Essays on Trade, CO2, and the Environment

by

Joseph S. Shapiro

B.A. Economics, Stanford University (2003)
M.Sc. Economics for Development, University of Oxford (2006)
M.Sc. Statistics, London School of Economics (2007)

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2013

Signature of Author ....
Department of Economics
May 15, 2013

Certified by..............................
Michael Greenstone
3M Professor of Environmental Economics
Thesis Supervisor

Certified by..............................
Arnaud Costinot
Pentti J.K. Kourri Associate Professor of Economics
Thesis Supervisor

Accepted by..............................
Michael Greenstone
3M Professor of Environmental Economics
Chairman, Departmental Committee on Graduate Studies

Essays on Trade, CO2, and the Environment

by

Joseph S. Shapiro

Submitted to the Department of Economics
on May 15, 2013, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

The first chapter of this thesis uses a general equilibrium model of trade and the environment to investigate two questions. First, how do the gains from trade compare against the environmental costs of trade? Trade can generate environmental costs by requiring long-distance transportation of goods and by relocating production to countries that emit a lot of CO2 to produce a given good. I find that trade's benefits exceed trade's environmental costs by two orders of magnitude. Second, what are the welfare consequences of proposed EU, US, and global climate change regulations on the carbon emissions from transporting goods? I find that the three proposed policies all increase global welfare. However, they decrease welfare in poor countries, and they provide economic benefits to the implementing region even ignoring environmental consequences.

The second chapter (coauthored with Olivier Deschênes and Michael Greenstone) quantifies the magnitude of one defensive investment that people undertake to protect themselves against air pollution. We analyze how the NOx Budget Trading Program, a large cap-and-trade market in the Eastern U.S., affected air pollution emissions, ambient air quality, medication purchases, hospitalizations, and mortality. We find that the market decreased medication expenditures by about $900 million annually. These defensive benefits have similar magnitude to the monetized effect of the market on preventing premature mortality.

The third chapter analyzes how the Clean Water Act affected U.S. water pollution levels. By almost any measure, water quality has improved since the 1972 Act. Nonetheless, water quality was improving at similar rates before 1972. The only exception is thermal pollution, which has worsened continually since 1969, presumably due to climate change. I find that the Act's two main activities – wastewater treatment grants and industrial permits – both improved water quality, as indicated by the omnibus measure of dissolved oxygen. At the same time, there is some evidence that the grants increased fecal coliform counts.

Thesis Supervisor: Michael Greenstone
Title: 3M Professor of Environmental Economics

Thesis Supervisor: Arnaud Costinot
Title: Pentti J.K. Kourri Associate Professor of Economics

Contents

[1]This chapter is coauthored with Olivier Deschênes and Michael Greenstone.

Acknowledgements

I would like to acknowledge some of the many people who helped in writing this thesis. I am extremely grateful to Michael Greenstone for pushing me to aim high and teaching me how to pursue those aims. Arnaud Costinot is a remarkable mentor who has provided extraordinary teaching and encouragement. David Autor and Esther Duflo both provided generous and insightful training and helped shape my interests. I am thank Josh Angrist, Dave Donaldson, Amy Finkelstein, and Jon Gruber for training and encouragement.

Fellow students often gave me reasons to laugh and think: thank you Emily Breza, Nathan Hendren, Brad Larsen, Nick Ryan, and Ashley Swanson. I also thank friends Kaitlin Anelauskas, Pat Bazinet, Chad Carr, Andrew Goldthorp, Kevin Gravina, Eduardo Ledesma, Nick Menzies, Bernat Olle, Jon Puz, Jay Salony, Hank Scollard, and Justin Verdirame for pacing and encouragement throughout the writing of this dissertation.

I also thank pre-MIT mentors who fostered my interest in economics and helped teach me how to think: Ken Arrow, Paul David, Stefan Dercon, Gillette Hall, Jack Huhtala, Harry Patrinos, Richard Petersen, Debra Satz, Emmanuel Skoufias, Bruce Wren, and Vic Wright.

Financial support from an EPA STAR Graduate Fellowship, a Martin Fellowship, an MIT Energy Fellowship, and the Schultz Fund has given me the luxury of time for this research. UCEI and LSE both provided supportive places to work during parts of this research.

My family has provided abundant love and support. My brothers Ari and Daniel have led by example and encouragement. My parents Elayne and Leonard have given me the gifts of incredible parenting and unconditional love. My wife Amanda's love, kindness, and wisdom have encouraged me throughout this work. Finally, I am grateful to our son Simon for frequent reminders about the fun and challenge of learning and discovery.

Chapter 1

Trade, CO2, and the Environment

1.1 Introduction

Economists have long argued that trade improves social welfare by increasing real income levels. This paper starts from the idea that trade may also generate a negative externality by contributing to climate change. Trade can generate this externality by requiring long-distance shipping of goods and by relocating production to regions which emit large amounts of CO_2 to produce a given good.

This paper builds a unified theoretical and empirical framework which can compare international trade's benefits against its environmental costs. The paper describes a model of trade and the environment, compiles new data on CO_2 emissions from international and intranational shipping, and estimates the model's key parameters via instrumental variables. The model and data are used to examine two types of counterfactuals.

The first counterfactual asks: how would welfare change if all international trade ceased? Autarky for all countries is (hopefully) not a realistic policy, but it provides a benchmark to use in thinking about policies that affect trade and the environment. I find that international trade increases CO_2 emissions by 6 percent. Several influential papers ask whether international trade is good for the environment (Antweiler, Copeland, and Taylor 2001; Copeland and Taylor 2003; Frankel and Rose 2005), and this conclusion provides a clear answer—no. However, the global gains from international trade, equal to $5.3 trillion annually, exceed the environmental costs of international trade by a factor of 149. The gains from trade exceed the environmental costs of trade in all 128 countries that I analyze.[1]

[1] The "environment" in this paper refers exclusively to climate change. Although this paper's framework could be

Second, the paper assesses the welfare consequences of several proposed climate change regulations. I focus on the EU's current effort to regulate the CO_2 emissions from domestic and international airplane flights, which the *New York Times* has described as "one of the most contested environmental initiatives ever undertaken" (Kanter 2011; Kanter 2012). I also analyze US regulations of the CO_2 emissions from all forms of shipping, which were part of the Waxman-Markey Bill of 2009 that passed the US House but not the Senate. Finally, I analyze a global tax on the CO_2 emissions from air and sea shipping, which the 1997 Kyoto Protocol required but which has not been implemented. Because these counterfactuals all represent incomplete regulation – they regulate only the CO_2 emitted on some trade routes, or on some modes of transportation, or from shipping but not production – they could distort trade and production towards unregulated areas. Their environmental and welfare effects therefore have theoretically ambiguous signs.

I find that these climate change policies all decrease international trade but increase global welfare. Additionally, I find that these policies have regressive incidence, and actually decrease welfare in many poor countries including much of Sub-Saharan Africa. This result is somewhat surprising because poor countries suffer the largest proportional damages from climate change. Moreover, these policies may be the only environmental regulations which increase welfare in the implementing region at the expense of its trading partners, even before accounting for environmental benefits. For example, regulating CO_2 emissions from US shipping under a policy similar to part of the Waxman-Markey bill would increase the US gains from trade by \$20 billion over a decade. The implementing region benefits because these policies resemble small unilateral tariffs, and it is well-established in trade theory that such tariffs can benefit a large country by decreasing the relative prices of imported goods (Bickerdike 1907).[2] I also find that the EU and US policies produce small amounts of leakage and trade diversion.[3] These effects represent about 1 and 3 percent of the total decrease in CO_2 emissions for the EU and US policies, respectively.

This paper analyzes regulation of the CO_2 emissions from transportation in part because international air and sea transportation represents the single fastest-growing anthropogenic source of greenhouse gas emissions (Figure 1 and Appendix 1.A.1). Between 1990 and 2008, CO_2 emis-

applied to any pollutant, due to data limitations I do not measure the effects of "local" pollutants like sulfur dioxide (SO_2) or nitrogen oxides (NO_x).

[2]That standard terms-of-trade argument assumes that a tariff is applied to international and not intranational trade. In my setting, the logic of the standard argument still applies because international shipping is more fuel-intensive than domestic shipping.

[3] "Leakage" describes the relocation of pollution emissions from a regulated region to an unregulated region. "Trade diversion" describes the related idea that regional trade policies may distort trade with countries outside the affected region (Viner 1950).

sions from international air and sea transportation grew by 65 percent, which is nearly double the growth of greenhouse gas emissions from the rest of the global economy. The reason for this rapid growth is globalization. Trade growth exceeds GDP growth, and in order to trade a good, a country must transport it. I show that international air and sea shipping represent 2.8 percent of global CO_2 emissions today. However, I emphasize that this paper's model and data account for all CO_2 emissions—from domestic shipping and production, as well as international shipping.[4]

This paper's model uses five standard assumptions which describe the form of preferences, production, trade costs, CO_2 emissions, and equilibrium. For expositional purposes, I describe a simplistic Armington (1969) trade model, in which each country produces one variety per sector and varieties are differentiated by country of origin. Analogous assumptions, however, would provide similar welfare calculations under other important trade theories (Arkolakis, Costinot, and Rodríguez-Clare 2012).[5] Moreover, although these models depend on numerous variables which are difficult to measure, such as price levels for each country and industry and nonpecuniary barriers to trade, a technique from Dekle, Eaton, and Kortum (2008) makes most of these variables unnecessary. This approach makes it possible to measure the impacts of counterfactual policies using only data on trade and production, measures of how each counterfactual policy affects trade costs, and estimates of one set of trade elasticities.

To conduct this analysis, I combine data from national commerce offices and public records to obtain what I believe is the most comprehensive set of files ever compiled on international and intranational shipping costs, transportation mode choice, pollution emissions, and trade flows. For 13 sectors of tradable goods, these data quantify the CO_2 emissions from transportation between and within 128 pairs of countries, representing nearly one million fuel consumption estimates. These data also measure the share of trade transported by five different shipping modes (air, sea, rail, road, other). Finally, I link these data to existing measures of bilateral trade flows, gross output, and the CO_2 emissions from production to provide a complete accounting of CO_2 emissions from

[4]I also study these counterfactual policies because changes in shipping fuel costs could have large effects on the global economy. Much trade research focuses on tariffs, while little focuses on carbon taxes. For the US and EU, however, mean fuel costs for international shipping roughly equal mean tariff rates—each represents 1 to 1.5 percent of the value of imported goods. Climate change research estimates that the social cost of carbon is roughly 20 dollars per ton of CO_2 (Greenstone, Kopits, and Wolverton 2011). Pigouvian taxes of this magnitude would increase fuel costs by 12 percent.

[5]Eaton and Kortum (2002), Bernard, Eaton, Jensen, and Kortum (2003), and this paper's Armington (1969) model provide numerically equivalent welfare measures. Because this paper analyzes multiple sectors, intermediate goods, and tariff revenues, Krugman (1980) and (Chaney 2008)'s application of Melitz (2003) would produce different measures. An emerging literature provides quantitative welfare comparisons between these models (Balistreri, Hillberry, and Rutherford 2011; Costinot and Rodríguez-Clare 2012).

the global economy.

The paper uses instrumental variable regressions to estimate trade elasticities separately for 13 sectors. These parameters are arguably the most important parameters in trade. They can guide trade negotiations, enable the quantitative evaluation of past policies like NAFTA (Caliendo and Parro 2011), measure the gains from trade (Arkolakis, Costinot, and Rodríguez-Clare 2012), and help explain why international trade is growing faster than global production (i.e., help explain globalization; see Baier and Bergstrand (2001) and Yi (2003)). These parameters represent the bilateral elasticity of trade in dollars with respect to bilateral trade costs. Trade costs include shipping fees but also tariffs, informational barriers, border effects, expropriation risk, and all other bilateral trade frictions (Anderson and van Wincoop 2004).[6] Like Hummels (2001), I use reported shipping costs. I also use panel data and obtain two measures of shipping costs for each observation. Using one of these measures as an instrumental variable for the other increases estimates of the trade elasticity in absolute value, which is consistent with the presence of attenuation bias due to classical measurement error. Using this instrumental variables regression, I obtain a trade elasticity of -7.91 for the global economy overall, -6.68 for manufacturing, and a range of -0.76 to -16.11 for thirteen specific sectors.

At each stage of the analysis, I test the paper's main results against independent estimates to determine whether this paper obtains reasonable findings. These comparisons provide encouraging evidence. For the data on the CO_2 emissions from shipping, my aggregate emissions measures resemble those of international organizations, which are able to use simpler methods but can only compile data at much coarser levels of aggregation. For my estimates of trade elasticities, classifications of product differentiation from Rauch (1999) show that my pattern of estimates across sectors follows economic theory—estimated bilateral demand is more elastic for more homogenous goods. For my analysis of the EU, US, and global carbon taxes, a simpler and distinct approach – Harberger (1964) triangles – obtains comparable aggregate global welfare measures, though Harberger triangles cannot answer many of the other detailed questions which this paper analyzes.

This paper seeks to advance research on the economic costs of regulation by using reduced-form methods to estimate the key parameters of a structural general equilibrium model, and by then

[6] In Armington (1969) models, the trade elasticity represents the elasticity of substitution across country-specific varieties. In Ricardian models with Fréchet-distributed technology (Eaton and Kortum 2002), it represents the inverse dispersion of productivity. In models of monopolistic competition with heterogeneous firms and Pareto-distributed technology (Melitz 2003; Chaney 2008), it represents the shape parameter of the Pareto distribution. In welfare analyses, this parameter plays similar roles in all of these models (Arkolakis, Costinot, and Rodríguez-Clare 2012).

applying the model to recover the full welfare consequences of proposed but untested policies.[7] I obtain estimates of parameters that are central to a large literature on the gains from trade. These regressions use fixed effects to control for key omitted variables and instrumental variables to address measurement error. The structural model uses these parameters to account for general-equilibrium price changes and to measure the effects of untested policies on social welfare.

This hybrid framework contrasts with most research on regulation, which generally uses either structural models, computable general equilibrium (CGE) models, or reduced-form regressions (Bovenberg and Goulder 1996; Greenstone 2002; Babiker 2005; Nordhaus 2008; Elliott, Foster, Kortum, Munson, Cervantes, and Weisbach 2010; Aldy and Pizer 2011; Balistreri and Rutherford 2012; Walker 2012; Fowlie, Reguant, and Ryan 2012). My joint analysis of transportation and production also contrasts with that literature's focus solely on production.[8] This contrast is potentially important both because transportation is essential for trade, and because international trade is the fastest growing global source of CO_2 emissions.

The paper also builds on a literature which develops robust approaches to measuring the gains from trade. The theory is developed in Eaton and Kortum (2002), Alvarez and Lucas (2007), Dekle, Eaton, and Kortum (2008), Caliendo and Parro (2011), and Arkolakis, Costinot, and Rodríguez-Clare (2012). Researchers have used this approach to study railroads (Donaldson 2010; Donaldson and Hornbeck 2012), optimal tariffs (Ossa 2011), trade imbalance (Dekle, Eaton, and Kortum 2008), and transportation mode choice (Harrigan 2010; Lux 2012). The broad differences between this more recent trade literature and longstanding CGE papers are an emphasis on using few assumptions; close ties between theory, econometrics, and data; and richer and more plausible microfoundations. I build on this literature by accounting for the environmental costs of trade and by calculating confidence regions for welfare estimates.[9]

This paper also builds on research in trade and the environment (Antweiler, Copeland, and Taylor 2001; Copeland and Taylor 2003; Frankel and Rose 2005) by analyzing real-world policies using data from many countries, using a structural "gravity" model of trade, and focusing on transportation. In contrast, most of the trade and the environment literature uses Heckscher-Ohlin

[7] Some trade research uses a similar framework (Dekle, Eaton, and Kortum 2008; Caliendo and Parro 2011; Arkolakis, Costinot, and Rodríguez-Clare 2012). Baldwin and Venables (1995) compare computable general equilibrium models and econometric regressions. I know no previous applications of this approach to the environment or energy.

[8] Several studies assess how the EU ETS would affect ticket prices and potential airline profits (Faber and Brinke 2011). Keen, Perry, and Strand (2012) measure potential revenue from taxes on international air and sea shipping.

[9] Lai and Trefler (2002), who use a different approach from the more recent literature, is the only trade paper I know of that reports a confidence region for the gains from trade.

or reduced-form models, analyzes pollutants like SO_2 which primarily affect the region where they are emitted, and focuses on production.

My estimates of trade elasticities and measurement of fuel costs also build on existing work. Some research uses reports of shipping costs (Hummels 2001), while others use theory to infer trade costs from trade flows (Anderson and van Wincoop 2004; Head and Ries 2001). One study estimates carbon emissions for international trade flows (Cristea, Hummels, Puzzello, and Avetisyan ming). Sections 1.4 and 1.5 contrast this paper's methodology and results with these literatures in more detail. My finding that measurement error biases conventional estimates of the trade elasticity towards zero may have broader applicability, and it suggests that the gains from trade are smaller than conventional estimates would imply.

The paper proceeds as follows. Section 2 outlines the theory. Section 3 describes the paper's data sources. Section 4 describes new measures of the CO_2 emissions from shipping. Section 5 reports new estimates of the trade elasticities. Section 6 measures the full welfare effects of international trade. Section 7 applies the model to evaluate EU, US, and global carbon taxes. Section 8 describes sensitivity analyses, and section 9 concludes.

1.2 Model of Trade and the Environment

This section explains the model's five assumptions, shows how they permit analysis of counterfactuals, and then explains their connection to the paper's regressions and data. The model describes a world of N countries, each with a fixed labor force L and a representative consumer. Production requires only one factor ("labor"). In this Armington (1969) model, each country produces one variety per sector, and varieties are differentiated by country of origin. The rationale for this simplistic model is expositional: Ricardian models with richer and more realistic microfoundations (Eaton and Kortum 2002; Bernard, Eaton, Jensen, and Kortum 2003) would generate numerically equivalent welfare calculations.

1.2.1 Primitive Assumptions

A1. Preferences. Consumers have constant elasticity of substitution (CES) preferences over varieties within a sector, Cobb-Douglas preferences across sectors, and experience quadratic damage from climate change:

$$U_d = \left[\prod_{j=1}^{J}\left(Q_d^j\right)^{\alpha_d^j}\right]\left[\frac{1}{1+\left(\mu_d^{-1}\sum_{o=1}^{N}E_o\right)^2}\right] \tag{1.1}$$

$$Q_d^j = \left(\sum_{o=1}^{N}\left(Q_{od}^j\right)^{\frac{\sigma^j-1}{\sigma^j}}\right)^{\frac{\sigma^j}{\sigma^j-1}}$$

The first bracketed term in (1.1) represents the utility from consuming goods, and the second bracketed term in (1.1) represents the disutility from climate change. The term Q_d^j is a CES aggregate of the varieties Q_{od}^j, each representing trade from origin country o to destination country d of sector j goods. The elasticity of substitution between sector j varieties is $\sigma^j > 1$. Due to the Cobb-Douglas preferences across sectors, country d spends the share α_d^j of its expenditure on sector j. Total CO_2 emissions due to country o are E_o, and the parameter μ_d dictates the social cost of CO_2 emissions. CO_2 emissions E_o are an externality that the representative agent takes as given when making consumption decisions. This model has no feedback loop from the environment to trade—the negative environmental externality of trade decreases utility, but climate change does not affect trade directly.

These preferences imply the following consumer price index for sector j in country d:

$$p_d^j = \left[\sum_{o=1}^{N}(p_{od}^j)^{1-\sigma^j}\right]^{\frac{1}{1-\sigma^j}}$$

Here p_{od}^j is the price for sector j varieties produced in country o and sold in country d. The expenditure required for one unit of utility is then $P_d \equiv \Pi_{j=1}^{J}(p_d^j)^{\alpha_d^j}$.

The functional form for climate damages is common in environmental economics (Nordhaus 2008; Weitzman 2012). However, most papers describe damages as a function of climate and use atmospheric science to determine how CO_2 emissions affect the climate. Assumption (1.1) approximates those models by using a single quadratic damage function to summarize both the effect of CO_2 emissions on climate and the effect of climate on utility. For this paper's counterfactuals, this specification provides a marginal social cost of carbon emissions which is nearly constant.

The parameter μ_d quantifies the magnitude of climate damages like diminished human health. I rely on the large climate change literature to measure μ_d, and it is the only parameter in this paper that cannot be determined within the model. The climate change literature assumes utility functions which are similar but not identical to equation (1.1). For example, the DICE model (Nordhaus 2008) assumes a utility function with constant relative risk aversion preferences where output is multiplied by a climate damage function which is quadratic in temperature.

A2. Production Technology and Market Structure. Firms have Cobb-Douglas production technology and trade costs take the "iceberg form," where $\tau^j_{od} \geq 1$ units must be shipped for one to arrive:

$$\begin{aligned} c^j_o &= (w_o)^{\beta^j_o} (p^j_o)^{1-\beta^j_o} && (2a) \\ p^j_{od} &= c^j_o \tau^j_{od} && (2b) \end{aligned}$$

Here labor has price w_o and share β^j_o, and intermediate goods have price p^j_o and share $1 - \beta^j_o$. Firms engage in perfect competition and arbitrage price gaps over space, so the product price at destination d equals the production cost c^j_o augmented by a trade cost τ^j_{od}. This cost function arises from assuming that output in each sector is combined into an intermediate good specific to that sector. Production uses the same CES price aggregator as consumption, so p^j_o represents both the consumer price index and the price of intermediate goods shown in (2a).

A3. Transportation Technology. Trade costs can be decomposed as follows:

$$\begin{aligned} \tau^j_{od} &= (1 + t^j_{od})(1 + f^j_{od}) \exp\left(\delta^j_{od}\right) && (1.3a) \\ t^j_{od} &= \sum_{m=1}^{M} D_{odm} \kappa^j_{odm} W^j_{odm} \xi_m \gamma_1 \left(t^{j,X}_{odm} + t^{j,M}_{odm}\right) && (1.3b) \\ f^j_{od} &= \sum_{m=1}^{M} D_{odm} \kappa^j_{odm} W^j_{odm} \xi_m \gamma_2 P^{oil} && (1.3c) \end{aligned}$$

Here t^j_{od} represents the carbon tax per dollar of expenditure, f^j_{od} represents the fuel cost per dollar of expenditure, and δ^j_{od} represents all other bilateral trade frictions. The costs δ^j_{od} are difficult to observe—they summarize tariffs, border effects, language differences, informational barriers, and other barriers to trade (Anderson and van Wincoop 2004). Equation (1.3a) summarizes two types of trade costs: an "iceberg" component $(1 + f^j_{od}) \exp(\delta^j_{od})$; and a carbon tax t^j_{od} which is rebated lump-sum to consumers.

Equations (1.3b) and (1.3c) relate carbon taxes and fuel consumption to observable data. The variable D_{odm} represents the distance between countries o and d via transportation mode m. Distances differ by transportation mode because ships cannot travel overland. The variable κ_{odm}

represents the share of o-d trade in dollars transported by mode m. The model abstracts from endogenous mode choice for a given sector and trading pair (Lux (2012) describes an alternative approach). The variable W_{odm} represents the weight-to-value ratio for goods traded between countries o and d by mode m. The variable ξ_m represents the fuel efficiency of transportation mode m, defined in gCO_2 emitted per ton-km transported. This functional form builds on Cristea, Hummels, Puzzello, and Avetisyan (ming). The variable P^{oil} represents the global petroleum price in dollars per barrel of crude. The variable $t^{j,X}_{odm}$ represents the carbon tax rate for exports, measured in dollars of tax per ton of CO_2. Similarly, $t^{j,M}_{odm}$ represents the carbon tax rate for imports. The constants γ_1 and γ_2 convert units of measurement.[10] I treat the global oil supply as perfectly elastic, so that shifts in oil demand due to counterfactuals do not affect global pre-tax oil prices.

I use (1.3a)-(1.3c) because they approximate reality and decompose t^j_{od} into terms that data report. In logs, δ_{od} becomes a bilateral fixed effect in panel regressions. These equations imply that non-fuel components of trade costs δ_{od} are proportional to fuel costs. Additionally, these equations embody two important restrictions: they assume perfect competition in the transport sector and imply that the counterfactual analyses will treat D_{odm}, κ_{odm}, W_{odm}, and ξ_{odm} as fixed parameters.

A4. Environment. Trade and production generate CO_2 emissions as follows:

$$E_d = \sum_{o,j} \left(\gamma_3 f^j_{od} + \chi^j_o \right) \frac{X^j_{od}}{p^j_{od}} \tag{1.4}$$

Equation (1.4) shows how trade contributes to climate change by affecting CO_2 emissions from both production and transportation. Here χ^j_o represents the CO_2 emissions per unit of output for sector j in country o, and the constant γ_3 represents tons of CO_2 emitted per dollar of fuel. The ratio X^j_{od}/p^j_{od} represents the units of goods produced in country o and consumed in country d. Trade generates an environmental externality through shipping (f^j_{od}) and through relocating production to countries with differing CO_2 emissions rates from production (χ^j_o). Because domestic trade (X_{oo}) plus international trade (X_{od}, $o \neq d$) accounts for all of country o's gross output, and because domestic shipping fuel ($f_{oo}X_{oo}$) plus international shipping fuel ($f_{od}X_{od}$, $o \neq d$) equals total shipping fuel consumption, equation (1.4) accounts for CO_2 emissions from all economic activity, and not merely from international trade.

[10] Specifically, $\gamma_1 = \frac{ton^2}{kg*g} = 10^{-9}$ and $\gamma_2 = \frac{ton*barrel}{kg*g} = 0.43^{-1} * 10^{-9}$, using the USEPA's standard value of 0.43 tons of CO_2 per barrel of crude oil. All tons in this paper are metric.

1.2.2 Competitive Equilibrium

A5. Market Clearing. Consumers maximize utility, firms maximize profits, and all markets clear.

Consumer utility maximization implies that demand can be separated into two stages. In the first stage, due to Cobb-Douglas preferences across sectors, each country chooses to spend the share α_d^j on sector j. In the second stage, countries allocate this expenditure across goods within a sector according to the following "gravity" demand structure:

$$\lambda_{od}^j = \left(\frac{c_o^j \tau_{od}^j}{p_d^j} \right)^{\theta^j} \tag{5a}$$

Here λ_{od}^j represents the share of country d's expenditure on sector j which is devoted to goods from producing country o. This gravity equation implies that bilateral trade is log-linear in the GDP of countries o and d and in bilateral trade costs.[11] I use the notation $\theta^j \equiv 1 - \sigma^j$ to highlight that the elasticity in equation (5a) does not merely represent a preference parameter (the Armington elasticity of substitution). Rather, it represents the key trade elasticity of a large family of gravity models, each of which has distinct microfoundations but all of which generate an equation like (5a).

Firms' profit maximization and consumer utility maximization imply the following expression for expenditure:

$$X_d^j = \left(1 - \beta_d^j\right) I_d^j + \alpha_d^j I_d$$

Economically, this equation states states that total expenditure on goods from a sector, $X_d^j \equiv \sum_{o=1}^{N} X_{od}^j$, equals the sum of two terms: expenditure on intermediate goods and expenditure on final goods. The income from sector j, $I_d^j = F_d^j X_d^j - T_d^j - \phi_d^j$, sums pre-tax imports and net exports. Here $F_d^j \equiv \sum_{o=1}^{N} \lambda_{od}^j / \left(1 + t_{od}^j\right)$ is a weighted measure of carbon taxes. Full income $I_d = w_d L_d + R_d + T_d$ sums labor earnings $w_d L_d$, carbon tax revenue R_d, and net imports T_d. Formally, $R_d = \Sigma_{o,j}[t_{do}^{j,X} X_{do}^j/(1 + t_{do}^{j,X}) + t_{od}^{j,M} X_{od}^j/(1 + t_{od}^{j,M})]$, where $t_{do}^{j,X}$ represents the carbon tax per dollar of d's exports and $t_{od}^{j,M}$ the carbon tax per dollar of d's imports.

[11] The "gravity" description comes from analogy to physics, where gravitational attraction between objects rises with their mass (analogously, GDP) and declines with their distance (trade cost). Most trade theories generate a gravity relationship. Dozens of authors who test this assumption in cross-sectional data obtain similar parameter estimates and an R-squared between 0.75 and 0.95 (Anderson 2011). Leamer and Levinsohn (1995, p. 1384) describe these estimates as "some of the clearest and most robust empirical findings in economics."

Market clearing implies that imports equal exports for each country:

$$\sum_{o,j} \frac{X^j_{od}}{1+t^j_{od}} = \sum_{o,j} \frac{X^j_{do}}{1+t^j_{do}} + T_d + \phi_d \tag{5b}$$

For a given country, trade is imbalanced sector-by-sector and a country's total net imports across sectors equals T_d, which is positive for a country with a trade deficit and negative otherwise. In this static model, one can think of net imports T_d as a transfer from the rest of the world to country d which does not change in counterfactuals.[12] ϕ_d measures international financial flows due to carbon taxes on exports. Formally, $\phi_d = \Sigma_{o,j} X^j_{do} t^{j,X}_{do}/(1+t^{j,X}_{do}) - \Sigma_{o,j} X^j_{od} t^{j,M}_{od}/(1+t^{j,M}_{od})$, where $t^{j,X}_{do}$ represents the carbon tax per dollar of d's exports, and $t^{j,M}_{od}$ the carbon tax per dollar of d's imports.

1.2.3 Counterfactual Calculations

Using this model to measure the welfare effects of a new policy involves simple calculations. Algebra shows that the indirect utility function for this model is

$$V_d = \left[\frac{I_d}{P_d}\right] \left[\frac{1}{1+\left(\mu_d^{-1} \sum_{o=1}^{N} E_o\right)^2}\right]$$

Here social welfare equals the product of two bracketed terms representing real income and the environment.

Because credible measures of prices, wages, and trade costs for all countries are difficult to obtain, I reformulate the model in terms of proportional changes (Dekle, Eaton, and Kortum 2008). Let x' denote the value of variable x after a policy is imposed and $\hat{x} \equiv x'/x$ represent the proportional change in x due to the policy. In this model, the equivalent variation is

$$\hat{V}_d = \left[\frac{\hat{I}_d}{\hat{P}_d}\right] \left[\frac{1+\left(\mu_d^{-1} \sum_{o=1}^{N} E_o\right)^2}{1+\left(\mu_d^{-1} \sum_{o=1}^{N} E'_o\right)^2}\right] \tag{1.6}$$

Equation (1.6) represents the amount a country would pay before introducing a policy in order to end up with the same utility level that the policy would provide.

[12] A dynamic model would recognize that these transfers represent intertemporal borrowing from the rest of the world.

I will evaluate counterfactuals by constructing empirical analogues to equation (1.6). Section 1.6 of the paper, which measures the full welfare effects of international trade, uses the fact that the effect of autarky on real incomes is a known function of the change in the share of expenditure which is purchased from domestic producers (Arkolakis, Costinot, and Rodríguez-Clare 2012). This fact implies that for the counterfactual of autarky, every term in (1.6) is observed in the data, and so no algorithm or optimization routine must be solved to measure the effect of autarky on social welfare. Section 1.7 of the paper, which analyzes EU, US, and global carbon taxes, uses the fact that once such taxes are introduced, a unique vector of wage changes $\hat{w}_d$ satisfies the trade balance condition (5b) and provides a counterfactual equilibrium. This fact implies that recovering the effect of carbon taxes on social welfare only requires solving a system of $N-1$ counterfactual trade balance conditions (one per country, excluding a numéraire due to Walras' Law) in $N-1$ unknown wage changes $\hat{w}_d$. Sections 1.6 and 1.7 of the paper describe these methods in more detail.

1.2.4 From Theory to the Data

In applying this model to counterfactuals, three economic objects will play key roles: measures of how carbon regulations affect trade costs ($\hat{\tau}^j_{od}$); trade elasticities (θ^j); and bilateral expenditure shares (λ^j_{od}).

The data $\hat{\tau}^j_{od}$ measure how a specific regulation changes trade costs for each sector and pair of countries. Under assumption (1.3a), measuring $\hat{\tau}^j_{od}$ requires data on the fuel cost per dollar of trade (f_{od}). Section 1.4 of the paper will describe these data.

The parameter θ^j represents the causal effect of log bilateral trade costs on log bilateral trade flows for sector j, holding wages and prices in each country fixed. θ^j identifies the effects of the counterfactuals I study because each counterfactual is equivalent to a change in trade costs. Although I estimate θ^j with data from specific countries and years, under assumption (1.1), θ^j describes the effects of trade costs for any countries and years. Section 1.5 of the paper estimates this parameter.

The data λ^j_{od} describe bilateral trade between all countries in one baseline year. This matrix summarizes key information on wages and prices. I obtain these data from readily-available public sources.

The key equations from the model which will be used throughout the rest of the paper are trade balance (5b), the gravity equation (5a), the structure of trade costs (1.3a), and the measure of welfare (1.6). Trade balance defines market equilibrium, which pins down the effect of a policy on wages. The gravity equation reveals how wages and prices affect trade flows and provides the regression equation to estimate θ^j. The trade cost assumption shows how carbon taxes affect trade costs. The measure of welfare is used to evaluate counterfactuals. I emphasize that while I have used the restrictive Armington assumption as an explanatory device to derive this model, these key results are common to a broader and more realistic family of trade models and do not depend on the Armington structure.

I bring this model to life using data on bilateral trade for the year 2007 between 128 countries. The data distinguish 13 tradable sectors and one non-tradable sector, which I assume to have infinite international trade costs. The paper treats the observed data as an equilibrium and perturbs trade costs by introducing autarky or carbon regulations. It then determines the wage changes which restore trade balance. Finally, I analyze how these changes in wages and trade costs affect welfare.

1.3 Data

1.3.1 CO_2 Emissions from Trade

Measuring CO_2 emissions as described in assumptions (1.3c) and (1.4) requires data on distances, transport mode shares, weight-to-value ratios, fuel efficiency, and production emission rates.

Distance. The Center for International Prospective Studies (CEPII) provides data on intranational and international distances (D_{odm}) for air, rail, and road trade (Mayer and Zignago 2005). These data account for population-weighted international distances, and intranational distances are defined as $0.67\sqrt{area/\pi}$ (Head and Mayer 2010).

Measuring transportation fuel costs requires estimates of distances for maritime trade. Geographic information system (GIS) files from ESRI describe the locations of all major global ports and land masses. Each non-landlocked country has at least one port, and in each country, I assume that all maritime trade flows through the port city with the greatest population. For the US, I assume that 80 percent of trade travels through New York and 20 percent through Los Angeles. To measure distances by sea, I create a one-degree grid spanning the globe. For each grid cell, I permit a ship to travel to any cell within three degrees of longitude or latitude, so long as that travel does not cross land. I then apply the Floyd-Warshall algorithm (Floyd 1963; Warshall 1962)—a standard method to find shortest paths in weighted graphs. The shortest path between two cities identifies the distance between their countries by sea.[13]

[13] Feyrer (2009b) constructs a similar measure.

Transportation Mode Shares. Data on the share of goods transported by each mode (κ^j_{odm}) are the most difficult to obtain since most public datasets do not identify transport modes. To compile these data, I obtained several files which together cover 83 percent of global trade by value and 74 percent of global trade by weight (see Appendix 1.A.2). I obtain data from US Imports and Exports of Merchandise (US air and sea); North American Freight (US truck and train); Trade Statistics of Japan (Japan); the Global Trade Atlas compiled by Global Trade Information Services (China); EU Secretariat (external trade is publicly available and internal trade I obtained by request); and the Latin America Integration Association (ALADI, for Argentina, Bolivia, Brazil, Chile, Columbia, Ecuador, Paraguay, Peru, Uruguay, and Venezuela). All data represent the year 2007 except EU internal trade, which is from the most recent year available (2000). I group unknown, post, pipeline, and self-propulsion transportation modes into one "other" category.

I impute transportation mode shares for the 17-26 percent of international trade flows where data do not report them (see Appendix 1.A.2). Mode shares have two important statistical properties: each share lies in $[0,1]$, and shares for each trade flow must sum across transportation modes to one. I use a fractional multinomial logit (Papke and Wooldridge 1996) to impute mode shares for trade flows where mode data are unavailable. I believe is the only statistical model that fits the requirements of the data.[14]

Weight-to-Value Ratios. The data sources recording mode of transportation also record weight-to-value ratios (W). These data report the total value and quantity of each trade flow, but not all quantities represent weights. I aggregate over the transportation mode datasets to obtain weight-to-values used to fill in missing data (see Appendix 1.A.2).

Fuel Efficiency. Fuel efficiency (ξ_{odm}, measured in gCO_2/ton-km) for air and sea shipping is measured from published data as follows. For airborne trade, data on global ton-km and global fuel consumption imply a fuel economy for air freight of 985.97 gCO_2/ton-km (IATA 2009).[15] For maritime trade, the CO_2 emissions due to international transportation (IEA 2011) and the international ton-km reported by the shipping industry imply fuel efficiency for sea freight of 11.11

[14] OLS and Tobit fitted values for each mode need not lie in the $[0,1]$ interval, even when adding-up constraints impose that shares sum to one. Beta and Dirichlet distributions exclude the extremal values 0 and 1, which appear frequently in the data. I use this imputation for all domestic mode shares since no data report them.

[15] Airplanes form atmospheric contrails which warm the climate. Airplane emissions also react to decrease atmospheric methane, which cools the climate (Schäfer, Heywood, Jacoby, and Waitz 2009). Because these effects have opposing signs, it is theoretically ambiguous whether the impact of airplanes on climate change exceeds the effect implied by CO_2 emissions. In practice, most research concludes that the contrail effect domaintes, and that airplanes contribute 1.5 to 3.0 times more to climate change than their CO_2 emissions would imply. For simplicity, I measure airplanes' climate change impact according to their CO_2 emissions.

gCO_2/ton-km.[16] This approach is not possible for rail and road shipping because I know of no data on total global fuel consumption for these transport modes. Instead, I compare across estimates in the transportation literature, each representing a specific region. This approach leads to fuel economy estimates of 23.0 gCO_2/ton-km for rail and 119.0 gCO_2/ton-km for road (Appendix Table 1). I impose a fuel consumption rate (ξ) of zero for the "other" transportation mode.

CO_2 from Production. I use estimates of the CO_2 emissions from production compiled by the Global Trade and Analysis Project (GTAP). GTAP uses the Tier 1 method of the Intergovernmental Panel on Climate Change (IPCC 1997) to compile these data. For each sector and country in the year 2007, these data report the tons of CO_2 emitted by producing the good. I use the GTAP data because it provides extensive country and sector detail. The Tier 1 method has the lowest data requirements and simplest methodology—it generally multiplies physical quantities of fuel consumption by mean emissions coefficients. GTAP obtains these data from input-output matrices and national accounts data for each country.

1.3.2 Shipping Costs for Estimating Trade Elasticities

I use quarterly reports of transportation costs and trade values for all US and Australian imports over the period 1991-2010.[17] The US data come from the US Imports of Merchandise dataset. I had the Australian Bureau of Statistics compile the Australian data. The US data report trade at the 10-digit Harmonized Commodity Description and Coding System (HS) level, while the Australian data report trade at the 6-digit HS level. Appendix 1.A.3 explains how I translate these data into the 13 sectors I analyze.

1.3.3 Counterfactuals: Welfare Effects of International Trade and EU, US, and Global Carbon Taxes

I use data on bilateral trade, gross output, and CO_2 emissions from production for the year 2007 from GTAP. These data report values for 128 countries and 57 sectors.[18] Their data are based on

[16] For air, IATA reports a global air fuel economy of 39.0 liters/ton-km. I convert this to gCO_2 using the US Energy Information Agency's reference rate of 9.57 kg CO_2 per gallon of jet fuel. For sea, IEA (2011) reports that international marine transportation emitted 624.5 $MtCO_2$. Freight accounts for 90 percent of civilian ship CO_2 emissions (IMO 2009, p. 160) and the IEA international maritime data generally exclude military ship emissions (Reece 2004). Dividing CO_2 emissions of 562.05 $MtCO_2$ by the 50.6 trillion ton-km of international freight reported to be traded by ship gives the 11.11 rate.

[17] Only these countries could provide panel data on transport costs for many sectors and years. 1991-2010 is the period for which I obtained quarterly data covering both countries.

[18] The data include all world production, but small countries are combined due to data limitations. Gross output equals value added (GDP) plus the value of intermediate goods. Total domestic purchases (X_{oo}) are calculated as

the UN's Comtrade data for goods. I aggregate these data to 14 sectors, including one non-tradable sector, which are comparable across all the datasets used in this paper.

The only parameter which cannot be estimated within the framework of the theory is μ_d, which is isomorphic to the social cost of CO_2 emissions. This cost represents the decrease in global welfare (measured in dollars and accounting for effects in future centuries) due to emitting one additional ton of CO_2. The paper cannot estimate this parameter because μ_d aggregates information on the damages of climate change (e.g., diminished human health), on the atmospheric processes translating CO_2 emissions into climate change, and on the trends and discounting of these values over future centuries to obtain their present value.

I choose the values of μ_d so that a one ton increase in CO_2 emissions decreases global GDP by \$19.96. This reflects the estimate of the marginal social cost of CO_2 emissions by an interagency panel of the US government (Greenstone, Kopits, and Wolverton 2011), which bases its analysis on three integrated assessment models (FUND, DICE, and PAGE). The calculation assumes a 3 percent social discount rate. These integrated assessment models quantify all of the aforementioned forces—the damages of climate change, the effects of CO_2 emissions on climate change, and aggregation and discounting over the next century. The value \$19.96 for the year 2007 linearly extrapolates the Interagency estimates of \$21.40 for the year 2010 and \$23.80 for the year 2015. For each counterfactual, I set the carbon tax rate equal to this social cost of carbon.[19] I also show results for each of the paper's counterfactuals under alternative values for the social cost of carbon or \$4.10 and \$1170. The low value of \$4.10/tonCO_2 reflects a 5% discount rate from the Interagency panel. The high value of \$1170 reflects the most extreme value estimated in a version of the DICE model with risk-averse policymakers and large and uncertain potential impacts of climate change (Cai, Judd, and Lontzek 2012). This is the largest value I have seen estimated in any study. For example, the ninety-fifth percentile of the distribution of social costs of carbon generated by the interagency panel is \$60.16. I use this large value to assess the sensitivity of conclusions to very extreme estimates of the social cost of carbon.

To assign this global cost of climate change to individual countries, I use the damage function of the RICE model (Nordhaus and Boyer 2000, p. 4-44). For each of 13 distinct regions of the

gross output minus international exports.

[19] Although this carbon tax rate summarizes the global damage from climate change, I also use it for the EU and US counterfactuals. I do so because the US government has chosen this global rate as its official measure of the social cost of carbon, and because this value is similar to forecasts of EU carbon allowance prices over the next decade (Point Carbon 2012).

globe, Nordhaus and Boyer report the damage due to a 2.5C warming, with all impacts monetized so they can be expressed as a proportion of GDP. I choose μ_d so that the global impact of marginal increases in CO_2 emissions is \$19.96/ton, but the country-by-country impact is proportional to GDP in the quantities documented in Nordhaus and Boyer.[20]

1.4 CO_2 Emissions from Trade

This paper compiles data on the CO_2 emissions from shipping because these data are necessary to apply the model and analyze counterfactuals. However, these data also merit interest because they provide novel evidence on how shipping contributes to climate change. This section describes salient facts from these data, then compares these data against independent estimates from international organizations. In total, the CO_2 emissions from production substantially exceed the CO_2 emissions from shipping. However, because latter data are new and the focus of this paper, this section focuses on CO_2 emissions from shipping while providing some comparisons between shipping and production emissions.

1.4.1 Results: CO_2 Emissions from Trade

Total Emissions. The CO_2 emissions due to production of traded goods and due to international transportation have similar orders of magnitude. International shipping emitted 1.4 gigatons of CO_2 in the year 2007, domestic shipping emitted 1.8 gigatons, the production of traded goods emitted 1.2 gigatons, and the production of nontraded goods emitted 25.3 gigatons (Table 1). In total, I calculate global CO_2 emissions in 2007 of 29.7 gigatons.[21]

---

[20] The functional form in assumption (1.1) does not allow for benefits from climate change, whereas in the Nordhaus and Boyer data, eight country-regions are projected to benefit from climate change: Australia, Canada, Hong Kong, Israel, New Zealand, Singapore, Russia, and the Rest of Europe. In the year 2007, these countries accounted for 3.6 percent of global population and 7.7 percent of global GDP. The Rest of Europe is an aggregated country which combines Andorra, Bosnia and Herzegovina, Faroe Islands, Gibraltar, Guernsey, Vatican City, Isle of Man, Jersey, Macedonia, Monaco, Montenegro, San Marino, and Serbia. For consistency with the standard quadratic damage function, I assume that each of these country-regions has zero damage from climate change.

