Trade From Space: Shipping Networks and The Global Implications of Local Shocks*

Inga Heiland† Andreas Moxnes‡
Karen Helene Ulltveit-Moe§ Yuan Zi¶

Preliminary and Incomplete, October 2019

Abstract

Container ships are the engines of global trade. In this paper, we use satellite data revealing the worldwide movements of all container ships to calculate the optimal travel routes of containerized goods. Few countries have direct shipping connections to their trade partners. An implication is that trade between two countries is determined by trade costs to and from third countries. We estimate the impact of local shocks to the shipping network on global trade by means of a natural experiment: the Panama Canal expansion in 2016. Based on a difference-in-difference estimation, we find that the trade between country pairs whose fastest shipping routes pass through the Panama Canal

*We would like to thank seminar participants at the University of Oslo, conference participants at NOITS in Copenhagen 2019, the CESifo Global Economy conference in Munich 2019, the ETSG 2019, and the Midwest Trade Meetings Fall 2019.for valuable comments and discussion. We thank Clarksons Platou for providing us with access to the Clarkson World Fleet Register. This project has received funding from the European Research Council under the European Union's Horizon 2020 research and innovation program (grant agreement 715147).

†University of Oslo & CESifo, inga.heiland@econ.uio.no
‡Princeton University, University of Oslo & CEPR, andreas.moxnes@econ.uio.no
§University of Oslo & CEPR, k.h.ulltveit-moe@econ.uio.no
¶University of Oslo, yuan.zi@econ.uio.no

increased by 9-10% after the expansion. We quantify the trade and welfare effects of the shock using a canonical Ricardian model of trade. The expansion increased world real income by USD 20 billion, and the gains per capita were concentrated among Central American countries which use the canal intensively. The results highlight the importance of trade networks for the quantification of the gains from trade.

1 Introduction

Container ships are the engines of global trade. Levinson (2010) and Bernhofen et al. (2016) detail the seismic changes that the worldwide adoption of container shipping technology has brought about in international trade. As documented by Rua (2014), by now nearly all countries have container ports, constituting the nodes of the global container shipping network. However, little is known about the structure of the shipping network, e.g. which route a container might travel from the dock of port i to port j. The structure of the shipping network is an essential determinant of the costs of trade. The networked environment also implies that a shock to a port, or a link, in the network, such as improvements in shipping infrastructure, can affect shipping costs and trade flows for many countries. In this paper, we use satellite data on the movement of container ships to establish novel evidence on routes that form the global shipping network. This in turn allows us to analyze how local shocks to the shipping network affect trade costs, global trade flows and real incomes.

Our contribution is threefold. First, we document salient features of the container shipping network based on a dataset covering the worldwide movements of all container ships in 2016. Second, we demonstrate how information about shipping routes can be used to investigate the impact of a local shock on global trade. We do so by using the Panama Canal expansion in 2016 as natural experiment, together with route information inferred from the satellite data, to provide reduced-form estimates of the Panama Canal expansion on global trade. Third, we quantify the trade and welfare effects of the shock using a canonical Ricardian model of trade along with the route information and reduced-form estimates.

Our empirical analysis of global container ship movements has become possible due to the rapid advent of the global Automated Identification System (AIS) over the last years. AIS reporting of vessel positions offers a degree of automation in data processing and aggregation that was not previously possible. Vessels send out AIS signals identifying themselves to other vessels or coastal authorities, and the International Maritime Organization (IMO) requires all international voyaging vessels with above 300 Gross Tonnage and all passenger vessels to be equipped with an AIS transmitter. This implies that all container ships carrying any significant amount of cargo are parts of our data universe. AIS messages include information regarding vessel identity, physical appearance, voyage-related information such as draught and destination. Simply put, AIS data offers real-time information on the whereabouts of all ocean-going vessels.

