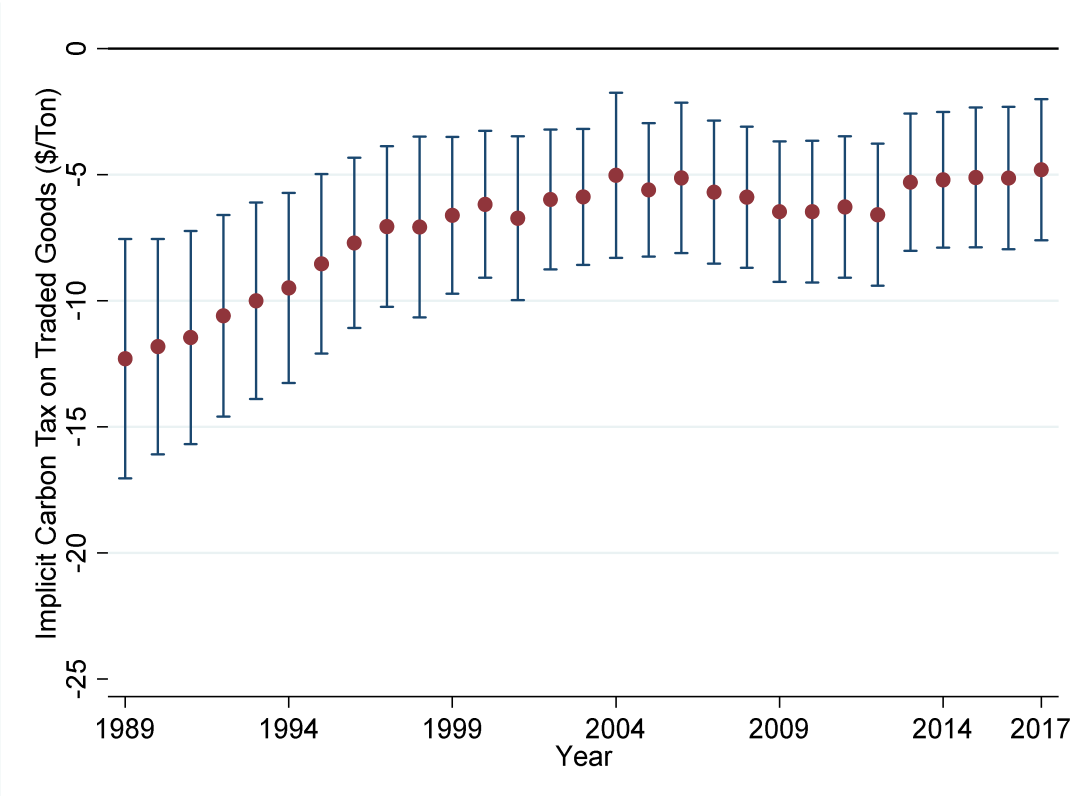
Panel F. Actual U.S. non-tariff barriers



Notes: Panels A and B plot a hypothetical carbon tariff of \$40/ton. Each point in global data is an importer×industry pair; each point in U.S. data is an industry. CO_2 rate is total (direct+indirect) emissions measured from inverting an input-output table. Line is linear trend; in Panels C and E, line is fitted from regressions including importer fixed effects. Each graph excludes the top 1% of CO_2 rates, tariffs, and NTB rate. Numbers for line slopes correspond to the specifications and values of Tables II and III, column 5. Standard errors are clustered by industry.