[21] Other sources report similar global estimates. The IEA (2011) estimates a total of 29.0 gigatons for this year. The World Bank WDI and USEPA use an estimate of 31.3 gigatons from Boden, Marland, and Andres (2010). This paper defines the non-traded sector to include all services and utilities other than transportation. This categorization is accurate for services like health, housing, education, and government, which represent most non-traded goods. However, although merchandise trade data do not record it, GTAP data indicate that about two percent of electricity is traded internationally. Because electricity production emits large amounts of CO_2, this trade accounts for about 250 $MtCO_2$. Accounting for trade of such non-merchandise goods in Table 1 would reveal production of traded goods to emit slightly more CO_2 than transportation of traded goods emits. However, such a breakdown would still show that CO_2 emissions from production and transportation have comparable magnitudes.

In aggregate, production emits far more CO_2 than transportation. Goods production is responsible for almost 90 percent of global CO_2 emissions. Although this measure of production includes direct household consumption (e.g., residential heating), even with a more detailed breakdown, the vast majority of CO_2 emissions come from production rather than transportation.

International trade increases CO_2 emissions from international shipping directly and affects production indirectly, by changing the location of production and by increasing aggregate output. Because production CO_2 emissions are much greater than transportation CO_2 emissions in aggregate, both trade and production emissions will be important for analyzing trade policies.

By Country. Differences in CO_2 emissions from shipping across countries presage this paper's finding that poor countries lose the most from rising prices of shipping fuels. Shipping fuel emissions depend on the weight-to-value ratio of goods, the distance goods are shipped, and the mode used for transportation, which all vary across the globe (Figure 2). Wealthy countries like the US and Europe disproportionately trade technological goods with low weight-to-value ratios. Poorer regions like Africa and the Middle East generally trade in heavy mining and agricultural goods like crude oil, grain, and iron ore. Distances to trading partners are greatest for Africa and lower for the US and EU. Wealthy countries are most likely to trade by airplane, while countries in Asia, Africa, and Latin America use more overland and maritime trade. Panel D of Figure 2 reports the fuel costs per dollar of trade (f_{od}) separately by country. This map makes clear that poor regions of the world, especially sub-Saharan Africa, are likely to experience the largest relative effects of rising shipping costs.

These differences in fuel intensity account for some differences in the total CO_2 emissions by region (Table 1). The EU and US account for 17 and 22 percent of global CO_2 emissions, respectively. The world's other large countries – Brazil, Russia, India, and China – collectively produce 30 percent of the world's CO_2 emissions. Sub-Saharan Africa, despite having almost three times the US population, emits only 2% of all global CO_2.

By Sector. Differences across sectors also predict the paper's regressivity finding. On average across all goods, shipping fuels represent only half a percent of the cost of goods (Appendix Figure 2). For internationally traded goods, this figure is three times as high, at 1.4 percent. Raw minerals (crude oil, iron ore, etc.) have twice the fuel requirements of any other sector, at 4.5 cents per dollar of goods. Agriculture and related sectors – foods, wood, paper, and manufactured petroleum products and mineral – have fuel costs of 1.5 to 2.6 cents per dollar of goods. Electrical goods, metals, textiles, and other sectors have low weight-to-value ratios and international fuel costs represent less than one percent of their values. These statistics reflect fuel costs use to transport the goods from each sector, and not to transport intermediates used to produce goods in each sector. Poor countries predominantly trade raw materials which are lower in the value chain and more fuel-intensive to transport.

By Transport Mode. Differences across transport modes provide one reason why ex ante one might expect rich countries to pay the highest relative costs of regulations which increase the price of shipping fuels (Table 1 and Appendix Table 1). Airplanes emit nearly 100 times as much CO_2 as ships do to move one ton-km. Rail shipment is nearly as efficient as sea shipment, and truck shipment has intermediate efficiency. Maritime trade accounted for nearly half of international shipping emissions, airborne trade for about a fourth, rail trade for about a fifth, and road trade for a small share. In part because air shipment is so costly, road trade accounts for most freight—trucks account for 67 percent of all CO_2 emissions due to shipping. Sea transportation is important for international trade and accounts for 48 percent of international shipping CO_2 emissions, but is relatively unimportant for domestic trade.

1.4.2 How Reasonable Are These CO_2 Estimates?

Comparison to International Organizations. This section has described detailed measures of CO_2 emissions from shipping for nearly a million specific trade flows. It is useful to assess the accuracy of these estimates by comparing them against independent sources which provide coarser measures of CO_2 emissions but which can therefore use simpler data and methods.

The totals implied by my data are close to totals of these published sources (Figure 3 and Appendix 1.B). For example, the EU collected data from every airline landing or departing in the EU about their fuel consumption in order to plan for the inclusion of airplane CO_2 emissions in the EU cap-and-trade system. I estimate total air freight CO_2 emissions involving the EU of 75.3 $MtCO_2$, while the European Commission (2011) implies a value of 78.9. To provide another comparison, I measure total sea freight of 7,900 tons, whereas UNCTAD (2009) provides an estimate of 7,882 tons. Overall, my estimates of total air freight emissions from the EU, total air freight emissions globally, total sea CO_2, and total sea tons shipped are extremely close to the estimates of international organizations. My estimates for international air freight, international air ton-km, and sea ton-miles are slightly larger than the estimates of international organizations. This suggests that my detailed data replicate the global stylized facts about these forms of shipping.[22]

1.5 Estimation of Trade Elasticities

This section describes a new approach to estimating the trade elasticities which are key to the model of this paper. I apply this approach using import data from the US and Australia; I then compare the results against existing approaches. Following the model, I estimate elasticities separately for each sector. I emphasize that these parameters do not merely represent the Armington elasticity of substitution; rather, they provide the trade elasticity in the "gravity" equation which appears in a large family of trade models.

1.5.1 Methodology: Trade Elasticities

Consider the following regression equation:

$$\log \lambda_{ody}^{j} = \theta^{j} \log \left(1 + s_{ody}^{j}\right) + c_{oy}^{j} + p_{dy}^{j} + \delta_{od}^{j} + \epsilon_{ody}^{j} \tag{1.7}$$

This equation is derived from the model by substituting equation (1.3a) into (5a) and taking logs, with the following additions. Because I use panel data, this equation builds on the model by allowing trade flows and prices to vary by year y. It also allows for idiosyncratic innovations ϵ_{ody} in the unobserved component of trade costs, δ_{od}. Finally, s_{ody} here represents the total shipping cost (including fuel and non-fuel components), which can be proportional to fuel costs under assumptions (1.3a) and (1.3c).[23]

I begin with a naive estimator of θ^j from an OLS regression of expenditure shares on shipping

[22] One other study has compiled sector-level CO_2 emissions for international shipping (Cristea, Hummels, Puzzello, and Avetisyan ming). Some differences are worth highlighting. Unlike their paper, I also calculate fuel consumption for domestic shipping. Domestic shipping data are important because according to the GTAP data I use, international trade only accounts for 12 percent of global production (domestic trade represents the remaining 88 percent). Domestic shipping data are also important because any change in international trade will affect intranational trade, and it would violate GATT Article III for a region to regulate international but not domestic shipping (Bartels 2012; Meltzer 2012). Compared to Cristea et al., this section also uses a different statistical methodology (fractional multinomial logit rather than OLS) and somewhat different assumptions and data. In aggregate these differences matter—for example, Cristea et al's estimate of CO_2 emissions from airplane trade is about three times the values implied by data from the International Energy Agency or other organizations, while my estimate is close to these organizations' estimates. The estimate in Cristea et al. matches published estimates of total CO_2 from all air transportation, but the estimate only represents freight, which accounts for only 33.6 percent of all air ton-km transportation (IATA 2009).

[23] The data used in this section do not distinguish fuel and non-fuel costs or have all the necessary data to use assumption (1.3c) to distinguish fuel from non-fuel costs. The end of section 1.5.2 discusses sensitivity analyses, including one based on estimated fuel costs.

costs, since discussing this estimator can clarify the obstacles to estimating θ^j. This regression has two econometric challenges. First, it suffers from omitted variables like distance, nonpecuniary trade barriers, wages, and prices. Countries with large nonpecuniary trade barriers may have greater shipping costs and lower trade flows, which will generate negative bias in the OLS estimate of θ^j. Omitting these terms actually means that OLS estimates the wrong elasticity: because OLS does not control for production costs and destination prices, it obtains a general equilibrium elasticity combining direct effects of trade costs on trade flows and indirect general equilibrium effects operating through the prices p^j_{dy} and c^j_{oy}. By contrast, θ^j in the model is a partial equilibrium elasticity which is purged of these general equilibrium effects.

A second econometric challenge is measurement error. Errors in variables are a common concern in trade data.[24] Measurement error in shipping costs s_{ody} occurs due to sector, exporter, and date misclassification, inaccurate currency conversion, and simple reporting errors. Classical measurement error will attenuate the OLS estimate of θ^j, forcing it closer to zero than the true parameter value.

To address omitted variables bias, a standard solution in related settings in the trade literature is to use fixed effects to control for the unobserved terms c^j_{oy}, p^j_{dy}, and δ^j_{dy}. I estimate equation (1.7) while including exporter-by-year fixed effects to control for production costs c^j_{oy}, importer-by-year fixed effects to control for destination prices p^j_{dy}, and country-pair fixed effects δ^j_{od} to control for time-invariant components of trade costs like distance and contracting environment. These detailed fixed effects in equation (1.7) remove many components of the error term ϵ^j_{ody} which data do not report, leaving less potential for omitted variables bias. Hummels (2001) reports a cross-sectional version of this model which relies on proxies for δ^j_{od} (distance, dummies for common languages, etc.), and Donaldson (2010) uses price gaps to measure trade costs then applies this approach to data from colonial India. However to my knowledge this estimator has not been applied with panel data, observed shipping costs, and the full fixed effects to recover θ^j.

Unfortunately, using fixed effects with panel data can decrease the signal-to-noise ratio and thereby exacerbate attenuation bias (Griliches and Hausman 1986). While the fixed effects can address omitted variables bias, they can make the consequences of measurement error even more severe.

[24] Many researchers document or attempt to address measurement error in trade volume data, unit cost data, or trade cost data which is estimated by comparing value reports across importers and exporters (Bowen, Leamer, and Sveikauskas 1987; Harrigan 1993; Feenstra 1994; Trefler 1995; Limão and Venables 2001; Hummels and Lugovskyy 2006).

Instrumental variables provide an appealing way to obtain consistent parameter estimates in the presence of classical measurement error (Durbin 1954; Freeman 1984; Ashenfelter and Krueger 1994).[25] I define the instruments as follows. For each year of data, I compile two measures of each variable: one measure containing data aggregated from quarters 2 and 3 of the year, and a second measure containing data aggregated from quarters 1 and 4. For each year, I then use mean reported shipping costs from quarters 2 and 3 as an instrumental variable for reported shipping costs from quarters 1 and 4. This approach essentially uses leads and lags of shipping cost variables as instruments in order to address measurement error, which has the same spirit as using additional lags as instruments in models with lagged dependent variables (Arellano and Bond 1991). Many papers use multiple contemporaneous reports of a variable to address measurement error (Black, Berger, and Scott 2000), and Dustmann and Soest (2002) use a similar approach to mine of constructing instruments from leads and lags of a mismeasured variable, although they do so in a labor economics setting.

If measurement error in the two samples is independent, then the following instrumental variables model will provide a consistent estimator of θ^j:

$$\log \lambda^j_{ody} = \theta^j \log\left(1 + s^{j,B}_{ody}\right) + \eta^{j,B}_{oy} + \zeta^{j,B}_{dy} + \delta^{j,B}_{od} + \epsilon^{j,B}_{ody} \tag{1.8}$$

$$\log\left(1 + s^{j,B}_{ody}\right) = \beta^j \log(1 + s^{j,A}_{ody}) + \eta^{j,A}_{oy} + \zeta^{j,A}_{dy} + \delta^{j,A}_{od} + \epsilon^{j,A}_{ody} \tag{1.9}$$

Here $s^{j,A}_{ody}$ represents a measure of shipping costs from quarters 2 and 3 and $s^{j,B}_{ody}$ represents a measure of shipping costs from quarters 1 and 4. Equation (1.9) describes the first stage while equation (1.8) describes the second stage. This estimator uses fixed effects to address omitted variables bias and uses instrumental variables to address attenuation bias.

This assumption of independence between measurement error in the two reports of shipping costs is strong. If misclassification and misreporting occur due to random errors in to each report, then (1.8) will provide a consistent estimator. If measurement error is systematically related across these two reports, however, then (1.8) will still suffer from some attenuation bias, albeit less severe than in the fixed effects estimates (Black, Berger, and Scott 2000).

Several possible leads and lags of shipping costs could be used to construct these instrumental

[25] A literature in labor and health economics examines the prevalence and implications of non-classical measurement error (Card 1996; Bound, Brown, and Mathiowetz 2002; Black and Kniesner 2003; Black and Smith 2006). In the simple case of negative covariance between measurement error and a variable's true value, the asymptotic value of a univariate regression lies between OLS and IV estimates (Black, Berger, and Scott 2000).

variables. I compare quarters 2 and 3 versus 1 and 4 because this comparison may have a stronger first stage in the presence of trends. However, Appendix Tables 2a-2c show other definitions of these instruments. As I discuss at the end of this section, those other definitions have sector-by-sector correlation with the main parameter estimates of over 90 percent.

In equation (1.8), the variation in s_{ody} which the fixed effects do not eliminate includes all components of the gravity equation varying at the *o-d-y* level. For example, fees for crossing the Suez canal and different efficiency growth of plane versus sea transportation are sources of variation which could contribute to identify s^B_{ody} (Feyrer 2009a; Feyrer 2009b).[26]

These panel data may also have autocorrelation over time within trade flows, so I report standard errors adjusted for clustering within trading partners (Bertrand, Duflo, and Mullainathan 2004). Finally, Appendix Tables 2a-2c report several variations on these estimates.

1.5.2 Results: Trade Elasticities

I begin with the naive OLS estimator, which includes no controls or fixed effects. The resulting estimate of $\theta = -21.0$ represents extraordinarily elastic demand (Table 2, columns 1-2). As discussed earlier, this estimate may suffer from negative omitted variables bias because it does not control for origin, destination, or bilateral characteristics.

Using detailed fixed effects as in equation (1.7) helps address omitted variables bias but may exacerbate attenuation bias due to classical measurement error (Table 2, columns 3-4). The fixed effect estimate treating all goods as equally differentiable is $\theta = -3.7$, or $\theta = -4.17$ for manufacturing only. Economically, this estimate implies that a 10 percent increase in bilateral trade costs causes a 37-42 percent decrease in bilateral trade flows.

Finally, I turn to the instrumental variables regressions, to address both measurement error and omitted variables bias. Most instruments are strong (Table 2, columns 5-6). The F-statistic for the instrument in the regression for all goods is 18, and for manufacturing is 30. These are above the rule-of-thumb cutoff of 10 for weak instruments (Staiger and Stock 1997). Twelve of the thirteen sector estimates have an F-statistic above 10, with an unweighted average F-statistic across sectors of 29.1. Only the metals sector (F=6.3) does not satisfy the rule-of-thumb cutoff. These first-stage estimates suggest that this research design provides a good way to measure trade elasticities while

[26] In principle, one could use these changes as instrumental variables for $1 + s_{ody}$. However, apart from the shipping cost instruments I use, I am not aware of strong instruments for the 13 tradable sectors and two importers in the 20 years I analyze.

addressing key sources of measurement error and omitted variables bias.

Instrumental variable estimates of the trade elasticity all have the expected negative signs and moderate magnitudes (Table 2, columns 7-8). Treating all goods as equally differentiable provides an estimate of $\theta = -7.91$ for manufactured goods only, or $\theta = -6.68$ for all goods. These values are calculated using aggregate data with a single observation per country pair *o-d*. Averaging over the 13 sector-specific elasticities gives roughly similar values of -7.83 for all goods or -8.75 for manufactured goods only. The literature reports conflicting evidence on the bias from aggregation over sectors—Imbs and Méjean (2011) conclude that aggregation over sectors can bias estimates of trade elasticities, while Feenstra, Obstfeld, and Russ (2010) find little bias from aggregation.

Alternative assumptions obtain similar patterns of point estimates and do not change the broad picture (Appendix Table 2c). Using generalized least squares provides an efficient response to heteroskedasticity but characterizes the average dollar of trade rather than the average country pair. These point estimates are all close to the main results though less precise. Including observations with zero trade flows obtains similar results to the main estimates.[27] Explicitly controlling for tariffs attenuates the point estimates, though their pattern across sectors is similar. The correlation across sectors between each of these sensitivity analyses and the IV estimates from Table 2 are all above 90 percent.

Finally, I consider using fuel costs rather than shipping costs (Appendix Table 2a, column 6). I set $\tau^j_{od} = 1$ for all trade in the year 2007, then measure trade costs in other years as $\hat{\tau}^j_{od}$. Here I calculate $\hat{\tau}^j_{od}$ by setting the global oil price in assumption (1.3c) to the observed real Brent crude oil price for each year, but holding all other components of the model fixed at the observed year 2007 values. These estimates allow me to use data from all countries. However, measurement error may be a more severe problem here than in the fixed effects estimates of Table 2. These results maintain the correct sign but vary more from Table 2—the correlation of these θ^j values with the IV estimates from Table 2 is positive but smaller, at 0.31.

1.5.3 How Reasonable are these Trade Elasticities?

Rauch Test. I evaluate the reliability of this paper's main IV estimates of these elasticities with a simple but informative test: theory predicts that demand should be more elastic for more

[27] The response variable in this Appendix Table 2c regression is the logged sum of a small constant plus the expenditure share. More sophistocated econometric approaches to adressing zero trade flow data in log gravity regressions (Helpman, Melitz, and Rubinstein 2008) impose a large computational burden with these panel data.

homogenous goods should be more elastic than is demand for differentiated goods.[28] I find that the pattern of elasticities across sectors is consistent with this theoretical prediction.

I implement this test using a classification from Rauch (1999), who separates traded goods into three classifications: goods traded on listed exchanges ("homogenous"); goods with reference prices; and all other goods ("differentiated"). Rauch classifies nearly all traded goods by product based on several printed volumes listing prices (e.g., the Knight-Ridder CRB Commodity Yearbook).

Separating goods according to this Rauch classification then using the instrumental variables model of equation (1.8) provides sensible results (Table 3). The instruments for all three types of goods are strong, with first stage F-statistics ranging from 26.4 to 90.3. I estimate that differentiated goods have the smallest trade elasticity in absolute value (-5.75) and homogenous goods have the largest elasticity (-9.18). Although theoretical predictions for reference-priced goods are less clear, those goods have an intermediate elasticity of -5.81. All three elasticities are precisely estimated.[29]

Comparison to Literature. Although researchers have used several strategies to estimate these elasticities, omitted variables and measurement error remain potentially important sources of bias.

Hummels (2001) uses cross-sectional shipping costs for the year 1992 with importer and exporter fixed effects and with proxies for distance and shared language. Caliendo and Parro (2011) compare tariffs and trade flows across two directions in triads of countries.[30] Donaldson (2010) uses price gaps to infer trade costs with panel data and then regresses trade flows on trade costs while including the full set of fixed effects—an analogue to equation (1.7). Eaton and Kortum (2002) measure the price gap between countries for manufactured goods to obtain an economy-wide estimate.

Anderson and van Wincoop (2004) discuss other methods to estimate these parameters. Most either observe a component of trade costs explicitly or measure price gaps between regions. Simonovska and Waugh (2011) provide additional evidence that data issues can substantially affect estimates of θ: they argue that finite samples available for the Eaton and Kortum (2002) methodology create bias of 50 percent.

[28]Broda and Weinstein (2006) implement a similar test for product-level elasticities using the Rauch classification.

[29]Table 3 reports the "conservative" classification of Rauch (1999), which minimizes the number of commodities classified as homogenous or reference priced. Using Rauch's "liberal" classification, which maximizes those numbers, obtains estimates (and standard errors) of -7.26 (2.37), -4.61 (1.58), and -10.5 (1.89) for the differentiated, reference-priced, and homogenous goods, respectively.

[30]For three countries A, B, and C, Caliendo and Parro (2011) regress $\log \frac{X_{AB}X_{BC}X_{CA}}{X_{BA}X_{CB}X_{AC}}$ on $\log \frac{\kappa_{AB}\kappa_{BC}\kappa_{CA}}{\kappa_{BA}\kappa_{CB}\kappa_{AC}}$ where X represents bilateral trade flows and κ represents the import tariff rate. Their main results estimate this regression with no additional controls, separately for each sector. Under the assumptions of no omitted variables, measurement error, or other asymmetric trade barriers, the coefficient from this regression equals θ. When applied to global GTAP data, this approach produced incorrectly-signed estimates for several sectors.

Existing estimates in the literature for the global economy or for manufacturing generally lie in the range -4 to -10 (Anderson and van Wincoop 2004). My estimate of $\theta = -6.68$ for manufactured goods lies in the middle of this range. This estimate is also roughly equidistant between the $\theta = -4.4$ value of Simonovska and Waugh (2011) and the $\theta = -8.28$ value of Eaton and Kortum (2002).

My sector-by-sector range of estimates vary somewhat from other estimates at similar aggregation levels (Caliendo and Parro 2011; Hummels 2001). My most homogenous sector has an elasticity of only -16.1, whereas Caliendo and Parro (2011) find elasticities of -51 to -69, and Hummels obtains an elasticity of -79. My most differentiated sector has an elasticity of -0.76, which is more negative than Caliendo and Parro's -0.37.

1.6 Counterfactual 1: Costs and Benefits of International Trade

This section uses the model together with the data described in the last two sections to measure the full welfare effects of international trade.

1.6.1 Methodology: Costs and Benefits of International Trade

I consider a counterfactual which closes off all countries from international trade. Although this is not a realistic policy, it represents an important benchmark which recurs throughout the trade literature because it provides a starting point for thinking about real-world policies. Recall that x' denotes the value of the variable x after a counterfactual policy is introduced, x denotes the initial value, and $\hat{x} \equiv x'/x$ denotes the proportional change due to a regulation. This counterfactual is equivalent to imposing infinite international trade costs ($\hat{\tau}^j_{od} = +\infty\ \forall o \neq d$) but changing no other variables.

The gains from international trade for country d equal the negative of the change in real income due to autarky (i.e., $\hat{I}_d/\hat{P}_d$) as shown in equation (1.6). The gains from trade can also be written as the change in the share of goods which are purchased from domestic producers (i.e., $\hat{\lambda}^j_{dd}$), weighted by the inverse intermediate share β and trade elasticity θ, and weighted across sectors by the Cobb-Douglas expenditure shares α (Arkolakis, Costinot, and Rodríguez-Clare 2012). Combining, we have $\hat{I}_d/\hat{P}_d = \prod_{j=1}^{J}(\hat{\lambda}^j_{dd})^{\alpha^j_d/\beta^j_d\theta^j_d}$.

Autarky would then produce the following proportional change in welfare for country d:

$$A_d = \left[\prod_{j=1}^{J}\left(\lambda_{dd}^j\right)^{-\frac{\alpha_d^j}{\beta_d^j\theta_d^j}}\right]\left[\frac{1+\left(\mu_d^{-1}\sum_{o=1}^{N}E_o\right)^2}{1+\left(\mu_d^{-1}\sum_{o=1}^{N}E_o'\right)^2}\right] \tag{1.10}$$

Equation (1.10) has a simple economic interpretation. The effect of moving a country to autarky equals the diminished gains from trade (the first bracketed term) multiplied by the change in the environmental costs of trade (the second bracketed term). I aggregate across countries to measure the global welfare effect that would occur if all of the world's countries went to autarky.

Equation (1.10) has an appealing feature: all terms in it are observed in the data. Therefore, calculating the full welfare effects of international trade does not require any kind of optimization algorithm. The counterfactual of autarky permits this straightforward calculation because the domestic expenditure share under autarky is one by definition ($\lambda_{dd}' = 1$). Hence the change in the domestic expenditure share due to autarky equals one divided by the baseline domestic expenditure share ($\hat{\lambda}_{dd}^j = 1/\lambda_{dd}^j$).

The only term in equation (1.10) which I have not previously explained is E_o', representing the CO_2 emissions from country o in autarky. However, E_o' can also be calculated as a function of observed data. This calculation for E_o' reflects the following algebra. By assumption (1.4) and the counterfactual of autarky, we have $E_d' = \Sigma_j(\gamma_3 f_{dd}^j + \chi_d^j)X_{dd}^j\hat{X}_{dd}^j/(p_{dd}^j\hat{p}_{dd}^j)$. Choosing the wage in country d as numéraire, we have $\hat{X}_{dd}^j = X_d^j/X_{dd}^j$. The proportional change in domestic prices due to autarky is $\hat{p}_{dd}^j = (\hat{w}_d)^{\beta_d^j}(\hat{p}_d^j)^{1-\beta_d^j}\hat{\tau}_{dd}^j$. The choice of numéraire and assumption that autarky does not change domestic trade costs imply $\hat{p}_{dd}^j = (\hat{p}_d^j)^{1-\beta_d^j}$. The methodology behind equation (1.10) implies $\hat{p}_{dd}^j = (\lambda_{dd}^j)^{-(1-\beta_d^j)/(\beta_d^j\theta_d^j)}$. Substituting $\hat{X}_{dd}^j$ and $\hat{p}_{dd}^j$ into equation (5b) gives

$$E_d' = \sum_j(\gamma_3 f_{dd}^j + \chi_d^j)(X_d^j)(\lambda_{dd}^j)^{(1-\beta_d^j)/(\beta_d^j\theta_d^j)}$$

Every term in this expression represents observed data—none requires an algorithm or optimization routine to solve.

It may clarify how this framework measures the welfare effects of international trade to walk through four mechanisms by which international trade can affect pollution emissions. First, international trade increases global output. This effect will tend to increase total CO_2 emissions. Second, because international trade requires allows goods to be produced in countries other than where they are consumed, trade changes each country's production overall and by sector. Because CO_2

emissions rates differ by country and sector, this effect can increase or decrease CO_2 emissions from production. Third, international trade requires international goods transportation, which increases CO_2 emissions. Fourth, international trade can increase or decrease domestic goods transportation and the associated CO_2 emissions. The empirical analysis accounts for all of these forces, and Tables 4-7 separately quantify the change in emissions due to production and due to transportation. I find that approximately half of the total CO_2 emissions due to international trade come from production and the remainder from transportation. However, for the EU, US, and global regulations of the CO_2 emissions from transportation, 69 to 98 percent of the change in CO_2 emissions comes from transportation and not production.

1.6.2 Inference: Costs and Benefits of International Trade

To perform inference, I conduct a bootstrap over the stochastic term of the model, θ^j, and report the resulting 95-percent confidence interval. This is useful because, as the introduction emphasized, the trade literature largely has not performed inference on welfare calculations. I report the bias-corrected bootstrap estimate, which can provide an accurate finite-sample approximation (Efron 1987). Appendix 1.C describes the bootstrap algorithm.

1.6.3 Results: Costs and Benefits of International Trade

The analysis provides several results (Figure 4 and Table 4). First, several papers in the trade-environment literature ask, "Is trade good for the environment?" This analysis shows that international trade harms the environment. International trade increases global CO_2 emissions by 6 percent (1.75 gigatons of CO_2 annually or about \$35 billion of global damages). Globally this effect is almost equally driven by production and transportation. This is notable since autarky only directly affects shipping. However, changing the location and level of production has almost equal magnitude as the direct environmental effects of shipping. The proportion of this effect due to transportation versus production varies somewhat by region.

Second, the gains from international trade exceed the environmental costs of international trade by a factor of 149 (i.e., by two orders of magnitude). The gains from international trade exceed the environmental costs of trade in every country (Appendix Table 3). The global gains from international trade, at \$5.3 trillion, equal 10 percent of global GDP. The environmental costs of international trade equal \$35 billion.

Third, a global analysis masks heterogeneity across countries. Not surprisingly, as a share of

GDP, the gains from trade are greatest in countries like Belgium where international trade is a large share of gross output, and smallest in relatively closed countries like the US. Also as a share of GDP, climate change is predicted to have the largest negative effects on poor regions like Sub-Saharan Africa and on India, and the smallest impacts on high-income countries like the US.

Finally, all of these results are precisely estimated. Since the literature on gains from trade generally has not performed inference, it is informative to observe that all the welfare measures are significantly greater than zero.

1.6.4 Discussion: How Sensitive are these Model Calculations?

One useful check on the sensitivity of these results is to ask: how large must the social cost of carbon be to overturn the finding that the gains from trade greatly exceed the environmental cost of trade? Using this paper's framework, I calculate that autarky decreases global welfare if the social cost of CO_2 emissions is any value below \$3,292 per ton of CO_2. This cutoff is over a hundred times the ninety-fifth percentile estimate of the social cost of carbon reported in Greenstone, Kopits, and Wolverton (2011), and several times larger than the estimate under a dynamic climate model that incorporates severe risk aversion, potential climate disasters, tipping points, and uncertain climate damages (Cai, Judd, and Lontzek 2012). It is larger than any estimate of the social cost of carbon I have seen in any source. This calculation provides fairly strong evidence that under any plausible value of the social cost of carbon, trade's benefits greatly exceed its environmental costs.

1.7 Counterfactual 2: EU, US, and Global Carbon Taxes

I now turn to an extremely different type of counterfactual—EU, US, and global environmental regulations which use targeted policy to address the environmental externalities of trade. The previous section found that the aggregate gains from trade substantially exceed the aggregate environmental costs of trade. However, this section shows that these targeted environmental regulations decrease the environmental costs of trade more than they decrease the gains from trade, and so these regulations increase global welfare. This section describes the methodology I use to analyze these regulations; it summarizes the real-world EU, US, and global policy proposals I analyze; it applies the paper's model to analyze these proposed regulations; and finally it compares these results to a simple Harberger triangle model of the demand for shipping fuels.

1.7.1 Methodology: Effects of EU, US, and Global Carbon Taxes

Measuring the effects of untested climate change regulations requires constructing an empirical analogue to the equivalent variation in equation (1.6). Algebra using the model's assumptions can express the model as the following system of $N-1$ nonlinear equations (one per country, excluding a numéraire due to Walras' Law) in $N-1$ unknown wage changes $\hat{w}_d$:

$$\sum_{o,j} \frac{X_{od}^{j\prime}(\hat{w}_d)}{1+t_{od}^{j}} = \sum_{o,j} \frac{X_{do}^{j\prime}(\hat{w}_d)}{1+t_{do}^{j}} + \phi_d^{j\prime} + T_d \tag{1.11}$$

Here the matrix $X_{od}^{j\prime}(\hat{w}_d)$ is a known function of observed data and of the wage changes $\hat{w}_d$ (see Appendix 1.D for details). Equation (1.11) states that we can determine the effect of counterfactual regulations by finding the wage changes $\hat{w}_d$ which restore trade balance in all countries. Every term in (1.11) is either reported in the data or is a known function of the wage change due to a carbon tax.

I use a trust-region dogleg algorithm (Nocedal and Wright 2006) to solve this system for equilibrium wages, with numéraire chosen so $\sum_d w_d L_d = \sum_d w_d' L_d$. Given the candidate wage vector $\hat{w}_d$ at a given iteration of the algorithm, I estimate the price vector $\hat{p}_d^j$ by using a contraction map which iterates over the price equation (1.D.1) described in the Appendix. Alvarez and Lucas (2007) prove that the equilibrium price and wage vectors are unique, and that this contraction map recovers the unique equilibrium price vector.[31] After recovering the $N-1$ values of $\hat{w}_d$ from this algorithm, calculating $\hat{X}_d$, $\hat{P}_d$, and E_d' just requires arithmetic using the model's assumptions. Finally, I substitute these values into (1.6) to find the effect of a counterfactual carbon tax on social welfare in each country.

For each carbon tax counterfactual, I introduce a shock by changing the value of the carbon tax t_{odm}^{j} from 0 to \$19.96 per $MtCO_2$. As discussed in section 1.3.3, I choose \$19.96 because it reflects a leading estimate of the social cost of CO_2 emissions. It is also similar to forecast prices of one-ton carbon allowances in the EU ETS in the years 2012-2020 (Point Carbon 2012). This carbon tax changes trade costs through assumption (1.3b). I then calculate $\hat{\tau}_{od}^{j}$, the change in trade costs due to the carbon tax. Finally, I feed this value of $\hat{\tau}_{od}^{j}$ into the model to estimate the effects of the carbon taxes.

[31] Their paper describes an exchange economy with different microfoundations than what I describe here. Because the same equations determine wages in equilibrium in the two economies, the uniqueness property also applies in this setting.

The paper reports the total effects of each policy over its first decade of implementation. This follows standard practice—the EU has planned aviation ETS allowances for the period 2012 through 2020, and many evaluations of proposed US regulations use budget scoring over a 10-year time horizon. The main results hold global aggregates fixed over the decade, so they equal ten times a policy's annual effects.[32]

In all these calculations, I conduct inference using a bias-corrected bootstrap with 200 replications over the distributions of θ^j estimated by instrumental variables in Table 2, as described in section 1.6.2.

1.7.2 Regulation Details

For each of the EU, US, and global climate change regulations, I use the model to analyze a counterfactual which represents a stylized version of the real-world policy.

EU ETS Counterfactual. The EU's Emissions Trading System (ETS) began in 2005 and represents the world's largest climate change regulation. The ETS sets an EU-wide cap for regulated CO_2 emissions, distributes CO_2 "allowances" to firms, then lets firms buy and sell those allowances. Each year, firms must provide the EU with allowances to cover their regulated CO_2 emissions. In 2011, the ETS regulated CO_2 emissions from five industries: electricity generation; oil refining; iron and steel; cement, glass, lime, brick, and ceramics; and pulp, paper, and boards. In January 2012, the ETS added a sixth industry, air transportation. The ETS regulates CO_2 emissions from an entire flight leg, and not only from the component which occurs within EU airspace. The EU distributed 85% of airline CO_2 allowances for free (a "grandfathering" system) based on each airline's year 2010 emissions, then auctioned the remaining 15 percent. Aviation CO_2 allowances may be traded one-for-one with allowances from the other regulated sectors.

This paper's EU counterfactual is similar but not identical to the EU ETS. Like the ETS, I consider the regulation of CO_2 emissions from airplane flights involving the 30 countries participating in the EU ETS. The ETS regulates all airplane transportation, whereas I include only shipping. Finally, the EU is initially distributing 85 percent of permits for free to airlines, whereas I treat the ETS as equivalent to a carbon tax, and so assume all permits are auctioned. Cap-and-trade systems and carbon taxes are economically equivalent if regulators have no uncertainty about marginal abatement and marginal cost schedules and firms have no transaction costs (Weitzman 1974).

[32] Allowing for trends in global aggregates would require credible forecasts of bilateral trade between all countries for each sector and year (λ^j_{od}), which are difficult to obtain.

US Waxman-Markey Counterfactual. The second counterfactual analyzes the regulation of CO_2 emissions from all US shipping. This analysis reflects the Waxman-Markey bill, which passed the US House but not Senate in 2009, and would have created a cap-and-trade system for US CO_2 emissions. The bill included refineries' petroleum products and fuel imports in the CO_2 emissions cap, though did not regulate shipping firms directly.

Like this bill, my US counterfactual analyzes the regulation of all shipping—by air, sea, rail, and road. Unlike the bill, I study a carbon tax which affects CO_2 emissions from both imports and exports. Moreover, I focus only on the regulation of goods transport and not passenger transportation, and I ignore other components of the Waxman-Markey bill.

Global Kyoto Protocol Counterfactual. Article 2.2 of the 1997 Kyoto Protocol called for UN agencies to develop a cap-and-trade policy for plane and sea emissions for 41 industrialized countries. This remains under negotiation but has never been implemented. My global counterfactual analyzes the regulation of all domestic and international airborne and maritime shipping.

1.7.3 Effects of EU, US, and Global Carbon Taxes

To explain the effects of these counterfactual policies, I organize findings by the main results of interest. Table 5 summarizes effects of the EU carbon tax, Table 6 the US carbon tax, and Table 7 the global carbon tax. Appendix Table 3 lists country-by-country results, and Figure 5 maps them.

Global Welfare

All three counterfactuals increase global welfare because they each decrease the global environmental costs of trade more than they decrease the global gains from trade. For the EU air regulation, the global gains from trade fall by \$200 million over a decade, but the global environmental costs of trade fall by \$1.8 billion over a decade. These effects are proportionally larger for the US and global policy because those regulations affect more trade – the US and global carbon taxes decrease the global gains from trade by \$1.2 billion and \$2.6 billion, and the global environmental costs of trade by \$5.8 billion and \$9.9 billion. In total, the EU, US, and global policies increase welfare by \$1.6, \$4.6, and \$7.3 billion over a decade, respectively.

These welfare results occur because all of these policies cause both international trade and CO_2 emissions fall. For the EU policy, they fall by \$76 billion and 93 $MtCO_2$ over a decade, respectively. Air trade involving the EU emitted 75 $MtCO_2$ in the year 2007, so this regulation causes annual CO_2 emissions to fall by 12 percent relative to the baseline regulated level. The US regulation decreases international trade involving the US by $233 billion over a decade, and it decreases CO_2 emissions involving the US by 319 MtCO_2 over a decade. US shipping emitted 901 MtCO_2 in the year 2007, so this change represents a decline of about 3.5 percent relative to the regulated baseline. The global regulation decreases global international trade by $591 billion over a decade.

Recall that because these policies regulate only transportation, they have theoretically ambiguous effects on CO_2 emissions from production: they could relocate output to countries where production is more CO_2-intensive, which would increase global CO_2 emissions; and they could decrease global production overall, which would decrease global CO_2 emissions. In practice, I find that regulating CO_2 emissions from transportation leads to no increase in the CO_2 emissions from production, and in fact decreases CO_2 from production by amounts which account for between 2 and 31 percent of the policies' total environmental benefits.

Regressivity

All of these regulations are regressive and actually decrease welfare in poor countries. The EU policy increases welfare in the richest third of countries (measured by GDP per capita) by 0.1 basis points, decreases welfare in the middle third of countries by the same proportion, and decreases welfare in the poorest third of countries by 0.3 basis points. These effects occur due largely to the fuel consumption patterns emphasized earlier—airplane trade generally takes place with distant countries (and countries far from the EU are poorer than countries near to the EU); and poor countries tend to trade goods that have high weight-to-value ratios, which are fuel-intensive to transport. Appendix Figure 3 plots each country's per capita income against the country-specific impact of an EU carbon tax. This figure makes clear that poor countries pay substantially more than richer countries for this tax. Countries like Uganda, Namibia, and Laos which pay the most for this tax in welfare terms are also among the world's poorest.

The US regulation generates similar patterns. This regulation increases welfare for the richest third of countries by a third of a basis point, decreases welfare for the middle third of countries by half a basis point, and decreases welfare for the poorest countries by three-fourths of a basis point. Because rail, road, and boat trade are concentrated among geographically proximate countries, this regulation disproportionately harms countries that are geographically close to the US. The global regulation generates similar patterns—this regulation increases welfare in the richest third of countries by 0.4 basis points, decreases welfare in the middle third by half a basis point, and decreases welfare in the poorest countries by nearly one basis point.

Private Gains from Trade

A third finding is that the EU and US regulations increase the implementing region's gains from trade, even before accounting for environmental benefits. This result occurs because these regulations act like a unilateral tariff which improves a country's terms of trade.

The EU regulation decreases real labor income in the EU by $8 billion over a decade, because consumers and producers face a price wedge in buying and selling goods. However, this regulation increases EU carbon tax revenues by $14 billion over a decade. In total, the EU's gains from trade rise by $6 billion. This policy decreases welfare for most countries outside the EU.

Appendix Figure 4 plots the welfare effects of the EU policy against the air carbon tax that the EU imposes, assuming that the true social cost of carbon is $20/t$CO_2$. These graphs makes clear that the optimal carbon tax from the EU perspective is about $3000/t$CO_2$, which is larger than any estimate of the social cost of carbon emissions. This is comparable to a 25 percent tariff.[33] Global welfare, however, is maximized at a carbon tax rate closer to $30/t$CO_2$. This globally optimal EU tariff is modestly higher than the social cost of carbon in part because the EU regulation decreases CO_2 from unregulated sources like production.

Similarly, the US regulation increases the US gains from trade, even before accounting for environmental benefits. This regulation decreases real US wages by $150 billion but generates $170 billion in US carbon tax revenue, leading to a $20 billion increase in the US gains from trade over a decade.

Trade Diversion and Leakage

A fourth finding is that these policies generate small amounts of trade diversion and leakage. The EU regulation decreases international trade involving the EU by $93 billion over a decade but increases trade not involving the EU by $17 billion over a decade. Thus, 23 percent of the EU's decrease in international trade is offset by international trade increases elsewhere in the world. Hardly any leakage occurs for the environment—transportation CO_2 emissions fall by 91 MtCO_2 for the EU nearly do not change elsewhere, and production CO_2 emissions actually fall elsewhere.

The US regulation generates slightly more trade diversion and leakage. This policy decreases international trade involving the US by $233 billion over a decade, but increases international trade not involving the US by $79 billion over a decade. This implies that a third of the US decline in

[33] Fuel costs represent about 1.4 percent of the value of goods for international trade, and a $19.96/t$CO_2$ carbon tax equals a 12 percent increase in fuel costs.. The tariff-equivalent comes from the calculation 0.014*.12*3000/19.96.

international trade is offset by increases in international trade elsewhere in the world. US CO_2 emissions from transportation fall by 207 $MtCO_2$ while CO_2 emissions from transportation elsewhere rise by only 6 $MtCO_2$, making leakage from transportation relatively unimportant. Production CO_2 emissions do rise in the rest of the world, and offset about 19 percent of the US decline in production CO_2 emissions.

1.7.4 How Robust Are These Model Calculations?

This section has used a gravity model to evaluate several potential climate change regulations. I now assess how the results compare against a much simpler model—a textbook Harberger (1964) triangle analysis of the demand for shipping fuels (the "Harberger model"). Overall, the two models estimate similar global effects of the taxes. This provides one piece of evidence that the gravity model obtains plausible estimates. The exact numbers differ, however, and the gravity model estimates many results which the textbook partial equilibrium model does not.

The classic Harberger model counts deadweight loss triangles in the area between supply and demand curves for shipping fuels. Suppose that a social planner imposes a tax rate of t on all shipping fuels, as in Appendix Figure 5. All shipping fuels have the demand elasticity η and have perfectly elastic supply. The effect of this tax on what I will call the gains from shipping is the area G under the demand curve; this is analogous to the gravity model's estimate of the gains from trade. The effect on the environmental costs of shipping is G+E, which is analogous to the gravity model's estimate of the environmental costs of trade. The effect of the tax on tax revenue is R, and the effect on social welfare is E.

The Harberger and gravity models have a close economic relationship because firms and consumers demand shipping fuels in order to access to varieties of goods. The Harberger model analyzes the demand for fuels, which are treated as a homogenous good described by one global demand elasticity. By contrast, the gravity model analyzes the demand for varieties of goods while accounting for the energy required for production and transportation.

Despite this close relationship, the Harberger model differs from the gravity model in five ways. First, the Harberger model applies the elasticity η to all shipping. Demand elasticities may differ by good and policy design. For example, taxing only airplane trade may be associated with a larger elasticity than taxing all modes of transportation.[34] Second, the Harberger model ignores potential

[34] The demand elasticity is also difficult to estimate credibly—the few papers which report a parameter like a global demand elasticity for shipping fuels use time-series variation in global oil prices for identification (Dargay and Gately

increases in consumption of unregulated fuels. For example, an EU tax on jet fuels could increase bunker fuel consumption. Third, the Harberger model ignores CO_2 emissions from production. Fourth, the Harberger model does not measure incidence. Finally, the Harberger model ignores general equilibrium effects, and does not ensure market clearing.

While there is a range of estimates of the demand elasticity for petroleum products, most of which do not distinguish bunker and jet fuel from other products, I consider two values which bracket most estimates in the literature: -0.02, and -0.50 (Dargay and Gately 2010; Faber, Markowska, Eyring, Cionni, and Selstad 2010; Mazraati 2011). Recent estimates of the demand elasticity for petroleum products are closer to -0.02 than to -0.50. For each statistic using the Harberger model, I report two separate results (one per demand elasticity).

The gravity estimate of how the EU, US, and global carbon taxes affect the global gains from trade is roughly equidistant from the two corresponding Harberger estimates (Figure 6 and Appendix Table 4). For the EU tax, the gravity model predicts a decline in the gains from trade of \$210 million over a decade, while the textbook model predicts a change in the gains from shipping of between \$20 million and \$440 million. For the US and global taxes, the gravity model again lies roughly in the middle of the two corresponding Harberger estimates. The gravity model predicts country-by-country incidence, but because the Harberger triangle model does not, I cannot compare incidence results.

The gravity and Harberger estimates of how these taxes affect CO_2 emissions are similar but not identical. For the EU policy, the gravity model calculates a larger decline in transportation CO_2 emissions than the Harberger model does. One possible explanation is that the EU tax only regulates one mode of transportation (airplanes), which emits 100 times as much CO_2 as sea shipping does. For the US and global taxes, which cover more modes of transportation, the textbook and Harberger models provide similar predictions.

1.8 Extensions and Robustness

This section of the paper assesses how the model's results change with its assumptions (Tables 8a and 8b).

2010; Faber, Markowska, Eyring, Cionni, and Selstad 2010; Mazraati 2011).