Using an exhaustive dataset of all port calls made by container ships in 2016, we document novel facts about the container shipping network. First, container ships typically operate on fixed routes, i.e. they serve a stable set of ports, akin to buses serving a fixed number of stops in a city. Second, shipping activity is highly concentrated across ports, with some nodes (ports) in the network handling almost two orders of magnitude more ships than the median port. Third, the network is very sparse in the sense that only few countries have direct shipping routes to their trade partners. Less than 6% of all 22,650 pairs of countries with container ports are directly connected.

While the AIS data provides unprecedented detail about the movement of ships, one cannot observe the movement of the cargo itself, i.e. the actual route of a shipment from country i to country j. To make progress, we use the observed shipping network along with actual travel times between all direct port-pair links and apply standard graph theory to calculate the fastest route between any potential port pair. Consider, for example, a shipping network with direct links between New York-London, New York-Hamburg, London-Oslo and Hamburg-Oslo. The fastest route between New York and Oslo might then be New York-London-Oslo if this route minimizes the sum of travel times of each leg of the journey. Of course, the actual route chosen might be determined by other factors than speed, such as port costs. As such, the calculated fastest route is an approximation to the actual unobserved route. The fastest path calculations reveal that 52% of all country-to-country connections involve stops in more than two other countries in between.

Besides adding to the distance traveled by a container, indirect routes expose bilateral flows to the shipping infrastructure of other countries. To demonstrate the importance of exposure to third-country infrastructure, we analyze the global trade effects of a large improvement in local shipping infrastructure in 2016: the expansion of the Panama Canal. After 10 years of construction, the extended Panama Canal opened on June 26th of 2016. The expansion nearly doubled the capacity of the canal by adding a third lane and by extending the width and depth of the existing two. We present direct evidence from the AIS data showing that shipping activity on the canal indeed increased significantly after the expansion date. Moreover, we employ our information on shipping routes to explore how exporters and importers worldwide were differentially affected by this local change in the shipping infrastructure. Using a difference-in-difference approach, we find that country pairs whose fastest connection passed through the Panama Canal prior to the expansion, traded 9-10% more after the expansion compared to other country pairs.

Finally, we use a canonical model of trade to quantify the general equilibrium effect of the Panama Canal expansion. The model is calibrated by using information on the changes in trade costs according to the reduced form estimates along with the fastest route information described above. The expansion increased world real income by 0.02% or USD 20 billion, and the gains per capita were concentrated among Central American countries which use the canal intensively.

Our paper is closely related to the growing number of studies using satellite data for economic analysis. Donaldson and Storeygard (2016) provide an overview of applications which so far has focused on environmental, development and spatial issues. This paper explores how shipping satellite data can be used within the field of international trade. There are only a few recent papers that have used shipping satellite data to explore issues related to trade. Brancaccio et al. (2017) study the role of the transportation sector in world trade focusing on search frictions and the endogeneity of trade costs. They use AIS data for dry bulk ships, which typically carry commodities such as iron ore, coal, grain and sugar. Our focus is instead on container ships, which typically carry manufactured goods and account for around two-thirds of world trade based on values.

Our paper also aims to contribute to literature on the effects of containerization. Besides having spurred global trade as documented by Bernhofen et al. (2016), new port technology has been shown to have significantly altered countries' economic geography (Brooks et al. (2018) and Ducruet et al. (2019)). Finally, this paper is related to the literature that studies the impact of canal openings or closings. Maurer and Rauch (2019) analyze how the Panama Canal changed U.S. population patterns, whereas Feyrer (2009) studies the relationship between trade and the closing and opening of the Suez Canal.

The rest of the paper is structured as follow. Section 2 documents the satellite data and the construction of the shipping network and presents salient features of the network. In Section 3 analyzes the global impact of the Panama Canal expansion on trade, while Section 4 presents the model and general equilibrium effects of the canal expansion. Section 5 concludes.

