

Impacts of Trade Friction and Climate Policy on the Global Energy Trade Network

Supplementary Information

Jun Ukita Shepard, Bas J. van Ruijven, Behnam Zakeri

1 Converting trade data to physical energy flows

To link the BACI trade and IEA world energy balance (WEB) databases, I rely on a correspondence table that defines specific energy (MJ/kg) for each representative energy resource in the MESSAGE model. This specific energy value is region specific for crude (CRU) and coal (COAL), for which quality can vary by geographic source. This concordance is built using the following method:

Coal: Use the share of (lignite + sub-bituminous) compared to share of (bituminous + anthracite) in reserves, sourced from the BP Statistical Workbook (2007). Use a representative specific energy for each type of coal and find the lignite-subbituminous average and bituminous-anthracite average. Calculate the weighted average for a country (where country-specific data are available) or a region (where country-specific data are not available). Representative specific energy values are sourced from the Indiana Center for Coal Technology Research at Purdue University

Coal Type	Representative specific energy value (MJ/t)
Anthracite	30,080
Bituminous	32,000
Lignite	16,000
Sub-bituminous	21,000
Mean (Lignite + Sub-bituminous)	18,500
Mean (Anthracite + Bituminous)	31,040

Table S1. Representative specific energy for coal, by type.

Crude: crude oil reserves vary in terms of weight (light to heavy) and sulfur content (sour to sweet). To differentiate specific energy values, we focus on variation in weight by region. We apply the following formula to obtain barrels of oil per ton using the representative API gravity for each benchmark crude, which are sourced from Petroleum.co.uk:

$$\text{barrels per ton} = \frac{API + 131.5}{141.5 \times 0.159}$$

Crude Benchmark	Representative API
West Texas Intermediate (WTI)	39.6
Brent	38.06
Dubai	31
Orb (OPEC)	32.7
Minas	35
Tapis	45.2
Bonny Light	32.9
Isthmus Light	33.74

Table S2. Representative API for crude, by type.

Petroleum (PET), nuclear-uranium (NUC), bioenergy-biodiesel and peat (BIO) are not region-specific.

2 Data validation for representative countries

Figure S1a–d show the IEA-BACI data validation for representative countries (China, Germany, Japan, Brazil).

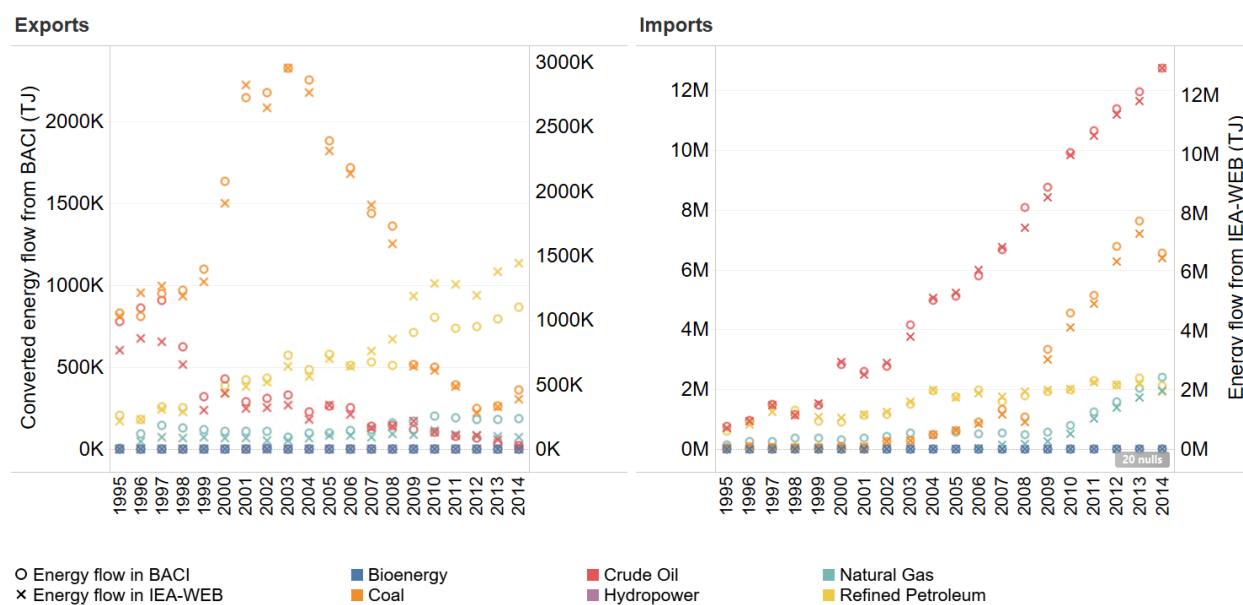


Figure S1a. Validating converted trade data (BACI) with energy data (IEA) for China.