1.8.1 Environmental Assumptions

I first assess how results change under low and high possible values for the social cost of CO_2: \$4.10/tonCO_2, and \$1170/tonCO_2. Under the lower social cost of carbon, the environmental costs of trade are \$8 billion/year; while under the greater social cost of carbon, the environmental costs of trade are much larger. Under either assumption, the gains from trade remain substantially larger larger than the environmental costs of trade.

The social cost of carbon is more relevant to evaluating climate change regulations. Under a low social cost of carbon (\$4.10/tonCO_2), a global tax on air and sea shipping produces a welfare gain over a decade of \$4.4 billion. Under a higher social cost of carbon, the welfare gain is far larger. Under all of these alternative assumptions, most environmental benefits are due to transportation and not production.

1.8.2 Trade Assumptions

I also consider how changing the model's trade assumptions affect the paper's main findings. First I consider an elasticity for each sector of $\theta = -4.12$, or alternatively an elasticity for each sector of $\theta = -8.28$. These represent leading estimates from the literature (Eaton and Kortum 2002; Simonovska and Waugh 2011), and they roughly bracket my estimates. The results reflect the fact that the gains from trade are exactly proportional to this parameter. Because an elasticity of $\theta = -8.28$ reflects a world where goods are more homogenous, the global gains from trade are smaller, at \$2.2 trillion per year. By contrast, the value $\theta = -4.12$ reflects a world with more differentiated goods, and the gains from trade are relatively larger, at \$4.4 trillion per year. These modifications do not substantially affect the counterfactual analysis of global air and sea shipping.

Finally, I consider a potentially important theoretical restriction—I modify the model to turn off general equilibrium effects. To be clear, turning off general equilibrium effects means that the results do not represent an equilibrium—supply does not equal demand. The purpose of this sensitivity analysis is to learn about the importance of model-estimated wage changes for the model's predictions. Overall, I find that accounting for general equilibrium effects is important for the magnitude of the model's results, though not for its qualitative conclusions. The model's key general equilibrium effect is that nominal wages w_d in each country change once a carbon tax is introduced, so as to restore trade balance. In row 7 of Table 8b, I account for each counterfactual's change in trade costs, while holding all nominal wages fixed. Turning off general equilibrium effects increases the estimated effect of the global carbon tax on the gains from trade by about 80 percent, but estimated effect on the environmental costs of trade by about 50 percent. Put another way, the wage changes required to clear markets tend to dampen these regulations' costs and benefits, making the regulations have smaller effects than they would if nominal wages were fixed. General equilibrium effects are less important for the EU and US carbon taxes.

1.9 Conclusion

This paper seeks to contribute to research on trade and the environment in three ways. First, it draws on trade theory to develop a new approach to evaluating regulation. The paper weds a structural general equilibrium model with reduced-form estimates of key parameters.[35] Although the full theory depends on numerous parameters which are difficult to identify and estimate, measuring the effects of policies on social welfare depends on only one set of elasticities which I estimate. This framework accounts for general equilibrium effects and analyzes untested policies while using straightforward econometric tools to estimate key parameters. A bootstrap allows for statistical inference on the model's welfare calculations.

Second, this paper compares international trade's benefits against its environmental costs in a unified theoretical and empirical framework. The gains from international trade exceed the environmental international costs of trade by two orders of magnitude, and they are precisely estimated.

Third, this paper analyzes the incidence and aggregate welfare effects of proposed climate change regulations. I study policies under the EU's Emissions Trading System, the US Waxman-Markey Bill, and the 1997 Kyoto Protocol, which would each regulate the CO_2 emissions from some forms of shipping. Poor countries specialized in trading raw materials like grain and iron ore, particularly those in Sub-Saharan Africa, lose the most from these policies. Because they regulate shipping for only some countries or modes of transportation, these policies slightly increase unregulated CO_2 emissions and divert trade to unregulated routes. These policies also create unequal incidence by increasing welfare in the implementing region and decreasing welfare elsewhere, even before accounting for environmental benefits. Nonetheless, all three of these policies increase global welfare because they decrease the environmental costs of trade more than they decrease the gains from trade.

[35] This approach has similar spirit to the literature on "sufficient statistics for welfare analysis" (Chetty 2009).

This paper demonstrates one way in which new ideas in international economics can clarify the links between trade and the environment. Such links are widespread—arguably, most environmental policies affect trade and most trade policies affect the environment. Similar approaches that may shed light on other questions in trade and the environment provide a good basis for building on this paper.

For example, a limitation of this paper is its focus on climate change to the exclusion of other pollutants like particulate matter. Accounting for these other pollutants could increase or decrease the estimated environmental costs of trade. Using a framework which accounts for heterogeneity across firms, it is possible to use extensive US plant-level data on these pollutants to analyze the benefits and costs of regulations which affect both trade and the environment.

Appendices

1.A Data Appendix

1.A.1 Greenhouse Gas Emissions by Sector, 1990-2008

The paper describes new data on CO_2 emissions for each sector and country pair in the year 2007. Here I discuss data behind the introduction's claim that CO_2 emissions from international shipping are growing faster than are greenhouse gas emissions from other sectors.[36]

IEA (2011) data are incomplete but support this claim (Appendix Table 5). These data combine CO_2 emissions from fuel combustion and greenhouse gas emissions from other activities. Greenhouse gas emissions from international air and sea transportation grew by 57 percent between 1990 and 2005, and by 71 percent between 1990 and 2008 (Panel A). This exceeded the growth rate of any other sector. Over this period, total greenhouse gas emissions grew by 30 percent, implying that international transportation grew at roughly twice the rate of other sectors. When greenhouse gases other than CO_2 are excluded, international shipping still had the fastest growth of any sector in the global economy (Panel B). For manufacturing and transportation only, the IEA reports CO_2 emissions at a more detailed industry level, which also support this claim (Panel C).[37]

1.A.2 Transport Modes

I use a few general rules to compile the data. I exclude observations with unknown trading partners or products. I convert all foreign currencies to dollars using the mean period exchange rate from the IMF's International Financial Statistics, then deflate values to the year 2007 using the US Bureau of Labor Statistics Consumer Price Index. I exclude small islands which cannot be linked to other data in the study.[38] Where possible, I use importer reports. When a trade flow reports currency but not weight, I impute weight using the mode-specific weight-to-value ratio at the most detailed level possible from all other countries reporting transportation modes (6 digit, or otherwise 5 digit, etc.).[39]

Some decisions are specific to each data source. For EU trade, I treat "inland waterway" trade

[36] IMO (2009) estimates that non-CO_2 greenhouse gases (CH_4, N_2O, and HFCs) account for less than 2 percent of the CO_2-equivalent of freight.

[37] The transportation equipment industry builds planes and ships for international transportation so is not distinct from international transportation. This calculation uses a concordance which I obtained from IEA staff which defines the hierarchical relationship of manufacturing and transportation industries in the published IEA data, so as to avoid double-counting.

[38] The exclusion includes St Martin (Dutch and French parts), the Canary Islands, and Ceuta.

[39] This typically occurs when a country uses a physical unit besides weight – number of t-shirts, pounds of steel, etc. – to quanity the volume of trade. HS codes have globally standard 6-digit definitions but each country chooses the physical unit(s) for reporting each 6-digit HS code.

as maritime trade. For US imports, I sum freight charges and product values to obtain the goods' value, since I use importer-reported trade values elsewhere (which include freight charges). Japan only distinguishes transport mode for airborne and container ship trade, so I assign additional Japanese trade values (obtained from the same Trade Statistics of Japan source) to sea shipment. I use the HS-to-14-sector concordance file described above to link these HS codes to the 14 sectors I analyze.

I obtain EU data at the 2-digit HS code level, so I use the procedure described above for the Australian freight data. In mapping 2-digit trade data to the 14 sectors I analyze, I apply value shares to the trade value data and weight shares to the trade weight data.

I impute transportation mode shares for remaining 17-26 percent of trade using fractional multinomial logit. Let x_{od} denote the covariates used to impute mode shares. x_{od} includes 11 variables: log importer and exporter GDP per capita and their squares, log bilateral distances by air and by sea and their squares, and dummy variables identifying landlocked, contiguous, and and island countries.

In a fractional multinomial logit, the share of the *o-d* trade flow transported by mode m is

$$\sigma_{odm} = \begin{cases} \frac{1}{1+\sum_{\tilde{m}=2}^{M} \exp(x_{od}\gamma_{\tilde{m}})} & \text{if } m = 1 \\ \frac{\exp(x_{od}\gamma_m)}{1+\sum_{\tilde{m}=2}^{M} \exp(x_{od}\gamma_{\tilde{m}})} & \text{if } m > 1 \end{cases} \tag{1.A.1}$$

One mode is arbitrarily chosen as the base category $m = 1$. The corresponding log likelihood for observation odm is

$$\ln(L_{odm}) = \sum_{\tilde{m}=1}^{M} \ln(\tilde{\sigma}_{od\tilde{m}})\, \sigma_{od\tilde{m}} \tag{1.A.2}$$

where $\tilde{\sigma}_{odm}$ represents the fitted value from (1.A.1).[40]

I use a secondary reference to impose mode shares in one case—UNECA (2010, p. 214) reports that 80 percent of intra-African freight transportation moves by road. I impose this statistic on all intra-African trade, then estimate the division of remaining trade between sea, rail, and air using equation (1.A.2).[41]

For equations (1.3b) and (1.3c), I impute weight-to-value ratios (W) for the quarter of world trade where weight is missing. Using data from global trade, I measure W separately for each of

[40] Equations (1.A.1) and (1.A.2) resemble a multinomial logit but with one important difference. Logit outcomes are binary, but here outcomes lie in the continuum $[0, 1]$. Hence, for each observation, the likelihood (1.A.2) sums the fitted mode shares, weighted by the observed shares.

[41] Cristea, Hummels, Puzzello, and Avetisyan (ming) make a similar adjustment.

the 13 tradable sectors and 4 observed transport modes, then apply these values to the missing data.

1.A.3 Freight Costs

I use a few rules to compile the Australian and US data. The Australian data report the value of goods at their port of origin and port of destination, and I define the shipping cost as the difference between these values. I exclude the few observations where shipping costs are negative, or outlying observations where the shipping cost exceeds the goods' value (which represent about a tenth of a percent of the aggregated data). For both importers, I exclude observations which list the exporter as unknown or where the exporter is not a country. I use a quarterly price deflator for these data from the Bureau of Labor Statistics Consumer Price Index.

Defining the 14 sectors in these data requires constructing one concordance file for the US data and a separate concordance file for the Australian data. The US data use different revisions of the Harmonized Commodity Description and Coding System (HS) codes (1992, 1996, 2002, and 2007). I construct a concordance file which links HS codes from each revision to the 14 sectors I analyze. For the 2002 revision, I use a dataset created by Thomas Hutcheson as part of GTAP which links each 6-digit HS code to the 14 sectors I analyze. For the 2007 HS revision, I invert a 2002-to-2007 concordance which the UN Statistics Division created, and I then apply the 2002 concordance described above. For the 1996 HS revision, I use a concordance file created by Robert McDougall and Mark Gehlhar as part of GTAP. For the 1992 revision, I invert the UN's 1992-to-1996 HS concordance file then apply the 1996 concordance described above. I find 21 6-digit HS codes which appear in the US trade data but not in this concordance file. I assign these codes to a sector based on the concordance for the same code in a different HS revision, or based on the assignment of adjacent HS codes.[42]

The Australian data are only available at the 2-digit HS code level. To link these data to the 14 sectors I analyze, I construct a concordance linking each year of the Australian data to a sector. Using 6-digit HS code trade value data from each year of UN-Comtrade (a source which reports trade value and weight but not mode) for Australian imports only, for each HS code-by-trading partner-by year cell, I measure the share of value which falls in each of the 14 sectors I analyze. I

[42]In inverting the 2002-to-2007 concordance file, for the few cases where one year 2007 code links to multiple year 2002 codes, I uniquely link it to the first 2002 code ordered numerically. I use the same procedure to invert the 1992-to-1996 HS concordance file. Unclassified trade (HS=999999) are mapped to the "Other" sector.

then apply these shares to the Australian data.

The Rauch (1999) classification of homogenous, reference-priced, and differentiated goods is defined in terms of SITC Revision 2 codes. To estimate trade elasticities according to this classification, I use concordances published by the UN Statistics Division for all four HS revisions which link 6-digit HS codes to SITC Revision 2 codes. I use this approach to define all the US trade data in SITC Revision 2 codes. I apply the same approach I use for the main data to define the Australian shipping cost data in terms of the three Rauch categories.

1.A.4 Global Gross Output, 1990-2009

Appendix Figure 1 compares gross output to international trade, summed over all countries, for the period 1990-2008. I am not aware of any panel data on gross output which includes all countries. So to obtain this series, I impute gross output from global GDP data as follows.

I first calculate the ratio of gross output to GDP for each country-year with data on both variables in OECD-STAN in the period 1990 to 2009, using the PROD and VALUE series. I then regress this ratio on country fixed effects and year fixed effects. I use the estimated constant and year fixed effects from this regression to calculate the global ratio of gross output to GDP. Finally, I multiply this fitted global ratio by the global GDP values from World Bank-WDI to obtain the global gross output measures plotted in Appendix Figure 1.

1.B Comparison to Estimates by International Organizations

This section discusses the assumptions required to compare my estimates of air and sea CO_2 emissions to the estimates of international organizations, as in Figure 3. For air travel, the global organization for airlines (IATA 2009) reports that all air transportation moved 498.7 billion ton-km in 2007 and that 167.7 billion of this represented freight. (The remainder is mail, passengers, and passenger baggage.) So globally, 33.6 percent of air transportation ton-km represents freight.

The EU air estimate is calculated as follows. To add air transportation to the ETS, the EU collected data indicating that in the years 2004-2006, flights to and from the EU emitted an annual mean of 221.4 $MtCO_2$. Applying the 0.336 freight/transportation ratio described above implies annual EU air freight emissions of 74.61 $MtCO_2$. Applying the 5.7 percent 2005-2007 growth in international air transportation CO_2 emissions (IEA 2011) implies an EU-reported total of 78.86 $MtCO_2$.

The IEA international air estimate is calculated as follows. The IEA reports that international air transportation emitted 431 $MtCO_2$ in the year 2007. Assuming again that 33.6 percent of this represents we have an IEA estimate for air freight of 145 $MtCO_2$.

The ICAO total air estimate is calculated as follows. The ICAO (2009) estimates that domestic and international aviation emitted a total of 632 $MtCO_2$ in the year 2006. I inflate this by the 2006-2007 5% growth in global air ton-km reported by IATA (2009), then multiply by the IATA freight/total ton-km ratio of 0.336 to obtain an ICAO estimate of 223 $MtCO_2$ for international plus domestic air freight.

1.C Inference

This section describes the methodology for bias-corrected bootstrap estimates of the 95-percent confidence intervals for counterfactual calculations. This bootstrap takes $B = 200$ draws of the 13×1 vector θ from the 13 independent normal distributions that have mean and standard deviation given by the instrumental variables parameter estimates and standard errors of Table 2, columns 7-8. If any element of the bth draw $\theta(b)$ is positive and so economically infeasible, I re-draw the $\theta(b)$ vector until I obtain negative values. This procedure consistently estimates the true confidence interval under the null hypothesis of assumption (1.1) that $\theta^j < 0$. For each draw $\theta(b)$, I calculate the model's estimate $\zeta(b)$ of the parameter of interest. $\zeta(b)$ for different table entries represents welfare, international trade, or pollution.

Given these draws, I report the bias-corrected bootstrap estimate of the 95-percent confidence region, which can provide an accurate finite-sample approximation (Efron 1987). The bootstrap estimate of the confidence region is given by the pair $(\zeta^{(\alpha_1)}, \zeta^{(\alpha_2)})$, where $\zeta^{(\alpha)}$ denotes the 100$\cdot\alpha$th percentile of the B estimates $\zeta(1), \ldots, \zeta(200)$. The unadjusted percentiles for the 95-percent confidence interval are $\alpha_1 = 0.025$ and $\alpha_2 = 0.975$. The bias-corrected percentiles are

$$\begin{aligned}
\alpha_1 &= \Phi\left(2z_0 + z^{(\alpha)}\right) \\
\alpha_2 &= \Phi\left(2z_0 + z^{(1-\alpha)}\right)
\end{aligned}$$

Here $\Phi(\cdot)$ represents the standard normal cumulative distribution function (CDF) and $z^{(\alpha)}$ represents the 100$\cdot\alpha$th percentile of a standard normal CDF. The bias correction coefficient z_0 is calculated from the share of bootstrap estimates $\zeta(1), \ldots, \zeta(200)$ which are less than the full-sample estimate ζ:

$$z_0 \equiv \Phi^{-1}\left(B^{-1}\sum_{b=1}^{B} 1[\zeta(b) < \zeta]\right)$$

Here $1[\cdot]$ represents the indicator function, which takes the value one if its argument is true and zero otherwise, and $\Phi^{-1}(\cdot)$ represents the inverse of a standard normal CDF.

1.D Welfare Effects of Climate Change Regulation

In equation (1.11), measuring the effect of climate change regulations requires calculating $X_{od}^{j\prime}$ as a function of $\hat{w}_d$ and known data. I calculate $X_{od}^{j\prime}$ from

$$X'(\hat{w}_d) = [I - F(\hat{w}_d)]^{-1}\,\Psi(\hat{w}_d)$$

To explain these matrices, I begin with three terms which are easiest to derive. The budget constraint, trade balance, and gravity equation imply the following three equations:

$$\hat{c}_o^j = (\hat{w}_d)^{\beta_d^j}\,(\hat{p}_d^j)^{1-\beta_d^j}$$

$$\hat{p}_d^j = \left[\sum_{o=1}^{N} \lambda_{od}^j(\hat{c}_o^j\hat{\tau}_{od}^j)^{\theta^j}\right]^{1/\theta^j} \tag{1.D.1}$$

$$\hat{\lambda}_{od}^j = \left(\frac{\hat{c}_o^j\hat{\tau}_{od}^j}{\hat{p}_d^j}\right)^{\theta^j} \tag{1.D.2}$$

These relationships represent the proportional effect of a regulation on production costs, prices, and trade flows.

I now turn to explain the main calculation. X' is an $NJ \times 1$ vector representing expenditures after a carbon tax is imposed.[43] The vector X' is ordered by country then sector, so that the first 14 entries represent the values for the first country; the second 14 entries represent values for the second country, etc. . I is an $NJ \times NJ$ identity matrix. $\Psi(\hat{w}_d)$ is an $NJ \times 1$ vector defined as follows:

$$\Psi(\hat{w}_d) \equiv \alpha_d^j\hat{w}_d\,(w_dL_d) - (1-\beta_d^j)T_d^j + \alpha_d^jT_d$$

Finally, the $NJ \times NJ$ matrix $F(\hat{w}_d)$ is the sum of four separate $NJ \times NJ$ matrices:

[43] As in the main text, x' represents the value of the variable x after a regulation is imposed. No vectors or matrices in this section are transposed.

$$F = A + B + C + D$$

These four matrices are defined as follows, where $G_{od} = 1$ if the importer receives the tariff revenue and $M_{od} = 0$ otherwise:

$$\begin{aligned}
A &= diag\left((1-\beta_d^j) \sum_{o=1}^{N} \frac{\lambda_{od}^{j\prime}}{1+t_{od}^j} \left[1 + t_{od}^j (1 - G_{od}) \right] \right) \\
&= \begin{bmatrix} \left(1-\beta_1^1\right) \sum_o \frac{\lambda_{o1}^{1\prime}\left[1+t_{o1}^1(1-G_{o1})\right]}{1+t_{o1}^1} & & & 0 \\ & \left(1-\beta_1^2\right) \sum_o \frac{\lambda_{o1}^{2\prime}\left[1+t_{o1}^2(1-G_{o1})\right]}{1+t_{o1}^2} & & \\ & & \ddots & \\ 0 & & & \left(1-\beta_N^J\right) \sum_o \frac{\lambda_{oN}^{J\prime}\left[1+t_{oN}^J(1-G_{oN})\right]}{1+t_{oN}^J} \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
B &\equiv diag\left(\begin{bmatrix} \rho_d^{j,k=1} & \cdots & \rho_d^{j=1,k=J} \\ \vdots & \ddots & \vdots \\ \rho_d^{j=J,k=1} & \cdots & \rho_d^{j=K,k=J} \end{bmatrix} \right) \\
&= \begin{bmatrix} \begin{bmatrix} \rho_{d=1}^{j=1,k=1} & \cdots & \rho_{d=1}^{j=1,k=J} \\ \vdots & \ddots & \vdots \\ \rho_{d=1}^{j=J,k=1} & \cdots & \rho_{d=1}^{j=K,k=J} \end{bmatrix} & & 0 \\ & \ddots & \\ 0 & & \begin{bmatrix} \rho_{d=N}^{j=1,k=1} & \cdots & \rho_{d=N}^{j=1,k=J} \\ \vdots & \ddots & \vdots \\ \rho_{d=N}^{j=J,j=1} & \cdots & \rho_{d=N}^{j=J,k=J} \end{bmatrix} \end{bmatrix} \\
\rho_d^{j,k} &= \alpha_d^j \sum_{o=1}^{N} \frac{t_{od}^k G_{od}^k \lambda_{od}^{k\prime}}{1+t_{odj}}
\end{aligned}$$

$$
\begin{aligned}
C &\equiv \left(\alpha_o^j \frac{t_{od}^k\left(1-G_{od}^k\right)\lambda_{on}^{k\prime}}{1+t_{od}^k}\right) \\
&= \begin{bmatrix}
\alpha_1^1 \frac{t_{11}^1(1-G_{11}^1)\lambda_{11}^{1\prime}}{1+t_{11}^1} & \dots & \alpha_1^1 \frac{t_{11}^J(1-G_{11}^J)\lambda_{11}^{J\prime}}{1+t_{11}^J} & \dots & \alpha_1^1 \frac{t_{1N}^J(1-G_{1N}^J)\lambda_{1N}^{J\prime}}{1+t_{1N}^J} \\
\vdots & & \vdots & & \vdots \\
\alpha_1^J \frac{t_{11}^1(1-G_{11}^1)\lambda_{11}^{1\prime}}{1+t_{11}^1} & \dots & \alpha_1^J \frac{t_{11}^J(1-G_{11}^J)\lambda_{11}^{J\prime}}{1+t_{11}^J} & \dots & \alpha_1^J \frac{t_{1N}^J(1-G_{1N}^J)\lambda_{1N}^{J\prime}}{1+t_{1N}^J} \\
\vdots & & \vdots & & \vdots \\
\alpha_N^J \frac{t_{N1}^1(1-G_{N1}^1)\lambda_{N1}^{1\prime}}{1+t_{N1}^1} & \dots & \alpha_N^J \frac{t_{N1}^J(1-G_{N1}^J)\lambda_{N1}^{J\prime}}{1+t_{N1}^J} & \dots & \alpha_N^J \frac{t_{NN}^J(1-G_{NN}^J)\lambda_{NN}^{J\prime}}{1+t_{NN}^J}
\end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
D &= \begin{bmatrix}
\begin{bmatrix} D_{11}^1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & D_{11}^J \end{bmatrix} & \begin{bmatrix} D_{12}^1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & D_{12}^J \end{bmatrix} & \begin{bmatrix} D_{1N}^1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & D_{1N}^J \end{bmatrix} \\
\vdots & \ddots & \vdots \\
\begin{bmatrix} D_{N1}^1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & D_{N1}^J \end{bmatrix} & \dots & \begin{bmatrix} D_{NN}^1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & D_{NN}^J \end{bmatrix}
\end{bmatrix} \\
D_{od}^j &= \left(1-\beta_d^j\right)\left[t_{od}^j\left(1-G_{od}\right)\right]\frac{\lambda_{od}^j}{1+t_{od}^j}X_d^j
\end{aligned}
$$

These matrices can be derived by solving the budget constraint after a regulation is imposed so as to obtain X' as a function of other parameters and variables of the model.

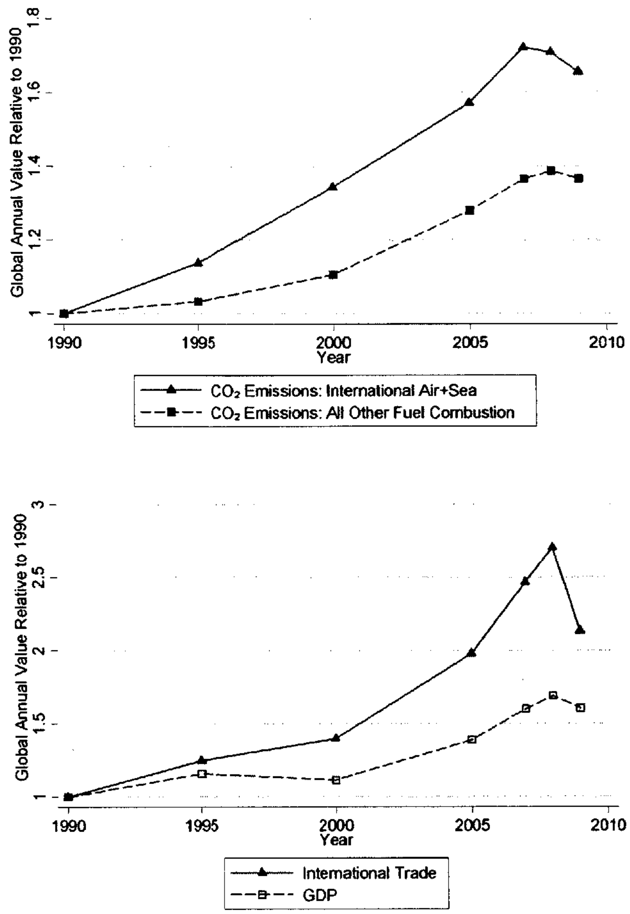


Figure 1. Environmental Costs of Trade Versus Environmental Costs of Other Economic Activity.

Notes: Air and Sea include international goods and passenger transportation. Data from IEA (2011) and World Bank WDI. Appendix Figure 1 shows the second graph with gross output in place of GDP.

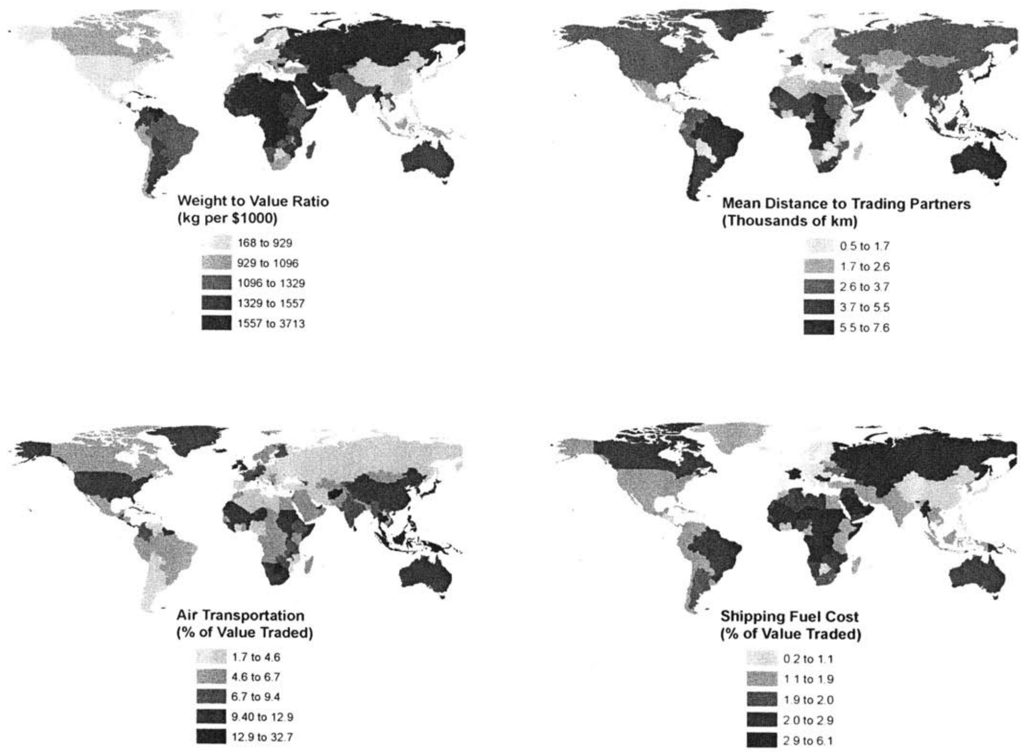


Figure 2: Shipping Fuel Intensity of Traded Goods

Notes: Data aggregate over intranational trade, international imports, and international exports. Mean distance is weighted by kg. Fuel cost is calculated by equation (3c).

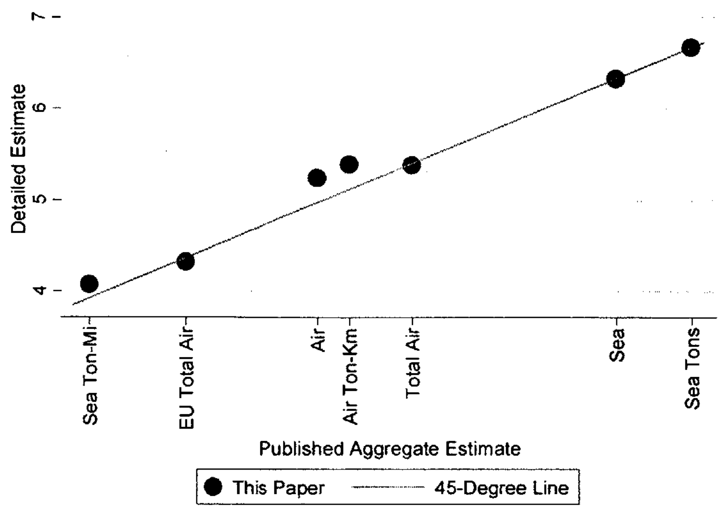


Figure 3: CO2 Emissions: Detailed Estimates from This Paper Versus Published Aggregate Values From International Organizations

Notes: Graph represents international shipping only, unles listed as "total." Data represent CO2 emissions unless unit is otherwise specified. All values in logs. See main text and Appendix B for data sources and details.

Figure 4: Gains from and Environmental Costs of International Trade, by Country.
Notes: Figures plot empirical analogues of equation (10).

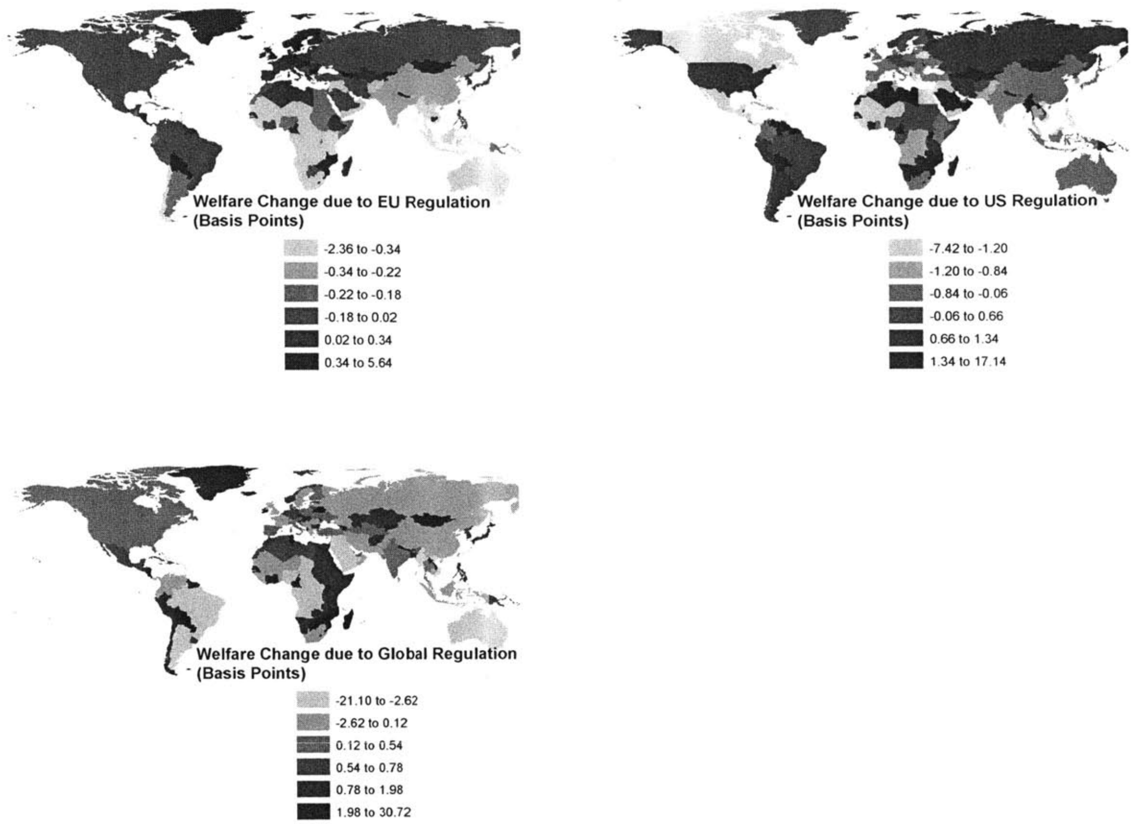


Figure 5: Impact of EU, US, and Global Climate Change Regulations on Social Welfare in Basis Points

Notes: one basis point equals a hundredth of a percentage point. Each regulation imposes a \$19.96/ton carbon tax on intranational and international shipping. Revenue is rebated to the country imposing the tariff (or, for the global tax, to the importer). EU tax applies only to air shipping; US tax applies to air, sea, rail, and road shipping; and global tax applies to air and sea shipping.

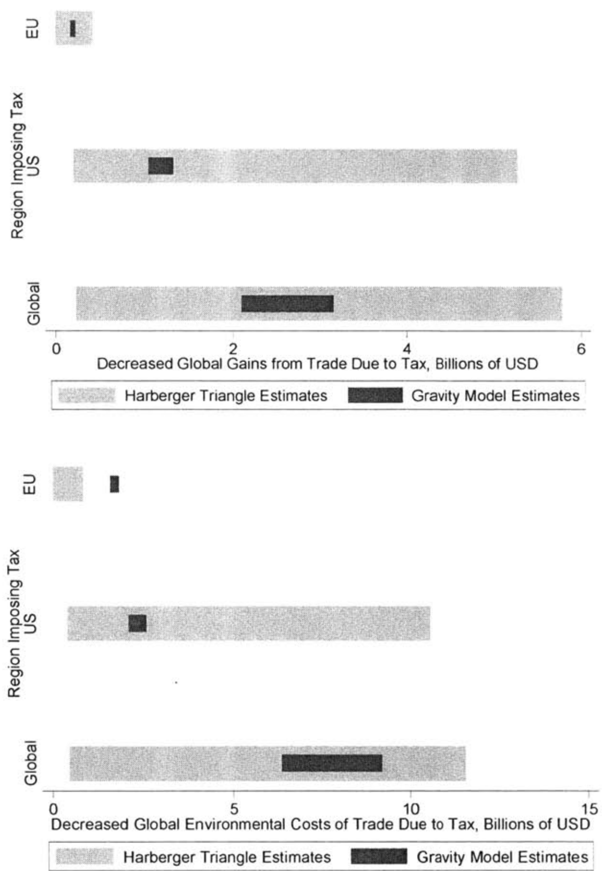


Figure 6. Comparing Estimates from the Gravity Model and from Harberger Triangles.

Notes: The gravity model is described in section 2 of the paper. Harberger triangles describe a partial equilibrium model of the demand for shipping fuels; see paper text for details. For the gravity model, the bar describes the 95% confidence interval, estimated using a bias-corrected bootstrap. For Harberger triangles, the bar represents the range implied by two possible demand elasticities: -0.02 and -0.50. Both models assume perfectly elastic oil supply. The EU, US, and Global carbon taxes of $19.96/tonCO2 are applied to international and intranational shipping; see paper text for details.

Source	International (1)	Domestic (2)	Total (3)
Panel A: CO_2 Emissions by Transport Mode and Type			
Shipping: Air	189	27	216
Shipping: Sea	653	110	763
Shipping: Rail	26	36	62
Shipping: Road	500	1,639	2,139
Shipping: Total	1,368	1,812	3,180
Production: Total	1,192	25,333	26,525
Global Total	2,560	27,145	29,705
Panel B: CO_2 Emissions by Sector			
Non-Manufacturing			
Agriculture, Forestry	160	757	917
Mining	782	601	1,383
Manufacturing			
Food, Beverages, Tobacco	117	618	735
Textiles	35	71	107
Apparel, Leather	29	26	55
Wood	44	115	159
Paper, Printing	68	297	365
Petroleum, Coal, Minerals	434	1,761	2,195
Chemicals, Rubber, Plastics	353	690	1,043
Metals	333	804	1,137
Machinery, Electrical	133	111	244
Transport Equipment	48	70	118
Other	21	49	70
Non-Tradable Goods	0	21,176	21,176
Panel C: CO_2 Emissions by Region			
US	355	6,106	6,462
EU	747	4,346	5,093
Brazil, Russia, India, China	360	8,438	8,798
Sub-Saharan Africa	66	528	595
Other	1,031	7,727	8,757

Table 1: Total Greenhouse Gas Emissions in 2007

Notes: All values represent MtCO2 in the year 2007. Section 3 of the paper describes data sources. International production represents production of internationally traded goods. Household consumption (e.g., passenger transportation) is included in production of non-tradable goods. Panel B combines production and shipping emissions.

Response Variable:	Log Freight Costs (OLS)		Log Trade Costs (FE)		Log Shipping Costs (FS)		Log Import Shares (IV)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
			Panel A: Economy-Wide Estimates					
Overall	-20.95***	(2.61)	-3.71**	(1.82)	0.21***	(0.05)	-7.91**	(4.23)
Overall: Manufacturing	-24.55***	(1.84)	-4.17***	(1.07)	0.22***	(0.04)	-6.68	(4.21)
			Panel B: Sector-Specific Estimates, Non-Manufacturing					
Agriculture, Forestry	-4.72**	(1.86)	-3.48***	(1.06)	0.30***	(0.07)	-2.41	(3.39)
Mining	-5.31***	(1.33)	-2.44**	(0.96)	0.43***	(0.05)	-4.07***	(1.22)
			Panel C: Sector-Specific Estimates, Manufacturing					
Food, Beverages, Tobacco	-15.82***	(1.82)	-5.07***	(1.28)	0.46***	(0.06)	-5.28**	(2.10)
Textiles	-19.78***	(1.40)	-6.47***	(0.93)	0.20***	(0.06)	-16.11***	(4.66)
Apparel, Leather	-18.26***	(1.47)	-3.52***	(1.27)	0.29***	(0.06)	-10.09***	(3.40)
Wood	-12.69***	(1.07)	-2.59***	(0.66)	0.32***	(0.05)	-5.83***	(2.16)
Paper, Printing	-14.44***	(1.38)	-1.79***	(0.61)	0.20***	(0.04)	-5.53*	(2.89)
Petroleum, Coal, Minerals	-12.57***	(1.21)	-2.28***	(0.87)	0.23***	(0.06)	-6.96*	(4.15)
Chemicals, Rubber, Plastics	-16.41***	(1.69)	-3.55***	(1.08)	0.35***	(0.05)	-0.76	(3.02)
Metals	-19.70***	(1.38)	-5.54***	(0.72)	0.20**	(0.08)	-12.99	(8.18)
Machinery, Electrical	-28.46***	(2.09)	-7.95***	(0.95)	0.24***	(0.04)	-10.84***	(2.82)
Transport Equipment	-23.28***	(2.52)	-4.43***	(1.08)	0.23***	(0.07)	-6.8	(3.57)
Other	-16.68***	(1.14)	-4.47***	(0.64)	0.16***	(0.05)	-13.06***	(4.51)
Exporter-by-Year Fixed Effects			x		x		x	
Importer-by-Year Fixed Effects			x		x		x	
Exporter-by-Importer Fixed Effects			x		x		x	

Table 2: Trade Elasticities, Instrumental Variables Estimates

Notes: Each table entry represents a separate regression. An observation represents a good-exporter-importer-time. The data include two importers: the US and Australia. Data ahve two observations per year: one aggregating quarters 2 and 3, and the other aggregating quarters 1 and 4. In IV, shipping costs measured from quarters 2 and 3 of a year are used as an instrument for freight costs measured in quarters 1 and 4 of that year. Standard errors clustered by importer-exporter pair. *** $p<0.01$, ** $p<0.05$, * $p<0.10$. Number of observations in each row is as follows: Overall (9,660); Overall: Manufacturing (9,600); Agriculture, Forestry (6,660); Mining (4,320); Food, Beverages, Tobacco (7,140); Textiles (7,116); Apparel, Leather (7,228); Wood (6,412); Paper, Printing (5,556); Petroleum, Coal, Minerals (5,932); Chemicals, Rubber, Plastics (7,172); Metals (6,408); Machinery, Electrical (7,820); Transport Equipment (5,088); Other (7,252).

	Log Shipping Costs		Log Import Shares		N
Differentiated	0.26***	(0.05)	-5.75**	(2.59)	4,750
Reference Priced	0.38***	(0.04)	-5.81**	(2.33)	4,104
Homogenous	0.36***	(0.07)	-9.18***	(2.77)	3,374

Table 3: Trade Elasticities, by Rauch (1999) Classification
Notes: see notes to Table 2.

	Gains from Trade (1)	Enviromental Costs of Trade: Total (2)	Enviromental Costs of Trade: Transport Share (3)	Welfare (4)	Ratio: (1)/(2) (5)
		Panel A: Global			
World	5292	35.4	0.52	5261	149
	[2725 , 21158]	[28.0 , 65.0]	[0.33 , 0.61]	[2622 , 19079]	[94 , 646]
		Panel B: By Region			
US	561	2.6	0.30	558	218
	[315 , 3935]	[2.0 , 4.7]	[0.15 , 0.47]	[313 , 3933]	[150 , 1655]
EU	2561	19.3	0.70	2545	133
	[1213 , 9630]	[15.2 , 35.3]	[0.51 , 0.76]	[1198 , 9622]	[81 , 540]
Sub-Saharan Africa	68	1.1	0.63	67	61
	[40 , 294]	[0.9 , 2.1]	[0.47 , 0.71]	[39 , 293]	[44 , 283]
		Panel C: By GDP Per Capita			
Richest Third	3979	25.1	0.64	3957	158
	[1977 , 16405]	[19.9 , 46.1]	[0.44 , 0.72]	[1907 , 14784]	[98 , 706]
Middle Third	1062	5.5	0.42	1057	192
	[575 , 4207]	[4.4 , 10.1]	[0.24 , 0.52]	[556 , 3807]	[133 , 778]
Poorest Third	251	4.8	0.36	247	53
	[142 , 1113]	[3.8 , 8.7]	[0.21 , 0.44]	[138 , 1109]	[39 , 193]

Table 4: Annual Effects of International Trade on Social Welfare

Notes: Social welfare measured in billions of US 2007\$. To convert proportional impacts into dollars, column (1) equals (GFT-1)*GDP, column (2) equals (ECT-1)*GDP, and column (4) equals (GFT*ECT-1)*GDP where GFT is gains from trade in percentage terms and ECT is environmental cost of trade in percentage terms. Bracketed numbers represent ninety-five percent confidence intervals, estimated using the bias-corrected bootstrap of Efron (1987) with B=200 draws from the θ^j distributions of Table 2, excluding draws of $\theta^j>0$. Transport share represents the proportion of the change in CO2 emissions due to international trade which comes from transportation. (The remainder comes from production.)