2 Data and Descriptives

2.1 Data

AIS data. We build a comprehensive and global data set for the container shipping network based on satellite data for ships. Our data set is based on AIS (Automatic identification System) data provided by Marine Traffic. AIS is an automatic tracking system used on ships and by vessel traffic services (VTS). AIS is intended to assist a vessel's watch-standing officers and allow maritime authorities to track and monitor vessel movements. AIS information supplements marine radar. The International Maritime Organization's International Convention for the Safety of Life at Sea requires AIS to be fitted aboard international voyaging ships with 300 or more gross tonnage (GT), and all passenger ships regardless of size. The coverage of AIS globally has increased rapidly over the last decade, and does now allow for a global coverage of all vessels and ports of significance.

Our data set is based on port calls. Every time a ship arrives or departs a port a signal is sent. This is referred to as a port call. We use data on all port calls made tracked by the AIS satellite system during the calendar year 2016. Every observation of a port call has a time stamp, which tells us when and where the call was made and by which ship. The observation contains information on the name of the port, country and geographical location (latitude and longitude). In addition, we get information on whether the ship is arriving or departing, whether it is in transit or not, as well as the draught of the ship at the time the port call is made. We merge the AIS data to the World Fleet register data base constructed by Clarkson, which has vessel specific information on a range of ship characteristics (see Appendix A).

Our point of departure is containerized trade.[1] Containerized seaborne trade captures the majority of merchandise world trade. Seaborne trade volumes accounted for over 80% of world merchandise trade in 2015 (UNCTAD, 2016).[2] Appendix Section A describes how we clean and process the port calls data. The final dataset includes 4,908 container ships and 515 ports for the year 2016.

Other data sets. The analysis in Section 3 requires data on trade flows, which we obtain from COMTRADE.[3] We aggregate monthly bilateral trade data to the quarterly level for the years 2015-2017. The analysis also requires variables such as distance and contiguity, which we obtain from the gravity database of CEPII. Data on free trade agreements come from the WTO's RTA databases. The analysis in Section 4 requires additional information about domestic absorption along with a few other variables, which we obtain from the Eora MRIO global supply chain database; we gather data for 189 countries for the 2015 cross-section.

[1]Based on the ship categories used by data supplier, Marine Traffic, we use the ships categorized as "container ship" and "Cargo/containership".

[2]Global seaborne container trade accounted for approximately 60 percent of the value of all seaborne trade in 2016 (Rajkovic et al., 2014).

[3]Downloaded from https://comtrade.un.org/api/get/bulk/C/M/201801/ALL/HS on March 15th, 2019.

Table 1: Ships and Ports

Variable:	Obs	Median	Mean	Sd	Min	Max
Ships:						
# ports passed	4,908	64	68	40	1	312
# distinct ports passed	4,908	11	12	7	1	46
Ports:						
# incoming ships	515	203	647	1,451	5	14,473
# outgoing ships	515	199	647	1,447	5	14,407
Port pairs:						
# ships	4160	38	80	168	5	2775
deadweight tonnes (in mio)	4160	1	4	9	.1	210

Note: Summary statistics are based on the port calls made by container ships in 2016. Only ships with non-zero duration are used. Summary statistics include only routes taken by at least 5 ships and only routes between ports that appear both as arrival and departure ports.

2.2 Stylized Facts: The Shipping Network

This section documents three key facts about the shipping network that will guide the subsequent analysis.

Fact 1: Container ships typically operate on fixed routes. Table 1 provides descriptive statistics on the numbers of ports passed per ship as well the number of ships that arrive and depart per port. A key feature of container ships is that they typically visit the same port many times. The table shows that the average number of distinct ports passed per ship is roughly one sixth of the total number of ports passed per ship (12 versus 68).

Fact 2: Shipping activity is highly concentrated in space. A few ports act as major hubs in the shipping network. While the median port only serves around 200 ships per year, the top ports serve close to 15,000 ships per year. The same pattern is observed at the port-pair level, i.e. there are a few links in the network that account for a large share of total shipping activity.