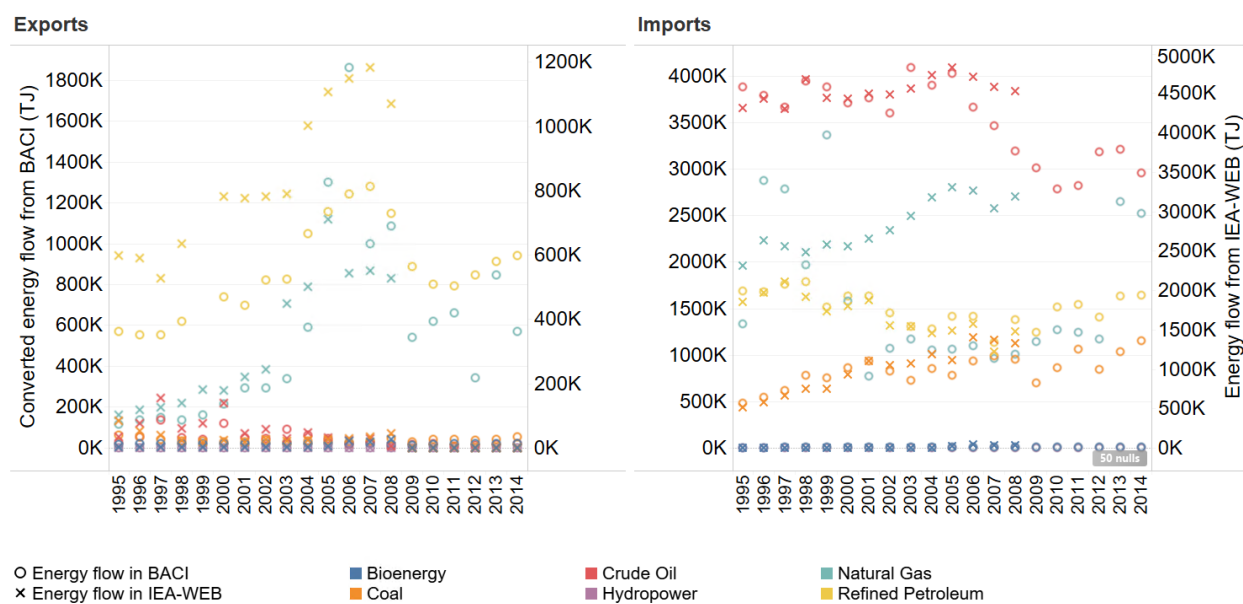


Figure S1b. Validating converted trade data (BACI) with energy data (IEA) for Germany.

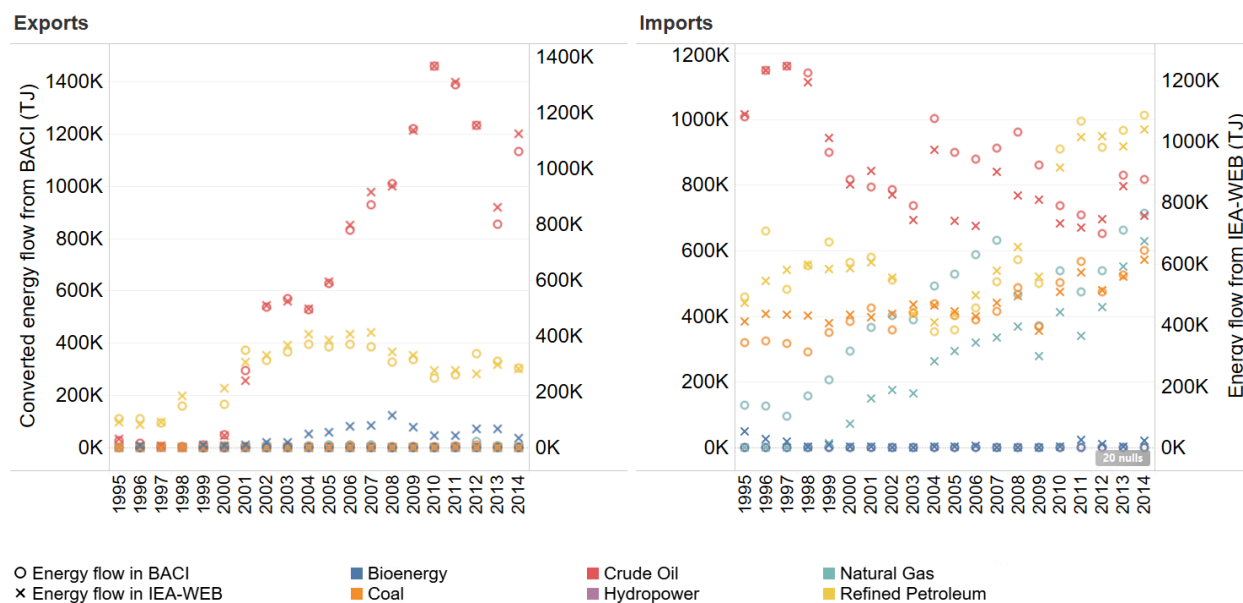


Figure S1c. Validating converted trade data (BACI) with energy data (IEA) for Japan.