	Intl Trade	CO_2 (Million Tons)		Gains from Trade			Environmental Costs of Trade			Welfare	
		Transport	Production	Real Labor Income	Tax Revenue	Total	Transport	Production	Total	Total	Basis Points
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
					Panel A: Global						
World	-76	-91	-2	-14.2	14.0	-0.2	-1.8	0.0	-1.8	1.6	0.03
	[-90 , -62]	[-95 , -87]	[-3 , 4]	[-14.4 , -14.1]	[13.9 , 14.1]	[-0.2 , -0.2]	[-1.9 , -1.7]	[-0.1 , 0.1]	-2.0 , -1.7]	[1.5 , 1.7]	[0.03 , 0.03]
					Panel B: By Region						
US				-1.9	0.0	-1.9	-0.1	0.0	-0.1	-1.8	-0.12
				[-2.0 , -1.9]	[0.0 , 0.0]	[-2.0 , -1.9]	[-0.1 , -0.1]	[0.0 , 0.0]	-0.2 , -0.1]	[-1.9 , -1.7	[-0.12 , -0.11]
EU	-93	-91	1	-7.7	14.0	6.3	-1.0	0.0	-1.0	7.3	0.41
	[-112 , -74]	[-96 , -87]	[1 , 1]	[-8.0 , -7.2]	[13.9 , 14.1]	[6.0 , 6.8]	[-1.0 , -1.0]	[0.0 , 0.0]	-1.1 , -0.9]	[7.0 , 7.8]	[0.39 , 0.44]
Sub-Saharan Africa				-0.6	0.0	-0.6	-0.1	0.0	-0.1	-0.5	-0.67
				[-0.7 , -0.5]	[0.0 , 0.0]	[-0.7 , -0.5]	[-0.1 , -0.1]	[0.0 , 0.0]	-0.1 , -0.1]	[-0.6 , -0.5	[-0.73 , -0.60]
					Panel C: By Baseline GDP Per Capita						
Richest Third				-10.8	13.7	2.9	-1.3	0.0	-1.3	4.2	0.10
				[-11.2 , -10.4]	[13.6 , 13.8]	[2.6 , 3.3]	[-1.3 , -1.2]	[0.0 , 0.1]	-1.4 , -1.2]	[3.9 , 4.6]	[0.09 , 0.11]
Middle Third				-2.1	0.3	-1.8	-0.3	0.0	-0.3	-1.6	-0.14
				[-2.4 , -2.0]	[0.3 , 0.3]	[-2.1 , -1.7]	[-0.3 , -0.3]	[0.0 , 0.0]	-0.3 , -0.3]	[-1.9 , -1.4	[-0.17 , -0.13]
Poorest Third				-1.3	0.0	-1.3	-0.3	0.0	-0.3	-1.0	-0.32
				[-1.4 , -1.2]	[0.0 , 0.0]	[-1.4 , -1.2]	[-0.3 , -0.2]	[0.0 , 0.0]	-0.3 , -0.2]	[-1.1 , -0.9	[-0.36 , -0.29]

Table 5: EU Carbon Tax on Air Shipping: Effects on Social Welfare

Notes: All columns except (2), (3), and (11) represent values in US 2007$ billions. All estimates summarize the total effect over a decade. The counterfactual policy applies a carbon tax of $19.96 per metric ton of CO2 to all airborne imports, exports, and intranational trade of the 30 countries that are part of the EU Emissions Trading System (ETS). Column (1) shows the change in international trade in US 2007$ billions, excluding purchases from domestic producers and excluding expenditures on the counterfactual carbon tax. Column (2) shows the change in CO2 emissions in millions of tons. The "US" row for columns (1) and (2) aggregates over all US trade. Bracketed numbers represent ninety-five percent confidence intervals, estimated using the bias-corrected bootstrap of Efron (1987) with B=200 draws from the θ^j distributions of Table 2, excluding draws of $\theta^j > 0$.

	Intl Trade	CO_2 (Million Tons)		Gains from Trade			Environmental Costs of Trade			Welfare	
		Transport	Production	Real Labor Income	Tax Revenue	Total	Transport	Production	Total	Total	Basis Points
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
				Panel A: Global							
World	-154	-168	-66	-171.2	170.0	-1.2	-3.3	-1.3	-4.7	3.5	0.06
	[-193 , -123]	[-183 , -157]	[-75 , -52]	[-171.7 , -170.7]	[169.6 , 170.4]	[-1.3 , -1.0]	[-3.6 , -3.1]	[-1.5 , -1.0]	[-4.9 , -4.4]	[3.1 , 4.0]	[0.06 , 0.07]
				Panel B: By Region							
US	-233	-174	-91	-150.3	170.0	19.8	-0.3	-0.1	-0.4	20.1	1.34
	[-273 , -197]	[-194 , -159]	[-116 , -68]	[-155.4 , -145.3]	[169.6 , 170.4]	[14.1 , 25.0]	[-0.3 , -0.2]	[-0.1 , -0.1]	[-0.4 , -0.3]	[14.5 , 25.4]	[0.97 , 1.69]
EU				-4.5	0.0	-4.5	-1.8	-0.7	-2.6	-2.0	-0.11
				[-5.4 , -3.6]	[0.0 , 0.0]	[-5.4 , -3.6]	[-2.0 , -1.7]	[-0.8 , -0.6]	[-2.7 , -2.4]	[-2.9 , -1.0]	[-0.16 , -0.06]
Sub-Saharan Africa				-0.5	0.0	-0.5	-0.1	0.0	-0.1	-0.3	-0.40
				[-1.0 , -0.1]	[0.0 , 0.0]	[-1.0 , -0.1]	[-0.1 , -0.1]	[0.0 , 0.0]	[-0.1 , -0.1]	[-0.9 , 0.0]	[-1.14 , 0.05]
				Panel C: By Baseline GDP Per Capita							
Richest Third				-161.5	170.0	8.5	-2.4	-0.9	-3.3	11.8	0.28
				[-164.3 , -158.6]	[169.6 , 170.4]	[5.4 , 11.8]	[-2.6 , -2.2]	[-1.1 , -0.7]	[-3.5 , -3.1]	[8.6 , 15.0]	[0.21 , 0.36]
Middle Third				-6.5	0.0	-6.5	-0.5	-0.2	-0.7	-5.8	-0.52
				[-9.6 , -3.7]	[0.0 , 0.0]	[-9.6 , -3.7]	[-0.5 , -0.5]	[-0.2 , -0.2]	[-0.7 , -0.7]	[-9.0 , -3.1]	[-0.81 , -0.28]
Poorest Third				-3.2	0.0	-3.2	-0.5	-0.2	-0.7	-2.6	-0.80
				[-4.6 , -2.5]	[0.0 , 0.0]	[-4.6 , -2.5]	[-0.5 , -0.4]	[-0.2 , -0.1]	[-0.7 , -0.6]	[-4.0 , -1.8]	[-1.25 , -0.58]

Table 6: US Carbon Tax on All Modes of Shipping: Effects on Social Welfare

Notes: All currency is in US 2007$ billions and represents the total effect over a decade. The counterfactual policy applies a carbon tax of $19.96 per metric ton of CO2 to all US imports, exports, and intranational trade. Column (1) shows the change in international trade in US 2007$ billions, excluding purchases from domestic producers and excluding expenditures on the counterfactual carbon tax. Column (2) shows the change in CO2 emissions in millions of tons. The "US" row for columns (1) and (2) aggregates over all US trade. Bracketed numbers represent ninety-five percent confidence intervals, estimated using the bias-corrected bootstrap of Efron (1987) with B=200 draws from the θ^j distributions of Table 2, excluding draws of θ^j>0.

	Intl Trade	CO_2 (Million Tons)		Gains from Trade			Environmental Costs of Trade			Welfare	
				Real Labor	Tax			Productio			Basis
		Transport	Production	Income	Revenue	Total	Transport	n	Total	Total	Points
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
					Panel A: Global						
World	-591	-345	-59	-178.2	175.6	-2.6	-6.8	-1.2	-8.0	5.4	0.10
	[-706 , -528]	[-405 , -294]	[-103 , -36]	[-179.3 , -177.4	174.2 , 176.7	[-3.2 , -2.1]	[-8.0 , -5.8]	[-2.0 , -0.7]	[-9.6 , -6.9]	[4.6 , 6.6]	[0.08 , 0.12]
					Panel B: By Region						
US				-30.0	32.2	2.1	-0.5	-0.1	-0.6	2.8	0.18
				[-30.6 , -29.7]	[32.1 , 32.2]	[1.6 , 2.5]	[-0.6 , -0.5]	[-0.2 , -0.1]	[-0.7 , -0.5]	[2.3 , 3.1]	[0.15 , 0.21]
EU				-28.8	33.8	4.9	-3.8	-0.6	-4.4	9.4	0.52
				[-30.5 , -26.9]	[33.5 , 34.1]	[3.3 , 6.7]	[-4.4 , -3.2]	[-1.1 , -0.4]	[-5.3 , -3.8]	[7.5 , 11.6]	[0.42 , 0.64]
Sub-Saharan				-6.9	4.1	-2.7	-0.2	0.0	-0.2	-2.5	-3.02
Africa				[-7.4 , -6.5]	[4.1 , 4.2]	[-3.2 , -2.4]	[-0.2 , -0.2]	[-0.1 , 0.0]	[-0.3 , -0.2]	[-2.9 , -2.2]	[-3.56 , -2.66]
					Panel C: By Baseline GDP Per Capita						
Richest Third				-99.1	107.6	8.4	-4.9	-0.8	-5.7	14.1	0.34
				[-100.7 , -97.6	106.8 , 108.2	[7.2 , 9.6]	[-5.7 , -4.2]	[-1.5 , -0.5]	[-6.9 , -4.9]	[12.2 , 15.8]	[0.29 , 0.38]
Middle Third				-57.6	50.7	-7.0	-1.0	-0.2	-1.2	-5.8	-0.52
				[-58.8 , -56.8]	[50.2 , 51.1]	[-8.3 , -5.9]	[-1.2 , -0.9]	[-0.3 , -0.1]	[-1.4 , -1.0]	[-7.0 , -4.8]	[-0.63 , -0.44]
Poorest Third				-21.4	17.3	-4.1	-1.0	-0.2	-1.2	-3.0	-0.93
				[-22.1 , -20.8]	[17.2 , 17.4]	[-4.7 , -3.6]	[-1.1 , -0.8]	[-0.3 , -0.1]	[-1.3 , -0.9]	[-3.6 , -2.5]	[-1.13 , -0.78]

Table 7: Global Carbon Tax on Air and Sea Shipping: Effects on Social Welfare

Notes: All columns except (2), (3), and (11) represent values in US 2007$ billions. All estimates summarize the total effect over a decade. The counterfactual policy applies a carbon tax of $19.96 per metric ton of CO2 to all airborne and maritime imports, exports, and intranational trade. Column (1) shows the change in international trade in US 2007$ billions, excluding purchases from domestic producers and excluding expenditures on the counterfactual carbon tax. Column (2) shows the change in CO2 emissions in millions of tons. The "US" row for columns (1) and (2) aggregates over all US trade. Bracketed numbers represent ninety-five percent confidence intervals, estimated using the bias-corrected bootstrap of Efron (1987) with B=200 draws from the θ^j distributions of Table 2, excluding draws of $\theta^j>0$.

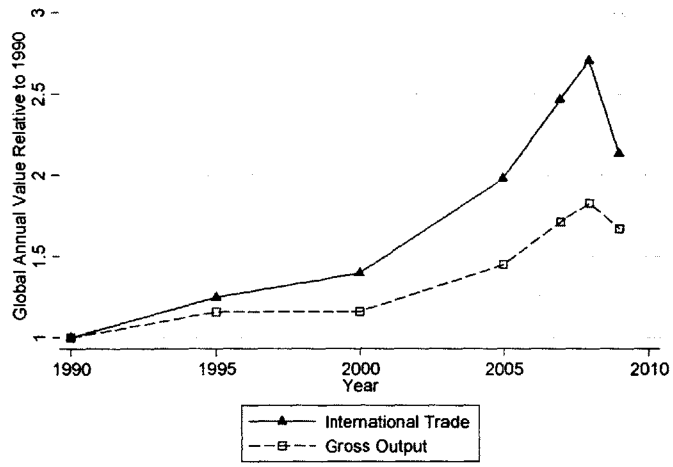
	GFT	ECT		Welfare ($)			Welfare (Percentage Points)		
		Transport	Production	Global	EU	US	Richest	Middle	Poorest
Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1. Detailed Model	5,292	19	17	5,261	2,545	558	10	9	8
2. Social Cost of CO_2=$4.10/ton	5,292	4	4	5,286	2,558	560	10	9	8
3. Social Cost of CO_2=$1170/ton	5,292	1,105	1,005	3,458	1,592	416	7	7	0
4. Homogenous Damage Function	5,292	19	17	5,260	2,552	552	10	9	8
5. Trade Elasticity = -4.12	4,358	18	14	4,330	1,560	607	7	8	12
6. Trade Elasticity = -8.28	2,238	16	8	2,216	799	305	4	4	6

Table 8a: Sensitivity to Model Assumptions, Full Welfare Effects of International Trade
Notes: GFT=Gains from Trade. ECT=Environmental Costs of Trade. See paper text for additional details on each row.

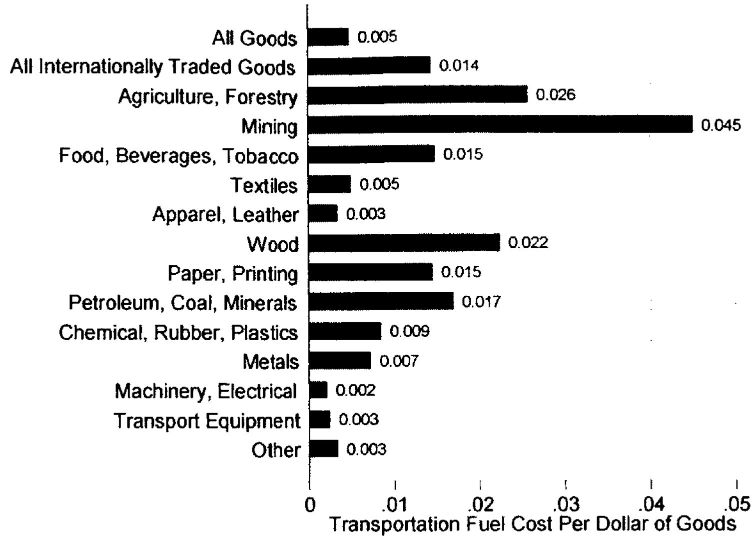
Model	GFT ($) Global (1)	GFT ($) EU (2)	GFT ($) US (3)	ECT ($) Transport (4)	ECT ($) Production (5)	Welfare ($) Global (6)	Welfare ($) EU (7)	Welfare ($) US (8)	Welfare (Basis Points) Richest (9)	Welfare (Basis Points) Middle (10)	Welfare (Basis Points) Poorest (11)
Panel A: EU Carbon Tax for Air Shipping											
1. Main Results	-0.21	6.32	-1.94	-1.74	-0.01	1.54	7.29	-1.80	0.10	-0.14	-0.32
2. Social Cost of CO_2=$4.10/ton	-0.15	1.19	-0.32	-3.64	0.00	3.49	3.14	-0.05	0.03	0.01	0.03
3. Social Cost of CO_2=$1170/ton	-10	184	-58	-393,037	51,821	346,466	331,700	-42,871	708.03	113.14	1,185.35
4. Homogenous Damage Function	-0.21	6.32	-1.94	-1.89	-0.01	1.70	6.93	-1.43	0.10	-0.13	-0.36
5. Trade Elasticity = -4.12	-0.23	6.46	-1.61	-0.16	0.03	-0.10	6.53	-1.61	0.03	-0.07	-0.13
6. Trade Elasticity = -8.28	-0.27	6.03	-1.59	-0.23	0.06	-0.10	6.12	-1.58	0.03	-0.06	-0.13
7. Turn Off General Equilibrium Effects	-0.27	6.25	-1.83	-1.92	-0.04	1.69	7.33	-1.68	0.10	-0.15	-0.19
	-1.33	7.68	-2.69	-3.32	0.03	1.96	9.17	-2.45	0.06	-0.08	-0.10
Panel B: US Carbon Tax for All Shipping											
1. Main Results	-1.18	-1.55	19.76	-3.34	-1.31	3.47	-1.98	20.12	0.28	-0.52	-0.80
2. Social Cost of CO_2=$4.10/ton	-0.05	-0.63	3.11	-0.33	-0.03	0.31	-0.43	3.14	0.02	-0.03	-0.03
3. Social Cost of CO_2=$1170/ton	-10	-13	50	-14,240	-2,858	-21,917	-5,490	-13,979	-462.40	-99.75	-460.86
4. Homogenous Damage Function	-1.18	-4.55	19.76	-3.63	-1.43	3.87	-2.93	21.11	0.29	-0.49	-0.91
5. Trade Elasticity = -4.12	-0.64	-1.14	21.43	-3.69	-0.57	3.62	1.15	21.76	0.11	-0.22	-0.13
6. Trade Elasticity = -8.28	-1.04	-1.22	18.67	-5.57	-0.60	5.13	2.08	19.14	0.11	-0.18	-0.06
7. Turn Off General Equilibrium Effects	-1.87	-7.33	36.84	-4.63	-1.92	4.68	-3.74	37.32	0.51	-1.22	-0.87
	-8.14	6.46	-40.99	-3.86	1.96	-6.24	7.49	-40.86	-0.39	0.97	0.01
Panel C: Global Carbon Tax for Air & Sea Shipping											
1. Main Results	-2.59	4.94	2.14	-6.84	-1.17	5.41	9.35	2.76	0.34	-0.52	-0.93
2. Social Cost of CO_2=$4.10/ton	-1.71	4.81	2.84	-6.04	-0.06	4.40	8.09	3.31	0.11	-0.18	-0.17
3. Social Cost of CO_2=$1170/ton	-5	6	3	-15,306	1,635	13,814	6,771	703	228.21	181.98	706.40
4. Homogenous Damage Function	-2.59	4.94	2.14	-7.43	-1.27	6.10	7.72	4.46	0.36	-0.47	-1.12
5. Trade Elasticity = -4.12	-2.08	5.72	3.11	-8.51	-0.19	6.62	10.39	3.77	0.15	-0.22	-0.25
6. Trade Elasticity = -8.28	-3.33	4.41	2.23	-12.67	0.23	9.10	11.11	3.18	0.15	-0.17	-0.14
7. Turn Off General Equilibrium Effects	-4.79	-3.33	-0.18	-10.65	-4.25	10.10	4.84	0.93	0.12	0.35	0.54
	-40.84	-25.53	-15.82	-24.02	3.03	-19.85	-14.14	-14.32	-0.29	0.07	0.11

Table 8b: Sensitivity to Model Assumptions, Carbon Taxes

Notes: GFT=Gains from Trade. ECT=Environmental Costs of Trade. See paper text for additional details on each row. Numbe
esent total effects over a decade. All table entries represent billions of 2007 US$.

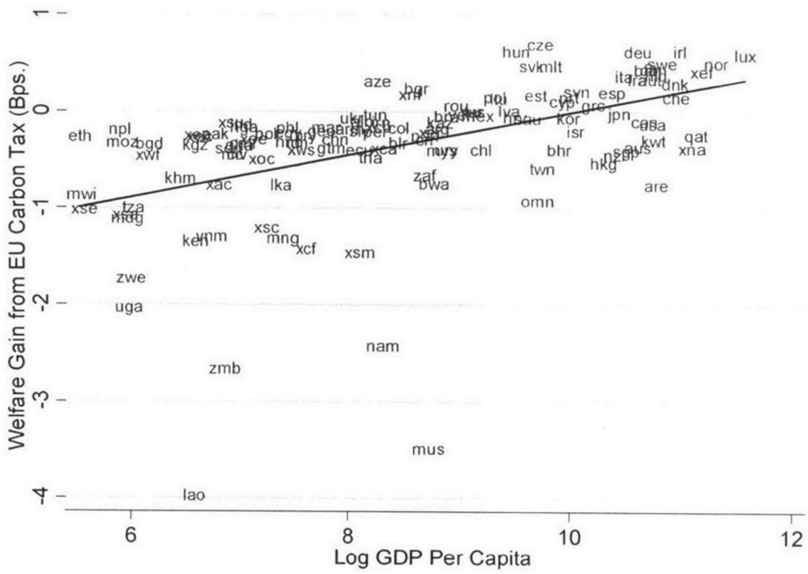


Appendix Figure 1: Gross Output Versus International Trade
Notes: Data from World Bank WDI and OECD STAN. See Appendix A.4 for details of measuring global gross output.



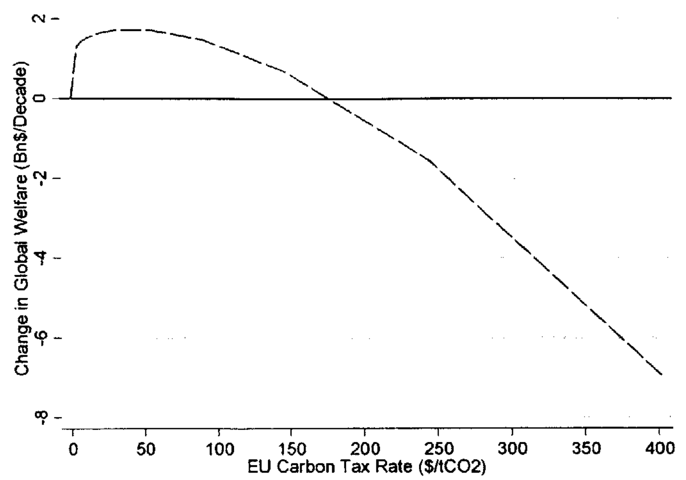
Appendix Figure 2: Transportation Fuel Cost per Dollar of Goods, by Sector

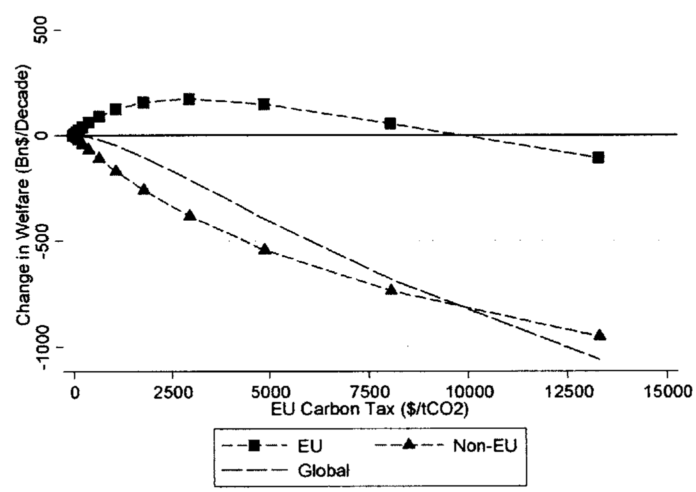
Notes: Graph does not count fuel costs used to transport intermediate goods. At 2007 oil prices, 5936 grams of CO2 are emitted per dollar of transportation fuels. The gCO2 emitted from shipping fuels per dollar of good in each sector are as follows: 30 (all goods); 108 (all internationally traded goods); 155 (agriculture and forestry); 285 (mining); 88 (food, beverages, tobacco); 31 (textiles); 21 (apparel, leather); 133 (wood); 86 (paper, printing); 102 (petroleum, coal, minerals); 50 (chemical, rubber, plastics); 43 (metals); 13 (machinery, electrical); 15 (transport equipment); 21 (other).



Appendix Figure 3: Welfare Effects of EU Carbon Tax, by Country

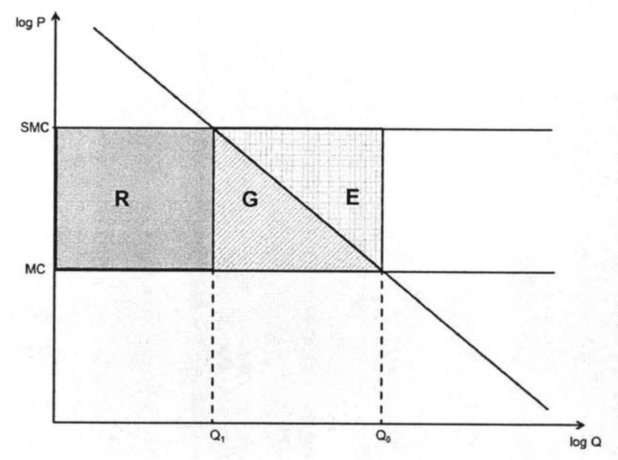
Notes: Graph is scatter plot of 128 countries, each represented by three letters. Each value represents the proportional impact for one country due to the counterfactual EU carbon tax. Black line shows linear trend. Welfare effects measured in basis points (Bps.).





Appendix Figure 4: Optimal EU Tariffs: Regional and Global Welfare as Function of EU Carbon Tax Rate

Notes: Graphs depict the welfare effect of an EU carbon tax at different tax rates, all assuming the marginal social cost of CO2 emissions is \$19.96/ton.



Appendix Figure 5: Graph of Harberger Triangles

Notes: R represents tax revenue, G represents change in gains from shipping, G+E represents change in environmental cost of shipping, SMC represents social marginal cost, MC represents marginal cost, and SMC-MC=t, the tax rate.

Mode	Value (gCO_2/ton-km)	Method	Region	Source
Air	985.97	Fuel consumption divided by ton-km	Global	This paper, IATA (2009)
Air	540	n.a.	Boeing 747	NTM (2012)
Air	912 to 963.45	Calculations from published data	US	Cristea et al. (2011)
Air	595-1916	Engineering estimates	UK	Defra (2009)
Sea	9.53	Global CO_2 emissions from IEA (2011) divided by original ton-km freight estimates	Global	This paper
Sea	4.5 to 16.3	Engineering estimates aggregated over ship fleet registeries	Global	Psaraftis and Kontovas (2009)
Sea	15 to 21	n.a.	n.a.	NTM (2012)
Sea	4 to 20	Engineering estimates	UK	Defra (2009)
Rail	23	Summary of studies listed below	Global	This paper
Rail	23	n.a.	Asia	ADB (2010)
Rail	22.7	n.a.	EU15	Giannouli et al. (2006)
Rail	7.3 to 26.3	Literature review	EU	Cefic (2011)
Rail	27.6	Fuel consumption divided by freight transport	UK	ORR (2009)
Rail	10-119	Literature review	Various	IMO (2009)
Road	119	Summary of studies listed below	Global	This paper
Road	119.7	n.a.	EU	Giannouli et al. (2006)
Road	61	n.a.	Asia	ADB (2010)
Road	118.6	Fuel consumption divided by freight transport	UK	Defra (2009)
Road	80-181	Literature review	Various	IMO (2009)
Other	0	Assumption	Global	This paper

Appendix Table 1: Review of Fuel Economy Estimates, by Transportation Mode
Notes: n.a.=not available.

	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: Economy-Wide Estimates			
Overall	-2.999***	-3.267***	-4.448*	-2.936*	-2.936***	-4.088***
	(1.067)	(0.979)	(2.453)	(1.629)	(0.756)	(0.806)
Overall: Manufacturing	-3.728***	-3.267***	-4.663***	-3.536***	-2.774***	-8.354***
	(0.700)	(0.979)	(1.525)	(0.857)	(0.709)	(0.989)
			Panel B: Sector-Specific Estimates, Non-Manufacturing			
Agriculture, Forestry	-3.356***	-2.692***	-3.394**	-3.066***	-3.447***	-1.721***
	(0.768)	(0.700)	(1.362)	(0.938)	(1.055)	(0.606)
Mining	-2.239***	-1.755***	-2.694**	-2.397***	-2.458**	-1.322***
	(0.569)	(0.566)	(1.301)	(0.917)	(0.958)	(0.438)
			Panel C: Sector-Specific Estimates, Manufacturing			
Food, Beverages, Tobacco	-4.535***	-4.503***	-5.480***	-4.624***	-5.055***	-4.568***
	(0.852)	(0.834)	(1.611)	(1.136)	(1.258)	(1.024)
Textiles	-6.213***	-6.100***	-6.563***	-6.164***	-6.563***	-3.719***
	(0.708)	(0.860)	(1.231)	(0.853)	(0.925)	(0.674)
Apparel, Leather	-3.853***	-4.614***	-4.374**	-3.547***	-3.613***	-2.733***
	(0.734)	(0.694)	(1.694)	(1.117)	(1.250)	(0.829)
Wood	-2.359***	-2.504***	-2.660***	-2.467***	-2.593***	-4.597***
	(0.461)	(0.536)	(0.919)	(0.614)	(0.658)	(0.974)
Paper, Printing	-1.798***	-1.730***	-1.880**	-1.918***	-1.785***	-1.616**
	(0.445)	(0.463)	(0.771)	(0.574)	(0.607)	(0.718)
Petroleum, Coal, Minerals	-2.824***	-3.191***	-2.638**	-2.159**	-2.279***	-1.907***
	(0.639)	(0.661)	(1.233)	(0.854)	(0.868)	(0.541)
Chemicals, Rubber, Plastics	-3.314***	-2.180***	-3.646***	-1.910**	-3.550***	-1.804
	(0.820)	(0.689)	(1.389)	(0.935)	(1.079)	(1.121)
Metals	-5.046***	-5.833***	-6.102***	-5.041***	-5.522***	-4.916***
	(0.529)	(0.643)	(1.026)	(0.605)	(0.719)	(1.172)
Machinery, Electrical	-7.027***	-6.774***	-8.503***	-6.063***	-7.928***	-4.250***
	(0.608)	(0.640)	(1.155)	(0.676)	(0.948)	(0.798)
Transport Equipment	-4.346***	-5.295***	-4.671***	-3.715***	-4.403***	-6.918***
	(0.898)	(1.027)	(1.528)	(1.118)	(1.070)	(1.657)
Other	-4.558***	-4.285***	-4.894***	-4.213***	-4.471***	-2.088**
	(0.474)	(0.611)	(0.888)	(0.596)	(0.640)	(0.823)
Quarters 2,3	x					
Quarters 1,4		x				
GLS weights			x			
log(x+0.00001)				x		
Include tariffs in freight cost					x	
Fuel costs						x

Appendix Table 2a: Trade Elasticities, Fixed Effects Estimates, Sensitivity Analysis

Notes: Columns (1)-(5) include two importers: the US and Australia. Column (5) uses log(1+s+k) as explanatory variable, where k is tariff rate reported in US and Australian data. Column (6) uses fuel cost data estimated from equation (3c) for all countries; see text for details. Each table entry represents a separate regression. An observation represents a good-exporter-importer-year. The data include two importers: the US and Australia. Standard errors clustered by importer-exporter pair. *** p<0.01, ** p<0.05, * p<0.10.

	(1)	(2)	(3)	(4)	(5)
Panel A: Economy-Wide Estimates					
Overall	0.244***	0.224***	0.231***	0.207***	0.460***
	(0.042)	(0.044)	(0.064)	(0.052)	(0.053)
Overall: Manufacturing	0.246***	0.224***	0.232***	0.219***	0.460***
	(0.036)	(0.044)	(0.056)	(0.043)	(0.054)
Panel B: Sector-Specific Estimates, Non-Manufacturing					
Agriculture, Forestry	0.299***	0.312***	0.304***	0.298***	0.225***
	(0.044)	(0.044)	(0.096)	(0.070)	(0.054)
Mining	0.442***	0.421***	0.468***	0.433***	0.401***
	(0.040)	(0.038)	(0.081)	(0.054)	(0.054)
Panel C: Sector-Specific Estimates, Manufacturing					
Food, Beverages, Tobacco	0.465***	0.363***	0.454***	0.460***	0.160***
	(0.052)	(0.046)	(0.073)	(0.061)	(0.032)
Textiles	0.206***	0.222***	0.204**	0.203***	0.154***
	(0.045)	(0.041)	(0.081)	(0.063)	(0.050)
Apparel, Leather	0.308***	0.332***	0.316***	0.288***	0.154***
	(0.042)	(0.037)	(0.082)	(0.061)	(0.045)
Wood	0.332***	0.184***	0.341***	0.322***	0.291***
	(0.035)	(0.048)	(0.068)	(0.047)	(0.043)
Paper, Printing	0.217***	0.122***	0.196***	0.198***	0.202***
	(0.030)	(0.041)	(0.051)	(0.038)	(0.039)
Petroleum, Coal, Minerals	0.229***	0.239***	0.219**	0.227***	0.191***
	(0.044)	(0.032)	(0.106)	(0.061)	(0.053)
Chemicals, Rubber, Plastics	0.350***	0.321***	0.357***	0.351***	0.331***
	(0.042)	(0.034)	(0.071)	(0.055)	(0.054)
Metals	0.185***	0.181***	0.209**	0.203**	0.205***
	(0.050)	(0.062)	(0.096)	(0.083)	(0.078)
Machinery, Electrical	0.269***	0.160***	0.228***	0.237***	0.205***
	(0.035)	(0.032)	(0.055)	(0.043)	(0.039)
Transport Equipment	0.242***	0.150***	0.189**	0.233***	0.172***
	(0.052)	(0.038)	(0.088)	(0.074)	(0.053)
Other	0.160***	0.124***	0.180***	0.155***	0.131**
	(0.039)	(0.028)	(0.069)	(0.054)	(0.051)
Quarters 1,4 instrument quarters 2,3	x				
Quarters 1,2 instrument quarters 3,4		x			
GLS weights			x		
log(x+0.00001)				x	
Include tariffs in freight cost					x

Appendix Table 2b: Trade Elasticities, First-Stage Estimates, Sensitivity Analysis

Notes: The data include two importers: the US and Australia. Column (5) uses log(1+s+k) as explanatory variable, where k is tariff rate reported in US and Australian data. Each table entry represents a separate regression. An observation represents a good-exporter-importer-year. The data include two importers: the US and Australia. Standard errors clustered by importer-exporter pair. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

	(1)	(2)	(3)	(4)	(5)
Panel A: Economy-Wide Estimates					
Overall	-10.099**	-5.376	-8.472*	-7.534**	-1.544
	(4.662)	(4.278)	(4.332)	(3.298)	(1.181)
Overall: Manufacturing	-6.698***	-6.039	-7.433*	-6.187**	-0.92
	(2.265)	(4.185)	(4.212)	(2.755)	(1.086)
Panel B: Sector-Specific Estimates, Non-Manufacturing					
Agriculture, Forestry	-2.598	-3.064*	-2.886	-1.074	-0.549
	(1.920)	(1.859)	(4.335)	(2.334)	(3.278)
Mining	-3.949***	-1.33	-3.837***	-3.682***	-4.381***
	(0.935)	(1.021)	(1.458)	(1.231)	(1.464)
Panel C: Sector-Specific Estimates, Manufacturing					
Food, Beverages, Tobacco	-6.138***	-9.268***	-5.705***	-5.193***	-2.587
	(1.419)	(2.456)	(2.207)	(1.657)	(2.192)
Textiles	-19.326***	-22.376***	-16.874***	-15.731***	-23.596***
	(3.620)	(5.729)	(5.682)	(4.220)	(5.998)
Apparel, Leather	-8.201***	-12.474***	-10.984***	-9.217***	-14.129***
	(1.788)	(2.736)	(3.550)	(2.523)	(4.573)
Wood	-5.891***	-6.828**	-4.438**	-5.769***	-7.679***
	(1.276)	(2.930)	(2.140)	(1.754)	(2.009)
Paper, Printing	-4.929**	-8.711	-5.348	-5.441**	-4.871*
	(2.039)	(10.231)	(3.634)	(2.702)	(2.755)
Petroleum, Coal, Minerals	-7.008***	-6.170***	-10.684	-6.598*	-9.569**
	(2.370)	(2.196)	(6.790)	(3.650)	(4.224)
Chemicals, Rubber, Plastics	-4.233**	-3.121	-1.419	0.385	-0.088
	(1.828)	(1.919)	(3.064)	(2.006)	(2.946)
Metals	-13.544***	-18.666***	-13.386*	-10.018**	-14.723**
	(4.217)	(6.786)	(7.565)	(5.012)	(6.504)
Machinery, Electrical	-10.924***	-19.976**	-11.051***	-7.789***	-13.273***
	(2.046)	(8.352)	(3.783)	(2.273)	(3.346)
Transport Equipment	-6.473**	-18.075*	-5.079	-4.491	-3.016
	(2.731)	(10.675)	(5.685)	(3.001)	(4.080)
Other	-9.937***	-23.401*	-12.185**	-11.342**	-16.749**
	(3.245)	(13.799)	(5.986)	(4.601)	(6.897)
Correlation with Table 2, Column 7	0.93	0.94	0.97	0.98	0.98
Quarters 1,4 instrument quarters 2,.	x				
Quarters 1,2 instrument quarters 3,4		x			
GLS weights			x		
log(x+0.00001)				x	
Include tariffs in freight cost					x

Appendix Table 2c: Trade Elasticities, Instrumental-Variables Estimates, Sensitivity Analysis

Notes: Data include two importers: the US and Australia. Columns (5) uses log(1+s+k) as explanatory variable, where k is tariff rate reported in US and Australian data. Each table entry represents a separate regression. "Correlation with Table 2" reports the correlation coefficient between the 13 sector-specific elasticities reported in a given column of this table and the 13 elasticities reported in Table 2, column 7. An observation represents a good-exporter-importer-year. The data include two importers: the US and Australia. Standard errors clustered by importer-exporter pair. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Country	2007 GDP	Welfare Effects of Trade				EU Air Carbon Tax			Global Air & Sea Carbon Tax		
		GFT	ECT	Welfare	%GDP	GFT	ECT	Welfare	GFT	ECT	Welfare
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
World Total	56,112.6	5,292.3	35.4	5,261.0	9.4%	-0.21	-1.84	1.63	-2.59	-9.93	7.33
Albania	14.1	0.7	0.01	0.7	5.0%	0.00	0.00	0.00	-0.02	0.00	-0.02
Argentina	241.6	19.7	0.26	19.5	8.2%	-0.05	-0.01	-0.04	-1.18	-0.07	-1.11
Armenia	11.1	0.7	0.01	0.7	6.0%	0.00	0.00	0.00	0.00	0.00	0.00
Australia	841.0	45.4	0.00	45.4	5.4%	-0.30	0.00	-0.30	-2.87	0.00	-2.87
Austria	375.4	73.9	0.43	73.5	19.7%	0.11	-0.02	0.13	0.01	-0.12	0.13
Azerbaijan	18.9	2.1	0.04	2.0	10.9%	0.00	0.00	0.01	-0.04	-0.01	-0.04
Bahrain	17.5	2.2	0.01	2.2	12.8%	-0.01	0.00	-0.01	0.01	0.00	0.02
Bangladesh	73.6	5.2	0.07	5.1	7.0%	-0.03	0.00	-0.02	0.02	-0.02	0.04
Belarus	55.3	10.3	0.01	10.3	18.7%	-0.02	0.00	-0.02	0.51	0.00	0.52
Belgium	515.0	287.1	0.54	286.9	55.7%	0.20	-0.03	0.23	0.52	-0.17	0.68
Bolivia	13.2	0.9	0.01	0.9	7.1%	0.00	0.00	0.00	0.02	0.00	0.02
Botswana	11.0	1.3	0.02	1.3	11.8%	-0.01	0.00	-0.01	0.01	0.00	0.01
Brazil	1,300.1	58.0	1.33	56.7	4.5%	-0.12	-0.07	-0.05	-3.55	-0.36	-3.19
Bulgaria	53.8	8.3	0.01	8.3	15.4%	0.01	0.00	0.01	0.04	0.00	0.04
Cambodia	9.2	1.2	0.01	1.2	13.3%	-0.01	0.00	-0.01	-0.05	0.00	-0.05
Cameroon	19.6	1.1	0.03	1.1	5.8%	-0.01	0.00	-0.01	0.00	-0.01	0.01
Canada	1,395.1	148.0	0.00	148.0	10.6%	-0.13	0.00	-0.13	0.32	0.00	0.32
Caribbean	291.9	19.0	0.26	18.7	6.5%	-0.06	-0.01	-0.05	0.18	-0.08	0.26
Chile	129.1	15.3	0.12	15.2	11.9%	-0.06	0.00	-0.05	0.10	-0.03	0.13
China	3,050.5	312.7	0.30	312.4	10.3%	-0.88	-0.01	-0.86	-0.19	-0.08	-0.11
Colombia	206.2	14.0	0.15	13.8	6.8%	-0.04	-0.01	-0.03	-0.31	-0.04	-0.27
Costa Rica	25.1	4.8	0.02	4.8	19.3%	-0.01	0.00	-0.01	0.00	-0.01	0.01
Cote d'Ivoire	17.8	2.3	0.03	2.2	12.8%	-0.01	0.00	-0.01	0.05	-0.01	0.06
Croatia	71.6	8.3	0.02	8.3	11.6%	0.00	0.00	0.00	-0.01	-0.01	-0.01
Cyprus	30.8	1.9	0.02	1.9	6.3%	0.00	0.00	0.00	-0.05	-0.01	-0.05
Czech Republic	173.3	41.7	0.05	41.7	24.1%	0.12	0.00	0.12	0.13	-0.01	0.14
Denmark	320.3	45.9	0.36	45.6	14.3%	0.08	-0.02	0.10	0.06	-0.10	0.16
Ecuador	44.8	4.3	0.03	4.3	9.7%	-0.02	0.00	-0.02	-0.06	-0.01	-0.05
Egypt	149.4	14.3	0.14	14.2	9.6%	-0.04	-0.01	-0.03	0.08	-0.04	0.13
El Salvador	23.8	2.4	0.01	2.4	10.2%	-0.01	0.00	0.00	0.01	0.00	0.01
Estonia	27.5	5.1	0.01	5.1	18.5%	0.00	0.00	0.00	-0.03	0.00	-0.03
Ethiopia	22.4	0.9	0.03	0.9	4.2%	-0.01	0.00	-0.01	-0.01	-0.01	0.00
Finland	246.3	29.2	0.29	28.9	11.9%	0.10	-0.01	0.11	0.07	-0.08	0.15
France	2,757.7	270.4	3.07	267.7	9.8%	0.79	-0.16	0.95	-0.04	-0.89	0.85
Georgia	13.9	1.2	0.01	1.2	8.8%	0.00	0.00	0.00	0.00	0.00	0.00
Germany	3,119.5	533.3	3.87	530.1	17.1%	1.79	-0.19	1.97	1.83	-1.00	2.83
Ghana	28.9	2.1	0.04	2.1	7.4%	-0.01	0.00	-0.01	0.01	-0.01	0.02
Greece	396.4	25.2	0.37	24.9	6.4%	0.01	-0.02	0.03	-0.32	-0.13	-0.19
Guatemala	38.0	4.6	0.04	4.6	12.1%	-0.02	0.00	-0.01	-0.02	-0.01	-0.01
Honduras	16.0	2.2	0.01	2.2	13.8%	-0.01	0.00	-0.01	-0.03	0.00	-0.03
Hong Kong	269.1	26.9	0.00	26.9	10.0%	-0.14	0.00	-0.14	-0.31	0.00	-0.31
Hungary	141.1	34.5	0.04	34.4	24.4%	0.09	0.00	0.09	0.10	-0.01	0.11
India	1,294.8	67.7	2.43	65.5	5.2%	-0.31	-0.13	-0.18	0.27	-0.73	0.99
Indonesia	386.8	37.3	0.46	36.9	9.7%	-0.15	-0.02	-0.12	-0.68	-0.12	-0.56
Iran	250.1	17.2	0.21	17.0	6.9%	-0.04	-0.01	-0.03	-0.14	-0.05	-0.09
Ireland	208.4	123.4	0.30	123.2	59.2%	0.12	-0.01	0.13	0.39	-0.07	0.46
Israel	166.2	29.2	0.00	29.2	17.6%	-0.03	0.00	-0.03	0.09	0.00	0.09
Italy	2,179.2	189.7	2.49	187.4	8.7%	0.69	-0.13	0.82	1.01	-0.70	1.71
Japan	4,303.5	167.5	0.89	166.6	3.9%	-0.08	-0.05	-0.04	3.74	-0.24	3.98
Kazakhstan	92.5	8.4	0.08	8.3	9.1%	-0.01	0.00	-0.01	0.04	-0.02	0.06
Kenya	31.5	2.7	0.04	2.7	8.6%	-0.04	0.00	-0.04	-0.01	-0.01	0.01
Korea	1,015.8	109.8	1.04	108.9	10.8%	-0.12	-0.05	-0.06	1.38	-0.28	1.67
Kuwait	69.0	5.6	0.09	5.5	8.1%	-0.02	0.00	-0.02	-0.88	-0.02	-0.86
Kyrgyzstan	5.2	0.4	0.00	0.4	8.0%	0.00	0.00	0.00	-0.01	0.00	0.00
Laos	4.5	0.4	0.00	0.4	8.2%	-0.02	0.00	-0.02	-0.02	0.00	-0.02
Latvia	37.7	4.9	0.01	4.9	12.9%	0.00	0.00	0.00	-0.06	0.00	-0.06
Lithuania	49.3	8.6	0.01	8.6	17.4%	0.01	0.00	0.01	-0.02	0.00	-0.02
Luxembourg	53.8	11.9	0.06	11.9	22.2%	0.03	0.00	0.03	-0.01	-0.02	0.00
Madagascar	7.5	0.5	0.01	0.5	6.9%	-0.01	0.00	-0.01	-0.01	0.00	-0.01
Malawi	3.5	0.5	0.01	0.5	15.2%	0.00	0.00	0.00	-0.02	0.00	-0.02
Malaysia	133.1	51.1	0.18	50.9	38.4%	-0.06	-0.01	-0.05	-0.03	-0.04	0.00
Malta	10.1	2.0	0.01	2.0	19.5%	0.00	0.00	0.00	-0.02	0.00	-0.02
Mauritius	7.9	1.0	0.01	1.0	12.4%	-0.03	0.00	-0.03	-0.04	0.00	-0.04
Mexico	981.0	92.0	0.75	91.4	9.4%	-0.07	-0.04	-0.03	0.47	-0.20	0.67
Mongolia	3.2	0.2	0.00	0.2	7.4%	0.00	0.00	0.00	0.04	0.00	0.04
Morocco	89.7	9.6	0.06	9.5	10.7%	-0.02	0.00	-0.01	0.02	-0.02	0.04