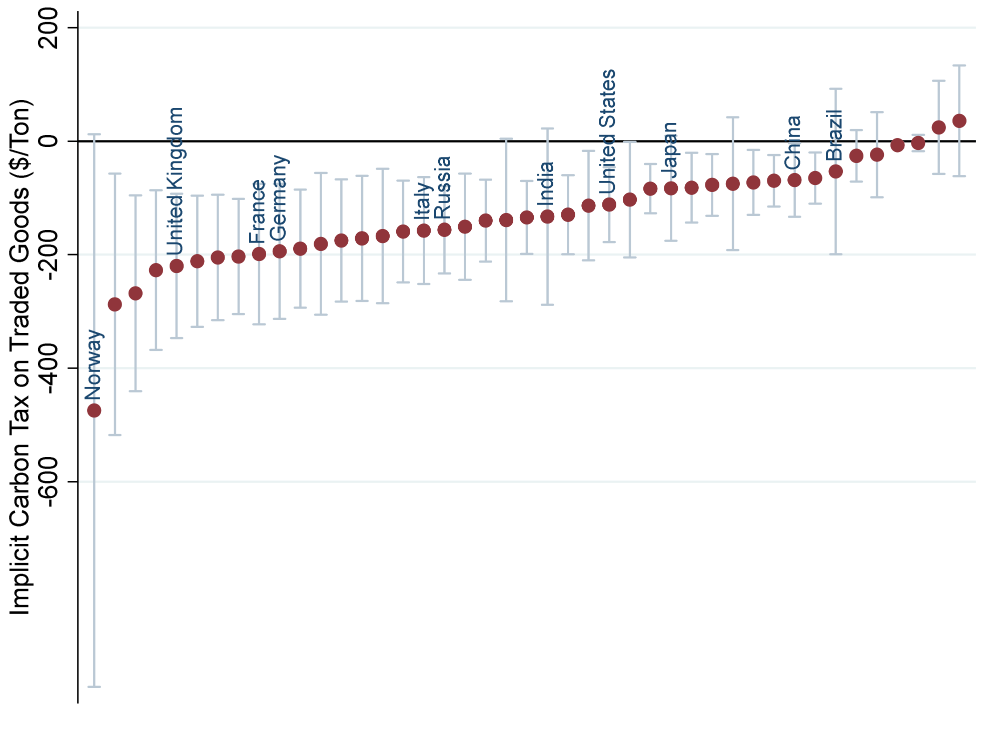
FIGURE III

Correlation Between U.S. Import Tariffs and CO_2 Emission Rates



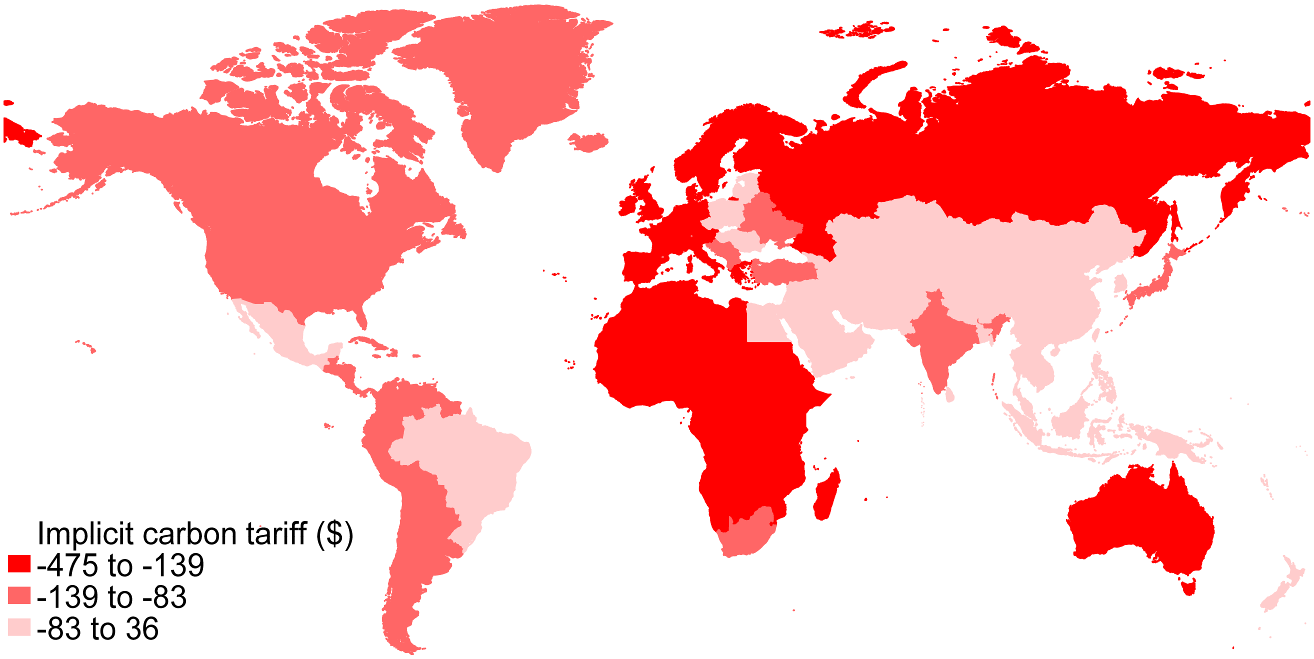
Notes: Implicit carbon tax is the coefficient from a regression of import tariffs on CO_2 emission rates, as in equation (1). Graph shows a separate regression for each year. Emissions intensity is estimated from 2007 input-output tables and applied to all years. Circles show the coefficient estimates, bars show robust 95% confidence intervals. Regressions use instrumental variables; total CO_2 is instrumented with direct CO_2.