Fact 3: Only 6 percent of all country pairs have a direct shipping connec-

Table 2: Port Networks

Variable:	Obs	Mean	Sd	Min	Max
All ports:					
Indegree	515	8.08	10.26	1	84
Outdegree	515	8.08	9.84	1	82
Top 10 ports:					
Indegree	10	54.10	12.03	42	84
Outdegree	10	50.10	13.88	37	82

Note: Summary statistics are based on the port calls made by container ships in 2016. Only ships with deadweight tonnes>15,800 and trips with non-zero duration are used. Summary statistics include only routes taken by at least 5 ships and only routes between ports that appear both as arrival and departure ports.

tion. We calculate the in-degree as the number of ports to which a port is directly connected based on incoming ships, and the out-degree as the number of ports to which a port is directly connected based on outgoing ships. Table 2 shows that on average most ports are connected to rather few other ports. However, there is great variation between ports in how well connected they are. Nevertheless, even the best connected ports are only directly connected to around one sixth of the total number of ports. The 515 ports in our data are allocated across 151 countries. Only 6 percent of all country pairs have a direct shipping connection. Trade between these countries accounts for only 54 percent of world trade. Therefore, a large share of global trade does not travel on direct routes, but on routes with multiple hops.

2.3 Calculation of Fastest Routes

While the AIS data provides unprecedented detail about the movement of ships, one cannot observe the movement of the cargo itself, i.e. the actual route of a shipment from country i to country j. This section documents our methodology to calculate routes and shows descriptive statistics on those routes. The route information will be an important part of the methodology in Section 3.

We calculate the routes as follows. First, based on the time stamps provided in our port-to-port data set, we observe *direct travel time* between two ports. The direct travel time between two ports is computed as the median over all ships' trip durations. Note that these travel times are *directed*, i.e. the travel time from D to A need not equal the travel time from A to D.[4] Second, for all port pairs, including the ones that are not directly connected, we calculate the *indirect minimum travel time* over all possible multihop routes. Using the port-to-port data and Dijkstra's algorithm, we compute the minimum travel time as the shortest path in a weighted directed network where edges reflect direct connections and weights are the direct travel times described just above.

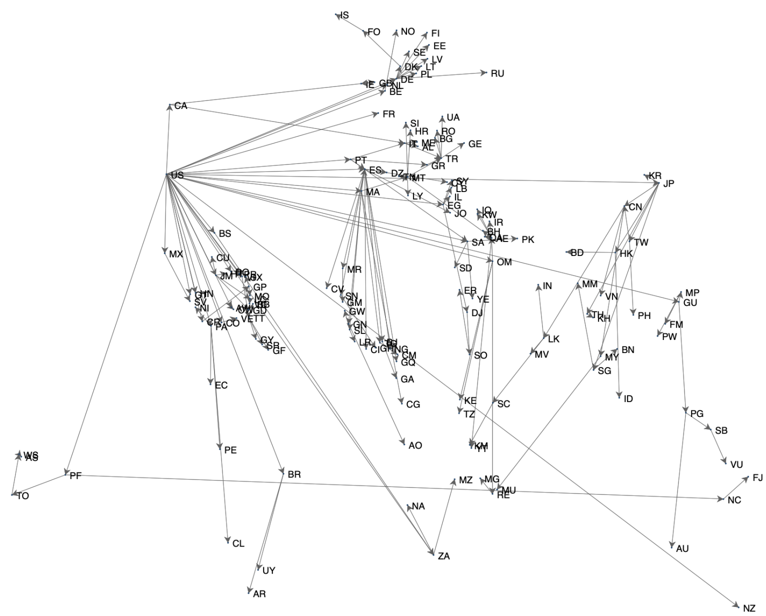
Figure 1 shows a subset of the fastest routes calculated by our algorithm. Focusing on the U.S., the graph plots the fastest routes for U.S. exports to all other countries. The fastest routes typically go through hubs. E.g., U.S. shipping to Europe tends to pass through Germany and the Netherlands, whereas U.S. shipping to Africa goes through a hub in Spain.