Appendix Table 3: Country-by-Country List, Counterfactual Total Impacts

Country	GDP	Welfare Effects of Trade				EU Air Carbon Tax			Global Air & Sea Carbon Tax		
		GFT	ECT	Welfare		GFT	ECT	Welfare	GFT	ECT	Welfare
	(1)	(2)	(3)	(4)		(5)	(6)	(7)	(8)	(9)	(10)
Mozambique	8.3	1.0	0.01	0.9	11.5%	0.00	0.00	0.00	-0.01	0.00	-0.01
Namibia	9.3	1.4	0.01	1.4	15.2%	-0.02	0.00	-0.02	-0.04	0.00	-0.03
Nepal	11.5	0.7	0.01	0.7	6.0%	0.00	0.00	0.00	-0.01	0.00	-0.01
Netherlands	757.3	216.2	0.92	215.6	28.6%	0.25	-0.05	0.30	0.63	-0.24	0.88
New Zealand	136.4	9.2	0.00	9.2	6.8%	-0.06	0.00	-0.06	-0.09	0.00	-0.09
Nicaragua	7.9	0.8	0.01	0.8	10.1%	0.00	0.00	0.00	-0.01	0.00	-0.01
Nigeria	131.9	7.1	0.26	6.9	5.4%	-0.03	-0.01	-0.02	-1.02	-0.06	-0.96
Norway	343.4	27.9	0.45	27.5	8.1%	0.16	-0.02	0.18	0.17	-0.11	0.28
Oman	32.7	3.4	0.03	3.4	10.4%	-0.03	0.00	-0.03	-0.06	-0.01	-0.05
Pakistan	164.6	9.9	0.16	9.8	6.0%	-0.05	-0.01	-0.04	-0.06	-0.05	-0.01
Panama	26.5	1.2	0.01	1.1	4.4%	-0.01	0.00	-0.01	-0.09	-0.01	-0.08
Paraguay	14.6	1.7	0.01	1.7	11.6%	0.00	0.00	0.00	0.00	0.00	0.00
Peru	93.4	8.0	0.08	7.9	8.6%	-0.02	0.00	-0.02	0.05	-0.02	0.06
Philippines	139.9	19.0	0.15	18.8	13.5%	-0.03	-0.01	-0.02	0.10	-0.04	0.14
Poland	468.0	61.1	0.13	60.9	13.0%	0.07	-0.01	0.07	0.03	-0.04	0.07
Portugal	262.3	26.5	0.27	26.3	10.1%	0.03	-0.02	0.04	0.00	-0.08	0.08
Qatar	68.9	5.8	0.06	5.8	8.5%	-0.02	0.00	-0.02	-0.02	-0.02	0.00
Other Country- Regions											
Central Africa	29.0	2.1	0.06	2.1	7.4%	-0.04	0.00	-0.04	-0.36	-0.01	-0.35
Central America	1.5	0.2	0.00	0.2	15.2%	0.00	0.00	0.00	0.00	0.00	0.00
East Asia	35.9	1.9	0.04	1.9	5.4%	-0.01	0.00	-0.01	-0.03	-0.01	-0.02
Eastern Africa	58.8	3.7	0.09	3.7	6.4%	-0.02	0.00	-0.02	0.01	-0.03	0.03
Eastern Europe	7.2	0.8	0.00	0.8	11.2%	0.00	0.00	0.00	-0.01	0.00	-0.01
EFTA	27.0	3.0	0.03	2.9	11.0%	0.01	0.00	0.01	-0.01	-0.01	0.00
Europe	114.6	11.9	0.00	11.9	10.3%	-0.02	0.00	-0.02	-0.08	0.00	-0.08
Former Soviet Union	28.3	4.2	0.04	4.2	14.8%	0.00	0.00	0.00	-0.04	-0.01	-0.03
North Africa	149.8	11.3	0.15	11.2	7.5%	0.02	-0.01	0.03	0.05	-0.03	0.08
North America	11.3	0.6	0.01	0.6	5.1%	0.00	0.00	0.00	-0.02	0.00	-0.02
Oceania	30.4	2.5	0.02	2.4	8.1%	-0.02	0.00	-0.01	-0.03	-0.01	-0.03
SACU	3.0	0.4	0.01	0.4	12.4%	0.00	0.00	0.00	-0.01	0.00	-0.01
South America	5.3	0.9	0.01	0.9	16.7%	-0.01	0.00	-0.01	-0.02	0.00	-0.02
South Asia	15.5	1.0	0.01	1.0	6.3%	-0.02	0.00	-0.02	-0.02	0.00	-0.02
South Central Africa	38.6	3.2	0.11	3.1	8.2%	-0.03	0.00	-0.03	-0.87	-0.02	-0.85
Southeast Asia	22.9	1.7	0.03	1.7	7.6%	-0.02	0.00	-0.02	-0.14	-0.01	-0.13
West Asia	129.8	17.5	0.14	17.3	13.5%	-0.06	-0.01	-0.05	-0.43	-0.04	-0.39
Western Africa	48.0	4.9	0.06	4.9	10.3%	-0.03	0.00	-0.02	-0.18	-0.02	-0.15
Romania	197.3	20.3	0.05	20.2	10.3%	0.01	0.00	0.01	-0.03	-0.02	-0.02
Russian Federation	1,128.7	79.8	0.00	79.8	7.1%	0.01	0.00	0.01	-2.07	0.00	-2.07
Saudi Arabia	220.4	37.9	0.30	37.6	17.2%	-0.02	-0.01	-0.01	-1.81	-0.05	-1.76
Senegal	14.2	1.6	0.02	1.6	11.2%	-0.01	0.00	-0.01	0.01	-0.01	0.01
Singapore	157.0	77.3	0.00	77.3	49.3%	-0.06	0.00	-0.06	1.02	0.00	1.02
Slovakia	87.3	19.4	0.02	19.3	22.2%	0.04	0.00	0.04	0.09	-0.01	0.10
Slovenia	51.3	15.8	0.01	15.8	30.8%	0.01	0.00	0.01	0.00	0.00	0.00
South Africa	284.7	26.3	0.21	26.1	9.2%	-0.20	-0.01	-0.19	-0.14	-0.06	-0.08
Spain	1,593.3	141.6	1.67	140.0	8.9%	0.23	-0.09	0.33	0.40	-0.51	0.92
Sri Lanka	36.4	3.6	0.03	3.6	10.0%	-0.03	0.00	-0.03	-0.06	-0.01	-0.05
Sweden	448.3	71.0	0.54	70.5	15.8%	0.20	-0.03	0.23	-0.06	-0.14	0.09
Switzerland	429.0	170.2	0.50	169.9	39.7%	0.04	-0.03	0.07	0.00	-0.14	0.14
Taiwan	335.0	92.5	0.39	92.2	27.6%	-0.21	-0.02	-0.19	0.79	-0.09	0.88
Tanzania	18.9	1.5	0.03	1.4	7.8%	-0.02	0.00	-0.02	-0.02	-0.01	-0.01
Thailand	213.4	58.9	0.18	58.8	27.6%	-0.11	-0.01	-0.10	0.01	-0.04	0.05
Tunisia	38.7	6.6	0.03	6.6	17.0%	0.00	0.00	0.00	0.03	-0.01	0.04
Turkey	722.0	52.9	0.49	52.5	7.3%	-0.02	-0.03	0.00	0.25	-0.15	0.40
Uganda	11.8	0.8	0.02	0.8	6.5%	-0.02	0.00	-0.02	-0.02	-0.01	-0.02
Ukraine	173.6	25.5	0.04	25.4	14.7%	-0.01	0.00	-0.01	-0.01	-0.01	0.01
United Arab Emirates	238.6	27.8	0.16	27.7	11.7%	-0.19	-0.01	-0.18	-0.01	-0.05	0.04
United Kingdom	3,017.7	261.8	3.26	258.8	8.7%	1.18	-0.18	1.36	0.14	-0.97	1.11
United States of Americ	14,982.4	560.9	2.57	558.4	3.7%	-1.94	-0.14	-1.79	2.14	-0.77	2.91
Uruguay	24.2	3.3	0.02	3.2	13.4%	-0.01	0.00	-0.01	-0.04	0.00	-0.03
Venezuela	204.9	12.5	0.16	12.4	6.1%	-0.01	-0.01	0.00	-0.55	-0.04	-0.51
Viet nam	73.5	17.8	0.07	17.8	24.3%	-0.10	0.00	-0.09	-0.45	-0.02	-0.43
Zambia	10.5	0.9	0.02	0.8	8.3%	-0.03	0.00	-0.03	-0.04	0.00	-0.03
Zimbabwe	4.3	0.7	0.01	0.7	16.4%	-0.01	0.00	-0.01	-0.01	0.00	-0.01

Appendix Table 3–Country-by-Country Results List (Continued)

Notes: GFT=gains from trade, ECT=environmental cost of trade. All values in US 2007$ billions. GDP is year 2007 annual value. Column (1) represents annual values. Columns (2)-(10) represent total values over a deacde. See GTAP documentation for countries grouped in regions like "Rest of SACU" and "Rest of West Asia."

	Gains from Trade			Environmental Cost of Trade			Welfare
	Real Labor Income (1)	Tax Revenue (2)	Total (3)	Transport (4)	Production (5)	Total (6)	(7)
	Demand Elasticity for Shipping Fuels: -0.02						
EU tax: air	-14.93	14.94	-0.02	-0.04	n.a.	-0.04	0.02
US tax: air, sea, rail, road	-178.89	179.10	-0.21	-0.42	n.a.	-0.42	0.21
Global tax: air, sea	-195.77	196.00	-0.23	-0.46	n.a.	-0.46	0.23
	Demand Elasticity for Shipping Fuels: -0.50						
EU tax: air	-15.35	15.79	-0.44	-0.88	n.a.	-0.88	0.44
US tax: air, sea, rail, road	-183.97	189.26	-5.29	-10.58	n.a.	-10.58	5.29
Global tax: air, sea	-201.33	207.12	-5.79	-11.58	n.a.	-11.58	5.79
	Reprinted from Tables 5-7: General Equilibrium Model Estimates						
EU tax: air	-14.20	13.99	-0.21	-1.80	-0.03	-1.84	1.63
US tax: air, sea, rail, road	-171.22	170.03	-1.18	-3.34	-1.31	-4.65	3.47
Global tax: air, sea	-178.17	175.57	-2.59	-6.84	-1.17	-8.00	5.41

Appendix Table 4: Comparing General Equilibrium Estimates to Harberger Triangles

Notes: All currency in US 2007$ billions and represent the total effect over a decade. EU includes all 30 ETS participating countries. The counterfactual policy applies a carbon tax of $19.96 per metric ton of CO2 to indicated intranational and international trade flows.

Sector	1990 Value (1)	2005 Value (2)	2008 Value (3)	Ratio: 2008/1990 (4)	Ratio: 2005/1990 (5)
Panel A: All Greenhouse Gas Emissions: By Sector					
International transportation	614	964	1,048	1.71	1.57
Electricity and heat	8,952	12,791	14,116	1.58	1.43
Manufacturing, construction, and mining	5,918	7,444	8,504	1.44	1.26
Domestic transportation	3,978	5,391	5,611	1.41	1.36
All	20,966	27,188	29,454	1.40	1.30
Agriculture and forestry	5,384	5,533	5,903	1.10	1.03
Trade and services	768	796	835	1.09	1.04
Residential	1,821	1,874	1,914	1.05	1.03
Other	11,980	12,948	11,188	0.93	1.08
Panel B: All CO_2 Emissions: By Sector					
International transportation	614	964	1,048	1.71	1.57
Electricity and heat	6,621	9,944	10,983	1.66	1.50
Domestic transportation	3,978	5,391	5,611	1.41	1.36
All	20,966	27,188	29,454	1.40	1.30
Manufacturing, construction, and mining	5,372	6,521	7,461	1.39	1.21
Trade and services	768	796	835	1.09	1.04
Residential	1,821	1,874	1,914	1.05	1.03
Agriculture and forestry	392	409	385	0.98	1.04
Other	10,114	10,871	9,283	0.92	1.07
Panel C: CO_2 Emissions From Fuel Combustion: By Industry					
Transport equipment (manufacturing)	17	46	53	3.10	2.69
International air (transportation)	256	408	440	1.72	1.59
Non-metallic minerals (manufacturing)	501	786	948	1.89	1.57
International sea (transportation)	358	556	608	1.70	1.55
Mining and quarrying	64	96	112	1.75	1.49
Wood and wood products (manufacturing)	17	25	31	1.83	1.49
Road (transportation)	3,288	4,673	4,900	1.49	1.42
Iron and steel (manufacturing)	932	1,237	1,472	1.58	1.33
Paper, pulp and printing (manufacturing)	140	184	182	1.30	1.31
Domestic navigation (transportation)	98	121	126	1.28	1.23
Food and tobacco (manufacturing)	189	229	252	1.33	1.21
Non-ferrous metals (manufacturing)	92	109	121	1.32	1.18
Pipeline (transportation)	127	145	150	1.18	1.14
Domestic aviation (transportation)	282	318	303	1.08	1.13
Construction	106	116	120	1.14	1.10
Chemical and petrochemical (manufacturing)	551	598	689	1.25	1.09
Other industry	1,663	1,573	1,722	1.04	0.95
Textile and leather (manufacturing)	99	87	90	0.91	0.88
Rail (transportation)	146	115	109	0.75	0.79
Machinery (manufacturing)	162	124	141	0.87	0.77
Other transport (transportation)	37	19	22	0.60	0.52

Appendix Table 5: Greenhouse Gas Emissions 1990-2008, by Sector and Industry

Notes: Values in MtCO2e. Data from IEA (2011) "Detailed CO2 Estimates" and "Emissions of CO2, CH4, N2O, HFC, PFC", and "SF6". International Transportation combines air & sea bunkers. Industry breakdown only available for manufacturing, construction, and mining, and for transportation. Within each panel, rows are ordered by 2005/1990 ratio.

Chapter 2

Defensive Investments and the Demand for Air Quality: Evidence from the NOx Budget Program and Ozone Reductions[1]

2.1 Introduction

Willingness to pay for wellbeing in many contexts depends on factors that enter the utility function directly and on costly investments that influence these factors. A prominent example is the canonical model of health production, wherein a person trades off the disutility of illness against the cost of actions to prevent illness (Grossman 1972). Anecdotal real-world examples abound: homeowners install burglar alarms, companies hire security guards, infants are vaccinated, builders install thick windows in noisy areas, and people take medications to protect against respiratory problems. Defensive investments have economic costs because they displace consumption of utility-generating goods. If consumers respond to externalities optimally, a standard neoclassical argument indicates that marginal cost of defensive expenditures would be equalized with the marginal disutility of the harm itself. Although it is unclear that consumers are fully informed about pollution (Currie, Davis, Greenstone, and Walker 2012), due in part to this argument it is widely believed that defensive investments constitute a significant portion of the welfare costs of negative externalities.

[1]This chapter is coauthored with Olivier Deschênes and Michael Greenstone.

However, empirical research on negative externalities largely analyzes how they directly affect utility, while leaving unanswered the empirical importance of compensatory behaviors (e.g., Chay and Greenstone (2003b), Chay and Greenstone (2003a), Currie and Neidell (2005)). Indeed, exclusively focusing on how a negative externality directly affects utility could substantially understate the externality's welfare costs.

This paper measures willingness to pay for clean air while accounting for both defensive expenditures and direct health outcomes. To measure defensive behavior, we investigate whether air quality affects medication purchases. This may be an especially important measure of defensive expenditures because the annual cost of asthma medications is reported to exceed the monetized value of any other component of asthma's social cost, including mortality, emergency department admissions, or lost productivity (Weiss and Sullivan 2001). The analysis also provides new evidence on how air quality affects mortality and hospital admissions, which lets us measure the share of air pollution's health costs caused by defenses.

The analysis exploits variation in space and time of an emissions market for nitrogen oxides (NOx) to construct a quasi-experiment. The NOx Budget Trading Program (NBP) operated a cap-and-trade system for over 2,500 electricity generating units and industrial boilers in the Eastern and Midwestern U.S. between 2003 and 2008. Because this market aimed to decrease ozone pollution, which reaches high levels in summer, the market operated only between May 1 and September 30. NOx is a primary ingredient in the complex function that produces ozone air pollution. The NBP provides seasonal variation in air pollution, which is a much longer time span than daily and monthly shocks analyzed in prior research, and which makes this study more directly relevant to future ozone regulation that is likely to vary by season but not by day.

Figure 1 shows the dramatic effect of the NBP on NOx pollution emissions in participating states. In 2002, daily NOx emissions were fairly flat throughout the calendar year, with a rise when electricity demand peaks in July. In 2005, emissions were also flat between January and April. But in May 2005, when the market's cap began to apply, NOx emissions dropped by 35 percent, practically overnight. Emissions remained lower throughout the summer of 2005 and then returned to their original level in October, when the cap stopped applying. Emissions dropped in May 2005 because many power plants began operating abatement technologies that substantially decreased their NOx emissions. This market lets us isolate the causal effects of air quality on health because it allows a simple research design. We use a triple-difference estimator that compares pollution and health outcomes in summer versus winter, before versus after 2003, and in the NBP participating and non-participating states.

The empirical analysis produces several key results. First, reductions in NOx emissions decreased mean ozone concentrations by 6% and reduced the number of summer days with high ozone levels (i.e., with ozone above 65 ppm) by 23%, or a third of a standard deviation. Second, these improvements in air quality produced substantial medium-run benefits. Drug expenditures decreased by about 1.9%, or roughly $900 million annually. These savings exceed an upper bound estimate of the market's abatement costs. Third, the summertime mortality rate declined by 0.5%, corresponding to 2,200 fewer premature deaths per summer, mainly among individuals 75 and older. The application of age-adjusted estimates of the value of a statistical life monetizes this reduced mortality at $900 million annually. The mortality estimates are less precise than the medication estimates, and should be interpreted accordingly. Fourth, there is little systematic evidence of an effect of the NBP on hospital admissions or charges.

Finally, it may be appropriate to conclude the reductions in ozone concentrations are the primary channel for these improvements in health. We find no association between the NBP and health conditions that are plausibly unrelated to air quality. Also, we find that the NBP did not affect ambient concentrations of carbon monoxide and sulfur dioxide, though there is mixed evidence about whether it led to reductions in particulate matter. Consequently, we cautiously utilize the NBP as an instrumental variable for ozone concentrations and find that the elasticity of medication purchases with respect to mean summer ozone is 0.28. The elasticity of mortality with respect to mean summer ozone is 0.22. When ozone is measured by the number of summer days where the concentration exceeds 65 ppb, instrumental variables regressions suggest that an extra high ozone day increases drug purchases by 0.23% and mortality by roughly 1 per 100,000.

In addition to providing new evidence on the empirical importance of defensive expenditures, this paper makes several contributions. First, we are unaware of other studies that use real-world data to assess the impact of an emissions market on ambient pollution and human health. Most studies of emissions markets combine engineering models of emissions abatement, atmospheric chemistry models of pollution transport, and epidemiological models of dose-response functions (e.g., (Muller and Mendelsohn 2009)). Such calibration could incorrectly estimate the market's effects—a limitation underscored by our failure to find consistent evidence that the NBP affected ambient particulate matter, which the models (and the EPA) projected as the primary channel for health benefits. Moreover, because atmospheric chemistry shows that decreasing NOx emissions can either decrease or increase ambient ozone pollution (i.e., it can cause perversely-signed changes in air quality), it is important to evaluate an emissions market directly rather than relying on ex ante simulations. Additionally, the NBP is important—it is among the largest pollution cap-and-trade markets ever implemented, and several other US NOx markets have had a similar design.

Second, the results may inform contentious academic and policy debates about ozone pollution. National Ambient Air Quality Standards for ozone have changed repeatedly since the Clean Air Act—more than for any other pollutant except particulates. In 2010, President Obama announced that the EPA would tighten ambient ozone standards. The EPA then missed four deadlines to decide on a new ozone standard, and in September 2011 announced that it would await 2013 to implement new standards. This announcement was followed by litigation by environmental groups. These ozone standards are contentious partly because there is substantial uncertainty about how ozone affects health ((Bell, McDermott, Zeger, Samet, and Dominici 2004); (Currie and Neidell 2005); (NRC 2008); (Jerrett, Burnett, III, Ito, Thurston, Krewski, Shi, Calle, and Thun 2009); (Neidell 2009); (Lleras-Muney 2010); (Moretti and Neidell 2011)).

Third, the analysis uses new identification together with the most comprehensive data file ever compiled on emissions, pollution concentrations, defensive expenditures, and mortality rates. We show that the NBP provides rich quasi-experimental variation in ambient ozone concentrations over five months periods, which reduced the ozone exposure of over 135 million individuals. Our results are therefore more informative about the impacts of new ozone regulation than is the existing literature, which focuses on daily or weekly variation in ozone and on individual states or cities. In addition, our use of medium-run seasonal variation decreases concerns about "harvesting" or temporal displacement of the drug expenditures and mortality.

The remaining text is organized as follows. Section II reviews ozone pollution and the NBP. Section III presents a simple economic model of defensive investments in response to pollution. Section IV describes the data. Section V discusses the econometric models. Section VI reports the results. Section VII uses the results to conduct a cost-benefit analysis of the NBP and to measure willingness to pay for ozone reductions. Section VIII concludes.

2.2 Ozone and the Emissions Market

2.2.1 Ozone

The Clean Air Act was designed to control ambient levels of ozone and five other pollutants that harm health. Ozone differs from the other pollutants in three ways that are important for our analysis. First, polluters do not emit ozone directly. Instead, ozone forms through a complex nonlinear function combining two chemical precursors – nitrogen oxides (NOx) and volatile organic compounds (VOCs) – with sunlight and heat. The market we study operates only in summer because winter ozone levels in the Eastern U.S. are low, and ozone spikes to high peaks on hot and sunny days.

Second, the health consequences of ozone are believed to occur from short-term exposure to high levels. Ozone regulation has targeted these peak exposures, rather than focusing on mean ozone levels. For example, the National Ambient Air Quality Standards for ozone primarily reflect the highest few readings of the year. Hence, this market is most likely to affect health if it truncates the right tail of the ozone distribution. Research has found negative effects of ozone on cardiovascular and particularly respiratory health (Lippman 2009).

Third, when this market began, national ozone levels changed relatively little since the Clean Air Act first regulated ozone in 1970. By contrast, concentrations of all five other "criteria" pollutants decreased by large amounts between 1973 and 2002 (USEPA 2008). During this period, the EPA imposed numerous regulations on businesses to decrease VOC and NOx emissions. This muted effect of existing ozone regulations set the stage for an emissions market as a new approach to decrease ozone.

2.2.2 The NO_x Budget Trading Program

The NOx Budget Trading Program (NBP) grew out of the Ozone Transport Commission (OTC), an organization of Northeast States which formed in the 1990s. Studies commissioned under the OTC found that ozone levels remained high in the Northeast U.S. partly because prevailing winds transported NOx from the industrial Midwest to the Northeast, where it produced ozone in the Northeast (OTC 1998). The OTC led to a version of the NOx Budget Program which operated in 1999-2002 and produced small declines in summer NOx emissions. The OTC then created a more stringent version of the NOx Budget Program which began in 2003 and operated until 2008. The market included 2,500 electricity generating units and industrial boilers, although the 700 coal-fired electricity generating units in the market accounted for 95 percent of all NOx emissions in the market (USEPA 2009).

The market was implemented partially in 2003 and fully in 2004. The 2003-2008 emissions market originally aimed to cover the eight Northeast states plus Washington DC (which were the focus of the OTC), plus 11 additional Eastern states. Litigation in the Midwest, however, delayed implementation in the 8 additional states until May 31, 2004. Appendix Figure 1 shows the division of states by NBP participation status in the subsequent analysis.

Accordingly, the EPA allocated about 150,000 tons of NOx allowances in 2003, 650,000 tons in 2004, and about 550,000 tons in each of the years 2005-2008. Many firms banked allowances: In each year of the market, about 250,000 tons of allowances were saved unused for subsequent years (USEPA 009a). Before the NBP began, about half of NOx emissions in the Eastern US came from electricity generation and industry—the rest were from mobile and other sources. About a fourth of NOx emissions in the East came from these stationary sources following the establishment of the NBP (USEPA 2005).

Each state received a set of permits and chose how to distribute those permits to affected sources. Once permits were distributed, affected sources could buy and sell them through open markets. A single emissions cap affected the entire market region, though firms could bank allowances for any future year. At the end of each market season, each source had to give the EPA one allowance for each ton of NOx emitted. Seventy percent of units complied by using emissions controls (e.g., low NOx burners or selective catalytic reduction), and the remainder complied exclusively by holding emissions permits (USEPA 2009).

The mean resulting permit price in the emissions market was \$2,080 per ton of NOx. This reflects the marginal abatement cost of the last unit of NOx abated. In the results below, we use it to develop an upper bound on the aggregate abatement cost associated with the NBP market.

2.3 Model of Willingness-to-Pay

We build upon the canonical Becker-Grossman health production function to highlight the role of defensive investments in the measurement of willingness-to-pay for clean air (Becker 1965; Grossman 1972). This model shows that measuring of willingness-to-pay requires knowledge of both how pollution affects health outcomes such as mortality and how it affects defensive investments that maintain health but otherwise generate no utility, such as medications.

Assume the sick days $s(d)$ which a person suffers depends on the dose d of pollution she is exposed to. The ingested dose $d(c, a)$ depends on the ambient concentration c of the pollutant and on the defensive behavior a. Substituting provides the following health production function:

$$s = s(c, a) \tag{2.3.1}$$

People gain utility from consumption of a general good X (whose price is normalized to 1), leisure f, and health. Budgets are constrained by non-labor income I, the wage rate pw, available time T, and the price pa of defensive investments:

$$max_{X,f,a}u(X, f, s) \text{ s.t. } I + p_w(T - f - s) \geq X + p_a a.$$

Assuming an interior solution to the maximization problem, we can rearrange the total derivative of the health production function (1) to give the following expression for the partial effect of ambient pollution on sick days:

$$\frac{\partial s}{\partial c} = \frac{ds}{dc} - \left(\frac{\partial s}{\partial a}\frac{\partial a^*}{\partial c}\right)$$

This expression is useful because it underscores that the partial derivative of sick days with respect to pollution is equal to the sum of the total derivative and the product of the partial derivative of sick days with respect to defensive behavior (assumed to have a negative sign) and the partial derivative of defensive behavior with respect to pollution (assumed to have a positive sign). In general, complete data on defensive behavior is unavailable, so most empirical investigations of pollution on health (see, e.g., Chay and Greenstone (2003b), Chay and Greenstone (2003a)) reveal $\frac{ds}{dc}$, rather than $\frac{\partial s}{\partial a}$. As equation (5) demonstrates, the total derivative is an underestimate of the desired partial derivative. Indeed, it is possible that virtually all of the response to a change in pollution comes through changes in defensive behavior and that there is little impact on health outcomes; in this case, an exclusive focus on the total derivative would lead to a substantial understatement of the health effect of pollution. The full impact therefore requires either estimation of $\frac{\partial s}{\partial a}$, which is almost always infeasible, or of $\frac{ds}{dc}$ and $\frac{\partial a^*}{\partial c}$.

To express the marginal willingness to pay for clean air w_c in dollars, we manipulate the previous expressions to obtain the following decomposition:

$$w_c = \left(p_w \frac{ds}{dc}\right) + \left(p_a \frac{\partial a^*}{\partial c}\right) - \left(\frac{\partial u/\partial s}{\lambda}\frac{ds}{dc}\right) \tag{2.3.2}$$

Expression (3) shows that the marginal willingness to pay for clean air includes three terms. , The first is the effect of pollution on productive work time, valued at the wage rate. The third is the disutility of sickness, valued in dollars. This third component includes mortality. The second is the cost of defensive investments, valued at their market price. This second component is the aspect of willingness-to-pay that existing research has not measured. It is important to note that medications are not a complete measure of defensive investments against air pollution. However, given that medications cost more than mortality, emergency visits, or any other components of asthma's social costs (Weiss and Sullivan 2001), they represent an important component of defensive investments. The paper's primary empirical goal is to develop a measure of marginal willingness to pay that is based on $\frac{ds}{dc}$ and $\frac{\partial a^*}{\partial c}$.

This neoclassical model assumes that markets are competitive, but the setting analyzed here has two important deviations from this benchmark: markups and moral hazard. Branded medications generally have low marginal cost and high markups that reflect intellectual property rights. Hence, it might seem that part of the price of medications is a transfer from consumers to drug firms, and not a social cost. One interpretation of our use of market prices for medications is that pharmaceutical firms must invest socially valuable resources to develop medications that treat conditions exacerbated by air pollution. With lower levels of air pollution, fewer resources would be spent to develop these medications—a similar induced innovation process as in Finkelstein (2004).

The second important deviation from the neoclassical benchmark is that consumers with insurance generally pay a copayment or deductible for medications. Hence the price exceeds the marginal cost to the consumer, generally by 80-90 percent in these data. Although we use data on the transacted price for medications (which is more accurate than the published or wholesale price), it remains likely that private willingness-to-pay for medications is smaller than the medication prices we analyze.

2.4 Data

This analysis has compiled an unprecedented set of data files to assess the impacts of the NOx Budget Program. Although market-based instruments are viewed as among the most important contributions of economics to environmental policy, to the best of our knowledge this study represents the first time any analysis has linked ex post health measurements directly to emissions and air quality measures in order to evaluate an emissions market. We compile high frequency data on medications, hospitalizations, mortality, pollution emissions, ambient pollution, and weather for the period 1997-2007. The analysis excludes Alaska, Hawaii, and states adjacent to the NBP participating states, which have ambiguous treatment status given the potential of pollution to cross state borders.

The U.S. has no national census of local medication purchases, and so we use the best available alternative: confidential data on medication and hospital admissions from the Thompson Reuters MarketScan Research Database. MarketScan contracts with large employers to obtain all insurance-related records for their employees, plus their insured spouses and dependents. Because the data include dependents, they cover children and teenagers who may be especially susceptible to air pollution's effects. The data report the county of the purchaser's home, the purchase date, the National Drug Code (NDC) of the medication, and the money paid from the consumer and insurer to the provider of each medication. An NDC is a unique identifier for a chemical compound, manufacturer, and package type, which helps us to identify the medical condition associated with each medication. Data on the transacted payment for medications, rather than the market price, provides useful information because few patients or insurers pay listed prices for medications.

We use data from all persons in the 16 covered firms which appear in all seven years, 2001-2007, of MarketScan, which is the largest panel the data allow us to obtain with these firms. This extract includes over 22 million person-season year observations, and over 100 million separate medication purchases. The MarketScan extract has persons in almost all U.S. counties. Because the distribution of persons across counties is skewed, we report all values as rates per 1,000 people, and use generalized least squares (GLS) weights equal to the square root of the relevant MarketScan population. Because the other datasets become available in 1997 but medication data become available in 2001, for non-medication results we report parameter estimates both with data for the period 1997-2007 and for the period 2001-2007.

Medications, unlike hospital visits or death counts, are not linked to a single International Classification of Disease (ICD) code. In the subsequent analysis, we follow the convention in the pollution-health literature and treat respiratory and cardiovascular related episodes as most likely to be affected by air pollution. We define an NDC as respiratory if it satisfies any of three criteria: (1) if it is listed in the Third Treatment Guidelines for Asthma (NHLBI 2007); (2) in a recent New England Journal of Medicine guide to asthma treatment (Fanta 2009); or (3) in the standard industry publication for medication characteristics (PDR 2006) as indicated for asthma, emphysema, bronchitis, or chronic obstructive pulmonary disorder. We identify cardiovascular and gastrointestinal medications by their corresponding therapeutic groups in Red Book (PDR 2006). The latter category is unlikely to be affected by air pollution and is used as a placebo test for the validity of the respiratory-cardiovascular results.

This broad approach to identifying respiratory and cardiovascular drugs is the most appropriate we can discern. Nonetheless, because doctors regularly prescribe medications to treat conditions for which the medications are not indicated, it remains likely that some of these medications were prescribed for non-respiratory and non-cardiovascular conditions. Moreover, it is also likely that medications prescribed for respiratory and cardiovascular conditions are not in this list.

We count hospital admissions as including all inpatient episodes plus all emergency outpatient episodes. We follow procedures in the MarketScan guide (Thompson Healthcare 2007, p. 59) to extract emergency department admissions from outpatient claims files. We define a hospital visit as respiratory or cardiovascular or external if the ICD9 diagnosis code applies to these categories. When a hospital visit has several associated procedures each with its own ICD9 code, we take the mode procedure. Our measure of hospital costs includes all charges from the hospital to the insurer and patient.

To measure mortality, we use restricted-access data on the universe of deaths in the 1997-2007 period. These Multiple Cause of Death files (MCOD) come from the National Center for Health Statistics (NCHS) and were accessed through an agreement between NCHS and the Census Research Data Centers. These files contain information on the county, cause of death, demographics, and date of each fatality.

To measure pollution emissions, we extract daily totals of unit-level NOx, SO2, and CO2 emissions for all states from the EPA's Clean Air Markets Division. The NOx emissions are the quantities for which firms must hold emissions permits in this cap-and-trade market, so they are the most accurate measure available. In 2008, ninety-seven percent of emissions came from units with continuous emissions monitoring systems. The EPA audits all of these data to verify their accuracy and internal consistency, and we believe the emissions data have little measurement error. Units which are part of the Acid Rain Program must report NOx emissions throughout the year, while units in the NBP must report NOx emissions only in the May 1 – September 30 period. Because we compare summer versus winter, estimates in the paper use only data from Acid Rain Units. However, in the examined period, units in the NBP and not in the Acid Rain Program represent a tiny share of NOx emissions.

We use a few criteria to select ambient pollution monitoring data from the EPA's detailed Air Quality System. Many pollution monitors operate for only part of a year and for part of the 1997-2007 period. Many ozone monitors operate only in the May-September months. Moreover, monitors operate more when ozone levels increase (Henderson 1996). Many monitors for fine particulates (PM2.5) record pollution only 1-2 times per week. To address the incompleteness of these measures, for each pollutant, the main analysis uses monitors which have valid readings for at least 47 weeks in all years 1997-2007. This fairly strenuous selection rule restricts our data to include only the most reliable monitors—it excludes monitors which operate only during summer, or which operate depending on weekly ozone and weather levels, or which have frequent technical problems. Appendix Table 1 shows that we obtain similar results with a weaker monitor selection rule. For ozone, we focus on a concentration measure the EPA regulates: For each day, we calculate an "8-hour value" as the maximum rolling 8-hour mean within the day.

We also compiled weather data from records of the National Climate Data Center Summary of the Day files (File TD-3200). The key control variables for our analysis are the daily maximum and minimum temperature, total daily precipitation, and dew point temperature. To ensure the accuracy of the weather readings, we construct our weather variables for a given year from the readings of all weather stations that report valid readings for every day in that year. The acceptable station-level data is then aggregated at the county level by taking an inverse-distance weighted average of all the valid measurements from stations that are located within a 200 km radius of each county's centroid, where the weights are the inverse of their squared distance to the centroid so that more distant stations are given less weight. This results in complete weather by county-day files that we can link with the other files in our analysis.

Table 1 shows that emissions, weather, and mortality data are available for all 2,539 counties in our sample. Medication and hospitalization data are available for 95 percent of these counties, which had a population of 261 million in 2004. Ambient ozone data are only available for 168 counties, but these counties are heavily populated and their 2004 population was 97 million. Data on particulates less than 2.5 micrometers (PM2.5) are available in 298 counties (population 144 million) and data on particulates less than 10 micrometers (PM10) are available for 39 counties (population of 26 million).

The summary statistics in Table 1 also provide a benchmark to measure the economic importance of medications and the emissions market. In summer, ozone averages 48 ppb. The 2010 proposed EPA air quality standard stipulated that a county could have no more than 3 days over a total of three years which exceed 60-70 ppb. Table 1 shows that during the sample period, 24 days every summer exceed 65 ppb in the typical county. On average during this time, the average person spent $339 per summer on medications, and about $500 on hospital admissions.

The summary statistics also show why the observational associations between ozone and health may reflect unobserved variables. Columns (4) through (10) of Table 1 divide all counties with ozone data into two sets—one set of counties with mean summer ozone above the national median ("high ozone"), and a second with mean summer ozone below the national median ("low ozone"). Row 1 shows that counties with high NOx emissions are slightly underrepresented in the high-ozone counties, which reflects the reality that NOx primarily creates ozone in counties other than where it is emitted. All pollutants except carbon monoxide have significantly higher levels in the high-ozone counties. Temperature, precipitation, and dew point temperature have lower levels in high-ozone counties. The finding that so many of these observed county characteristics covary with ozone suggests that an observational association of ozone with health is likely to reflect the contributions of other unobserved variables and may explain the instability of the estimated health-ozone relationship that has plagued the previous literature. This implication of Table 1 underscores the need to distinguish the effect of ozone on health from the effects of the other possible confounders.

2.5 Econometric Model

We use a differences-in-differences-in-differences (DDD) estimator to isolate the causal effects of the emissions market on pollution and health, and use an instrumental variables approach to measure the "structural" effect of ozone on health. The DDD estimator exploits three sources of temporal and geographical variation in the emission and health data. First, we compare the years before and after the NBP's operation. Eight states plus Washington DC initiated this market in 2003, while 11 other states joined in 2004. This market did not operate before 2003. Second, twenty states participated in the NBP while twenty-two other states did not participate and were not adjacent to a NBP state (see Appendix Figure 1). Third, the NBP market only operated during the summer, so we compare summer versus winter.

Specifically, we estimate the following model:

$$Y_{cst} = \gamma_1 1\left(NBP\ Operating\right)_{cst} + W'_{cst}\beta + \mu_{ct} + \eta_{st} + \upsilon_{cs} + \varepsilon_{cst}$$

Here, c references county, s indicates season, and year is denoted by t. The year is divided into two seasons, summer and winter: Summer matches the NBP's operation period of May 1-September 30. The outcome variables, Y_{cst}, are pollution emissions, ambient pollution concentrations, medication costs, hospitalization costs, and mortality rates. Because the NBP market started partway in 2003, we define Post=0.5 in 2003 and Post=1.0 in 2004 through 2007. All regressions limit the sample to a balanced panel of county-season-years.

Ozone formation is a complex function of ambient NOx, ambient volatile organic compounds and temperature. Since there is a nonlinear relationship between health and temperature, it is important to adjust for weather flexibly. The matrix of weather controls, W_{cst}, includes measures of precipitation, temperature, and dew point temperature (a measure of humidity). For temperature and humidity, we calculate 20 quantiles of the overall daily distribution. For each county-season-year observation in the data, we then calculate the share of days that fall into each of the 20 quantiles.

To operationalize the DDD estimator, the specification includes all three sets of two-way fixed effects. The vector μ_{ct} is a complete set of county by year fixed effects, which account for all factors common to a county within a year (e.g., local economic activity and the quality of local health care providers). The season-by-year fixed effects, η_{st}, control for all factors common to a season and year: For example, it would adjust for the development of a new drug to treat asthma that was sold in NBP and non-NBP states. Finally, the county-by-season fixed effects, υ_{cs}, allow for permanent differences in outcomes across county-by seasons.

The parameter of interest is γ_1 associated with the variable $1\,(NBP\ Operating)_{cst}$. This variable is assigned a value of 0.5 in 2003 for all NBP states when the market was operating in 9 of the 20 states and a value of 1 in 2004 and all subsequent years in these states. The 2003 value was assigned to all NBP states, rather than just the implementing states, because NO_x and ozone travel great distances and emissions reductions in one NBP state affected ozone concentrations in many other NBP states. After adjustment for the fixed effects, γ_1 captures the variation in outcomes specific to NBP states, relative to non-NBP states, in years when the NBP operated, relative to before its initiation, and in the summer, relative to the winter. Importantly, this only leaves variation in the outcomes at the level at which the market operated. We also report variants on equation (7) that change the level of county, year, and season controls, and the detail of weather controls.

Given the potential for temporal and spatial autocorrelation, we use a few approaches for inference. Pollution and health data are available for each county. States decided whether to enter the market, but the market only affected pollution in summer. As a result, we report standard errors that allow clustering at the state*season level in the main tables. The appendix reports standard errors that allow for arbitrary autocorrelation within counties, states, state-years, and county-seasons; but in general the conclusions are unaffected by these alternative assumptions about the variance-covariance matrix.

Although the tables focus on the triple-difference parameter γ_1 from equation (7), separate measures of the market's effect in each year provide additional useful information. Hence, for most outcomes, we also report the parameters $\alpha_{1997} \ldots \alpha_{2007}$ from the following model:

$$Y_{cst} = \sum_{t=1997}^{2007} \alpha_t 1\left(NBP\ State\right)_{cs} + W'_{cst}\beta + \mu_{ct} + \eta_{st} + \upsilon_{cs} + \varepsilon_{cst}$$

where $1\left(NBP\ Operating\right)_{cs} = 1$ for all summer observations from NBP states, regardless of the year.

A threat to the validity of triple-difference estimators like the one we use is that differences in pre-trends could cause bias the estimated treatment effect. In our setting a pre-NBP trend in outcomes that was specific to NBP states during the summer after nonparametrically adjusting for all county by year and season by year factors could produce spurious result. The event study style graphs reported in the Appendix permit a visual test for evidence of pre-trends effects. (We also report on formal statistical tests for the presence of such pre-trends.)

We also exploit the NBP-based DDD design to obtain instrumental variables estimates of the impacts of ozone on medication purchases and mortality rates. Specifically, $1\left(NBP\ Operating\right)_{cst}$ serves as an instrumental variable for ozone concentrations. In this framework, the version of equation (7) where ozone is the dependent variable is the first-stage, and the versions with medication purchases or mortality rates as the outcomes are the reduced-form relationships between the instrument and the outcomes of interest. The validity of the required exclusion restriction is explored below.

2.6 Results

This section reports estimates of the effects of the NBP on pollution emissions, ambient concentrations of pollution, medication purchases, mortality rates and hospital admissions. Additionally, it implements the instrumental variables strategy outlined above to obtain estimates of the effect of ozone concentrations on medication purchases and mortality rates. The results are organized into separate subsections.

2.6.1 Emissions

The NOx Budget Trading Program legally required units to reduce NOx emissions, so it is unsurprising that the market decreased NOx emissions. Nonetheless, many analyses of pollution regulations compare emissions levels in a recent year against levels that would be present without the 1990 Clean Air Act Amendments (e.g., (USEPA 2009)). Such comparisons make it difficult to identify the contribution of a specific recent policy to total emissions.

Figure 2 illustrates the tremendous impact of the NBP on NOx emissions. The figure shows the unadjusted summer-equivalent NOx emissions, by year (before and after NBP operation) by season (winter and summer) and by NBP status (NBP participating states and non-participating states). The first key point shown in Figure 2 (B) is that summer and winter NOx emissions in non-participating states evolve very smoothly over time, with similar downward trends and with no evidence of any discernible change in 2003 and 2004 when NBP was implemented. In contrast, Figure 2 (A) shows that the NBP led to a sharp reduction in summer emissions, starting in 2003 when the emissions market began in 8 Northeastern states and Washington DC. As a result, summer NOx emissions declined by nearly 20 percent in the summer of 2003, and another 15-20% starting in May 2004, when the market added 11 more Eastern states. Additionally, winter emissions continued their gradual downward pre-2003 trend, with perhaps a modest slowing of that trend post-2003. In short, NOx emissions declined in exactly the areas, months, and years that the market design would predict.

Regression analogues of these graphs in Table 2 similarly show that the NBP market decreased NOx emissions by 34-38%. Table 2-5 are based on similar sets of specifications and so we explain them here. Column (1) includes no weather controls and includes three sets of two-way fixed effects—it uses state-by-year rather than county-by-year controls. Column (2) adds the full set of binned weather controls. The weather controls increase the point estimates slightly, although the estimates remain precise. Column (3) replaces state-by-year fixed effects with county-by-year fixed effects, which is the most precise control the data allow. The point estimates remain unchanged. Column (4) drops the years 1997-2000 and forces the sample to begin in 2001, since the medication and hospitalization data are only available for 2001-2007. Finally, column (5), report estimates from models where the regression is weighted by population (as opposed to weights reflecting the number of pollution/emission monitor). This is because the remainder of the paper is focused on explaining per capita defensive expenditures and hospitalization costs and the mortality rate; these equations will naturally be weighted by the relevant population to obtain estimated impacts on the average person

Based on the richest specification in Panel A (column 4), the estimates in row 1 indicate that NBP lead to a decrease in NOx emissions in the average county by 330 tons per summer, or 38% relative to baseline average NOx level. These results for NOx emissions are unchanged in alternative specifications (see Appendix Table 1)..

We also measure whether the NBP affected emissions of pollutants other than NOx. Two reasons explain why the market might have affected emissions of such co-pollutants. If permits for NOx emissions cost enough that the market caused relatively clean natural gas units to displace electricity generation from relatively dirty coal-fired units, then the market could have decreased emissions of pollutants other than NOx. Second, complementarity or substitutability of NOx with other pollutants in electricity generation could lead units to change emissions of other pollutants. Any effect of the market on ambient levels of co-pollutants, however, would imply that the market could have affected health through channels other than ozone.

The data do not provide strong evidence that the market affected emissions of co-pollutants. Rows 2 and 3 in Panel A of Table 2 show that NBP had no impact on emissions of SO2 or CO2. Further, the estimated size effects (point estimate over the mean of the dependent variable in the pre NBP years) for the co-pollutants are all nearly zero: In the preferred specifications of columns (3) and (4), they imply a statistically insignificant decrease in SO2 or CO2 emissions of about 1-3%. By comparison, the NOx reductions (row 1) range 34%-38%.