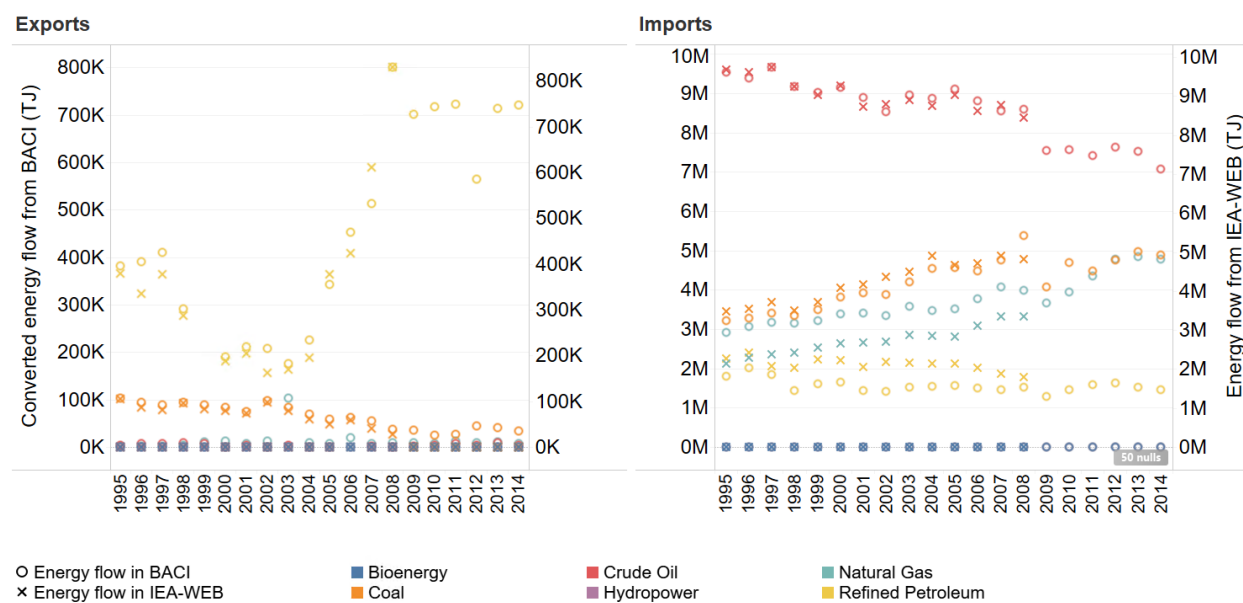


Figure S1d. Validating converted trade data (BACI) with energy data (IEA) for Brazil.

3 Country to region correspondence

The following table presents the country to region correspondence used for this study. Note that we use the most disaggregated MESSAGE global model to date, which includes 14 representative regions.

MESSAGE Region	Countries
Africa (AFR)	Angola, Benin, Burkina Faso, Burundi, Cameroon, Cabo Verde, Central African Republic, Chad, Comoros, Cote d'Ivoire, Djibouti, Equatorial Guinea, Eritrea, Ethiopia, Gabon, Gambia, Ghana, Guinea, Guinea-Bissau, Kenya, Liberia, Madagascar, Malawi, Mali, Mauritania, Mauritius, Mozambique, Niger, Nigeria, Rwanda, Sao Tome and Principe, Senegal, Seychelles, Sierra Leone, Somalia, South Africa, Sudan, Tanzania, Togo, Uganda, Zambia, Zimbabwe
Central Asian States (CAS)	Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan, Uzbekistan
Centrally Planned Asia (CPA)	Cambodia, China, Laos, Mongolia, North Korea, Vietnam, Taiwan
Eastern Europe (EEU)	Albania, Bosnia Herzegovina, Bulgaria, Croatia, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Montenegro, Poland, Romania, Serbia, Slovakia, Slovenia
Latin America (LAM)	Antigua and Barbuda, Argentina, Bahamas, Barbados, Belize, Bermuda, Bolivia, Brazil, Chile, Colombia, Costa Rica, Cuba, Dominican Republic, Ecuador, El Salvador, Falkland Islands, Guatemala, Guyana, Haiti, Honduras, Jamaica, Mexico Nicaragua, Panama, Paraguay, Peru, Suriname, Trinidad and Tobago, Uruguay, Venezuela
Middle East (MEA)	Algeria, Bahrain, Egypt, Iran, Iraq, Israel, Jordan, Kuwait, Lebanon, Libya, Morocco, Oman, Qatar, Saudi Arabia, South Sudan, Syria, Tunisia, United Arab Emirates, Yemen
North America (NAM)	Canada, United States
Pacific OECD (PAO)	Australia, Japan, New Zealand
Pacific Asia (PAS)	Brunei Darussalam, Fiji, French Polynesia, Indonesia, Malaysia, Myanmar, New Caledonia, Republic of Korea, Papua New Guinea, Philippines, Samoa, Singapore, Solomon Islands, Thailand, Timor-Leste, Vanuatu
Russia (RUS)	Russia
South Asia (SAS)	Afghanistan, Bangladesh, Bhutan, India, Maldives, Nepal, Pakistan, Sri Lanka
South Caucasus States (SCS)	Armenia, Azerbaijan, Georgia
Belarus, Moldova, Ukraine (UBM)	Belarus, Moldova, Ukraine

Table S3. Region to country correspondence in MESSAGE R14.