Figure 2 plots the fastest travel times between all port pairs against geodetic distance. Distance is strongly correlated with direct travel time, represented by the light blue dots in the figure. However, we observe that for indirect routes, represented by the dark blue dots, geodetic distance is much less informative for travel times.

To understand further the role of shipping hubs, we examine the number of hops on the fastest shipping routes between all ports in the network. Figure 3

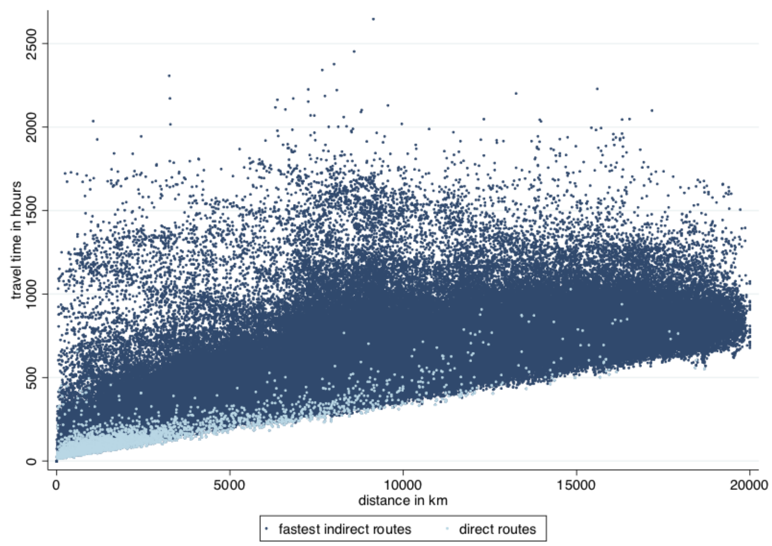
[4]Note, that as we calculate travel time, we exclude trips which involve crossings of anchorages where the ship is sailing in ballast and does not indicate that it is *in transit*. Moreover, we exclude port-to-port connections where less than 5 ships were observed over the whole year. Note also that the travel time reflects the time it takes to get from port D's geofence to port A's geofence plus the time spent traveling, waiting or lading/unlading within the port area. Since we do not observe the time when ships arrive at the dock, we account for the latter by adding one half of the median time that ships spent within the geofence of port D and port A, respectively to the travel time between the geofences.

Figure 1: Fastest Travel Times.



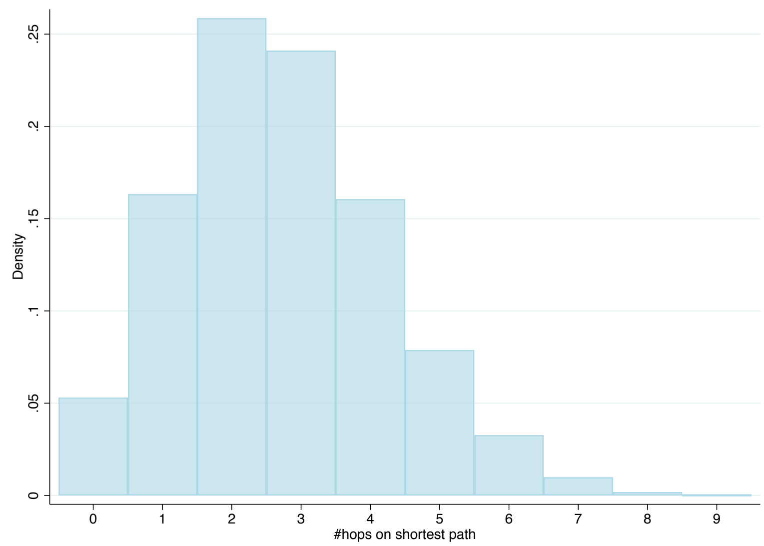
Note: The figure plots the fastest routes from the U.S. to other countries. All computations are based on observed travel times between all regular (non-anchorage) ports in the AIS. Routes with less than 5 ships are dropped. Indirect travel time computed as shortest path in port network where edges, reflecting direct connections, are weighted by direct travel times.