2.6.2 Ambient Pollution

Panel B in Table 2 shows how NBP affected ambient concentrations of ozone and the other pollutants that are most heavily regulated under the Clean Air Act. Panel B, rows 1-2, reveals large and precisely estimated effects of the emissions market on ground-level ozone concentrations (as measured by the maximum 8-hour value). The richest specifications in columns (3) - (5) indicate that the NBP decreased mean summer ozone by about 3 ppb (or 6-7 percent relative to the baseline mean). The NBP market also decreased the number of high-ozone days (days with average concentration above 65 ppb) by 7.5 to 8.6 days per summer, or 23%-28% of the baseline average. Thus NBP's impacted the distribution of daily ozone concentration non-uniformly, with larger reduction in the upper part of the distribution.

Appendix Figure 2A shows the corresponding event study estimates for average daily summer ozone concentrations. Again, the results suggest that the NBP decreased average ozone concentrations by 3 ppb. However, there is some evidence of differential pre-existing trends in summer-winter ozone concentrations across NBP and non-NBP states. Accounting for these differences increases the magnitude of the NBP's estimated reduction on ozone concentrations, though these models are demanding of the data and so the estimates are less precise.

We also analyze the market's impact on the density function for daily ozone concentrations to explore where in the daily ozone distribution the NBP affected concentrations (Appendix Figure 2C(. The main result is that the market reduced the number of summer days with high-ozone concentrations (i.e. greater than 60 ppb) and increased the number of days with ozone concentrations less than 60 ppb. It is noteworthy that the EPA has recently experimented with daily ozone standards of 65, 75, and 85 ppb. The variation in ozone concentrations comes from the part of the distribution where there is great scientific and policy uncertainty.

Rows 3-5 in Panel B of Table 2 test for impacts on carbon monoxide (CO), sulfur dioxide (SO2), and nitrogen dioxide (NO2). CO emissions come primarily from transportation, so it is not surprising that the regressions fail to find evidence that the NBP affected CO concentrations. Further, there is little evidence of an impact on SO2. Thus, it appears that any impacts of ozone will not be confounded with changes in CO or SO2 and this supports the use of the NBP as an instrumental variable to identify the effects of ozone on health.

NOx is a term used to describe a mix of two compounds—nitric oxide (NO) and NO2, but NO2 is a pollutant subject to its own regulations. Row 5 shows that the market decreased ambient NO2 levels by 6-7 percent. Because NO2 has limited or possibly no effect on health, this finding is unlikely to threaten the exclusion restriction necessary to identify the impact of ozone.

Air quality models show that atmospheric NOx can undergo reactions that transform it into a component of particulates. The impact of the NBP on particulates concentrations is of especial interest because particulates are widely believed to be the most dangerous air pollutant for human health (Pope, Ezzati, and Dockery 2009; Chay and Greenstone 2003b; Chay and Greenstone 2003a; Chen, Ebenstein, Greenstone, and Li 2011). Further, before its implementation, the EPA projected that 48-53 percent of the projected health benefits from the NBP would come through the channel of reduced particulates concentrations (USEPA 1998). On the other hand, the appendix describes air quality model simulations in more detail and provides an explanation for why the NBP might not affect the particulates concentrations. We examine the impact of the NBP on the concentrations of particles smaller than 10 micrometers (PM10) and 2.5 micrometers (PM2.5), both of which are small enough to be respirable, in rows 6 and 7 in Panel B of Table 2.

The results about particulates are mixed. In column (4), where the equation is weighted by the number of monitor observations, there is little evidence that the NBP affected airborne particulate matter concentrations. Alternatively, when the equation is weighted by population, as is the case in the preferred defensive expenditures and health outcomes equations, the entries indicate that the NBP is associated with a 6% reduction in PM2.5. However in the smaller sample of counties with PM10 monitors, there continues to be no evidence of a meaningful change in PM10. Overall, these results are inconclusive about whether the NBP affected particulates concentrations. These mixed results mean that the subsequent two-stage least squares results of the effects of ozone on defensive expenditures and health outcomes should be interpreted cautiously, because they may reflect the impact of ozone or particulates, or a combination of the two pollutants. Nevertheless, the evidence in Table 2 indicates that the first-order impact of NBP on ambient pollution is through its effect on high ozone days (e.g., row 2 in Panel B).

2.6.3 Defensive Investments

This section explores the relationship between the NBP market and the resources people devote to defending themselves against air pollution through medication purchases. This relationship is important in its own right as a partial measure of the NBP's social benefits, and in the instrumental variables framework where it may also be valid to interpret it as the reduced-form relationship between the instrument and each outcome.

Table 3, Panel A, reports the reduced-form effect of the market on log medication costs. The richest specification in columns (3) and (4) indicates that the NBP reduced total medication costs by 1.9 percent. This estimate is precise with the full set of controls, and has similar magnitude but less precision with less detailed controls. The theoretical model discussed earlier implies that this reduction in defensive expenditures is a key component of total willingness-to-pay for air quality, but it is one that previous research had not measured empirically.

We also measure medication purchases separately by cause. As discussed earlier, the allocation of medications to causes is inexact—doctors can prescribe a medication for many purposes, and the MarketScan data do not identify the cause for which a specific medication was prescribed. The goal of this exercise is to test whether the decline in medication purchases was especially evident among respiratory and cardiovascular medications (although the imprecision of the assignment of causes to medications means that there are good reasons to expect an impact in other categories). The estimates in row 2 of Panel A indicate that the NBP decreased expenditures on respiratory and cardiovascular medications by 2.3 percent in column (3). This estimate would be judged statistically significant at conventional levels. The corresponding estimate in column (4) would not. We also use medication costs for gastrointestinal conditions as a placebo test, because we are unaware of evidence linking air pollution exposure to these conditions. Although the column (3) estimate is marginally significant, these results together suggest that the NBP had little impact on medications for gastrointestinal problems.

Appendix Table 2 reports results from several robustness checks, none of which alter the qualitative conclusions. We investigate changing the level of clustering, estimating models where the dependent variable is the log number of medications (rather than log medication costs), changing the sample to be balanced, and estimating models that add differential pre-existing trends in the NBP states during the summer. Appendix Table 2 also explores heterogeneity in the log medication results in several ways. First, we separately estimated these regressions for children and obtained results with similar magnitude though less precision. Second using National Drug Codes, we also attempted to distinguish "maintenance" respiratory medications that are taken regularly to treat chronic respiratory conditions, from "rescue" respiratory medications that are taken once acute respiratory symptoms appear. We again obtained similar negative parameter estimates for both categories though with less precision.

2.6.4 Hospital Visits and Mortality

Hospital Visits

Because we seek to compare defensive costs against direct health costs, we also measure how the market affected hospital visits and mortality. Due to the large number of county-year-season observations with 'zeros' for hospitalization costs, we focus on the level rather than the log of per capita hospitalization costs.

Due to space limitation the results are presented in Appendix Table 3. Row (1) in the table reports that the market decreased hospitalization costs by about $6.00 per person-year, or roughly 1%, and that this estimate is not statistically significant. We find a similar result with the log of hospitalization costs as the dependent variable. The sign of the parameter estimates suggest that the market decreased cardiovascular and respiratory hospitalizations, but this result also is imprecise. Overall, we judge based on the evidence in Appendix Table 3 that the balance of evidence suggests that the NBP did not have a meaningful impact on hospitalization costs, and we do not pursue this outcome further. We emphasize however that the MarketScan data exclude uninsured, Medicare, and Medicaid patients whereas these groups are included in some studies which find effects of ozone on hospitalization (Currie and Neidell 2005; Lleras-Muney 2010).

Mortality

Finally, we assess the NBP's impact on mortality. In most analyses of air pollution, mortality accounts for the largest share of the regulatory benefits. Tables 4 and 5 report the results. In the full sample, the emissions market decreased the all-cause, all-age summertime mortality rate by 1.6 to 3.0 deaths per 100,000 population, depending on the sample, and are generally statistically significant. The effect in the subsample of counties with ozone monitors is larger, indicating a reduction of 5.4 deaths per 100,000 population.

Across the columns, an analysis of cause-specific mortality rates reveals that 35% to 56% of the decline in overall mortality is concentrated among cardiovascular/respiratory deaths. We find that the market had no effect on external (primarily accidents) deaths, which is a reassuring placebo test. Further, the impacts on neoplasms are small and statistically insignificant. This result was unknown ex ante since the relationship between ozone and cancer remains uncertain (NRC 2008). We also consider all causes of mortality other than respiratory, cardiovascular, and neoplasm. There is evidence that deaths from all other causes significantly declined also.

Appendix Table 5 reports on a series of specification checks that leave the qualitative findings unchanged. The addition of a separate time trend for summer observations from NBP states is easily rejected at conventional significance levels for all-cause, respiratory-cardiovascular, and external mortality rates. Further, the addition of this variable caused the standard errors for the estimates of γ_1 to roughly triple. We conclude that this model is over-determined and that the data do not support the inclusion of these NBP by summer trends.

Table 5 breaks the entire population into four age groups and separately estimates the effect of the NBP on each group's mortality rate using the full sample and the preferred specification (i.e., column (3) in Table 4). We detect no meaningful effect on the mortality of persons aged 74 and below, although taken literally, the point estimates imply that the market prevented about 375 deaths within this group. The largest impact on mortality occurs among people aged 75 and older. These results suggest that the NBP prevented about 1,800 deaths each summer among people 75 and older. As with the entire population, respiratory and cardiovascular deaths explain much of the effects on elderly mortality.

The age-group decomposition implies that the NBP prevented 2,175 summer deaths annually. About 80 percent of these were among people aged over 75. By contrast, the overall share of all summer deaths which occur among people aged over 75 is 55%. These comparisons suggest that the market had larger effects on the mortality rates of the elderly than of the non-elderly. These results suggest that the NBP market prevented deaths which would have occurred during the summer. But the people who die from ozone pollution may have pre-existing respiratory, cardiovascular, or other medical conditions which cause them to have limited life expectancies, even for people in this age category. In the extreme, the market could merely have moved the date of these deaths to the winter months immediately following the market. Because the regressions reflect a triple-difference estimator which compares summer and winter deaths within a year, any deaths displaced from summer to October-December of the same year would cause the triple-difference estimator to overstate the impact by a factor of two.

We explored two approaches to investigate the empirical relevance of this possibility of short-term 'seasonal' displacement hypothesis. First, we experimented with redefining each "year" to begin on May 1 of one calendar year and conclude on April 30 of the following calendar year. This redefined "year" compares each summertime season against the seven following months. Second, we estimated differences-in-differences regressions where each observation represents a calendar year (as opposed to a calendar-season-year), and where we measure the change in mortality rates by NBP status pre vs. post. We also combined these two approaches to estimate differences-in-differences models with the restructured year.

These approaches did not strongly support for the short-term displacement hypothesis. In most cases, the estimated effect of the market on mortality was negative and had similar magnitude to the models reported in the paper. Nonetheless, these estimates were imprecise and could not reject the null hypothesis that the market had no long-run impact on mortality.

This paper's focus on summertime mortality is an advance from the previous literature which has primarily estimated how ozone affects same-day or same-week mortality rates. Because the approaches described above obtained imprecise estimates of the market's long-run effect on mortality, however, we conclude that this research design lacks power to measure the effect of ozone on life expectancy beyond the five month length of the NBP's summer season.

2.6.5 Instrumental Variables

The preceding sections measure the reduced- form effects of the NBP market on pollution, defenses, and health. We now turn to an IV approach to measure the "structural" effect of ozone on health. This parameter – the social cost of marginal reductions in ozone – is widely used in economic and policy analysis. However, we want to underscore that these results should be interpreted cautiously due to the mixed evidence of an impact of the NBP on particulates concentrations. Definitive evidence of an impact of NBP on particulates concentrations would violate the IV approach's exclusion restriction.

Table 6, Panel A, reports the association of ozone with medication purchases and with elderly mortality rates. The OLS and IV regressions use the same data, so each observation represents a county-year-season. These models regress the indicated outcome on measures of ozone concentrations while adjusting for county and year fixed effects and detailed weather controls. The OLS medication regressions have varying signs. The only statistically significant OLS results imply that ozone increases the purchase of gastrointestinal medications, which are expected to have no relationship to pollution. We take this as evidence against the reliability of OLS to infer the ozone-health relationship. These estimates may reflect the feature highlighted in Table 1 that counties with high ozone differ substantially from counties with low ozone.

The OLS mortality regressions are more consistent across the two measures of ozone and suggest a positive association between ozone concentrations and mortality rates. Instrumental variables estimates use the same sample as OLS and the Table 4 (column 4) specification with fixed effects for county by season, summer by year, and county by year, as well as detailed weather controls. This specification reveals significant effects of ozone on medication purchases, with a semi-elasticity of 0.007 for average 8-hour ozone and 0.002 for the number of days when the ozone concentration exceeds 65 ppb. These estimates imply that a 10% increase in mean ozone leads to a 2.2% increase in medication purchases. There is also a significant effect for respiratory/cardiovascular related medication purchases but no effect for gastrointestinal medications. Table 2 implies that NBP is a strong instrument for ozone.

The mortality estimates also imply large direct effects of ozone. The 2SLS point estimates suggest that a 1 ppb increase in ozone pollution leads to 2.6 additional summertime deaths per 100,000 people. This implies an elasticity of mortality with respect to ozone of 0.22. These elasticities may seem large. However, recall that the reduced form relationship between the NBP and mortality rates is substantially larger in the counties with ozone monitors than in the full sample of counties with mortality data, which could reflect smaller reductions in ozone in the non-monitored counties. Further, it is worth underscoring that the counties with ozone monitors account for an important share of the country as they have a population of 97 million, which is 37% of the 262 million people in the counties covered by the mortality data. The 2SLS estimates reveal no effect of ozone on external deaths, which supports the estimates' internal validity.

If interpreted causally, these results would substantially change our understanding of the welfare consequences of ozone. The most prominent ozone-mortality study (Bell, McDermott, Zeger, Samet, and Dominici 2004) finds an elasticity of weekly ozone with respect to daily mortality rates that is smaller than ours. Further, we are unaware of evidence on the relationship between ozone and defensive expenditures measured by medication purchases.

2.7 A Cost-Benefit Analysis of the NBP and Cautious Estimates of Willingness to Pay for Ozone Reductions

The results of this paper let us report a simple cost-benefit analysis for the entire NOx Budget Trading Program, with the caveat that we only calculate some of the health benefits of this market. Nevertheless, as emphasized earlier, our analysis includes a larger set of health outcomes than most of the previous literature. Table 2 implies that the NBP market decreased NOx emissions by 365,750 tons per summer. The average cost of a NOx permit during the market was $2,080/ton. Because firms should only use abatement technologies which cost less than the permit price, the permit price represents an upper bound on the abatement cost. Specifically, this approach implies that the market required firms to spend some amount less than $759 million per year to abate NOx. Defining 2003 to have half a year of typical abatement costs, we obtain an upper bound on 2003-2007 total abatement costs of $3.4 billion (759 * 4.5).

We can now turn to estimating the social benefits of the NBP. As we discussed above, it is tempting to assume that a change in pharmaceutical purchases are simply a transfer from consumers to pharmaceutical firms and thus have zero social cost. However, lower levels of air pollution and the resulting decline in medication purchases that protect individuals from air pollution will free resources used to develop these types of drugs and allow them to be applied to more productive uses. Monetizing the social value of these freed resources is not straightforward, so we use the value of the drug purchases as a proxy. The calculations reported in Appendix Table 6 suggest that the

NBP market let Americans decrease medication expenditures by about $900 million per year, or $4 billion when summed over the 4.5 years during which NBP was implemented. It is not immediately evident whether this extrapolation from the MarketScan population is an under- or over-statement of the effect on the full population.

If the mortality estimates are taken literally, they imply that the market prevented about 2,200 deaths annually. The monetary value assigned to these deaths depends on the value of a statistical life (VSL). We use Ashenfelter and Greenstone (2004)'s upper bound VSL of $1.93 million ($2006$) for a prime age person and use Murphy and Topel (2006)Murphy and Topel's (2006) method to develop estimates of the VSL for each age group in our analysis. This is especially important in this setting where the avoided fatalities are largely coming from individuals 75 and over. The implied VSLs are as follows: $1.9 million (infants), $1.5 million (ages 1-64), $0.6 million (ages 65-74), and $0.2 million (ages 75+). The application of this approach implies that the value of the mortality avoided by the NBP is $900 million per year, or $4 billion in the period 2003-2007.

The final column of Appendix Table 6 allows for a comparison of the costs and benefits. An upper bound on the NBP's aggregate abatement costs is $3.4 billion, but the value of the reduced drug purchases alone ($3.9 billion) exceeds these costs. This finding demonstrates that defensive investments are economically important here. Once the value of the reduced rates of mortality is added in, the benefits of the market exceeded the upper-bound of its abatement costs by 232% (7.9 / 3.4). It appears that the NBP's social benefits easily exceeded its abatement costs.

Finally, estimates of willingness to pay for a reduction in ozone would have tremendous practical importance as the EPA is currently considering revising the ozone standard with an expected announcement in 2013. Noting that they must be interpreted cautiously due to uncertainty about the validity of the exclusion restriction, the IV ozone results suggest that each 1 ppb decrease in the mean summer ozone concentration in the Eastern U.S. is worth approximately $1.3 billion in social benefits. Similarly, one fewer day per summer nationally with an ozone concentration exceeding 65 ppb would yield roughly $500 million of benefits.

2.8 Conclusions

Theoretical models make clear that willingness to pay (WTP) for well-being in many contexts is a function of factors that enter the utility function directly (e.g., the probability of mortality, school quality, local crime rates) and the costly investments that help to determine these factors. One approach to measuring WTP is to find a single market that captures individuals' full valuation, as can be the case with property markets under some assumptions (e.g., (Chay and Greenstone 2005; Greenstone and Gallagher 2008)). All too frequently though, the data and/or a compelling research design for the key market are unavailable, making it necessary to develop measures of WTP by summing its components.

However, across a wide variety of applied literatures, the empirical evidence on WTP has almost exclusively focused on the factors that enter the utility function directly. The resulting measures of WTP are thus generally underestimated and the extent of this underestimation is unknown. This paper has demonstrated that defensive expenditures are an important part of willingness to pay for air quality. Indeed in the context of the NOx Budget Program, the improvement in air quality generates reductions in medication purchases that in monetized terms are as large as the value of the observed reduction in mortality rates. A fruitful area for research is to explore whether individuals' compensatory behavior and resulting defensive investments account for such a large fraction of willingness to pay in other settings.

Appendices

2.A The NOx Budget Trading Program and Particulate Matter

This note provides one explanation based in atmospheric chemistry as to why the NOx Budget Trading Program might have little or no effect on particulate matter. We begin by defining the relevant compounds:

- PM10 and PM2.5: particulate matter
- NOx: nitrogen oxides
- NO: nitric oxide, a component of NOx
- NO2: nitrogen dioxide, a component of NOx
- NH4NO3: ammonium nitrate, the component of PM2.5 and PM10 which NOx can form.
- NO3: nitrate, a derivative of NOx
- NH4: ammonium
- SO4: sulfate, formed as a byproduct of electricity generation.
- NH4e: excess ammonium, i.e., ammonium which remains after NH4 has bonded with SO4
- NH3: ammonia
- HNO3: nitric acid, a derivative of NOx

A summary is that excess ammonium (NH4e) is the necessary ingredient for nitrate (NO3) to become ammonium nitrate (NH4NO3), which is a component of particulates. In the absence of NH4e, NOx and NO3 do not form particulate matter. NH4e levels were low in the Eastern US during the operation of the NOx Budget Trading Program because levels of sulfate (SO4) were high enough to absorb much of the available NH4 so that little sulfate remained to bond with nitrate.

A more detailed explanation follows. For NOx to become a component of PM10 or PM2.5, NOx must decompose to nitrate (NO3). Nitrate then must undergo a reaction with excess ammonium (NH4e) to form ammonium nitrate (NH4NO3). Ammonium nitrate is a component of particulate matter but nitrate is not. So a necessary condition for NOx to increase particulate matter is the presence of sufficient excess ammonium to convert nitrate into ammonium nitrate.

To assess the empirical relevance of this explanation, we calibrated an air quality model (CRDM) using the 2002 National Emissions Inventory, as in Muller and Mendelsohn (2012). According to calculations from CRDM, the Eastern US had relatively low levels of NH4e during the operation of the NOx Budget Trading Program. Excess ammonium levels were low in part because NH4 preferentially bonds with SO4, which is a byproduct of sulfur emissions. Even with the Acid Rain program, sulfur levels were high enough in the Eastern US in 2003-2007 that little NH4 remained as NH4e after the NH4-SO4 reaction occurred.

According to calculations using CRDM, in the period 2003-2007, the Eastern US had relatively low levels of excess ammonium, which could explain why we fail to consistently that the NOx Budget Program affected particulate levels. Pandis and Seinfeld (2006, p. 483), a widely-cited atmospheric chemistry text, note that this phenomenon is well-established:

> The formation of ammonium nitrate is often limited by the availability of one of the reactants. Figure 10.24 shows the ammonium concentration as a function of the total available ammonia and the total available nitric acid for a polluted area. The upper left part of the figure (area A) is characterized by relatively high total nitric acid concentrations and relatively low ammonia. Large urban areas are often in this regime. The isopleths are almost parallel to the y-axis in this area, so decreases in nitric acid availability do not affect significantly the NH4NO3 concentration in this area.

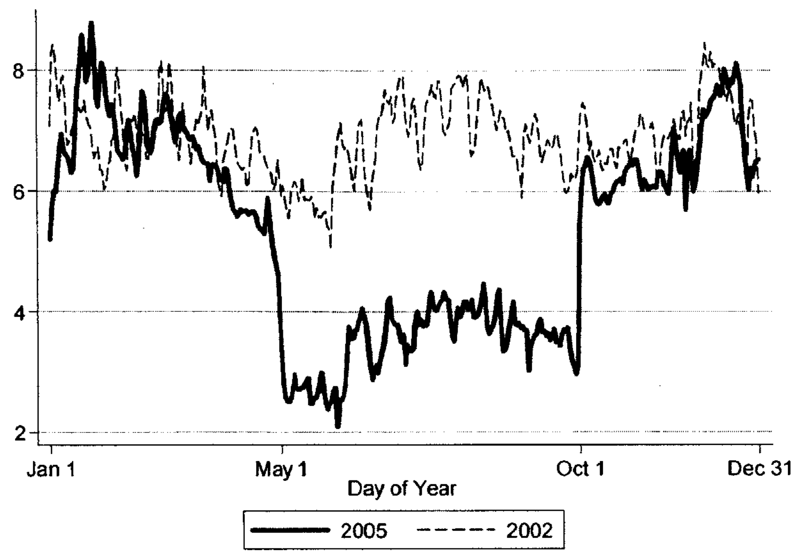
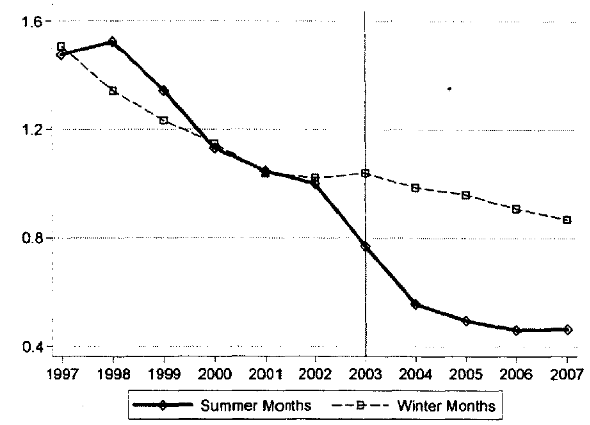


Figure 1. Total Daily NOX Emissions in NBP-Participating States

Notes: Graph depicts values from an OLS regression of NOx emissions on 6 day-of-week indicators and a constant. We control for day-of-week fixed effects since additional electricity generation on weekdays adds visible weekly cycles to the image, although the overall picture is unchanged in the raw data. The values in the graph equal the constant plus the regression residuals, so that the graph depicts fitted values for the reference category (Wednesday). Y-axis is measured in thousands of tons. Data include Acid Rain Units. NBP participating states include: Alabama, Connecticut, Delaware, District of Columbia, Illinois, Indiana, Kentucky, Maryland, Massachusetts, Michigan, Missouri, New Jersey, New York, North Carolina, Ohio, Pennsylvania, Rhode Island, South Carolina, Tennessee, Virginia, and West Virginia. See the text for more details.



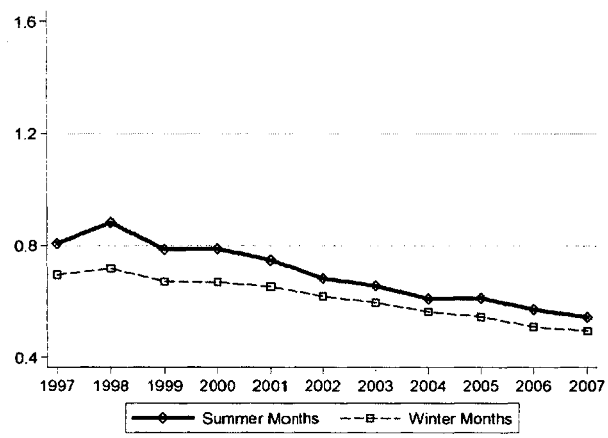


Figure 2. Summer-Equivalent Seasonal NOx Emissions (Mil. Tons)
(A) States Participating in NBP

(B) States Non-Participating in NBP
Notes: The data show raw, unadjusted emissions totals. y-axis is in millions of tons of summer-equivalent NOx emissions. Summer defined as May-September, winter as January-April and October-December. Summer-equivalent multiplies the winter total by 5/7. NBP participating states include: Alabama, Connecticut, Delaware, District of Columbia, Illinois, Indiana, Kentucky, Maryland, Massachusetts, Michigan, Missouri, New Jersey, New York, North Carolina, Ohio, Pennsylvania, Rhode Island, South Carolina, Tennessee, Virginia, and West Virginia. States not participating in NBP include: Arkansas, Arizona, California, Colorado, Florida, Idaho, Kansas, Lousiana, Minnesota, Montana, Nebraska, Nevada, New Mexico, North Dakota, Oklahoma, Oregon, South Dakota, South Carolina, Texas, Utah, Washington, Wyoming. Alaska, Georgia, Hawaii, Iowa, Maine, Mississippi, Missouri, New Hampshire, Vermont, and Wisconsin are excluded from the main analysis sample. See the text for more details.

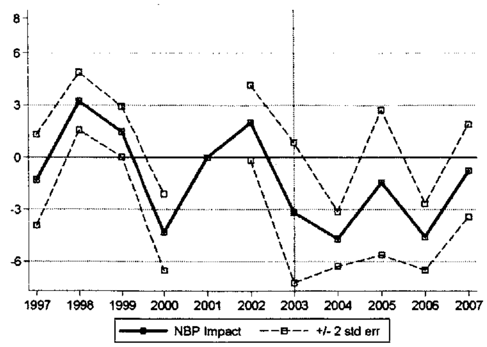
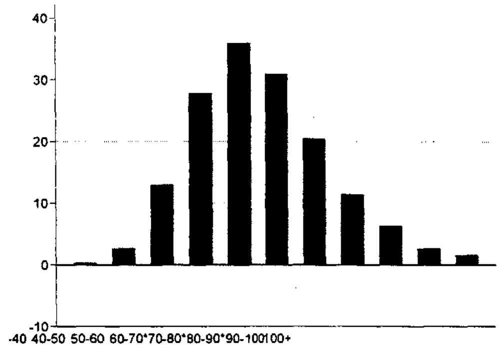


Figure 3. NBP Market Impact on Ambient Ozone Pollution

(A) Event Study for Daily Ozone 8-Hour Values, 1997-2007

Notes: Ozone 8-hour value is measured as the maximum rolling 8-hour mean of hourly values within in each day, which is the statistic used in EPA nonattainment designations. Estimate for year 2001 restricted to take a value of 0. Regression models include detailed weather controls, and a full set of county*year, season*year, and county*season fixed effects. Regression is GLS weighted by square root of number of underlying pollution readings. Standard errors based on covariance matrix allows arbitrary autocorrelation within each state-season. See Figure 2 notes or text for NBP participation status designation. See Appendix Table 5 for the full set of estimates underlying this figure.



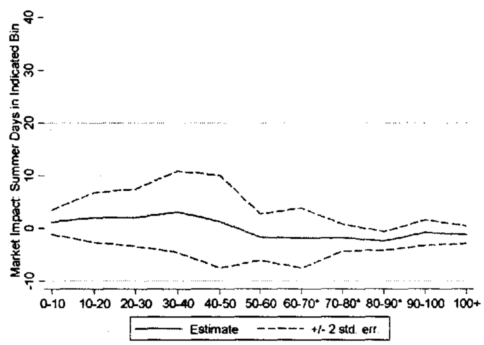


Figure 3. NBP Market Impact on Ambient Ozone Pollution (Continued)
(B) Number of Summer Days in 11 Ozone Bin, NBP Participating States, 2001-2002

(C) NBP Market Impact on Number of Summer Days in 11 Ozone Bins

Notes: Ozone 8-hour value is measured as the maximum rolling 8-hour mean of hourly values within in each day, which is the statistic used in EPA nonattainment designations. Panel B shows the average number of summer days (out of a possible 153 days) in 11 categories for daily ozone 8-hour value in the NBP states in 2001-2002 (pre-NBP period). Panel C shows the estimated impact of NBP on the number of summer days in 11 categories for daily ozone 8-hour value. Asterisks in Panel C represent EPA nonattainment standards in ppb: 85 (1997 standard), 75 (2008 standard), and 60-70 (2010 proposed standard). Estimates in Panel C are based on regression models that include detailed weather controls, and a full set of county*year, season*year, and county*season fixed effects. Regression in Panel C is GLS weighted by square root of number of underlying pollution readings. Standard errors based on covariance matrix allows arbitrary autocorrelation within each state-season. See Figure 2 notes or text for NBP participation status designation.

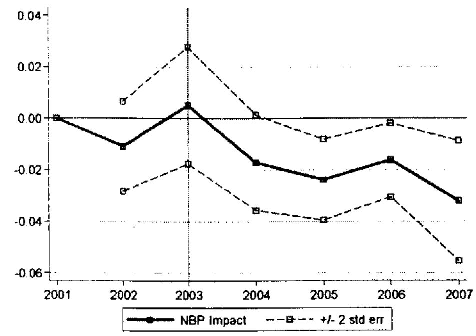


Figure 4. Impact of NBP Market on Log Medication Costs ($2006)

Notes: Log medication cost is the log of total medication costs per person-season in a county. All medication and hospital costs are in 2006 dollars, deflated using the BLS CPI for urban consumers. Estimate for year 2001 restricted to take a value of 0. Regression models include detailed weather controls, and a full set of county*year, season*year, and county*season fixed effects. Regression is GLS weighted by the square root of MarketScan population in a given county-year-season. Standard errors based on covariance matrix allows arbitrary autocorrelation within each state-season. See Figure 2 notes or text for NBP participation status designation.

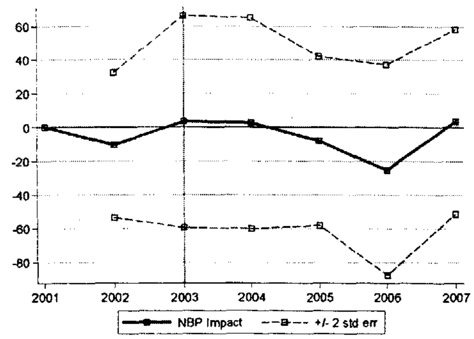


Figure 5. Impact of NBP Market on Hospital Costs ($2006)

Notes: Hospitalization costs are total hospitalization costs per person-summer in a county. All medication and hospital costs are in 2006 dollars, deflated using the BLS CPI for urban consumers. Estimate for year 2001 restricted to take a value of 0. Regression models include detailed weather controls, and a full set of county*year, season*year, and county*season fixed effects. Regression is GLS weighted by the square root of MarketScan population in a given county-year-season. Standard errors based on covariance matrix allows arbitrary autocorrelation within each state-season. See Figure 2 notes or text for NBP participation status designation.

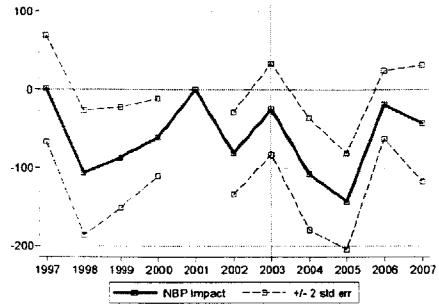


Figure 6. Impact of NBP Market on Elderly Mortality Rates

Notes: The dependent variable is the all-cause mortality rate for persons aged 75+ per 100,000 persons aged 75+. Estimate for year 2001 restricted to take a value of 0. Regression models include detailed weather controls, and a full set of county*year, season*year, and county*season fixed effects. Regression is GLS weighted by the square root of the relevant population in a given county-year. Standard errors based on covariance matrix allows arbitrary autocorrelation within each state-season. See Figure 2 notes or text for NBP participation status designation.

	All Counties			Low Ozone			High Ozone			
	Counties With Data	Mean	s.d.	Counties With Data	Mean	s.d.	Counties With Data	Mean	s.d.	p-value of H_0: (8)-(5)=0
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Pollution Emissions (000's of tons/summer)										
NO_x Emissions	2,539	0.52	(1.99)	84	1.67	(3.26)	84	1.30	(4.14)	[0.09]
SO_2 Emissions	2,539	1.50	(6.52)	84	2.92	(6.20)	84	1.41	(4.04)	[0.00]
CO_2 Emissions	2,539	384	(1,299)	84	1,263	(1,896)	84	918	(2,030)	[0.00]
Ambient Pollution										
Ozone 8-Hour Value	168	48.06	(9.28)	84	41.28	(6.10)	84	54.85	(6.58)	[0.00]
Ozone Days ?65 (ppb)	168	23.60	(22.64)	84	10.93	(9.41)	84	36.28	(24.81)	[0.00]
NO_2 (ppb)	110	11.45	(5.39)	34	8.67	(4.57)	37	12.15	(4.85)	[0.00]
CO (ppm)	125	0.44	(0.24)	35	0.46	(0.22)	33	0.42	(0.17)	[0.06]
$PM_{2.5}$ ($\mu g/m^3$)	298	13.33	(4.19)	47	10.70	(3.01)	45	11.63	(4.45)	[0.00]
PM_{10} ($\mu g/m^3$)	39	27.28	(6.26)	4	25.14	(3.85)	6	29.70	(6.86)	[0.00]
SO_2 (ppb)	150	3.26	(2.27)	32	2.04	(1.49)	33	2.60	(1.97)	[0.00]
Weather										
Temperature (°F)	2,539	70.59	(5.79)	84	73.82	(7.40)	84	72.40	(5.90)	[0.00]
Precipitation (1/100")	2,539	11.46	(5.37)	84	13.91	(8.59)	84	7.35	(6.12)	[0.00]
Dew Point Temp. (°F)	2,539	58.31	(7.58)	84	62.36	(8.59)	84	55.28	(9.57)	[0.00]
Medication Costs ($ Per Person)										
All	2,435	338.53	(302.10)	84	269.69	(84.92)	84	284.89	(107.62)	[0.01]
Respiratory + Cardio.	2,435	87.84	(97.86)	84	69.33	(28.66)	84	70.94	(30.18)	[0.35]
Hospitalizations ($ Per Person)										
All	2,435	502.62	(2120.44)	84	474.77	(418.56)	84	484.25	(703.12)	[0.78]
Respiratory + Cardio.	2,435	99.69	(768.61)	84	92.47	(250.19)	84	73.58	(142.45)	[0.11]
Mortality (Deaths Per 100,000 People)										
All	2,539	402.42	(121.32)	79	331.26	(89.47)	79	316.25	(76.94)	[0.00]
Respiratory + Cardio.	2,539	180.80	(69.93)	79	144.31	(45.37)	79	137.08	(39.59)	[0.00]

Table 1. Mean Summer Values of Pollution, Weather, and Health, by Ozone Level

Notes: All currency in 2006 dollars deflated using the US CPI for urban consumers. Emissions, medications, and deaths are totals per summer. Ambient pollution and weather are mean summer values. Low and High ozone are based on comparisons to the county with median summer ozone. Means are across counties (i.e., not weighted). All data 2001-2007.

	(1)	(2)	(3)	(4)
1. NO_x	-0.36***	-0.38***	-0.37***	-0.33***
	(0.05)	(0.05)	(0.07)	(0.07)
Effect / Mean	-0.34	-0.36	-0.35	-0.38
2. SO_2	-0.08**	-0.120	-0.070	-0.07**
	(0.04)	(0.07)	(0.05)	(0.03)
Effect / Mean	-0.03	-0.04	-0.02	-0.03
3. CO_2	-3.340	-19.040	-6.190	-12.65*
	(4.38)	(16.07)	(6.13)	(6.61)
Effect / Mean	-0.01	-0.04	-0.01	-0.03
County-by-Season FE	x	x	x	x
Summer-by-Year FE	x	x	x	x
State-by-Year FE	x	x		
County-by-Year FE			x	x
Detailed Weather Controls		x	x	x
Data Begin in 2001				x

Table 2. Effect of NBP Emissions Market on Emitted Pollution

Notes: Each observation represents a county-year-season. Winter emissions are multiplied by 5/7 so all values are summer-equivalent. Response variable measured in thousands of tons. Mean represents 2001-2002 summer in NBP areas. Covariance matrix allows arbitrary autocorrelation within each state-season. Unless otherwise noted, the sample period begins in 1997. Number of observations is 55,858 for columns (1) to (3) and 35,546 for column 94). Asterisks denote p-value < 0.10 (*), <0.05 (**), <0.01 (***).

	(1)	(2)	(3)	(4)	(5)
1. Ozone 8-Hour Value	-2.91***	-4.22***	-2.97***	-3.25***	-3.43***
	(0.77)	(1.24)	(0.75)	(0.60)	(0.60)
Effect / Mean	-0.06	-0.08	-0.06	-0.06	-0.07
2. Ozone Days ? 65	-7.40***	-8.26***	-7.46**	-8.40***	-8.62***
	(2.50)	(2.75)	(2.96)	(2.55)	(2.51)
Effect / Mean	-0.23	-0.26	-0.23	-0.25	-0.28
3. CO: Carbon Monoxide	-0.05**	-0.04	-0.04	-0.02	0.00
	(0.02)	(0.03)	(0.04)	(0.03)	(0.03)
Effect / Mean	-0.09	-0.07	-0.08	-0.03	0.00
4. SO_2: Sulfur Dioxide	0.16	0.16	0.10	0.11	0.12
	(0.12)	(0.25)	(0.18)	(0.16)	(0.15)
Effect / Mean	0.03	0.03	0.02	0.02	0.03
5. NO_2: Nitrogen Dioxide	-1.13***	-0.02	-1.21***	-1.00***	-1.25**
	(0.21)	(0.90)	(0.40)	(0.37)	(0.49)
Effect / Mean	-0.07	0.00	-0.07	-0.06	-0.07
6. $PM_{2.5}$: Particulates Less than	n.a.	n.a.	n.a.	-0.38	-1.01***
2.5 Micrometers	n.a.	n.a.	n.a.	(0.28)	(0.28)
Effect / Mean	n.a.	n.a.	n.a.	-0.02	-0.06
7. PM_{10}: Particulates Less than 10	n.a.	n.a.	n.a.	-0.90	0.11
Micrometers	n.a.	n.a.	n.a.	(1.02)	(1.25)
Effect / Mean	n.a.	n.a.	n.a.	-0.03	0.00
County-by-Season FE	x	x	x	x	x
Summer-by-Year FE	x	x	x	x	x
State-by-Year FE	x	x			
County-by-Year FE			x	x	x
Detailed Weather Controls		x	x	x	x
Data Begin in 2001				x	x
Weighted by Population					x

Table 3. Effect of NBP Emissions Market on Ambient Pollution

Notes: Each observation represents a county-year-season. Pollution readings are mean values. Regressions are GLS weighted by square root of number of underlying pollution readings unless otherwise noted. Insufficient PM data are available for the 1997-2007 period. Mean is for 2001-2002 summers in NBP States. Covariance matrix allows arbitrary autocorrelation within each state-season. Unless otherwise noted, data begin in 1997. Number of observations for each pollutant based on 1997-2007 sample (2001-2007 sample for PM) is 3,124 (Ozone); 2,244 (CO); 4,172 (PM2.5); 546 (PM10); 2,684 (SO2); 1,782 (NO2). Asterisks denote p-value < 0.10 (*), <0.05 (**), <0.01 (***).

	(1)	(2)	(3)	(4)
1. All Medications	-0.008	-0.026	-0.019***	-0.019***
	(0.011)	(0.021)	(0.006)	(0.006)
2. Respiratory + Cardiovascular	-0.005	-0.019	-0.023***	-0.015
	(0.014)	(0.023)	(0.006)	(0.010)
3. Gastrointestinal	0.012	-0.004	-0.011*	-0.001
	(0.014)	(0.027)	(0.006)	(0.014)
County-by-Season FE	x	x	x	x
Summer-by-Year FE	x	x	x	x
State-by-Year FE	x	x		
County-by-Year FE			x	x
Detailed Weather Controls		x	x	x
Counties With Ozone Monitors				x

Table 4. Effect of NBP Emissions Market on Log Medication Costs

Notes: All currency in 2006 dollars deflated using BLS CPI for urban consumers. Dependent variable is log of medication costs per person-season-year in a county. Regressions are GLS with weight equal to square root of MarketScan population in a given county-year-season. Covariance matrix allows arbitrary autocorrelation within each state-season. Asterisks denote p-value < 0.10 (*), <0.05 (**), <0.01 (***). Number of observations is as follows: Row 1 columns (1) to (3): 30,926. Row 1 column (4): 2,338. Row 2 columns (1) to (3): 28,784. Row 2 column (4): 2,324. Row 3 columns (1) to (3): 24,080. Row 3 column (4): 2,296.

	(1)	(2)	(3)	(4)
1. All Hospitalizations	-5.32	-0.47	-6.00	-78.51***
	(17.13)	(17.44)	(18.95)	(23.76)
2. Respiratory + Cardiovascular	-8.15*	-8.26	-8.70	-44.87***
	(4.73)	(5.23)	(5.72)	(9.82)
3. External	-2.75	-2.93	-3.63	-15.49
	(3.76)	(4.43)	(6.49)	(9.37)
County-by-Season FE	x	x	x	x
Summer-by-Year FE	x	x	x	x
State-by-Year FE	x	x		
County-by-Year FE			x	x
Detailed Weather Controls		x	x	x
Counties With Ozone Monitors				x

Table 5. Effect of NBP Emissions Market on Hospitalization Costs

Notes: All currency in 2006 dollars deflated using BLS CPI for urban consumers. Dependent variable is dollars per person-season-year in each county-year-season cell. Regressions are GLS with weight equal to square root of MarketScan population in a given county-year-season. Covariance matrix allows arbitrary autocorrelation within each state-season. Asterisks denote p-value < 0.10 (*), <0.05 (**), <0.01 (***).Number of observations is 31,822 for columns (1) to (3) and 2,352 for column (4). Number of observations differs from Table 4 since the log response variable of Table 4 excludes cells with no drug purchases.

	(1)	(2)	(3)	(4)	(5)
1. All Deaths	-2.15**	-3.03	-1.56*	-5.41***	-2.67*
	(0.94)	(3.47)	(0.81)	(1.83)	(1.54)
2. Respiratory + Cardiovascular	-0.75	-1.70	-0.55	-2.28*	-1.11
	(0.49)	(1.81)	(0.68)	(1.23)	(1.00)
3. Neoplasm	0.09	0.15	0.10	-0.17	-0.14
	(0.28)	(0.75)	(0.27)	(0.40)	(0.40)
4. External	0.31	-0.07	0.12	-0.66	0.17
	(0.21)	(0.37)	(0.31)	(0.66)	(0.38)
5. All Other	-1.49***	-1.49	-1.11**	-2.96***	-1.41*
	(0.38)	(1.09)	(0.43)	(0.78)	(0.72)
County-by-Season FE	x	x	x	x	x
Summer-by-Year FE	x	x	x	x	x
State-by-Year FE	x	x			
County-by-Year FE			x	x	x
Detailed Weather Controls		x	x	x	x
Counties With Ozone Monitors				x	
Data Begin in 2001					x

Table 6. Effect of NBP Emissions Market on Mortality Rates

Notes: Dependent variable is deaths per 100,000 population in each county-year-season cell. Regressions are GLS with weight equal to square root of population in a given county-year-season. Covariance matrix allows arbitrary autocorrelation within each state-season. "All Other" row corresponds to all causes of death other than respiratory, cardiovascular, and neoplasm. Unless otherwise noted, data begin in 1997. Asterisks denote p-value < 0.10 (*), <0.05 (**), <0.01 (***). Asterisks denote p-value < 0.10 (*), <0.05 (**), <0.01 (***). Number of observations is 55,858 for columns (1) to (3); 3,124 for column (4); and 35,546 for column (5).