FIGURE IV
Covariance of Trade Protection and CO_2 Emission Rates, by Country



Notes: Implicit carbon tax is the coefficient from a regression of import tariffs plus NTBs (ad valorem equivalent) on a constant and on total CO_2 emission rate (tons/$), measured from inverting the input-output matrix, which accounts for both primary fossil fuels used in an industry and emissions embodied in intermediate goods used in the industry. A separate regression is run for each country. Total CO_2 is instrumented with the direct CO_2 emissions rate from the input-output table, measured in the same industry but in the ten smallest other countries. Data from year 2007. Graph excludes five Exiobase countries missing NTB data: Bulgaria, Cyprus, Malta, Slovakia, and Taiwan. Red circles are point estimates, vertical bars are robust 95% confidence intervals.

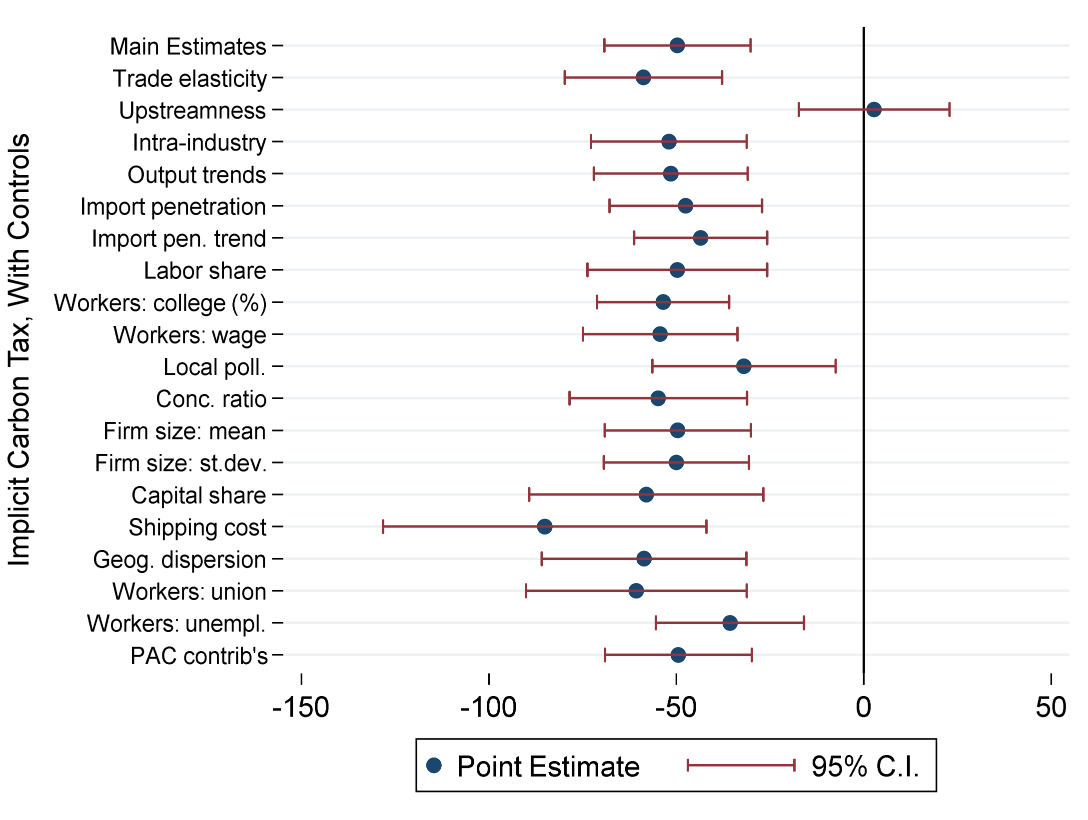
FIGURE V
Implicit Carbon Tax on Traded Goods, by Country