Figure 2: Travel time and distance across port-pairs.



Note: The figure plots travel times between two ports against their geodetic distance. All computations are based on observed travel times between all regular (non-anchorage) ports in the AIS data. Routes with less than 5 ships are dropped. Indirect travel time is computed as the shortest path in port network where edges, reflecting direct connections, are weighted by direct travel times.

Figure 3: The distribution of the number of hops across country-pairs.



Note: The figure shows the distribution of the number of hops along the fastest route between all country pairs in the sample. The average (median) is 2.7 (3). Computations are based on port-to-port shipments from the AIS data and a shortest-path algorithm using travel-time-weighted edges. For countries with multiple ports, the number of hops refers to the connection with lowest number of hops.

shows the frequency of hops after aggregating ports by country.[5] Most country pairs are connected by routes involving one to four hops.[6]

3 The Panama Canal Expansion

Guided by the stylized facts presented in Section 2, we investigate how a shock to a particular link in the shipping network not only affects trade between ports/countries on either side of the link, but also trade between any

[5]For countries with multiple ports, we use the minimum number of hops across multiple connections to the partner country.

[6]Figure 3 shows that for roughly 5% of country pairs the direct route is also the fastest. This is slightly lower than the 6% percent of directly connected country pairs reported above, due to the fact that for a small number of pairs there exists an indirect connection that is faster than the direct route.

port/country that is using that link indirectly.

We use the the Panama Canal expansion as a natural experiment. The expansion was formally proposed on 24 April 2006, and the construction began in 2007. The expanded canal began commercial operation on 26 June 2016. It doubled the capacity of the Panama Canal by adding a new lane of traffic allowing for more ships, and it increased the width and depth of the lanes and locks allowing larger ships to pass.[7] In 2017, the Canal container tonnage increased by 22%.[8]

We perform two complementary empirical analyses. First, we use the route information calculated above to estimate the impact on trade between country-pairs using the Panama Canal versus countries not using the Panama Canal. Second, we perform an event study estimating the effect on container traffic for countries that are using a direct shipping link through the Panama Canal.

3.1 The Effect of the Panama Canal Expansion on Global Trade

We investigate how the Panama Canal expansion affected global trade. We do so by employing a simple differences-in-differences analysis:

$$y_{ikt} = \beta Post_t \times PanExposure_{ik} + \delta \cdot Z_{ikt} + \delta_{ik} + \delta_{it} + \delta_{kt} + \varepsilon_{ikt}, \qquad (1)$$

where y_{ikt} is log quarterly exports (from COMTRADE, see Section 2) from country i to country k at time t. The variable $Post_t$ is a dummy that takes on the value one if the date is after Jun 2016, and zero otherwise. Exposure to the Panama Canal expansion is captured by $PanExposure_{ik}$, which takes on the value one if the fastest route between countries i and k passes the Panama canal and zero otherwise.[9] Z_{ikt} refers to a set of bilateral controls:

[7] https://en.wikipedia.org/wiki/Panama_Canal_expansion_project

[8] https://www.moodys.com/research/Moodys-Upgrades-Panama-Canal-Authority-to-A1-Outlook-stable--PR_396338

[9] For country pairs with multiple ports we average over the binary indicator across all port-to-port connections using the source and destination port size as weights.

Table 3: Panama Canal Exposure: Summary Statistics for 2016.