4 Tariff rate aggregation

Tariffs are endogenous to trade flow; while tariffs can impact the amount of imports to a country, the amount of imports can also influence whether and how much a tariff is imposed. This poses an issue when aggregating tariff rates so that they are identified by MESSAGEix-represented region. To address this endogeneity, we follow the aggregation methodology put forth in Guimbard et al. (2012) and Bouët et al. (2008) [29,30]. We first cluster countries into reference groups by GDP and trade openness. This gives us an exogenous group of countries to compare variations in tariff rates. We then assign weights to each observation in the WTO, defined as:

$$w_{jpet} = M_{jpet,REF} \left[\frac{M_{j...}}{M_{.....REF}} \right]$$

Where w_{ijpet} is the weight for a given importer j , HS6 product code p which is associated with energy commodity e , in year t ; $M_{jpet,REF}$ are imports of p to j in year t ; $M_{j...}$ is the total imports to country j , and $M_{.....REF}$ is the total imports to the reference group associated with j . Finally, we aggregate the data by calculating the importer-energy-year level mean, weighted by w_{jpet} .

5 MESSAGE parameters in global trade schema

The following table lists all parameters included in the global trade schema that required re-parameterization for the bilateral trade schema.

Parameter(s)...	Represents...	Used for...
Activity bounds (lower and upper)	How much an activity can be conducted by a region (e.g. how much a region can export oil)	Activity constraints
Capacity factor	Capacity factor of a technology (e.g. 0.75 for coal power plants)	Determining the usage of an activity
Emission factor	Per-unit emissions associated with an activity	Assigning costs associated with an emissions tax; estimating regional/global emissions
Fixed cost	Fixed costs of a technology (e.g. exports)	Cost minimization
Growth bounds (lower and upper)	How much an activity can be increased/decreased by a region	Dynamic constraints
Historical activity	Historic data on the activity of a region (e.g. historic exports)	Setting initial conditions for the model
Historical new capacity	Historic investment into new capacity of a technology (e.g. coal power plants)	Setting initial conditions for the model
Initial activity bounds (lower and upper)	How much an activity can be increased/decreased in absolute terms	Dynamic constraints
Input and output	The input required per unit output (e.g. input of coal into power plant is > 1 , while output is < 1 , due to losses)	Building relationships among technologies
Investment cost	Investment costs of a technology	Cost minimization
Levelized cost soft activity bounds (lower and upper)	Allows relaxation of cost activity bounds	Regional cost accounting
Activity relations	Relationships among parameters	Builds constraints for optimization
Soft activity bounds (lower and upper)	Allows relaxation of activity bounds that does not exceed investment	Regional capacity accounting
Technical lifetime	Lifetime of a technology (e.g. 20 years for power plant)	Determining when a technology is phased out
Variable cost	Variable costs of a technology	Cost minimization

Table S4. Parameters used for global trade schema in MESSAGE.

6 Port selection for shortest sea routes

Shortest sea routes are selected based on a user-defined list of ports. For this study, we use the “Seaports of the World by Country” list published by the Virginia Economic Development Partnership, which includes 835 of the most active sea/inland ports in the world.

We first define the nodes to be included in the shortest sea route calculations. We set the distance between nodes for uniform nodes to two degrees and overlay this map with a map of the active ports.

Figure 2 shows both uniform and port nodes:

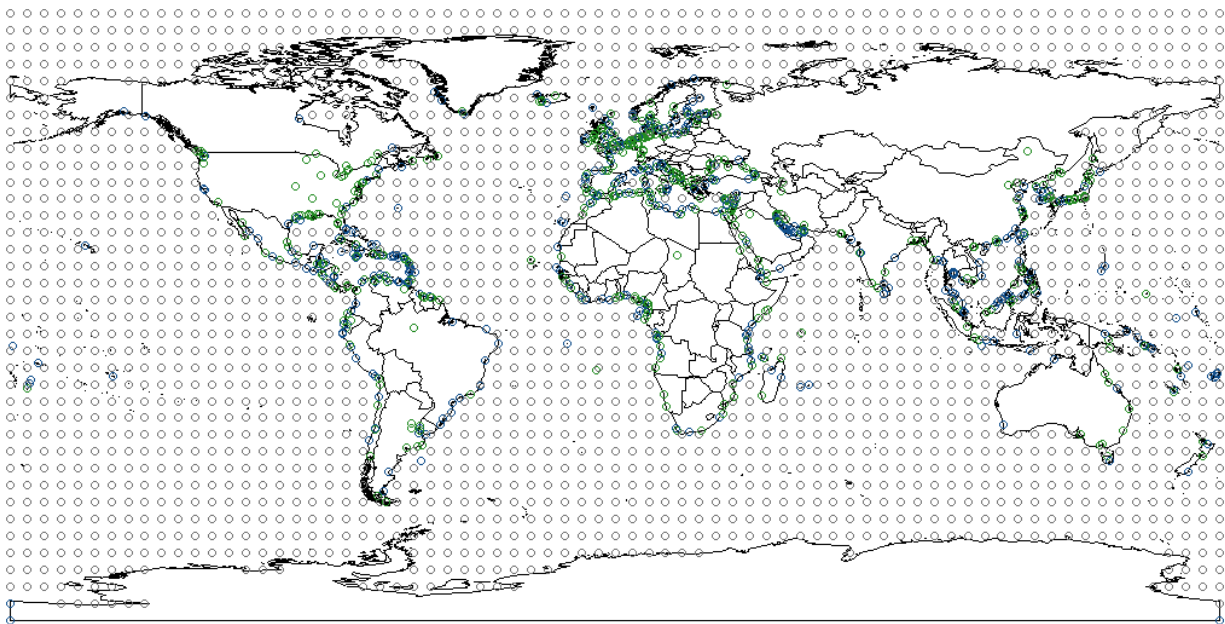


Figure S2. Map of nodes used in shortest path calculation. Uniform nodes are in grey, inland water ports are in green, and sea ports are in blue.

Based on this list of n nodes, the program then generates a dataset of all combinations of all nodes. Based on this dataset, we calculate the Haversine distance between the nodes and put these distances into an $n \times n$ matrix. We use this matrix to run the Floyd-Warshall algorithm which, for each pair of nodes, compiles a list of intermediate nodes that produce the shortest route between them. For countries with multiple ports, we take the port-port combination that allows the shortest route for each country-country combination. Finally, we link these data to bilateral trade data for each commodity k .

7 Effect of sea distance on trade cost (details)

The following section should be read with the corresponding section in the main text (Section 2.3).

7.1 Measuring sea distance

We apply the Floyd-Warshall Algorithm on a set of nodes that cover major bodies of water, combined with a set of nodes at the locations of major sea and inland water ports. The set of uniform nodes are separated by 2 degrees.¹ Additionally, there can be more than one port per country; for instance, the United States has ports in the northeast (New York), southeast (New Orleans), and the west (California). The Floyd-Warshall Algorithm finds the shortest path between two nodes by adding or subtracting intermediate points based on whether they shorten or lengthen the path. We aggregate to the 14 regions represented in MESSAGEix by selecting the primary import and export ports for each region, for each energy commodity in each year. This means that the distance between regions is time-variant; it is based on the country/port that is most crucial for each commodity in a given year. A detailed framework for port selection can be found in the Supplementary Material.

7.2 Marginal effect of sea distance on trade cost

We estimate the marginal effect of sea distance by first building a dataset of trade flows and distances between countries for 1995-2014. To this, we add the gravity terms derived from USITC data (see Data). Using this data, we run the following ordinary least squares (OLS) specification:

$$P_{ijkt} = \alpha + \beta_1 d_{ijt} + \beta_2 g_{it} + \beta_3 g_{jt} + \beta_4 c_{ij} + \beta_5 l_{ij} + \beta_{6:19} \Gamma_{R_i} + \beta_{20:33} \Theta_{R_j} + \beta_{34:53} \Psi_t + \epsilon$$

¹ A map of nodes (both uniform and sea/inland ports) can be found in the Supplementary Material.

Where:

- P_{ijkt} is the per-unit cost of energy commodity k that is exported from country i to country j in year t ;
- d_{ijt} is distance between country i and country j ;
- g_{it} and g_{jt} are GDP of countries i and j in year t ;
- c_{ij} is an indicator equal to 1 if the two countries are contiguous;
- l_{ij} is an indicator equal to 1 if countries i and j share a common language;
- Γ_{R_i} are fixed effects for the MESSAGEIX region corresponding to country i ;
- Θ_{R_j} are fixed effects for the MESSAGEIX region corresponding to country j ;
- Ψ_t are year fixed effects; and
- ϵ is the residual.

We run this specification across all energy commodities and by energy commodity. Table S2 presents the results of these six models. Note that our coefficient of interest is β_1 , or the marginal effect of each 1000km increase in bilateral distance.

	β_1	Std. Err.	t-value	Pr(> t)	Mean(Y)
All energy commodities	0.0031	0.0001	26.4550	0.0000	326.3215887
Coal	0.0000	0.0002	0.0973	0.9225	255.1141421
Crude Oil	0.0055	0.0002	22.7109	0.0000	181.8138803
Petroleum Products	0.0036	0.0002	21.1065	0.0000	395.9311973
LNG	0.0032	0.0003	10.8340	0.0000	333.1324012

Table S5. Results of gravity model-based regression. The first column displays the coefficient on distance, or the marginal effect of distance on trade cost. The second column displays the standard error on this coefficient, the third column displays the t-value, the fourth column displays the p-value, and the fourth displays the mean Y-value (\$1000/t of the given commodity).

Model results suggest that there is significant heterogeneity across energy commodity. For instance, the trade costs of LNG and fuel oil tend to be more sensitive to distance than crude oil. We

therefore use an energy-specific marginal effect to parameterize the distance-based variable cost.² For ethanol and methanol, we assume the same distance-based variable cost coefficient as petroleum products. For liquid hydrogen, we assume the same distance-based cost coefficient as LNG.