Notes: Implicit carbon tax is the coefficient from a regression of import tariffs plus NTBs (ad valorem equivalent) on CO_2 emission rates and a constant, separately for each country. Data correspond to Figure IV. Graphs include five rest-of-world groups, one per continent.

FIGURE VI

Political Economy Explanations for CO_2 Subsidies Implicit in U.S. Imports



Notes: Each blue circle represents the coefficient on total CO_2 intensity, instrumented by direct CO_2 intensity, from a regression of tariffs+NTBs on CO_2 intensity. The red bar depicts the robust 95 percent confidence interval. Each regression includes one additional political economy control, indicated at the left part of the graph. Regressions are weighted by the value of imports.

TABLE I

Cleanest and Dirtiest Manufacturing Industries in Global Data

	CO_2 Rate (Tons/\$)×1000 (1)	Import Tariff Rate (2)	Non-Tariff Barriers (3)
Panel A. Cleanest industries			
Pork processing	0.34	0.10	0.37
Meat products n.e.c.	0.36	0.10	0.37
Sugar refining	0.37	0.20	0.42
Wood products	0.37	0.01	0.03
Motor vehicles	0.40	0.03	0.05
Mean of cleanest 5 industries	*0.37*	*0.09*	*0.25*
Panel B. Dirtiest industries			
Bricks, tiles	1.54	0.02	0.02
Coke oven products	1.64	0.01	0.01
Iron and steel	1.74	0.01	0.02
Phosphorus fertilizer	1.93	0.02	0.11
Nitrogen fertilizer	2.53	0.02	0.11
Mean of dirtiest 5 industries	*1.88*	*0.02*	*0.05*

Notes: CO_2 rates are measured in metric tons of CO_2 per thousand dollars of output, calculated by inverting a global multi-region input output region from Exiobase. Dollars are deflated to real 2016 values using U.S. GDP deflator. Global refers to the mean value across all countries, weighted by the value of output; industries ordered based on global emissions; n.e.c. means not elsewhere classified. Import tariffs are ad valorem and measured in year 2007 CEPII Macmap data. Non-tariff barriers are ad valorem, from Kee et al. (2009).

TABLE II

Association of Import Tariffs and CO_2 Emissions Rates

	FS		RF		OLS		IV	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. All global trade (global input-output table)								
CO_2 rate	1.38***	1.54***	-44.69***	-17.19**	-28.28***	-4.48	-32.31***	-11.17**
	(0.09)	(0.08)	(13.24)	(8.16)	(8.42)	(6.17)	(8.41)	(5.40)
N	2,021	2,021	2,021	2,021	2,021	2,021	2,021	2,021
Dependent Var. Mean	0.001	0.001	0.052	0.028	0.052	0.028	0.052	0.028
K-P F Statistic	—	—	—	—	—	—	228.96	352.59
Panel B: U.S. Imports (U.S. data)								
CO_2 rate	1.32***	1.58***	-7.52***	-10.35***	-4.89***	-3.23***	-5.69***	-6.55***
	(0.19)	(0.51)	(2.00)	(3.71)	(1.40)	(0.94)	(1.44)	(2.29)
N	379	379	379	379	379	379	379	379
Dependent Var. Mean	0.001	0.001	0.018	0.016	0.016	0.018	0.018	0.016
K-P F Statistic	—	—	—	—	—	—	50.33	9.77
Weighted		X		X		X		X