Cause of Death	All (1)	Respiratory + Cardio. (2)
1. Age 0 (Infants)	-4.61	-1.85
	(6.28)	(1.21)
Response Var Mean	306	13
Estimated Change in 2005 Deaths	-81	-33
2. Ages 1-64	-0.14	0.24
	(0.50)	(0.26)
Response Var Mean	104	30
Implied 2005 Deaths	-168	281
3. Ages 65-74	-1.49	-3.18
	(6.00)	(3.51)
Response Var Mean	964	417
Estimated Change in 2005 Deaths	-132	-282
4. Ages 75+	-20.70*	-11.20
	(10.85)	(9.84)
Response Var Mean	3,182	1,795
Estimated Change in 2005 Deaths	-1,794	-970

Table 7. Effect of NBP Emissions Market on Mortality Rates, by Age

Notes: Dependent variable is deaths per 100,000 population in each county-year-season cell. Regressions are GLS with weight equal to square root of population in a given county-year-season. Covariance matrix allows arbitrary autocorrelation within each state-season. In 2005, market-area population levels in millions were 1.8 (infants), 116.5 (1-64), 8.9 (65-75), and 8.7 (75-99). Sample includes 1997-2007 data. Asterisks denote p-value < 0.10 (*), <0.05 (**), <0.01 (***).

	Log Medication Costs			Mortality			
		Respiratory			Respiratory		
	All	+ Cardio.	Gastrointestinal	All	+ Cardio.	External	All Other
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: OLS							
8-Hour Ozone	-0.002	-0.002	-0.003*	0.27***	0.08*	0.05**	0.13***
	(0.001)	(0.001)	(0.001)	(0.08)	(0.05)	(0.02)	(0.03)
Days ?65 ppb	0.000	0.000	0.000	0.11***	0.04**	0.01**	0.06***
	(0.000)	(0.000)	(0.000)	(0.03)	(0.02)	(0.01)	(0.01)
Panel B: 2SLS							
8-Hour Ozone	0.007***	0.005**	0.001	2.60**	1.19	0.23	1.40***
	(0.001)	(0.002)	(0.003)	(1.18)	(0.77)	(0.18)	(0.32)
Days ?65 ppb	0.002***	0.002**	0.000	1.03*	0.48	0.09	0.56***
	(0.001)	(0.001)	(0.001)	(0.58)	(0.35)	(0.08)	(0.19)

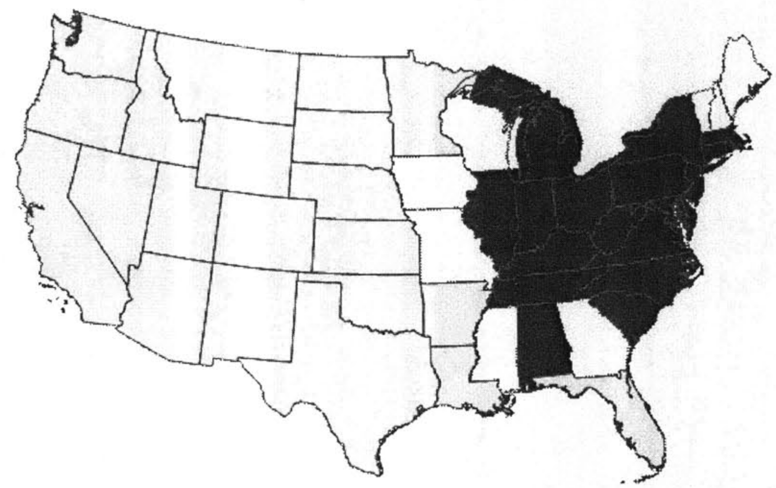
Table 8. Effect of Ambient Ozone On Medication Purchases and Mortality:
Ordinary Least Squares and Instrumental Variables Estimates, 2001-2007

Notes: Endogenous variable is ozone. Excluded instrument is Summer*Post*NBP. OLS includes county fixed effects, year fixed effects, and detailed weather control variables. Data includes population of all ages. IV regression includes specification of Table 6, column (5). GLS weights equal square root of the relevant population. Regressions use counties with ozone monitors. Covariance matrix allows arbitrary autocorrelation within each state-season. Number of observations is 2,212 for all mortality regressions. Number of observations is 2,338 for All Medications; 2,324 for Respiratory+Cardiovascular Medications, and 2,296 for Gastrointestinal medications. Asterisks denote p-value < 0.10 (*), <0.05 (**), <0.01 (***). P-values are in brackets

	Medications ($ Million)	Mortality		Total ($ Million)
		Number of Deaths	Monetized Value ($ Million)	
Panel A. An Upper Bound Estimate of NBP's Social Costs				
Upper Bound Per Year				$759
Upper Bound, 2003-2007 Total				$3,414
Panel B. Estimates of the NBP's Benefits				
Total Per Year	$873	2,175	$883	$1,756
Total 2003-2007	$3,929	9,788	$3,973	$7,902
Panel C: The Social Benefits of Ozone Reductions in the Eastern US				
1 ppb Ozone Decrease	$312	3,524	$1,431	$1,743
1 Less Day With Ozone > 65 ppb	$106	1,402	$569	$675

Table 9. The Welfare Impacts of the NBP and the Social Benefits of Ozone Reductions

Notes: All currency in 2006 dollars deflated using BLS CPI for urban consumers. Mortality table entries without dollar signs are number of deaths. Mortality dollar impact uses the VSL of $1.93 million (2006 dollars) from Ashenfelter and Greenstone (2004) and the age adjustments from Murphy and Topel (2006, p. 888). The implied VSLs are as follows: $1.9 million (infants); $1.5 million (age 1-64); $0.6 million (age 65-74); $0.2 million (age 75+). Total 2003-7 decrease due to NBP assumes impact is for half of 2003 summer and for all of summers 2004-2007. NBP cost upper bound is based on the mean permit price of $2080/ton and estimated total abatement quantity of 412,380 tons. Panel C takes the IV estimates from Table 8 and applies them to the full population of the NBP region.



Appendix Figure 1. Participation in NBP by State

Notes: Dark blue states are participating in NBP during the 2003-2007 period (NBP states). Light blue states are not participating (non-NBP states). White states are excluded from the main analysis sample.

	Emitted Pollution			Ambient Pollution						
	NO_x	SO_2	CO_2	Ozone	Ozone?65	CO	$PM_{2.5}$	PM_{10}	SO_2	NO_2
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1. Original	-0.33	-0.07	-12.65	-3.25	-8.40	-0.02	-0.38	-0.90	0.11	-1.00
State-Season Clusters	(0.07)***	(0.03)**	(6.61)*	(0.60)***	(2.55)***	(0.03)	(0.28)	(1.02)	(0.16)	(0.37)***
County Clusters	(0.08)***	(0.05)	(7.60)*	(0.54)***	(2.44)***	(0.03)	(0.31)	(1.23)	(0.24)	(0.47)**
State Clusters	(0.09)***	(0.05)	(9.41)	(0.84)***	(3.59)**	(0.04)	(0.39)	(1.44)	(0.22)	(0.52)*
State-Year Clusters	(0.05)***	(0.04)*	(6.47)*	(1.21)***	(3.77)***	(0.03)	(0.49)	(1.40)	(0.18)	(0.41)**
County-Season Clusters	(0.05)***	(0.04)*	(5.37)**	(0.38)***	(1.75)***	(0.02)	(0.22)*	(0.87)	(0.17)	(0.34)***
2. Counties With Ozone Monitors	-0.23*	-0.25	-69.21	-3.25***	-8.40***	-0.02	-0.58	-4.13	0.15	-1.11*
	(0.12)	(0.20)	(45.35)	(0.60)	(2.55)	(0.03)	(0.41)	(5.81)	(0.25)	(0.57)
3. Including ME, NH, and VT	-0.33***	-0.07**	-12.37*	-3.25***	-8.40***	-0.02	-0.38	-1.07	0.11	-1.00***
	(0.07)	(0.03)	(6.42)	(0.60)	(2.55)	(0.03)	(0.27)	(1.05)	(0.16)	(0.37)
4. Monitors Operating ? 30 weeks				-2.96***	-10.87***	-0.02	-0.52**	-0.06	0.10	-0.65*
				(0.45)	(1.90)	(0.02)	(0.26)	(1.18)	(0.14)	(0.39)
5. Summer*Post*NBP *VOC-Constrained				0.22	1.03					
				(1.18)	(4.63)					
6. Summer*Post*NBP* (High Weekend O_3)				1.54***	4.94**					
				(0.57)	(2.29)					

Appendix Table 1. Sensitivity Analysis: Emitted and Ambient Pollution

Notes: Unless otherwise noted, each table entry shows the coefficient on Summer * Post * NBP from a separate regression. Regression uses specification and sample of Tables 2-3 column (4) unless otherwise noted. The entries after row 1 present different levels of clustering for standard errors. "Allow Summer-by-East Time Trend; Weighted by County Pop." adds Summer*East*Year as a regressor. "Including ME, NH, and VT" redefines the regression sample to include data from these three states. "Monitors Operating $\geq$ 30 weeks" uses a monitor selection rule which requires each monitor to have valid readings in 30 weeks of each year in the data, rather than the 47-week rule used in the main results. "Summer*Post*NBP*VOC-Constrained" reports the interaction of the main triple-difference term with an MSA indicator for being VOC constrained based on Blanchard (2001). "Summer*Post*NBP*(High Weekend O3) interacts the main triple-difference term with an indicator for whether the weekend/weekday ozone ratio of a county exceeds 1.05. This provides an alternative indicator of VOC-constrained regions. Regressions use 2001-2007 data. Asterisks denote p-value < 0.10 (*), <0.05 (**), <0.01 (***).

	All	Respiratory + Cardio.	Gastrointestinal
	(1)	(2)	(3)
1. Original	-0.019	-0.023	-0.011
State-Season Clusters	(0.006)***	(0.006)***	(0.006)*
County Clusters	(0.006)***	(0.006)***	(0.011)
State Clusters	(0.008)**	(0.009)**	(0.008)
State-Year Clusters	(0.007)***	(0.008)***	(0.010)
County-Season Clusters	(0.005)***	(0.005)***	(0.008)
2. Including ME, NH, and VT	-0.018***	-0.023***	-0.009
	(0.006)	(0.006)	(0.006)
	-0.015***	-0.022***	-0.019***
3. Log Medications (Not Costs)	(0.005)	(0.005)	(0.005)
4. Panel of People	-0.013*	-0.018**	-0.001
	(0.007)	(0.007)	(0.010)
5. Levels (Not Logs)	-10.129***	-2.542***	-1.260***
	(2.115)	(0.642)	(0.316)
6. Purchase-Specific Costs	-0.016***	-0.022***	-0.023***
	(0.006)	(0.005)	(0.008)

Appendix Table 2. Sensitivity Analysis: Medications

Notes: Each table entry shows the coefficient on Summer * Post * NBP from a separate regression. Regressions use specification of Table 4 column (3) unless otherwise noted. The entries after row 1 present different levels of clustering for standard errors. "Including ME, NH, and VT" redefines the regression sample to include data from these three states. "Allow Summer-by-East Time Trend" adds Summer*East*Year as a regressor. Medications uses counts of medication purchases, rather than cost measures. "Panel of People" uses the much smaller panel of persons who appear in all observations of the MarketScan sample. "Levels (Not Logs)" specifies the response variable in levels rather than logs. "Purchase-Specific Costs" uses the raw reported prices, rather than averaging across national drug codes to deal with outliers as in the main analysis. "Counties with Ozone Data" restricts the analysis to include only counties with ozone monitors satisfying the monitor selection rule. "Private Costs" measures costs as purchase-level patient expenditures. Regressions use 2001-2007 data. Asterisks denote p-value < 0.10 (*), <0.05 (**), <0.01 (***).

	All (1)	Respiratory + Cardio. (2)	External (3)
1. Original	-6.00	-8.70	-3.63
State-Season Clusters	(18.95)	(5.72)	(6.49)
County Clusters	(21.94)	(8.81)	(7.01)
State Clusters	(26.94)	(8.13)	(9.22)
State-Year Clusters	(20.32)	(7.73)	(6.67)
County-Season Clusters	(15.53)	(6.24)	(4.96)
2. Including ME, NH, and VT	-1.54	-6.08	-3.22
	(18.20)	(5.47)	(6.21)
3. Hospitalizations (Not Costs)	0.00	-0.00**	0.00
	(0.00)	(0.00)	(0.00)
4. Panel of People	1.08	3.01	0.64
	(7.18)	(4.14)	(2.64)
5. Logs (Not Levels)	0.01	-0.12	-0.11
	(0.04)	(0.09)	(0.10)

Appendix Table 3. Sensitivity Analysis: Hospitalization Costs

Notes: Each table entry shows the coefficient on Summer * Post * NBP from a separate regression. The entries after row 1 present different levels of clustering for standard errors. Regressions use specification and sample of Table 5 column (3) unless otherwise noted. "Allow Summer-by-East Time Trend" adds Summer*East*Year as a regressor. "Including ME, NH, and VT" redefines the regression sample to include data from these three states. "Hospitalizations (Not Costs)" uses counts of hospitalizations, rather than cost measures. "Panel of People" uses the much smaller panel of persons who appear in all observations of the MarketScan sample. "Logs (Not Levels)" specifies the response variable in logs rather than levels. "Counties with Ozone Data" restricts the analysis to include only counties with ozone monitors satisfying the monitor selection rule. "Private Costs" measures costs as purchase-level patient expenditures. Regressions use 2001-2007 data. Asterisks denote p-value < 0.10 (*), <0.05 (**), <0.01 (***).

	All	Respiratory + Cardio.	External
	(1)	(2)	(3)
1. Original	-1.56	-0.55	0.12
State-Season Clusters	(0.81)*	(0.68)	(0.31)
County Clusters	(1.16)	(0.78)	(0.34)
State Clusters	(1.16)	(0.96)	(0.44)
State-Year Clusters	(1.65)	(1.12)	(0.36)
County-Season Clusters	(0.82)*	(0.55)	(0.24)
2. Including ME, NH, and VT	-1.70**	-0.67	0.15
	(0.79)	(0.66)	(0.30)
3. Logs (Not Levels)	-0.01***	-0.01**	0.01
	(0.00)	(0.00)	(0.01)
4. Age-Adjustment	-1.50*	-0.76	0.12
	(0.85)	(0.67)	(0.31)

Appendix Table 4. Sensitivity Analysis: Mortality

Notes: Each table entry shows the coefficient on Summer*Post*NBP from a separate regression. Regressions show specification and sample of Table 6 column (3) unless otherwise noted. The entries after row 1 present different levels of clustering for standard errors. "Allow Summer-by-East Time Trend" adds Summer*East*Year as a regressor. "Including ME, NH, and VT" redefines the regression sample to include data from these three states. "Logs (Not Levels)" specifies the response variable in logs rather than levels. Age-adjustment modifies the response variable to use age-adjusted mortality counts, rather than total deaths per population. Regressions use 1997-2007 data. Asterisks denote p-value < 0.10 (*), <0.05 (**), <0.01 (***).

Chapter 3

The Clean Water Act and U.S. Water Pollution

3.1 Introduction

Since the U.S. passed the Clean Water Act in 1972, public and private sources together have spent roughly a trillion dollars to decrease water pollution, or about 100 dollars per person per year.[1] A 1969 fire on the Cuyahoga River in Cleveland, Ohio, provided the immediate impetus for the Clean Water Act:

> Chocolate-brown, oily, bubbling with subsurface gases, it oozes rather than flows. "Anyone who falls into the Cuyahoga does not drown," Cleveland's citizens joke grimly. "He decays." The Federal Water Pollution Control Administration dryly notes: "The lower Cuyahoga has no visible life, not even low forms such as leeches and sludge worms that usually thrive on wastes." It is also – literally – a fire hazard (Time 1969).

The Cuyahoga lit on fire about every decade beginning in 1868, though it has had no fire since 1969.

[1]Expenditures refer to 2007 dollars, inflated using the annual Plant Cost Index of *Chemical Engineering*. Total 1972-2007 expenditure is approximately \$847 billion, mostly focused on controlling municipal and industrial discharges. Most regulations and expenditure began soon after the 1972 Federal Water Pollution Control Amendments. Although the 1977 law has the title of "Clean Water Act," I follow the practice of most writers in referring to the 1972 law as the "Clean Water Act." The \$847 sum includes the following: \$133 billion of federal construction grants; \$63 billion from a federal Revolving Loan Fund; \$547 billion in private operating and capital costs through 1994; and \$8.1 billion annually since then (U.S. Census Bureau 2008). US Census annual population estimates show an average population of 251 million between 1972 and 2007.

A classic story in economics describes a factory which dumps waste into a river, leading people downstream to face more polluted water, and this sort of externality can justify public action.[2] But this story does not help design policy. This paper compiles new data on US water quality and uses it to investigate two questions that help evaluate this $1 trillion in expenditure.

First, I quantify levels of and trends in U.S. river water quality over the last 35 years. I find that mean pollution over the last 35 years has been low, as measured by any pollutant. Nonetheless, a small proportion of waters are severely polluted. More importantly, national pollution levels have substantially declined since the Clean Water Act. At the same time, most pollutants were declining at similar rates in years before the Act.

Second, I evaluate how the Clean Water Act affected river water quality. I analyze the Act's two main components: EPA grants to build or improve municipal wastewater treatment plants; and industrial permits given for all facilities which directly discharge pollution into U.S. waters which limit their allowable water pollution effluent. I find that both policies improved U.S. water quality, as measured by the oxygen content of rivers. Industrial regulations for the largest polluters have especially large (though imprecisely estimated) water quality benefits. At the same time, I obtain some evidence that the municipal grants increase bacteria concentrations of local rivers, which is consistent with a story where expanding treatment plants attracted waste from nearby municipalities.

Overall, these results suggest that the quality of U.S. rivers has dramatically improved in the last four decades. Policies and trends which preceded the Clean Water Act contributed to this improvement, and the Clean Water Act furthered it. Nonetheless, direct evidence of the benefits of improvements in water quality, caused by the Clean Water Act or other forces, remains weak.

To conduct this analysis, the paper arrays the most comprehensive set of files ever compiled on U.S. water quality and its determinants. I extract water quality information from three repositories which together provide data on several hundred thousand water quality monitors and almost all targeted pollutants of the Clean Water Act over four decades. I supplement these with the results of a Freedom of Information Act request, which identify the location and timing of wastewater treatment grants and industrial regulations promulgated as part of the Clean Water Act.

[2] This story appears in many economics textbooks. The earliest mention I have found is George Stigler's The Theory of Price (Stigler 1952). This is the source that Coase (1960, p 2) cites in mentioning the story. Even Milton Friedman (1962, p. 30) writes, "The man who pollutes a steram is in effect forcing others to exchange good water for bad. The others might be willing to make the exchange at a price. But it is not feasible for them, acting individually, to avoid the exchange or to enforce the appropriate compensation."

This paper builds on several literatures. Some other studies quantify trends in water pollution for a single pollutant or handful of sites. To the best of my knowledge, however, no study uses any research design besides measuring post-1972 trends to evaluate the Clean Water Act. Smith and Wolloh (2012) find that dissolved oxygen levels in freshwater lakes did not change over the period 1975 to 2011, using data from 8,000 sites. USEPA (2000) compares dissolved oxygen in 311 river segments between the two periods 1961-1965 and 1986-1990. This analysis explicitly uses "data mining" to find the largest possible estimate of improvements in water quality. Smith, Alexander, and Wolman (1987) find that concentrations of some organic pollutants decreased between 1974 and 1981 at 388 monitors. Hayward (2011) reviews EPA reports and other sources and summarizes that water quality has improved in the last 40 years. The EPA Annual Report to Congress summarizes the share of rivers, lakes, and estuaries which are "good," "threatened," or "impaired," though definition of "good" can vary across states or within states. Many studies lament the dearth of good data.[3] Finally, several studies document historic river temperature increases in individual sites or watersheds in Colorado, Austria, East Africa, the UK, and elsewhere.[4]

This paper also builds on recent research on the causes and consequences of water pollution. Water pollution in many developing countries has extreme levels. Ebenstein (2012) finds that China's water pollution has elevated digestive cancer rates. Greenstone and Hanna (2011) find that India's water pollution regulations have been largely ineffective. Duflo, Greenstone, Pande, and Ryan (2012)'s experiment of incentives for Indian pollution auditors also suggests that the failure to implement written regulations contributes to high Indian pollution levels.

3.2 The Clean Water Act and its Predecessors

Background on U.S. water policy may clarify the role of this paper. The Clean Water Act originally sought to eliminate pollutant discharges entirely by 1985, and to make waters "fishable and swimmable" by 1983. Its two main components were the construction grants, which paid 75

[3]Harrington (2004, p. 82) write, "Thirty years (1972-2002) is certainly enough time to observe the effects of the Clean Water Act . . . [But] the relevant data, when collected at all, are scattered. . . " Similarly, Knopman and Smith (1993) write, "Unfortunately, the existing information base [on water quality] is fragmentary at best." Even in 1972, a Senate report accompanying the Clean Water Act stated, "[M]uch of the information on which the present water quality program is based is inadequate and incomplete. . . . The fact that may industrial pollutants continue to be discharged in ignorance of their effect on the water environment is evidence of the information gap" (cited in Knopman and Smith p. 19).

[4]A partial list of papers measuring water temperature trends in a handful of sites includes IPCC (2001), Christensen, Wood, Voisin, Lettenmaier, and Palmer (2004), Webb and Walling (1992), and Liu, Yang, Ye, and Berezovskaya (2005).

percent of the capital cost for new municipal wastewater treatment plants through 1983 and 50 percent thereafter, and industrial permits (the National Pollutant Discharge Elimination System, or "NPDES").[5]

3.2.1 Predecessors to the Clean Water Act

Although the 1972 Clean Water Act provided the legal basis for ensuing water quality policies, it was not the first U.S. water quality regulation. The federal government passed major water pollution control laws in 1948, 1956, 1961, 1965, 1966, and 1970. U.S. states implemented local water pollution regulation, particularly in the 1960s.[6]

The predecessor national laws had the same general structure as the 1972 Clean Water Act, but they had far lower levels of investment and enforcement. The 1948, 1956, and 1961 rules all encouraged municipal wastewater treatment and industrial abatement, but provided grant funds far below municipal demand (and an order of magnitude below the investments of the Clean Water Act). They also provided limited means for federal enforcement of industrial pollution limits. The 1965 and 1966 laws increased grant money and enforcement powers, though the 1965 law still resulted in only fourteen industrial violation notices nationally (Andreen 2003).

State and local water pollution regulations before 1972 were more scattered. By 1966, all 50 states had passed some type of water pollution legislation. Some states took these laws seriously and actively enforced violations, while many others had written policies with little enforcement (Hines 1967). If there is limited evidence on the consequences of the Clean Water Act, there is even less evidence on the consequences of pre-1972 state and local policies.

3.2.2 Wastewater Treatment Grants

The Clean Water Act grants represented a significant outlay—for parts of the 1980s, they were the federal government's second-largest infrastructure program. Through these grants, the EPA gave money to municipalities to build or improve publicly owned facilities which treat wastewater. In 1973, these facilities predominately treated sewage, but since then these facilities have increasingly treated effluents from industry. In total, the grants program distributed $61.1 billion

[5] Although the federal government was directed to pay 75 percent of the cost, if "special aid" was provided, some localities paid as little as 5 percent of the cost.

[6] The laws were the Federal Water Pollution Control Act of 1948; the Federal Water Pollution Control Act Amendments of 1956; the Federal Water Pollution Control Act Amendments of 1961; the Water Quality Act of 1965; and the Clean Water Restoration Act of 1966.

to municipalities (USEPA 2000).

The EPA designates three levels of wastewater treatment – "primary" treatment allows solids to settle and removes 35-50 percent of biochemical oxygen demand (BOD_5), a standard measure of pollution; "secondary" treatment exposes water to bacteria to let the bacteria degrade sludge and organic matter and removes 85-90 percent of BOD_5; and "tertiary" treatment uses more complicated chemicals to remove over 90 percent of the BOD_5. The CWA mandated that all wastewater treatment plants which were already built or had approval for construction granted before June 1974 had to add secondary treatment by July 1977. While only a third of major facilities met this deadline, the technology used in each facility did improve considerably over time. In 1978, 4,278 facilities failed to provide secondary treatment. By 1996, only 200 facilities did, even while the total number of facilities increased by almost 40 percent (USEPA 000b, p. 2).

Why did some municipalities receive grants but others did not, and why did some receive grants before others? Initially, the major determinant of grant receipt was whether a facility had only primary treatment. States created priority lists based on facility needs then the EPA distributed grants directly to municipalities (CBO 1985). Although political issues may have affected the distribution of grants – areas with more influential political representation may have received more money – grants were officially distributed according to objective criteria. It is likely, then, that municipalities receiving grants had worse water treatment infrastructure, though it is unclear if they had worse water quality.

3.2.3 Industrial Permits

The Clean Water Act's second main component was industrial discharge permits. Every facility which discharged pollutants directly into U.S. surface waters was required to have a permit. Permits typically specify the total amount or concentration of each pollutant which the facility is permitted to emit.[7] Facilities which received permits also had to monitor their own effluent and to report it to the EPA each quarter. The EPA distributed the first round of permits between 1972 and 1976.

Initially, numerous permit writers decided what requirements each facility would face, based on the writer's judgment of what was feasible.[8] The Clean Water Act stated that the EPA should

[7] Facilities which sent effluents to wastewater treatment plants did not receive these permits, but instead had to satisfy a different set of "pretreatment" criteria.

[8] The EPA uses several formal standards for the technical requirements they impose on firms. In 1973, most grants were based on Best Professional Judgment (BPJ). The Clean Water Act set the goal that all industries should adopt Best Practicable Technologies (BPT) by 1977, and Best Available Technology (BAT) by 1983.

write guides for 30 different industries, but the complexity of establishing uniform standards made it difficult for the EPA to comply quickly.

The EPA estimated that full compliance with these permits would decrease the direct discharge of priority pollutants by 97 percent Harrington (2004, p. 79). But these permits require firms to monitor and report their own effluents, so firms could falsify reports. Enforcement of permits has been reasonable since the 1970s. In the pulp and paper industry in the 1980s, for example, the average "major" pollution source received one EPA inspection every year (Magat and Viscusi 1990).[9]

3.3 Data on Water Quality and the Clean Water Act

The structure of the Clean Water Act has been well-known since its passage in 1972. However, it has been difficult to evaluate the Act due to a dearth of data on water pollution and on the location and timing of the Act's main activities.

In order to provide clearer evidence on the consequences of the Clean Water Act, this study compiles the first-ever national annual longitudinal dataset of U.S. rivers. I summarize key facts about these data here.

3.3.1 Water Pollution Data

I extract water pollution measurements from three large repositories: the EPA's Legacy Database Center, which has historic data collected by many organizations; the EPA's Modernized Storage and Retrieval System (STORET), which has more recent data; and the National Water Information System, which is collected by state USGS affiliates. The combined data cover the lower 48 states and most counties. The Legacy Database Center has the most data, with over 200 million readings and several hundred thousand monitors. The EPA continuously updates STORET and it contains more recent pollution data. The USGS data include readings from several small pollution monitoring networks. The National Stream Quality Assessment Network (NASQAN), which provides the basis for Smith, Alexander, and Wolman (1987), the most prominent of these networks, is included in the USGS data used in the present paper.

Because the data come from various sources and organizations, I impose several sample restrictions to enhance the data's internal consistency. I keep observations with non-missing observation

[9]EPA inspectors define the largest sources as "major" according to composite measures of their emissions.

date. I also restrict the data to include only "Ambient" monitoring (which describes typical environmental conditions). Additionally, I exclude data from monitoring sites which indicate any of the following types: industrial monitoring; tissue samples; industrial effluent; nonambient data; ocean samples; land samples; fish samples; industrial intake; and runoff.

Each measure of water quality corresponds to a single code which describes the method of measuring the pollutant. For example, codes may identify the temperature at which the lab analysis must be undertaken or the time between the water quality sample and the lab analysis. A given measure of water quality can have multiple codes—for example fecal coliforms are measured 11 different ways in the data. For each measure of water quality, I analyze only data from the most common code, so that the analysis compares only readings taken under similar conditions.[10] To limit the influence of outliers, the main sample excludes the top percentile of readings in the 1969-1998 period. The sample also excludes monitors where the listed county code conflicts with the county corresponding to the monitor's listed longitude and latitude. The sample size diminishes rapidly before 1969 and after 1998, so I use this 30-year window for analysis. I relax several of these restrictions in sensitivity analysis.

I link these measures to rivers using the National Hydrography Dataset (NHD). NHD provides geocoded information on every river and stream in the U.S. Few pollution measurements list the river where they are located in interpretable form. So I use geographic information system software to assign monitors designated as being on a "river" to the longest river which is within 0.2 miles. The analysis excludes data from monitors which are more than 0.2 miles from a river.

3.3.2 Clean Water Act: Construction Grants

The Clean Water Act grants distributed 61.1 billion in nominal dollars (USEPA 000b). I filed a Freedom of Information Act request to obtain a list of all grant amounts, dates, and recipients. The data report about 15,000 grants given to 8,000 different municipal agencies. Some agencies received multiple grants. The data include only \$58 billion in grants, and EPA authorities were unable to explain the \$3.1 billion difference. These grants were exclusively for municipal wastewater treatment facilities.

[10]The Legacy Data Center and NWIS share a coding system. I link each measurement in modernized STORET to one of these codes using alphanumeric fields in modernized STORET. A unique link is infeasible for fecal streptococcus, fecal coliforms, total coliforms, or pH, so I do not use data from Modernized STORET on these parameters. Modernized STORET has hardly any data on ammonia, so I exclude Modernized STORET from the sensitivity analysis which compares across data sources.

The data provide alphanumeric strings describing grant recipients. I use three sequential steps to identify the exact location of each facility which received each grant. First, I extract information on water treatment facilities (SIC codes 4941 and 4952) from the EPA's Permit Compliance System. Second, for grants which did not link to the Permit Compliance System, I use the EPA's 1996 Clean Water Needs Survey, which is a periodic census of all U.S. wastewater treatment facilities. Third, because the Freedom of Information Act data list zip code and city for most grants, for facilities which still remain unlinked, I identify the county by linking zip code and city to non-wastewater facilities in the Permit Compliance System. This algorithm identifies the destination county for 81 percent of funds distributed under the grants program.

Because these grants represent investment in physical capital, in regressions I measure the cumulative investment up to a given year. So the paper measures the effect of having received a wastewater treatment investment in the current or any previous year on current water pollution.

3.3.3 Clean Water Act: Industrial Permits (NPDES)

The EPA's Permit Compliance System (PCS) provides public data on facilities which emit water pollution. The PCS database includes a field identifying when a plant first received a permit. This field is not available in the public version of the PCS, so I obtained it also from a Freedom of Information Act request. I identify plants with an initial permit date between 1973 and 1976. This identifies about 15,000 plants, some of which are "major emitters" and many of which are water treatment facilities.

3.3.4 Data Characteristics

These data provide extremely detailed information on rivers and pollution in America since 1968 (Table 1). The data come from nearly half a million monitors. The average pollution monitor records 25-30 measurements and operates for 3-4 (not necessarily consecutive) years. For most pollutants, two-thirds of the data comes from the oldest database (Legacy Database Center). Figure 1 depicts each monitoring station. The map shows that some states like Massachusetts have very dense monitoring networks, while others like Texas have very few monitors. The differences between monitoring intensity are so large that many state boundaries are visible simply from the difference in density of monitors. Since these monitors only record river pollution, they exclude almost the entire Florida Everglades and the Great Lakes.

The data represent an imbalanced panel, and few monitors appear in all four decades (Appendix

Table 1). While the imbalance does not affect consistency of the estimates in this paper, it does underscore the importance of including monitor fixed effects in regressions to address the sample composition.

3.3.5 Data Reliability

Because these data are new and untested, I assess their reliability in two ways. One is to compare readings across pollutants. While the main pollutants in this study capture different aspects of water, a reading of "high pollution" according to one measure almost always means that other measures also indicate high pollution (Appendix Table 1). For example, more oxygen in the water associates with fewer bacteria (fecal coliforms) and less suspended solids. Most of the pairwise correlations reported in Appendix Table 1 are statistically significant at the 99 percent level and have the expected signs.

I also explore whether dissolved oxygen measurements follow standard chemical predictions. Dissolved oxygen typically decreases in hot summer months; in early morning hours; and at greater depths. Appendix Figure 1 plots regressions of dissolved oxygen on binned indicators for each of these physical factors, while including monitor fixed effects. The patterns closely fit standard chemistry predictions: reported dissolved oxygen levels are lowest in summer, in early morning, and at greater depths.

3.4 Econometrics

The paper first asks a purely statistical question: how has U.S. river water quality changed over time? This analysis requires care because even for a single methodology, water quality measurements depend on the season, depth, and temperature at which they are taken.

To account for these measurement differences, I estimate water quality trends from the following statistical model:

$$P_{it} = \kappa + Y_{it}\alpha + \gamma_i + \gamma_s + \gamma_h + \gamma_d + \epsilon_{it} \tag{3.4.1}$$

The water quality P at monitor i and time t covaries with the year Y of the measurement. Regressions either specify the year Y as a linear trend or a vector of indicators. The fixed effects γ_i for each monitor, γ_s for each month, γ_h for each hour of the day, and γ_d for each possible reading depth nonparametrically adjust for all time-invariant differences specific to these possible measurement variables. The parameter α represents the water quality trend. The constant κ represents mean water quality, evaluated at the reference categories of the other explanatory variables.

This regression includes several features to address standard challenges in water quality data. First, because water quality data displays spatial and temporal autocorrelation, I allow for arbitrary autocorrelation of any form within each watershed. Watersheds are demarcated by 8-digit Hydrologic Unit Codes (HUCs, also called Catalogue Units).

Second, different organizations record data at different monitors. The monitor fixed effects γ_i adjust for all river attributes, idiosyncrasies of monitoring equipment, and non-anthropogenic sources of pollution which differ across monitors but are constant over time.

Third, monthly fixed effects γ_s address seasonality by adjusting for variation in water quality which is specific to a particular month. For example, dissolved oxygen levels in July are typically lower than in January.

Fourth, using a vector of year indicators in some specifications rather than a linear trend allows for transparent, graphical, and nonlinear depiction of trends.

Finally, the raw water quality data do not constitute a random sample of U.S. rivers—the data disproportionately represent areas with dense population or dense river networks (Figure 1). Hence, the statistical model assigns each observation a weight of $1/N_{wy}$, where N_{wy} represents the number of observations in the corresponding watershed and year. I define watersheds according to 2,264 different Hydrologic Unit Codes (HUCs). I also explore results which are weighted by the population in each watershed, and thereby summarize the water quality that the average American faced.

The paper's second research question investigates the extent to which the Clean Water Act caused any changes in water quality. I estimate the following statistical model separately for each pollutant:

$$Z_{jt} = \gamma b_{jt} + X'_{jt}\beta + \eta_j + \alpha_t + \epsilon_{jt} \tag{3.4.2}$$

Here Z_{jt} measures the pollution concentration in county or watershed j and year t. b_{jt} measures regulations imposed under the Clean Water Act – either grants for municipal wastewater treatment or permits for industrial effluent. X_{jt} includes controls for the share of readings taken from each hour of the day, each month of the year, and each of 12 categories of river depth.

The county fixed effects η_j in equation (3.4.2) control for all time-invariant determinants of water pollution which are specific to each county and could be correlated with the Clean Water Act investments X'_{jt}. For example, counties which received grants may have greater population levels, worse wastewater treatment technology, or greater initial pollution concentrations. These fixed effects nonparametrically adjust for time-invariant components of these confounding variables.

The year fixed effects α_t control for secular trends which are common across the entire U.S. For example, the scale of the wastewater treatment grants increased over time, whereas mean U.S. pollution levels declined. Including these fixed effects makes it possible to separate the effects of these secular trends from the grants' direct effects.

Equation (3.4.2) provides an unbiased estimate of the parameter γ if there is no covariance between the regulations b_{jt} and the regression error ϵ_{jt}, after controlling for the other explanatory variables:

$$\mathbb{E}[b_{jt}\epsilon_{jt}|X'_{jt}, \eta_j, \alpha_t] = 0$$

This assumption would be violated if, for example, grants or permits responded to unobserved shocks to variables like population which affect pollution concentrations.

3.5 Levels and Trends in Water Pollution

The level of water pollution in the U.S. is an economically important statistic. To the extent that water quality is valuable, a summary of water pollution measures one component of U.S. well-being. Similarly, measuring trends in water pollution helps quantify changes in well-being.

3.5.1 Measures of Water Quality

This analysis reports levels and trends for four types of water quality indicators: the "conventional" pollutants which the Clean Water Act targeted; broad measures of water quality; microbiological indicators, and nutrients. I analyze trends in all of these outcomes rather than just the conventional pollutants because it is informative to know whether levels of pollutants like nutrients have diminished even though the Clean Water Act did not regulate them.

The Clean Water Act targeted five conventional pollutants. Biochemical oxygen demand (BOD) measures the quantity of oxygen which aerobic microorganisms consume in five days while decomposing organic matter. It constitutes an omnibus measure of the oxygen-demanding pollution in water from sewage, industry, and other sources. Fecal coliforms come primarily from sewage and can indicate the presence of pathogenic bacteria, viruses, and protozoans which harm human health.

Pathogens remain the most common reason why water quality violates state standards. Fecal coliforms produce cloudy and malodorous waters. pH reflects a river's overall health—excessively acidic pH (below 6.5) or basic pH (above 9.0) both impair animal life. Total suspended solids, the fourth conventional pollutant, includes farm runoff, industrial effluent, and sewage, so provides another composite water quality requirement. The Clean Water Act also targeted a fifth pollutant – oil and grease – but because I have little data on this, I do not analyze it.

Besides these conventional pollutants which the Clean Water act targeted, I also quantify trends in several general measures of quality. Alkalinity measures the buffering capacity of rivers, which can protect a river against extreme pH even if for example highly acidic effluent enters the water. Specific conductance measures the ability of water to carry an electrical current and proxies for a variety of inorganic dissolved solids. Dissolved oxygen responds to almost all forms of pollution, it provides essential sustenance for most aquatic life, it prevents the growth of anaerobic bacteria which produce malodorous waters, and it is sometimes considered as the single most important measure of water quality. Oxygen depletion and associated organic enrichment is the third-most common reason why rivers violate state standard. Temperature is an important measure of both climate change and thermal pollution from power plants and industrial sites. Scientists predict that climate change will cause warming of rivers, and that this increased temperature may lead to lower dissolved oxygen, increased algal blooms, and other decreases in water quality.

The third category of water quality indicators includes two bacteriological markers: fecal streptococcus and total coliforms. Bacteria are the indicator of water quality which is most important to human health. Although the most data is available on fecal coliforms, in some cases fecal streptococcus more closely reflects human sewage than fecal coliforms do.

The final category of water quality indicators are nutrients. These are one of the most active topics of water quality discussion because they predominantly come from agriculture and nonpoint sources which the Clean Water Act does not regulate, because they are a leading cause of poor water quality, and because they generate hypoxic zones in the Chesapeake Bay and the Gulf of Mexico. Phosphorus in particular is a leading cause of eutrophication. Both phosphorus and ammonia come from fertilizer, though ammonia also comes from sewage.

3.5.2 Results

The mean represented in these data is unpolluted, though between 5 and 10 percent of readings come from heavily polluted waters (Table 1 and Appendix Figure 1). U.S. water quality standards set a target for dissolved oxygen to exceed 5.5 or 6.5, and the average dissolved oxygen level is 9.0. Nonetheless, more than 5 percent of readings have a dissolved oxygen level below 4.6, at which point aquatic life beings to suffer. Similarly, while the mean fecal coliform measurement of 100 ($\approx e^{4.6}$) is safe, at least 5 percent of observations have values above 5,000, which is unfit for human contact. pH data can be difficult to interpret since unpolluted water has a pH level between 6 and 9, and water pollution can either make water too acidic (low pH) or too basic (high pH). U.S. authorities express the greatest concern about acidic water due to acid rain associated with sulfur dioxide air pollution, however. Again, while most readings represent a benign pH level, and essentially none are too basic, the pH distribution has a long left tail. Fish eggs rarely hatch at a pH value below 5, and sensitive species cannot survive at this level. Similar patterns appear for total suspended solids.

The most striking summary statistic from these data is that by almost any measure, water pollution has fallen dramatically in the last 40 years. Linear models which include a fixed effect for every pollution monitor, along with controls for environmental and technical factors which affect measurements, show statistically significant and meaningful declines in pollution.

All four conventional pollutants have declined. For fecal coliforms and total suspended solids, the decline was bigger before 1972, whereas for BOD and pH, the decline was larger after 1972. The overall BOD trend implies a decrease from about 3.7 mg/L to 2.2 mg/L. Relative to mean BOD levels of 2.8, this change represents a decline of over 50 percent. Fecal coliform counts have fallen by 1.2 percent annually since 1972. In the U.S., acidic waters are a common problem while basic waters are rare (Appendix Figure 1). River pH increased steadily until 1989 then became roughly constant. These trends correspond to sharp declines in sulfur dioxide air pollution – an important cause of acid rain – which slowed in the 1990s (USEPA 2008). Total suspended solids declined at 5.7 percent annually before 1972, though the trend slowed to 1.8 afterwards and was near zero in the 1990s.

Excluding temperature, other measures of water quality all show either significantly improving water quality or insignificant trends. Alkalinity was slightly decreasing before the Act but began gradually increasing after 1972. Specific conductance has more data than any other indicator, though reveals no trend. Mean water improved from having a dissolved oxygen level of 7.7 mg/L to a level of roughly 8.0. The two additional bacteriological indicators – fecal streptococcus and total coliforms – also declined at rapid and statistically significant rates after 1972. And despite the current regulatory concern over nutrients, the mean levels of two nutrients has declined at 1.7

percent per year (phosphorus) and 2.7 percent per year (ammonia), and these declines became more rapid after 1972.

These data provide a good first opportunity to calculate long-run temperature trends in U.S. rivers. Since 1972, the mean temperature of U.S. rivers has increased by 0.04°F annually. Of all water quality indicators that I examine, temperature is the only one which demonstrates significant worsening both before and after 1972. Over 25 years, the estimated trend implies a temperature increase of 1.0°F. Many EPA industrial regulations – particularly for thermal power plants – limit the extent to which plants may change the temperature of nearby water. Hence, climate-induced increases in water temperature may necessitate stricter limits on those plants, or at least may undermine the efforts of those regulations to prevent increases in water temperature.

No aspect of water pollution implies that these trends should be linear. To estimate water pollution trends without imposing linearity, I estimate models which include indicators for each year (with 1970 as reference category), along with site fixed effects and controls for relevant environmental and technical factors. The graph of year coefficients (Figure 2) gives a semi-parametric estimator of time trends. The data show that beginning in about 1970, improvements in most of these pollutants were roughly linear. The log of total suspended solids shows more abrupt declines.

3.5.3 Sensitivity Analysis

I also investigate the sensitivity of all the estimated trends to alternative methods. For each sensitivity analysis, I re-estimate the model for each parameter. Appendix Figure 3 plots point estimates, while Table 2 presents results and compares them against the original.

Trend estimates are similar when using different samples or statistical assumptions. Controlling for the monitoring protocol matters very little for these trend estimates. Although each water quality indicator uniquely identifies a parameter code, monitors may change the method of calculating water quality even while keeping the same code. I re-estimate the main statistical model, but redefine γ_i to include a fixed effect for every combination of monitor and analytical protocol. If a monitor begins listing an unknown or a new protocol, then this estimator calculates the trend only within use of a single protocol. The Legacy Database Center does not record the analytical protocol used for analysis, so I exclude it from this comparison. In the other two datasets, however, the correlation of trends estimated with and without this adjustment is above 0.99.

The three separate repositories also all imply similar trends. These correlations are less precise, but estimates from each of the three sources individually is highly correlated with the overall estimates reported in the paper. In no case can I reject at 95% confidence that any one of the repositories implies different trends than the combined estimates do.

Aggregating measures to the mean county-year or watershed-year also has little effect on estimated trends. In this sensitivity analysis, I use only county or watershed fixed effects, rather than monitor fixed effects. The correlation of these estimates with the original, monitor-level regressions is over 0.9

Few of the monitors operate for the full 30-year period. I investigate restricting the sample only to monitors with readings both from before 1972 and also after 1996. With this restricted sample, trend estimates have a correlate with the main estimates of over 0.95.

Some readings have extremely large values, so the main analysis excludes the top percentile for each pollutant. Using a more liberal rule – excluding the top 0.01 percentile of observations on each pollutant – has limited effect on trends.

Some variables which have approximately log-normal density functions report readings of zero. The log of zero is undefined, so the main analysis excludes these observations. The "log" entry of Table 2 instead uses $\log\left(10^{-9}+P_{it}\right)$. The correlation of these trends with the original estimates is 0.93.

Finally, I investigate weighting observations by the population, rather than by the number of readings per watershed-year. This alternative weighting calculates the trend of rivers near the average American, rather than calculating the trend in the mean river. The population-weighted estimates are similar to though larger than the original estimates, suggesting that if anything, the original estimates understate river pollution trends in highly populated areas.