Finally, to build the variable cost parameter used in MESSAGEix, we multiply the region-to-region sea distances by the marginal effect:

$$varcost_{R_i R_j k t} = \beta_{1k} \times d_{R_i R_j k t}$$

Note that variable cost is identified by exporting region (R_i), importing region (R_j), energy commodity k , and year t . This is because the sea distance we calculated is differentiated by region pairs and energy commodity. Figure 4 below presents the distribution of distance-based variable costs.

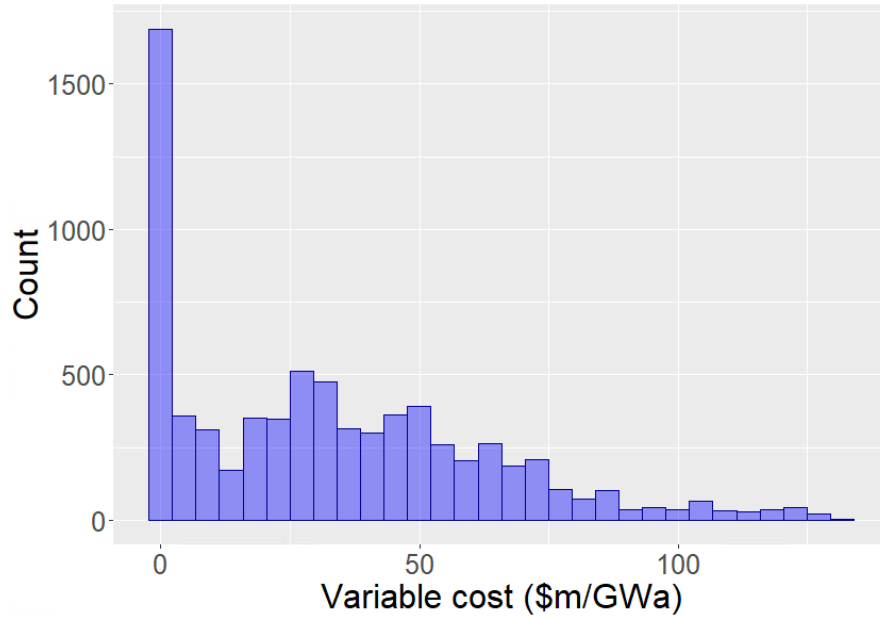


Figure S3. Distribution of estimated variable costs based on regression analysis. This is the distance-based variable cost, so the product of β_1 and bilateral distance, where β_1 is differentiated by energy commodity. Variable costs will therefore be uniquely identified by region, energy commodity, and year.

² We also run a specification by exporting region but find that there is too much uncertainty in our results to use a region-specific effect. Results of the region-specific effect can be found in the Supplementary Material.

8 Shipping constraints, assumptions, and costs

We add the following constraints in MESSAGEix to represent global shipping capacities:

$$\begin{aligned} \sum_{k^{liquid}} (liquid_shipping_{Rift}) - (hv_{Rift} \times ACT_{R_i R_j k^{liquid}}) &\geq 0 \forall t \\ \sum_{k^{solid}} (solid_shipping_{Rift}) - (hv_{Rift} \times ACT_{R_i R_j k^{solid}}) &\geq 0 \forall t \\ \sum_{k^{LNG}} (LNG_shipping_{Rift}) - (hv_{Rift} \times ACT_{R_i R_j k^{LNG}}) &\geq 0 \forall t \\ \sum_{k^{H_2}} (H_2_shipping_{Rift}) - (hv_{Rift} \times ACT_{R_i R_j k^{H_2}}) &\geq 0 \forall t \end{aligned}$$

These equations dictate that the capacity of shipping (e.g. *liquid_shipping*) summed across all exporting regions (i.e. global capacity) must be greater than the global trade activity for a given type of commodity k^* . Here we differentiate shipping capacity by what they carry (k^*) and the fuel they use (f). In this study, k^{liquid} includes crude oil, light oil, and fuel oil; k^{solid} includes coal; and k^{LNG} includes LNG. f can be fuel oil (i.e. diesel), LNG, or electricity. Heavy crude has historically been the primary fuel for maritime shipping. Interest has increased more recently in LNG-fueled vessels [27,28]. Electricity-powered vessels are still not in development, but we include this as an option (albeit an expensive one) for the latter period of the model horizon (post-2030). We include hydrogen-based transport as another possible but still prohibitively expensive option. hv_{kt} is the conversion factor for a given energy commodity k from the exporting region (R_i) in time t ; we derive these values from the IEA NCV dataset. This allows us to have consistent units.

The following tables include costs, parameters, and assumptions used to build the constraints related to global shipping capacity in the re-parameterized MESSAGE model. More information on the shipping constraints can be found in the Methods section of the main text.

Type of shipping	Type of fuel	Fuel input (kg/bton-km-y)		
		By weight (kg/bton-km-y)	By energy (GWa/bton-km-y)	Emissions factor (Mt/bton-km-y)
Liquid	Diesel	6894865.04	0.01	0.02
Liquid	LNG	3678751.33	0.00	0.01
Liquid	Electricity	NA	0.00	0.00
Solid	Diesel	7283748.84	0.01	0.02
Solid	LNG	3052216.32	0.00	0.01
Solid	Electricity	NA	0.00	0.00
LNG	Diesel	16810656.90	0.02	0.05
LNG	LNG	5594115.78	0.01	0.01
LNG	Electricity	NA	0.00	0.00

Table S6. Fuel input required by type of shipping. Based on author calculations using data from Johansson et al. (2017), International Energy Agency, International Maritime Organization, and DNV-GL [1–5].