Notes: Table shows regressions of import tariffs on CO_2 rates. Weights are the value of imports. Panel A uses global Exiobase data; Panel B uses U.S. data. Each observation in Panel A is an importer×industry; each observation in Panel B is an industry. Panel A includes importer fixed effects. All regressions include a constant. The endogenous variable is the total CO_2 emissions rate (tons/$) measured from inverting the input-output matrix, which accounts for both primary fossil fuels used in an industry and emissions embodied in intermediate goods used in the industry. For Panel A, the instrument is the direct CO_2 emissions rate from the input-output table,, measured in the same industry but in the 10 smallest other countries. For Panel B, the instrument is the CO_2 emissions rate measured from MECS and CM, which accounts for primary fossil fuels used in an industry and electricity consumed in the industry. Odd-numbered columns are unweighted, even-numbered columns are weighted by the value of the trade flow. Emissions rates measured in metric tons of CO_2 per dollar of output. Output is measured in 2016 US$, deflated with the U.S. GDP deflator. FS is first-stage, RF is reduced-form, OLS is ordinary least squares, IV is instrumental variables. All data from year 2007. Standard errors clustered by industry are in parentheses. Asterisks denote p-value * < 0.10, ** < 0.05, *** <0.01.

TABLE III
Association of Non-Tariff Barriers and CO_2 Emissions Rates

	FS		RF		OLS		IV	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. All global trade (global input-output table)								
CO_2 rate	1.38***	1.54***	-124.15***	-116.51**	-85.58***	-73.22*	-89.78***	-75.67**
	(0.09)	(0.08)	(40.92)	(43.81)	(24.33)	(36.75)	(26.75)	(29.38)
N	2,021	2,021	2,021	2,021	2,021	2,021	2,021	2,021
Dep. Var. Mean	0.001	0.001	0.126	0.088	0.126	0.088	0.126	0.088
K-P F Statistic	—	—	—	—	—	—	228.96	352.59
Panel B. U.S. imports (U.S. data)								
CO_2 rate	1.32***	1.58***	-63.34***	-59.13***	-41.04***	-17.98***	-47.96***	-37.41***
	(0.19)	(0.51)	(16.68)	(20.78)	(7.44)	(4.15)	(10.03)	(12.33)
N	379	379	379	379	379	379	379	379
Dep. Var. Mean	0.001	0.001	0.109	0.079	0.109	0.079	0.109	0.079
K-P F Statistic	—	—	—	—	—	—	50.33	9.77
Weighted		X		X		X		X

Notes: Table shows regression of NTB rates on CO_2 rates. Columns 1 through 4 are weighted by the value of imports. Panel A uses global Exiobase data; Panel B uses U.S. data. Each observation in Panel A is an importer×industry; each observation in Panel B is an industry. Panels A includes importer fixed effects. All regressions include a constant. The endogenous variable is the total CO_2 emissions rate (tons/$) measured from inverting the input-output matrix, which accounts for both primary fossil fuels used in an industry and emissions embodied in intermediate goods used in the industry. For Panel A, the instrument is the direct CO_2 emissions rate from the input-output table, measured in the same industry but in the 10 smallest other countries. For Panel B, the instrument is the CO_2 emissions rate measured from MECS and CM, which accounts for primary fossil fuels used in an industry and electricity consumed in the industry. Odd-numbered columns are unweighted, even-numbered columns are weighted by the value of the trade flow. Emissions rates measured in metric tons of CO_2 per dollar of output. Output is measured in 2016 US$, deflated with the U.S. GDP deflator. FS is first-stage, RF is reduced-form, OLS is ordinary least squares, IV is instrumental variables. All data from year 2007. The dependent variable is the ad valorem NTB rate from Kee et al. (2009). Standard errors clustered by industry are in parentheses. Asterisks denote p-value * < 0.10, ** < 0.05, *** <0.01.