3.6 Did the Clean Water Act Cause These Trends?

These trends show that water quality has improved by almost any measure. But many factors could have caused these changes: improving technology, declining manufacturing, Coasian bargaining with neighbors, or others. This section reports estimates of equation (3.4.2) which exploit variation in the timing and magnitude of Clean Water Act activities to assess the extent to which the Act caused these changes.

3.6.1 Grants

Estimates of how the grants affected water pollution give mixed results (Table 3). For parsimony, I focus on three pollutants which depend the most on sewage and which provide the most informative omnibus measures of water quality: dissolved oxygen, fecal coliforms, and total suspended solids. There is at least some measurement error in the grants records, since only 80 percent of the Freedom of Information Act data can be linked to a county. Hence I report both an annual specification and a long-differenced style model which uses only data from every decade.

The regressions imply that the Clean Water Act increased the oxygen content of water (Table 3). In particular, these estimates imply that \$100 million in grants increased the oxygen content of water by about 0.1 in the recipient county in every subsequent year. In total, this linear specification would suggest that the \$60 billion in grants increased dissolved oxygen by 0.1 mg/L in 600 counties. The trend in the previous section implies that dissolved oxygen increased by 0.3 in all 3,000 counties. Under the linearity assumptions of the statistical model used here, this implies that the construction grants program was responsible for about 7 percent of the improvement in dissolved oxygen between 1969 and 1998.

On the other hand, the estimates are consistent with the idea that the grants slightly increased fecal coliform concentrations. One possible explanation which echoes anecdotal stories in the literature is that investment in sewage treatment capacity attracts wastewater from nearby municipalities. While these investments decrease the pollution concentration from effluent, they could increase the total local effluent, and thereby worsen fecal coliform counts. The regression estimates imply moderate-sized through statistically insignificant increases in total suspended solids due to the grants.

3.6.2 Industrial Permits (NPDES)

I also examine the effects of industrial discharge permits (NPDES). Two types of variation identify the effects of these permits. First, these permits were first distributed between 1973 and 1976, with the exact year varying by plant. Second, some counties had many plants with effluent regulated by the permits, whereas other counties had few or no plants with effluent affected by the permits. The regressions combine both types of variation by measuring the number of total or large plants in a county which were affected by a discharge permit in a given year.

For these regressions, I analyze the same three omnibus pollution measures: dissolved oxygen, fecal coliforms, and total suspended solids. Because many industrial discharge permits regulate pH, I also measure how the permits affected the acidity of river water.

The regressions suggest that the permits increased the oxygen content of water (Table 3). Each regulated facility is estimated to increase the oxygen content of water by about 1.0 mg/L. The regressions measure no statistically significant effects of the industrial permits on fecal coliforms or total suspended solids. Finally, the regressions imply moderate effects of the grants on the acidity of water. While the estimates are only significant at the 90 percent level, they do suggest that the industrial discharge permits helped diminish water acidity.

Looking at major facilities only gives rise to a larger estimate than looking at all facilities. If interpreted literally, these benefits are extremely large: each facility increased the dissolved oxygen content of water by 1 to 2. While this impact is large, anecdotal evidence is consistent with pronounced effects of particularly polluting facilities.[11]

3.7 Discussion

3.7.1 Benefits of Surface Water Quality

The changes in water quality measured here only increase social welfare to the extent that Americans value clean water. The most natural way to assess this value is a hedonic study of how water pollution affects home values. However, few studies estimate the effects of water pollution on home values. Leggett and Bockstael (2000), using home-level data from the Chesapeake Bay, find some property value benefit of having lower fecal coliform pollution. Other papers estimate the effect of dissolved oxygen or other pollutants in a specific lake or river on nearby property values and find many different estimates, some with perverse signs.[12]

The Clean Water Act regulations analyzed here provide a plausible setting for measuring the effect of water pollution on property values. Regressions of median county-level home prices from the decennial censuses on either water pollution itself or on the Clean Water Act investments produced imprecise estimates which cannot be distinguished from a conclusion of no effect, but

[11] Pulp and paper facilities, for example, are among the most important contributors to low oxygen. In the Lower Willamette River in Oregon, for example, they were the dominant contributor to poor water quality. In 1929, the Willamette near Portland, Oregon, had an oxygen content of 0.5 mg/L. A state official explained, "Fish died. The threat of disease put a stop to safe swimming. Rafts of sunken sludge, surfacing in the heat of summer, discouraged water-skiing and took the pleasure out of boating." A conservation organization in 1967 described it as a "stinking slimy mess, a meance to public health, aesthetically offensive, and a biological cesspool." (USEPA 000b, pp. 404-45)

[12] Leggett and Bockstael (2000) review this literature.

which also cannot rule out moderate-sized benefits. While the willingness to pay for clean surface waters remains an open question, this setting lacks statistical power to answer it.

3.7.2 Air Versus Water Regulation

This paper's mixed conclusions about the Clean Water Act contrast substantially with analysis of the U.S. Clean Air Act, which caused immediate and large declines in U.S. air pollution in the 1970s, and which substantially improved human health and well-being, as measured by home prices or human health (USEPA 1997; Chay and Greenstone 2005; Sanders and Stoecker 2011). Existing estimates suggest that the Clean Air Act and Clean Water Act have required similar levels of investment. This paper reports the extremely rough figure of 1 trillion dollars in expenditure due to the Clean Water Act. The USEPA (1997) reports costs of $500 billion over the period 1970 to 1990 for the Clean Air Act. These figures clearly reflect different methodologies, but they suggest that the costs of the air and water policies have the same order of magnitude.

It is worth emphasizing two differences between air and water regulations which might contribute to this contrast. First, the Clean Air Act imposes more stringent regulation on counties with high pollution levels ("nonattainment counties"). By contrast, the Clean Water Act requires the same regulatory standard for all economic activity in the U.S., regardless of local water quality.[13] For a given level of expenditure on reducing pollution, it is plausible that concentrating this expenditure in areas with the greatest initial pollution levels could produce larger benefits.

Second, surface water is typically purified through municipal water filtration systems before people drink it. Even if factories dump effluent in a river, then, little of this effluent will be ingested. By contrast, air pollution which a factor emits is typically breathed without filters. People are exposed to surface water pollution through recreation (fishing, boating, swimming), through consumption of seafood, and in rare cases through drinking surface water which is not filtered. Hence, surface water pollution may have less direct effects on human health than air pollution.

[13] Section 303(d) of the Clean Water Act targeted river segments with high pollution levels which national regulations were insufficient to address. This section of the Act directs local regulators to design a "pollution budget" for the area and adjust local industrial discharge permits acordingly. Regulators almost completely ignored this section of the Clean Water Act until at least 2002.

3.8 Conclusion

The Clean Water Act is among the largest environmental investments in U.S. history. This paper assembles an array of data to assess whether U.S. water quality has improved since the years before the Clean Water Act, and whether the Clean Water Act has caused any improvements.

I find that by almost any measure, U.S. water quality has been improving. The only exception is thermal pollution, which is presumably worsening due to climate change. At the same time, most of these beneficial trends had similar magnitude before the Clean Water Act. Microdata on municipal wastewater treatment grants and industrial discharge permits indicate that these two main activities of the Clean Water Act improved an omnibus measure of water quality (dissolved oxygen). However, there is some evidence that the wastewater treatment grants increased local fecal coliform counts.

This paper focuses on trends in river water quality. It leaves open questions about other trends. For example, it would be useful to compare this paper's analysis of rivers against Smith and Wolloh (2012)'s conclusion that dissolved oxygen in U.S. freshwater lakes has not increased since 1975. Additionally, it would be informative to determine whether recent policies, such as the focus on particularly polluted river segments (Total Maximum Daily Load Requirements), or small-scale tradable permits for water pollution, have continued these beneficial trends.

Figure 1. Water quality monitors
Notes: Each pixel represents one monitor used in the analysis.

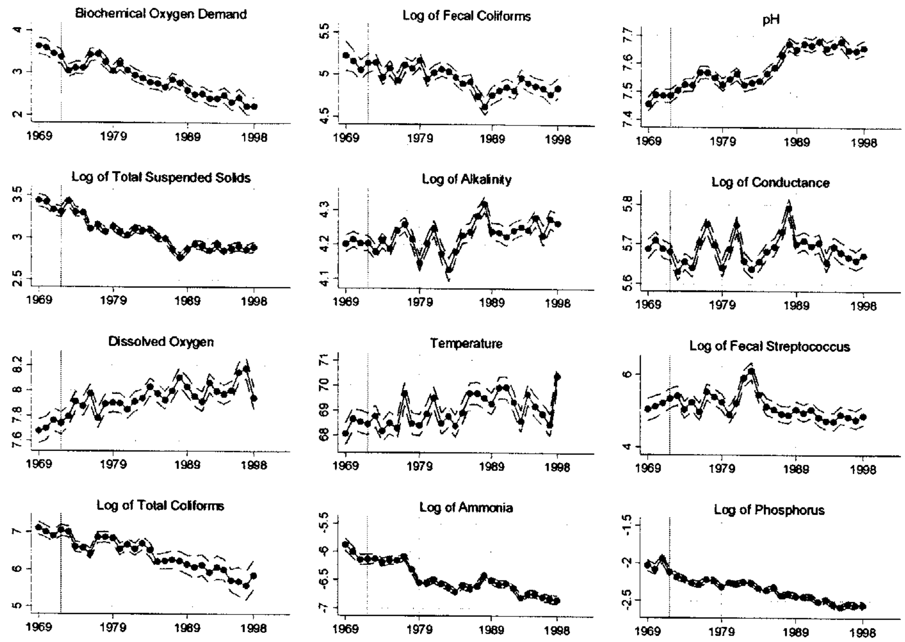


Figure 2. Nonlinear trends in water quality.

Notes: Each point depics an element of the coefficient vector plus the constant from equation (1).

Dashed lines depict the 95% confidence interval. The vertical grey line marks the 1972 Clean Water Act. Reference year is 1969, reference month is June, reference depth is 0-10 feet, and reference time is noon.

	1969-1971			1972-1988			1969-1998				
	Trend (1)	Standard Error (2)	Number of Monitors (3)	Trend (4)	Standard Error (5)	Number of Monitors 6	Chow Test p-Value (7)	Mean (8)	Median (9)	5th Pctile (10)	95th Pctile (11)
Conventional Pollutants											
Biochemical Oxygen Demand (mg/L)	-0.059	(0.037)	6,260	-0.150	(0.004)	34,517	0.224	2.77	1.80	0.50	8.00
Log of Fecal Coliforms (MPN/100mL)	-0.127	(0.037)	4,086	-0.012	(0.002)	30,562	0.046	4.63	4.61	0.69	8.56
pH	-0.009	(0.004)	12,398	0.070	(0.000)	81,549	0.074	7.57	7.70	6.20	8.50
Log of Total Suspended Solids (mg/L)	-0.057	(0.020)	3,802	-0.018	(0.001)	42,789	0.026	2.77	2.71	0.00	5.42
General Water Quality											
Log of Alkalinity as $CACO_3$ (mg/L)	-0.002	(0.005)	11,556	0.003	(0.000)	54,606	0.635	4.24	4.58	2.08	5.62
Log of Specific Conductance (µmhos/cm at 25 C)	-0.002	(0.004)	12,112	0.001	(0.000)	95,019	0.779	5.76	5.83	3.69	7.95
Dissolved Oxygen (mg/L)	0.003	(0.019)	9,045	0.010	(0.001)	59,391	0.042	8.99	9.00	4.60	13.00
Temperature (°F)	0.226	(0.054)	7,680	0.040	(0.003)	45,821	0.000	56.49	56.30	32.90	79.70
Microbiological											
Log of Fecal Streptococcus (MPN/100mL)	0.044	(0.058)	591	-0.028	(0.003)	9,428	0.040	4.91	4.79	1.10	8.88
Log of Total Coliforms (MPN/10mL)	-0.145	(0.034)	4,116	-0.041	(0.006)	14,142	0.065	6.34	6.33	1.95	10.62
Nutrients											
Log of Un-Ionized Ammonia (NH_3, mg/L)	-0.147	(0.028)	3,182	-0.027	(0.002)	35,842	0.000	-6.93	-6.83	-10.49	-3.70
Log of Phosphorus, total as P (mg/L)	0.034	(0.019)	3,994	-0.017	(0.001)	58,854	0.000	-2.44	-2.53	-4.61	-0.07

Table 1–Long-Run Trends in Water Quality

Notes: Chow test (7) reports the p-value of the null hypothesis that the 1969-1971 trend in column (1) and the 1972-1988 trend in column 94) are equal.

Method 1	Method 2	Regression Coefficient	Std. Error	R-Squared	p-value of H_0: Coeff = 1	Water Quality Indicators
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Adjust for Analytical Protocol						
Modernized, Adjusted	Modernized, Unadjusted	0.989	0.013	0.999	0.41	8
NWIS, Adjusted	NWIS, Unadjusted	1.000	0.007	0.998	0.98	11
Compare Data Sources						
Modernized STORET	Original Estimates	0.971	0.092	0.957	0.76	7
Legacy Database Center	Original Estimates	0.676	0.184	0.734	0.11	12
NWIS, Adjusted	Original Estimates	1.09	0.143	0.873	0.54	12
Geographic Averages						
County	Original Estimates	0.944	0.079	0.96	0.50	12
HUC	Original Estimates	0.924	0.077	0.962	0.35	12
Others						
Balanced Panel	Original Estimates	0.941	0.06	0.961	0.35	12
Include Outliers	Original Estimates	1.131	0.067	0.966	0.08	12
Log	Original Estimates	0.928	0.048	0.974	0.17	12
Population Weight	Original Estimates	1.659	0.272	0.789	0.04	12

Table 2–Water Quality Trends, Sensitivity Analysis

Notes: column (1) describes one way of calculating trends for each pollutant, and column (2) describes a separate way of plotting trends for each pollutant. Column (3) shows the coefficient from an OLS regression of the results from the method in column (1) on the results from method (1). Each observation in this regression is one pollutant.

Water Quality Measure	Dissolved Oxygen	Log of Fecal Coliforms	Log of Total Suspended Solids
County-Year			
Total Grants ($100 millions)	0.124***	0.029***	0.050
	(0.043)	(0.009)	(0.031)
N County-Years	59,633	43,427	41,352
Counties	2,980	2,760	2,757
County-Decade			
Total Grants ($100 millions)	0.088***	0.034*	0.056
	(0.025)	(0.014)	(0.058)
N County-Years	8,911	6,695	6,748
Counties	2,919	2,627	2,638
Temperature, depth, season contrc	yes	yes	yes
County fixed effects	yes	yes	yes
Year or decade fixed effects	yes	yes	yes

Table 3–Did the Clean Water Act Decrease Water Pollution? Construction Grants

Notes: standard errors adjusted for clusteirng within the relevant geographic unit (county or watershed). Annual data uses only years 1968-1982. Greater dissolved oxygen represents better water quality. Greater levels of fecal coliforms and other pollutants represent worse water quality.

Pollutant	Dissolved Oxygen	Log of Fecal Coliforms	pH	Log of Total Suspended Solids
County-Year				
# Facilities	1.139**	0.162	0.148*	0.024
	(0.404)	(0.128)	(0.063)	(0.033)
Major Facilities	1.898**	0.224	0.363**	0.011
	(0.686)	(0.193)	(0.115)	(0.058)
N County-Years	59,633	43,427	65,095	41,352
Counties	2,980	2,760	2,997	2,757
County-Decade				
# Facilities	1.149	0.247	0.182*	0.012
	(0.621)	(0.187)	(0.084)	(0.075)
$ Major Facilities	1.868	0.350	0.402*	0.019
	(1.084)	(0.327)	(0.168)	(0.136)
N County-Years	8,911	6,695	9,630	6,748
Counties	2,919	2,627	2,943	2,638
Temperature, depth, season controls	yes	yes	yes	yes
County fixed effects	yes	yes	yes	yes
Year or decade fixed effects	yes	yes	yes	yes

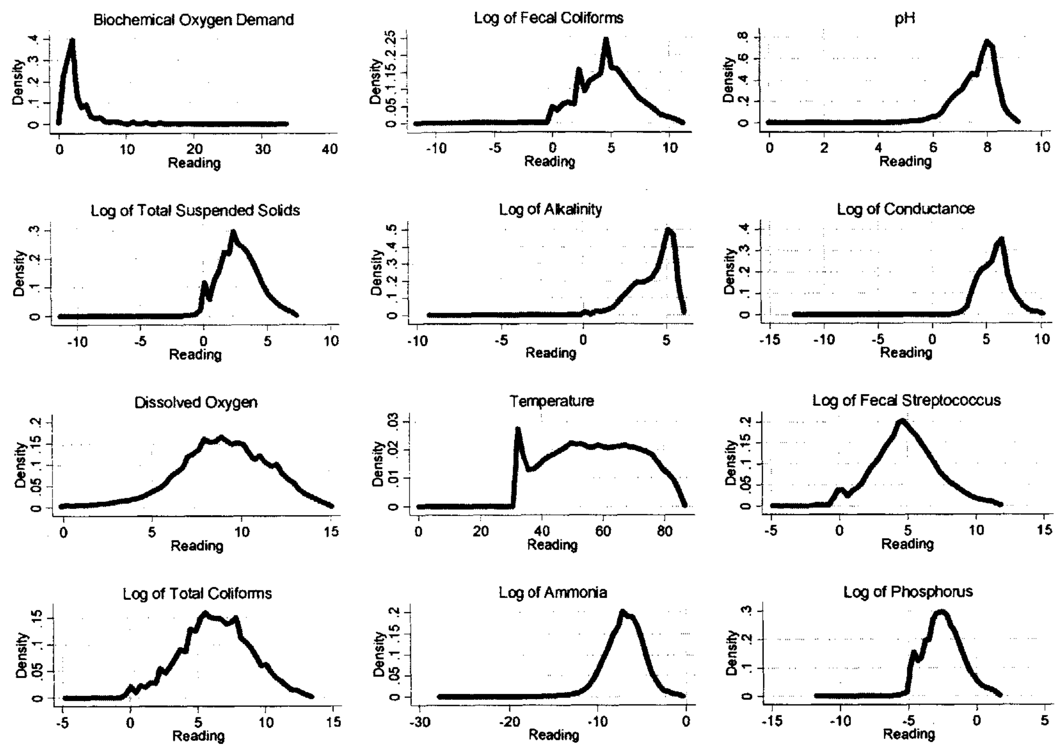
Table 4–Did the Clean Water Act Decrease Water Pollution? Effluent Permits

Notes: standard errors adjusted for clusteirng within the relevant geographic unit (county or watershed). Annual data uses only years 1968-1982. Greater dissolved oxygen represents better water quality. Greater levels of fecal coliforms and other pollutants represent worse water quality.

	Dissolved Oxygen	Fecal Coliforms	pH	Total Suspended Solids	Temperature
Dissolved Oxygen	1.00				
Fecal Coliform	-0.12*	1.00			
pH	-0.28*	0.00	1.00		
Total Suspended Solids	-0.02*	0.39*	0.20*	1.00	
Temperature	-0.29*	0.09*	-0.09*	-0.08*	1.00

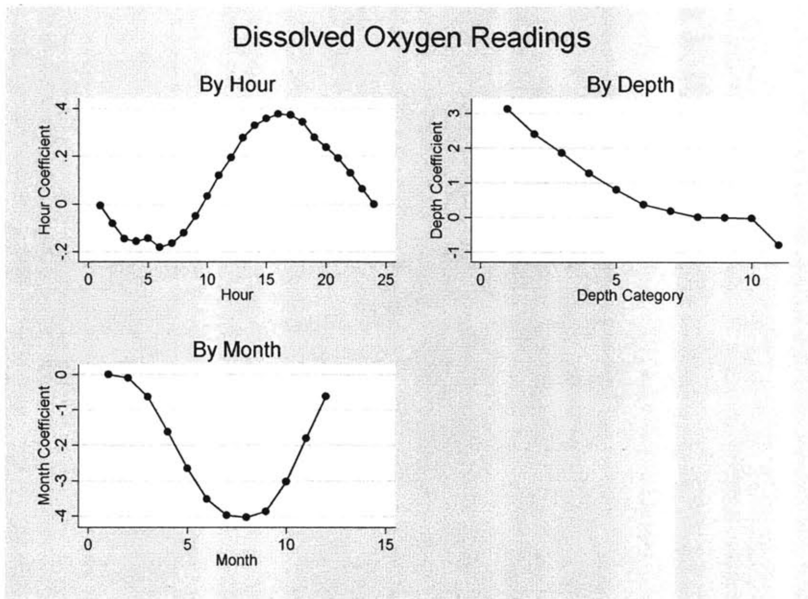
Appendix Table 1–Pairwise Correlations Between Water Characteristics

Notes: Uses river subsegment-year data. * indicates statistical significance at 99%. Greater dissolved oxygen and secchi depth represent better water quality. Greater levels of other pollutants represent worse water quality.



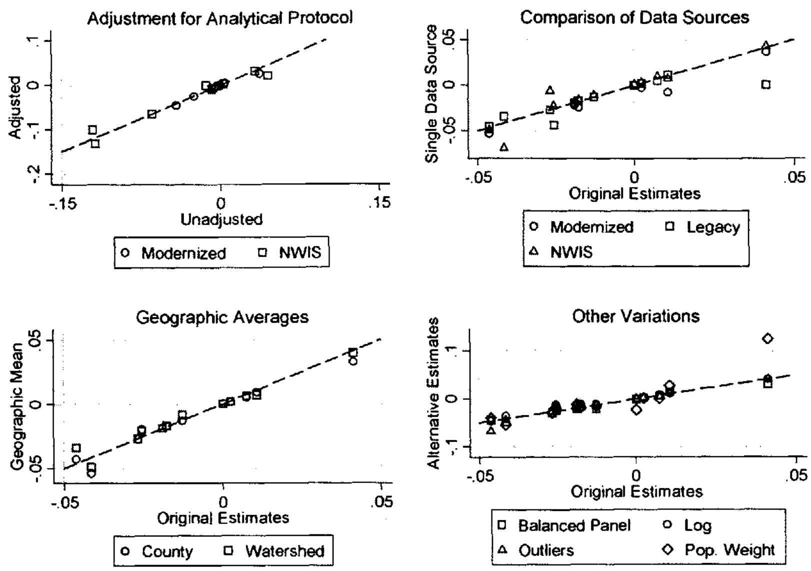
Appendix Figure 1. Density of Water Quality Readings

Notes: All graphs represent 1969-1998, weighted to represent all watersheds in the data equally. Graphs use an epinechnikov kernel evaluated at 50 points.



Appendix Figure 2. Effects of Time, Season, and Depth on Dissolved Oxygen Readings

Notes: Figure plot fixed effects from estimating equation (1), excluding the year variable Y. Regressions include monitor fixed effects.



Appendix Figure 3–Sensitivity Analysis of Trend Estimates

Notes: Each ponit compares the original estimate of the 1969-1998 linear trend for one water quality indicator against the alternative estimate for that trend. The dashes represeent the 45-degree line.

Bibliography

ADB (2010). Reducing carbon emissions from transport projects. Technical report, ADB.

Aldy, J. E. and W. A. Pizer (2011). The competitiveness impacts of climate change mitigation policies. NBER Working Paper 17705.

Alvarez, F. and R. E. Lucas (2007). General equilibrium analysis of the eaton-kortum model of international trade. *Journal of Monetary Economics 54*, 1726–1768.

Anderson, J. E. (2011). The gravity model. *Annual Reviews of Economics 3*, 133–60.

Anderson, J. E. and E. van Wincoop (2004). Trade costs. *Journal of Economic Literature 42(3)*, 691–751.

Andreen, W. L. (2003). The evolution of water pollution control in the united states - state, local, and federal efforts, 1978-1972: Part ii. *Stanford Environmental Law Journal 22*, 215–294.

Antweiler, W., B. R. Copeland, and M. S. Taylor (2001). Is free trade good for the environment? *American Economic Review 91(4)*, 877–908.

Arellano, M. and S. Bond (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *Review of Economic Studies 58(2)*, 277–297.

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

Armington, P. S. (1969). A theory of demand for products distinguished by place of production. *Staff Papers - International Monetary Fund 16(1)*, 159–178.

Ashenfelter, O. and M. Greenstone (2004). Using mandated speed limits to measure the value of a statistical life. *Journal of Political Economy 112(S1)*, S226–S267.

Ashenfelter, O. and A. Krueger (1994). Estimates of the economic return to schooling from a new sample of twins. *American Economic Review 84(5)*, 1157–1173.

Babiker, M. H. (2005). Climate change policy, market structure, and carbon leakage. *Journal of International Economics 65(2)*, 421–445.

Baier, S. L. and J. H. Bergstrand (2001). The growth of world trade: tariffs, transport costs, and income similarity. *Journal of International Economics 53(1)*, 1–27.

Baldwin, R. E. and A. Venables (1995). *Handbook of International Economics*, Chapter Regional Economic Integration. Elsevier Science Publishing.

Balistreri, E. J., R. Hillberry, and T. F. Rutherford (2011). Structural estimation and solution of international trade models with heterogeneous firms. *Journal of International Economics 83(2)*, 95–108.

Balistreri, E. J. and T. F. Rutherford (2012). Subglobal climate policy and the competitive selection of heterogeneous firms. Working Paper 2012-01, Colorado School of Mines.

Bartels, L. (2012). The inclusion of aviation in the eu ets: Wto law considerations. ICTSD Issue Paper No. 6.

Becker, G. S. (1965). A theory of the allocation of time. *Economic Journal 75(299)*, 493–517.

Bell, M. L., A. McDermott, S. L. Zeger, J. M. Samet, and F. Dominici (2004). Ozone and short-term mortality in 95 us urban communities, 1987-2000. *Journal of the American Medical Association 292(19)*, 2372–2378.

Bernard, A. B., J. Eaton, J. B. Jensen, and S. Kortum (2003). Plants and productivity in international trade. *American Economic Review 93(4)*, 1268–1290.

Bertrand, M., E. Duflo, and S. Mullainathan (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics 119(1)*, 249–75.

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

Black, D. A., M. C. Berger, and F. A. Scott (2000). Bounding parameter estimates with nonclassical measurement error. *Journal of the American Statistical Association 95(451)*, 739–748.

Black, D. A. and T. J. Kniesner (2003). On the measurement of job risk in hedonic wage models. *Journal of Risk and Uncertainty 27:3*, 205–220.

Black, D. A. and J. A. Smith (2006). Estimating the returns to college quality with multiple proxies for quality. *Journal of Labor Economics 24(3)*, 701–728.

Boden, T., G. Marland, and R. Andres (2010). Global, regional, and national fossil-fuel co_2 emissions. Technical report, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Department of Energy.

Bound, J., C. Brown, and N. Mathiowetz (2002). *Handbook of Econometrics*, Chapter Measurement Error in Survey Data, pp. 3705–3843. Springer-Verlag.

Bovenberg, A. L. and L. H. Goulder (1996). Optimal environmental taxation in the presence of other taxes: General-equilibrium analyses. *American Economic Review 86(4)*, 86(4): 985–1000.

Bowen, H. P., E. E. Leamer, and L. Sveikauskas (1987). Multicountry, multifactor tests of the factor abundance theory. *American Economic Review 77(5)*, 791–809.

Broda, C. and D. E. Weinstein (2006). Globalization and the gains from variety. *Quarterly Journal of Economics 121(2)*, 541–585.

Cai, Y., K. L. Judd, and T. S. Lontzek (2012). The social cost of abrupt climate change. Mimeograph, Stanford University.

Caliendo, L. and F. Parro (2011). Estimates of the trade and welfare effects of nafta. Mimeograph, Princeton University.

Card, D. (1996). The effect of unions on the structure of wages: A longitudinal analysis. *Econometrica 64(4)*, 957–979.

CBO (1985). Efficient investment in wastewater treatment plants. Technical report, Congressional Budget Office.

Cefic (2011). Guidelines for measuring and managing co2 emission from freight transport operations. Technical report, Cefic.

Chaney, T. (2008). Distorted gravity: The intensive and extensive margins of international trade. *American Economic Review 98(4)*, 1707–1721.

Chay, K. Y. and M. Greenstone (2003a). Air quality, infant mortality, and the clean air act of 1970. NBER Working Paper No. 10053.

Chay, K. Y. and M. Greenstone (2003b). The impact of air pollution on infant mortality: Evidence from geographic variation in pollution shocks induced by a recession. *Quarterly Journal of Economics 118(3)*, 1121–1167.

Chay, K. Y. and M. Greenstone (2005). Does air quality matter? evidence from the housing market. *Journal of Political Economy 113(2)*, 376–424.

Chen, Y., A. Ebenstein, M. Greenstone, and H. Li (2011). The long-run impact of air pollution on life expectancy: Evidence from china's huai river policy. Mimeograph, Beijing University.

Chetty, R. (2009). Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. *Annual Review of Economics 1*, 451–488.

Christensen, N. S., A. W. Wood, N. Voisin, D. P. Lettenmaier, and R. N. Palmer (2004). The effects of climate change on the hydrology and water resources of the colorado river basin. *Climatic Change 62(1-3)*, 337–363.

Coase, R. H. (1960). The problem of social cost. *Journal of Law and Economics 3*, 1–44.

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

Costinot, A. and A. Rodríguez-Clare (2012). Trade theory with numbers: Quantifying the consequences of globalization. Mimeograph, MIT.

Cristea, A. D., D. Hummels, L. Puzzello, and M. G. Avetisyan (forthcoming). The contribution of international transport to global greenhouse gas emissions. *Journal of Environmental Economics and Management.*

Currie, J., L. Davis, M. Greenstone, and R. Walker (2012). Do housing prices reflect environmental health risks? evidence from more than 1600 toxic plant openings and closings. Mimeo, MIT.

Currie, J. and M. Neidell (2005). Air pollution and infant health: What can we learn from california's recent experience? *Quarterly Journal of Economics 120(3)*, 1003–1030.

Dargay, J. M. and D. Gately (2010). World oil demand's shift toward faster growing and less price-responsive products and regions. Mimeograph, University of Leeds.

Defra (2009). 2009 guidelines to defra / decc's ghg conversion factors for company reporting. Technical report, Defra.

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

Donaldson, D. (2010). Railroads of the raj: Estimating the impact of transportation infrastructure. Mimeograph, MIT.

Donaldson, D. and R. Hornbeck (2012). Railroads and american economic growth: New data and theory. Mimeograph, MIT.

Duflo, E., M. Greenstone, R. Pande, and N. Ryan (2012). Truth-telling by third-party auditors: Evidence from a randomized field experiment in india. Working Paper, MIT.

Durbin, J. (1954). Errors in variables. *Review of the International Statistical Institute 22(1-3)*, 23–32.

Dustmann, C. and A. V. Soest (2002). Language and the earnings of immigrants. *Industrial and Labor Relations Review 55(3)*, 473–492.

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

Ebenstein, A. (2012). The consequences of industrialization: Evidence from water pollution and digestive cancers in china. *Review of Economics and Statistics 94(1)*, 186–201.

Efron, B. (1987). Better bootstrap confidence intervals. *Journal of the American Statistical Association 82(397)*, 171–185.

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

European Commission (2011). Allocation of aviation allowances in an eea-wide emissions trading system.

Faber, J. and L. Brinke (2011). The inclusion of aviation in the eu emissions trading system. Technical report, ICTSD.

Faber, J., A. Markowska, V. Eyring, I. Cionni, and E. Selstad (2010). A global maritime emissions trading system: Design and impacts on the shipping sector, countries and regions. Technical report, CE Delft.

Fanta, C. H. (2009). Asthma. *New England Journal of Medicine 360(10)*, 1002–1014.

Feenstra, R. C. (1994). New product varieties and the measurement of international prices. *American Economic Review 84(1)*, 157–177.

Feenstra, R. C., M. Obstfeld, and K. N. Russ (2010). In search of the armington elasticity. Mimeograph, UC Davis.

Feyrer, J. (2009a). Distance, trade, and income – the 1967 to 1975 closing of the suez canal as a natural experiment. Mimeograph, Dartmouth College.

Feyrer, J. (2009b). Trade and income – exploiting time series in geography. Mimeography, Dartmouth College.

Finkelstein, A. (2004). Static and dynamic effects of health policy: Evidence from the vaccine industry. *Quarterly Journal of Economics 119(2)*, 527–564.

Floyd, R. W. (1963). Algorithm 97: Shortest path. *Communications of the ACM 5(6)*, 345.

Fowlie, M., M. Reguant, and S. P. Ryan (2012). Market-based emissions regulation and industry dynamics. Unpublished Mimeography, UC Berkeley.

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.

Freeman, R. B. (1984). Longitudinal analyses of the effects of trade unions. *Journal of Labor Economics 2(1)*, 1–26.

Friedman, M. (1962). *Capitalism and Freedom*. University of Chicago Press.

Giannouli, M., Z. Samaras, M. Keller, P. deHaan, M. Kallivoda, S. Sorenson, and A. Georgakaki (2006). Development of a database system for the calculation of indicators of environmental pressure caused by transport. *Science of the Total Environment 357*, 247–270.

Greenstone, M. (2002). The impacts of environmental regulations on industrial activity: Evidence from the 1970 and 1977 clean air amendments and the census of manufactures. *Journal of Political Economy 110(6)*, 1175–1219.

Greenstone, M. and J. Gallagher (2008). Does hazardous waste matter? evidence from the housing market and the superfund program. *Quarterly Journal of Economics 123(3)*, 951–1003.

Greenstone, M. and R. Hanna (2011). Environmental regulations, air and water pollution, and infant mortality in india. NBER Working Paper 17210.

Greenstone, M., E. Kopits, and A. Wolverton (2011). Estimating the social cost of carbon for use in u.s. federal rulemakings: A summary and interpretation. NBER Working Paper 16913.

Griliches, Z. and J. A. Hausman (1986). Errors in variables in panel data. *Journal of Econometrics 31*, 93–118.

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy 80(2)*, 223–255.

Harberger, A. C. (1964). The measurement of waste. *American Economic Review 54(3)*, 58–76.

Harrigan, J. (1993). Oecd imports and trade barriers in 1983. *Journal of International Economics 35*, 91–111.

Harrigan, J. (2010). Airplanes and comparative advantage. *Journal of International Economics 82*, 181–194.

Harrington, W. (2004). *Choosing Environmental Policy: Comparing Instruments and Outcomes in the United States and Europe*, Chapter Industrial Water Pollution in the United States: Direct Regulation or Market Incentive?, pp. 67–90. Resources for the Future.

Hayward, S. F. (2011). *2011 Almanac of Environmental Trends.* Pacific Research Institute.

Head, K. and T. Mayer (2010). *The Gravity Model in International Trade: Advances and Applications*, Chapter Illusory Border Effects: Distance Mismeasurement Inflates Estimates of Home Bias in Trade. Cambridge University Press.

Head, K. and J. Ries (2001). Increasing returns versus national product differentiation as an explanation for the pattern of u.s.-canada trade. *American Economic Review 91(4)*, 858–876.

Helpman, E., M. Melitz, and Y. Rubinstein (2008). Estimating trade flows: Trading partners and trading volumes. *Quarterly Journal of Economics 123(2)*, 441–487.

Henderson, V. J. (1996). Effects of air quality regulation. *American Economic Review 86(4)*, 789–813.

Hines, N. W. (1967). Nor any drop to drink: Public regulation of water quality. part i: State pollution control programs. *Iowa Law Review 186.*

Hummels, D. (2001). Toward a geography of trade costs. Mimeograph, Purdue University.

Hummels, D. and V. Lugovskyy (2006). Are matched partner trade statistics a usable measure of transportation costs? *Review of International Economics 14(1)*, 69–86.

IATA (2009). World air transport statistics. Technical report, IATA.

ICAO (2009). Global aviation co_2 emissions projections to 2050. Technical report, ICAO.

IEA (2011). Co2 emissions from fuel combustion. Technical report, IEA.

Imbs, J. and I. Méjean (2011). Elasticity optimism. Mimeograph, Paris School of Economics.

IMO (2009). Second imo ghg study 2009. Technical report, IMO.

IPCC (1997). Ipcc guidelines for national greenhouse gas inventories. Technical report, IPCC.

IPCC (2001). *Climate Change 2001: Impacts, Adaptation, and Vulnerability.* Cambridge University Press for IPCC.

Jerrett, M., R. T. Burnett, C. A. P. III, K. Ito, G. Thurston, D. Krewski, Y. Shi, E. Calle, and M. Thun (2009). Long-term ozone exposure and mortality. *New England Journal of Medicine 360*, 1085–1095.

Kanter, J. (2011). The battle over aviation emissions. *New York Times December 18.*

Kanter, J. (2012). E.u. rebuffs china's challenge to airline emission system. *New York Times February 6.*

Keen, M., I. Perry, and J. Strand (2012). Market-based instruments for international aviation and shipping as a source of climate finance. World Bank Policy Research Working Paper 5950.

Knopman, D. S. and R. A. Smith (1993). Twenty years of the clean water act: Has u.s. water quality improved? *Environment 35(1)*, 17–41.

Krugman, P. (1980). Scale economies, product differentiation, and the pattern of trade. *American Economic Review 70(5)*, 950–959.

Lai, H. and D. Trefler (2002). The gains from trade with monopolistic competition: Specification, estimation, and mis-specification. NBER Working Paper 9169.

Leamer, E. E. and J. Levinsohn (1995). *Handbook of International Economics*, Chapter International Trade Theory: The Evidence, pp. 1339–1394. Elsevier.

Leggett, C. G. and N. E. Bockstael (2000). Evidence of the effects of water quality on residential land prices. *Journal of Environmental Economics and Management 39*, 121–144.

Limão, N. and A. J. Venables (2001). Infrastructure, geographical disadvantage, transport costs, and trade. *World Bank Economic Review 15(3)*, 451–479.

Liu, B., D. Yang, B. Ye, and S. Berezovskaya (2005). Long-term open-water season stream temerature variations and changes over lena river basin in siberia. *Global and Planetary Change 48(1-3)*, 96–111.

Lleras-Muney, A. (2010). The needs of the army: Using compulsory relocation in the military to estimate the effect of air pollutants on children's health. *Journal of Human Resources 45(3)*, 549–590.

Lux, M. (2012). Defying gravity: The substitutability of transportation in international trade. Mimeograph, Cornerstone Research.

Magat, W. A. and W. K. Viscusi (1990). Effectiveness of the epa's regulatory enforcement: The case of industrial effluent standards. *Journal of Law and Economics 33(2)*, 331–360.

Mayer, T. and S. Zignago (2005). Market access in global and regional trade. CEPP Working Paper 2.

Mazraati, M. (2011). Challenges and prospects of international marine bunker fuels demand. *OPEC Energy Review 35(1)*, 1–26.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica 71(6)*, 1695–1725.

Meltzer, J. (2012). Climate change and trade–the eu aviation directive and the wto. *Journal of International Economic Law 15(1)*, 111–156.

Moretti, E. and M. Neidell (2011). 2011. *Pollution, Health, and Avoidance Behavior: Evidence from the Ports of Los Angeles 46(1)*, 154–175.

Muller, N. Z. and R. Mendelsohn (2009). Efficient pollution regulation: Getting the prices right. *American Economic Review 99(5)*, 1714–1739.

Murphy, K. M. and R. H. Topel (2006). The value of health and longevity. *Journal of Political Economy 114(5)*, 871–904.

Neidell, M. (2009). Information, avoidance behavior, and health: The effect of ozone on asthma hospitalizations. *Journal of Human Resources 44(2)*.

NHLBI (2007). Expert panel report 3: Guidelines for the diagnosis and management of asthma. Technical report, Washington, DC: NHLBI.

Nocedal, J. and S. J. Wright (2006). *Numerical Optimization.* Springer Verlag.

Nordhaus, W. (2008). *A Question of Balance: Weighing the Options of Global Warming Policies.* Yale University Press.

Nordhaus, W. D. and J. Boyer (2000). *Warming the World: Economic Models of Global Warming.* MIT Press.

NRC (2008). *Mississippi River Water Quality and The Clean Water Act: Progress, Challenges, and Opportunities.* National Academies Press.

NTM. Basic freight calculator. Available at http://www.ntmcalc.org/index.html. Accessed August 5, 2012.

ORR (2009). National rail trends 2007-2008 yearbook. Technical report, ORR.

Ossa, R. (2011). Trade wars and trade talks with data. Mimeograph, University of Chicago.

Papke, L. E. and J. M. Wooldridge (1996). Econometric methods for fractional response variables with an application to 401(k) plan participation rates. *Journal of Applied Econometrics 11(6)*, 619–32.

PDR (2006). *Red Book: Pharmacy's Fundamental Reference.* Thomson PDR.

Point Carbon (2012, July 31). 600 million allowances may be withdrawn from eu ets. Press Release, Point Carbon.

Pope, C. A., M. Ezzati, and D. W. Dockery (2009). Fine-particulate air pollution and life expectancy in the united states. *New England Journal of Medicine 360*, 376–86.

Psaraftis, H. N. and C. A. Kontovas (2009). Co2 emission statistics for the world commercial fleet. *WMU Journal of Maritime Affairs 8(1)*, 1–25.

Rauch, J. E. (1999). Networks versus markets in international trade. *Journal of International Economics 48(1)*, 7–35.

Reece, M. (2004). Reporting of international aviation and marine bunkers in the joint iea/iece/eurostat questionnaires. Presentation, Workshop on Emissions of Greenhouse Gases from Aviation and Navigation, Copenhagen, May 2004.

Sanders, N. J. and C. F. Stoecker (2011). Where have all the young men gone? using gender ratios to measure fetal death rates. NBER Working Paper No. 17434.

Schäfer, A., J. B. Heywood, H. D. Jacoby, and I. A. Waitz (2009). *Transportation in a Climate-Constrained World.* Massachusetts Institute of Technology.

Simonovska, I. and M. E. Waugh (2011). The elasticity of trade: Estimates and evidence. NBER Working Paper 16796.

Smith, R. A., R. B. Alexander, and M. G. Wolman (1987). Water-quality trends in the nation's rivers. *Science 235(4796)*, 1607–1615.

Smith, V. K. and C. V. Wolloh (2012). Has surface water quality improved since the clean water act? NBER Working Paper 18192.

Staiger, D. and J. H. Stock (1997). Instrumental variables regression with weak instruments. *Econometrica 65(3)*, 557–586.

Stigler, G. J. (1952). *The Theory of Price.* Macmillan.

Thompson Healthcare, I. (2007). *MarketScan Research Databases User Guide and Database Dictionary. Commercial Claims and Encounters Medicare Supplemental and COB. Data Year 2006 Edition.* Ann Arbor, Michigan: Thompson Healthcare, Inc.

Time (1969). The cities: The price of optimism. *Time Magazine August 1.*

Trefler, D. (1995). The case of the missing trade and other mysteries. *American Economic Review 85(5)*, 1029–1046.

UNCTAD (2009). Review of maritime transport. Technical report, UNCTAD.

UNECA (2010). Assessing regional integration in africa iv: Enhancing intra-african trade. Technical report, UNECA.

U.S. Census Bureau (2008). *Pollution Abatement Costs and Expenditures: 2005.* U.S. Government Printing Office.

USEPA (1997). *The Benefits and Costs of the Clean Air Act, 1970 to 1990.* USEPA.

USEPA (1998). Regulatory impact analysis fo rthe nox sip call, fip, and section 126 petitions. volume 2: Health and welfare benefits. Technical report, EPA-452/R-98-003B. Washington, DC.

USEPA (2000). Liquid assets 2000: America's water resources at a turning point. Technical report, Office of Water.

USEPA (2000b). *Progress in Water Quality: An Evaluation of the National Investment in Municipal Wastewater Treatment.* USEPA.

USEPA (2005). Evaluating ozone control programs in the eastern united states; focus on the nox budget trading program, 2004. Technical report, Washington, DC: USEPA.

USEPA (2008). National air quality status and trends through 2007. Technical report, Research Triangle Park, North Carolina: USEPA.

USEPA (2009). The nox budget trading program: 2008 highlights. Technical report, Washington, DC: USEPA.

USEPA (2009a). The nox budget trading program: 2008 emission, compliance, and market data. Technical report, Washington, DC: USEPA.

Viner, J. (1950). *The Customs Union Issue.* Carnegie Endowment for International Peace.

Walker, R. (2012). The transitional costs of sectoral reallocation: Evidence from the clean air act and the workforce. Mimeograph, Columbia University.

Warshall, S. (1962). A theorem on boolean matrices. *Journal of the ACM 9(1)*, 11–12.

Webb, B. W. and D. E. Walling (1992). Long term water temperature behaviour and trends in a devon, uk, river system. *Hydrological Sciences 37*, 567–580.

Weiss, K. B. and S. D. Sullivan (2001). The health economics of asthma and rhinitis. i. assessing the economic impact. *Current Reviews of Allergy and Clinical Immunology 107(1)*, 3–8.

Weitzman, M. L. (1974). Prices vs. quantities. *Review of Economic Studies 41(4)*, 477–491.

Weitzman, M. L. (2012). Ghg targets as insurance against catastrophic climate damages. *Journal of Public Economic Theory 14(2)*, 221–244.

Yi, K.-M. (2003). Can vertical specialization explain the growth of world trade? *Journal of Political Economy 111(1)*, 52–102.