Type of shipping	Type of fuel	Capital cost		
		Overnight cost (\$M)	Annualized, 10%DR (\$M/y)	Annualized (\$M/bton-km-y)
Liquid	Diesel	76.00	8.37	6.24
Liquid	LNG	96.00	10.58	7.88
Liquid	Electricity	192000.00	21152.27	15758.19
Solid	Diesel	51.60	5.68	4.24
Solid	LNG	71.60	7.89	5.88
Solid	Electricity	1432000.00	157760.68	117529.83
LNG	Diesel	180.00	19.83	14.77
LNG	LNG	200.00	22.03	16.41
LNG	Electricity	4000000.00	440672.29	328295.62

Table S7. Capital costs assumptions by type of shipping and fuel. Based on author calculations using data from Johansson et al. (2017), International Energy Agency, International Maritime Organization, and DNV-GL [1–5].

Type of shipping	Type of fuel	Assumptions				
		DWT (ton)	Annual distance (10 ⁹ km/y)	Number of ships	Mean payload per ship (bton-km/y)	Mean distance traveled (km/y)
Liquid	Diesel	100000	5.05	376219	1.34	13423.03
Liquid	LNG	100000	5.05	376219	1.34	13423.03
Liquid	Electricity	100000	5.05	376219	1.34	13423.03
Solid	Diesel	100000	5.05	376219	1.34	13423.03
Solid	LNG	100000	5.05	376219	1.34	13423.03
Solid	Electricity	100000	5.05	376219	1.34	13423.03
LNG	Diesel	100000	5.05	376219	1.34	13423.03
LNG	LNG	100000	5.05	376219	1.34	13423.03
LNG	Electricity	100000	5.05	376219	1.34	13423.03

Table S8. Distance and payload assumptions by type of shipping and fuel. Based on author calculations using data from Johannson et al. (2017), International Energy Agency, International Maritime Organization, and DNV-GL [1–5].

Type of shipping	Type of fuel	NCV	Annualized (\$M/GWa)
Liquid	Diesel	0.0007	5.86094E-08
Liquid	LNG	0.0007	7.40329E-08
Liquid	Electricity	0.0007	0.000148066
Solid	Diesel	0.0014	7.95854E-08
Solid	LNG	0.0014	1.10432E-07
Solid	Electricity	0.0014	0.00220865
LNG	Diesel	0.0007	1.38812E-07
LNG	LNG	0.0007	1.54235E-07
LNG	Electricity	0.0007	0.003084706

Table S9. Author-calculated costs converted into per GWa units. GWa are the energy units used in the MESSAGE model. Based data from Johannson et al. (2017), International Energy Agency, International Maritime Organization, and DNV-GL [1–5].

9 HS6 and IEA to MESSAGE energy commodities

The following table lists the correspondences of HS6 product code and IEA fuels to MESSAGE-represented energy commodity. Note that the IEA values are only used for data validation.

MESSAGE Energy Commodity	HS6 Code	IEA Fuel
Coal	270111, 270112, 270119, 270120, 270210, 270220, 270400	"Coal and coal products"
Crude oil	270900	"Crude, NGL, and feedstocks"; "Oil shale and oil sands"
Fuel oil	27071, 270720, 270730, 270740, 270750, 270760, 270791, 270799, 271000, 271012, 271019, 271020, 271311, 271312, 271320, 271390	"Oil products"
Light oil	N/A	N/A
LNG	270500, 271111, 271112, 271113, 271119, 271121, 271129	"Natural gas"

Table S10. Correspondence between Harmonized System (HS) 6-digit codes with energy commodities represented in the MESSAGE global energy model. Note that light oil is not represented in bilateral trade data, so we assume it follows a similar pattern with fuel oil.