TABLE IV

Political Economy Explanations for Implicit Carbon Taxes in Trade Policy

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. All global trade						
CO_2 rate	-120.55***	-32.90	-120.76***	-121.42***	-120.92***	-120.44***
	(33.73)	(25.60)	(33.17)	(35.50)	(34.12)	(33.62)
N	1,990	1,990	1,990	1,990	1,990	1,990
Panel B. All global trade, instrument for political economy						
CO_2 rate	-120.55***	34.61	-111.66***	-125.64***	-101.50**	-119.33***
	(33.73)	(38.88)	(40.04)	(47.61)	(43.86)	(33.95)
K-P F Statistic	—	43.29	27.20	41.97	10.15	21.05
N	1,990	1,990	1,990	1,990	1,990	1,990
Panel C. U.S. imports						
CO_2 rate	-49.72***	2.74	-51.99***	-47.50***	-49.75***	-54.32***
	(9.90)	(10.19)	(10.54)	(10.32)	(12.19)	(10.45)
N	358	358	358	358	358	358
Upstreamness		X				
Intra-industry			X			
Import pen. ratio				X		
Labor share					X	
Mean wage						X

Notes: Dependent variable in all regressions is sum of tariffs and NTBs. Each observation is a country*industry (Panels A and B) or industry (Panel C). In all regressions, the CO_2 rate is the total CO_2 rate (tons/$) from inverting an input-output table, which is instrumented with the direct CO_2 rate. In panel B, political economy variables (upstreamness, intra-industry share, etc.) are also treated as endogenous. The Panel B regressions use a second instrument equal to the mean of each political economy variable in the industry of interest across the ten smallest other countries in the data, measured by gross manufacturing output. Panels A and B use Exiobase data, panel C uses U.S. data. Panels A and B include country fixed effects. All regressions include a constant. Standard errors clustered by industry are in parentheses. Asterisks denote p-value * < 0.10, ** < 0.05, *** <0.01.

TABLE V
Effects of Setting Tariffs and NTBs to Mean, Model-Based Estimates

	Change in CO_2 Emissions (%) (1)	Change in Real Income (%) (2)	Change in CO_2 Intensity = (1) - (2) (3)	Climate benefits (4)	Social welfare (5)
Panel A. Global Total					
Global Total	-3.59%	0.65%	-4.24%	0.08%	0.57%
Panel B. By region					
Pacific Ocean	33.31%	1.02%	32.29%	—	—
Western Europe	23.33%	0.90%	22.43%	—	—
Eastern Europe	0.77%	0.99%	-0.22%	—	—
Latin America	-3.36%	0.74%	-4.10%	—	—
North America	-3.80%	0.26%	-4.06%	—	—
China	0.03%	0.22%	-0.19%	—	—
Southern Europe	54.67%	0.64%	54.03%	—	—
Northern Europe	26.96%	1.06%	25.90%	—	—
Indian Ocean	-5.15%	0.31%	-5.46%	—	—
Rest of World	-14.96%	0.93%	-15.89%	—	—
Panel C. Decomposition					
Scale	0.20%	—	—	—	—
Composition	-1.29%	—	—	—	—
Technique	-2.50%	—	—	—	—
Panel D. By Fossil fuel					
Coal	-2.63%	—	—	—	—
Oil	-4.65%	—	—	—	—
Natural gas	-3.97%	—	—	—	—

Notes: Global change in real income refers to the weighted mean percentage change in countries' real incomes due to a counterfactual policy, where weights equal each country's baseline income. In all baseline and counterfactual scenarios, intra-national tariffs and NTBs are assumed to equal zero.