10 Regression results for distance-based variable cost

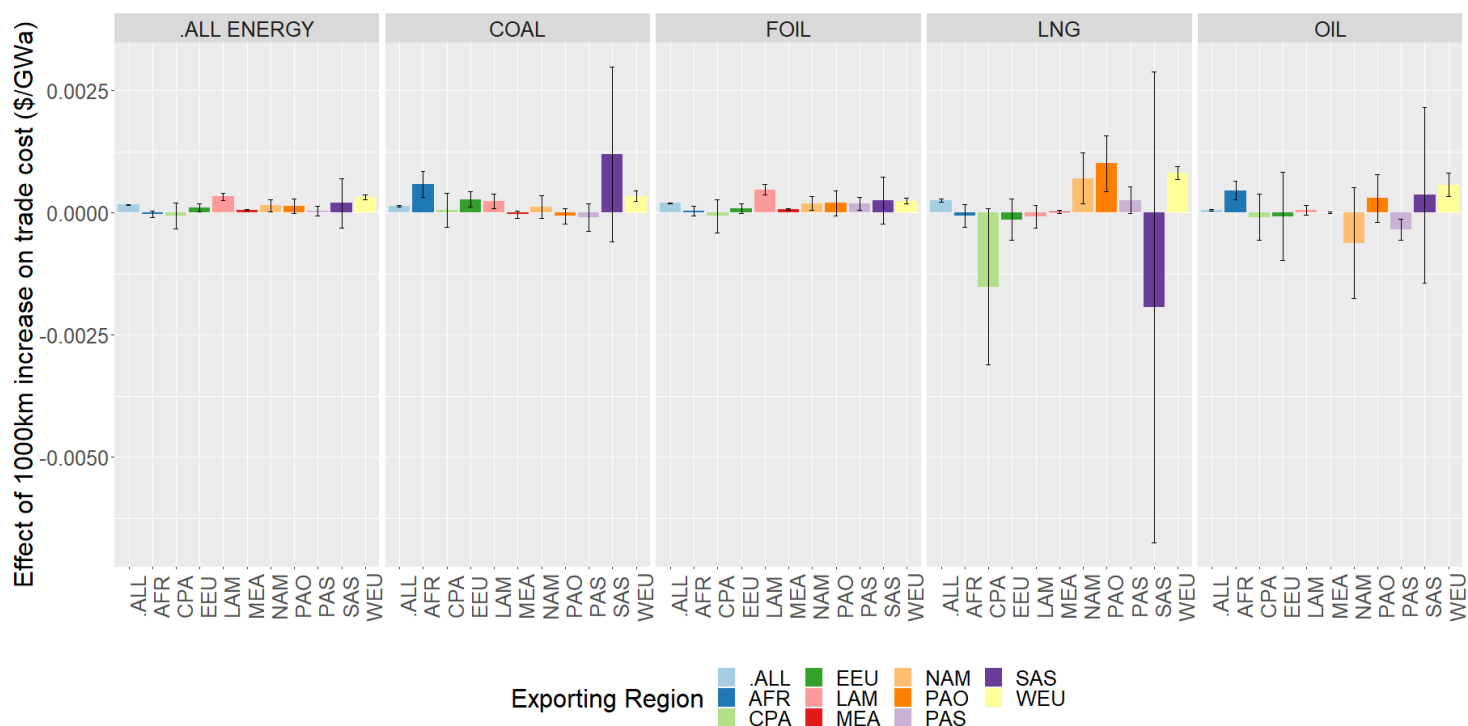


Figure S4. Results of a gravity model-based regression of trade cost (\$/GWh) on distance. The regression is run across region ("ALL"), and then by region. Different colors represent different regions. The regression is also run across all energy commodities (first column), and by energy commodity. FOIL represents fuel oil, OIL represents crude oil. Errors bars represent the 95% confidence interval on the results. The bars presented here are equivalent to β_1 of the regression specification in the main text.

11 Table: Policy effects on the energy trade network

The table below presents the effects of the six examined scenarios discussed in the main text. This table corresponds to the discussion in Section 3.1. This table outlines key network metrics for 2050. In this network, nodes represent regions and edges represent directional trade flows.

	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline	High tariffs	Low tariffs	Emissions tax + baseline tariffs	Emissions tax + high tariffs	Emissions tax + low tariffs
Global energy system cost	2528475.28	2563606.78	2517104.10	128160362.83	128201829.80	128149153.81
Total number of exporters	11	11	11	11	11	11
Total number of importers	12	12	12	12	12	12
Total number of links	111	106	113	183	185	180
Total amount of trade (EJ)	178.15	158.12	187.6	96.36	88.42	98.04
Mean trade flow (EJ)	1.6	1.49	1.66	0.53	0.48	0.54
	3.99	3.9	4.11	2.14	1.85	2.08
Minimum trade flow (EJ)	0	0	0	0	0	0
	PAO to UBM	PAO to UBM	PAO to UBM	UBM to PAO	UBM to PAO	PAO to EEU
Maximum trade flow (EJ)	20.7	20.39	22.98	24.54	18.54	23.57
	MEA to SAS	MEA to PAS	MEA to SAS	LAM to SAS	LAM to SAS	LAM to SAS
Mean # of edges per node	18.5	17.67	18.83	30.5	30.83	30
	4.7	4.83	4.99	8.48	8.61	8.81
Outdegree centrality	0.21	0.2	0.21	0.3	0.31	0.3
	0.1	0.08	0.1	0.1	0.1	0.1
Indegree centrality	0.16	0.16	0.17	0.26	0.26	0.25
	0.05	0.05	0.06	0.08	0.08	0.08
Size of energy-specific trade						
Coal	19.62	18.49	20.13	4.1	4.1	4.1
Ethanol	0.02	0.42	0.01	1.65	1.68	1.63
Fuel oil	3.97	4.07	4.02	2.15	1.76	2.32
Light oil	0.03	0.01	0.08	40.54	34.88	40.21
LNG	56.7	48.33	62.48	8.16	8.16	8.16
Crude oil	97.81	86.79	100.87	39.75	37.83	41.61

Table S11. Network metrics for the global fuel trade network in 2050.