Country pairs with exposure		Global trade exposed		Importers with exposure	
(1)	(2)	(3)	(4)	(5)	(6)
# pairs	% of total	value in trn $	% of total	# importers	% of total
3,085	12 %	1.8	12 %	141	65 %

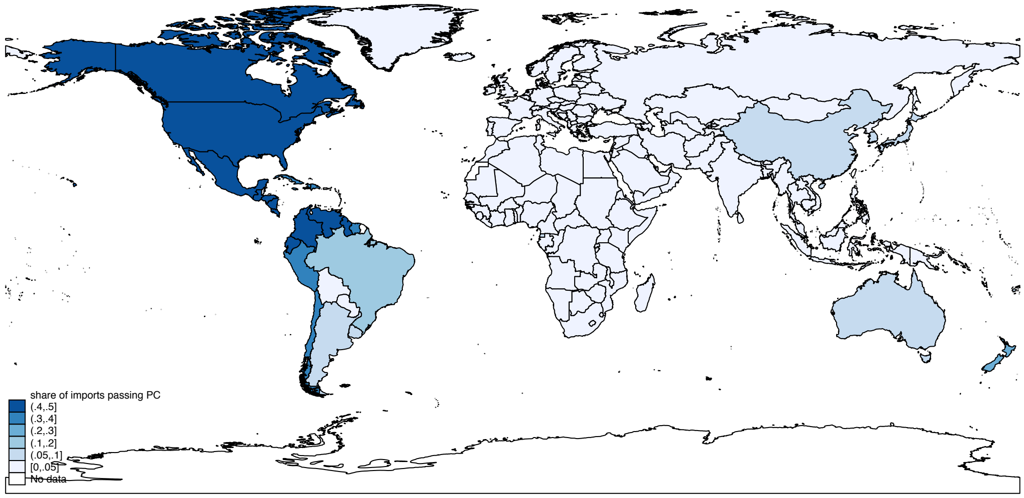
Note: The table shows in column 1 (2) the number (share) of country pairs with a fastest connection passing the Panama Canal; in column 3 (4) the value of (share of global) trade between country pairs whose fastest connection passes the Panama Canal; in column 5 and 6, respectively, the number of importers with at least one fastest connection passing the Panama Canal and their share in the total number of importers.

a dummy for joint membership in a free trade agreement, as well as bilateral geographical variables (distance, contiguity and common language) interacted with the $Post_t$ dummy. Hence, we allow for trends in trade that may differ according to observed geographical characteristics. We also include a large set of fixed effects: source country-time δ_{it} and destination country-time δ_{kt} fixed effects will control for trends in overall exporting and importing for each country, while source-destination country fixed effects δ_{ik} control for time-invariant country pair characteristics.

Panama Canal exposure. Table 3 presents summary statistics for the Panama Canal exposure measure in 2016. There are 3,085 country pairs (12% of 25,025 pairs with positive trade flows) which are connected by a fastest route passing the Panama Canal. The value shipped between these countries accounts for 12% of global trade. The table also shows that the majority of countries are in some way exposed to the Panama Canal: 65% of all importers have at least one fastest connection to a trade partner that passes through the canal. Across all importers, the average share of imports exposed to the Panama Canal is 7%. Figure 4 shows the share of imports passing through the Panama Canal by country.

Empirical Results. We estimate the empirical specification in equation (1)

Figure 4: Panama Canal Exposure by Country.



Note: The figure shows the share of imports passing through the Panama Canal in total imports by country.

using quarterly Comtrade trade data as the dependent variable. Our preferred time period is one year before and one year after the expansion of the canal (2015Q3 to 2017Q2). Appendix Table 10 summarizes the estimation sample.[10] Estimation results are reported in Table 4. Across specifications, we find that bilateral trade between country pairs whose fastest route passes the Panama Canal increased by 8-9% after the expansion. Columns (1)-(3) report results for the full sample of country pairs and quarters, while columns (4)-(6) show results using a balanced panel where all country-pairs are observed in every quarter. The inclusion of the vector of controls in columns (2) and (5) does not change the results significantly, underscoring the robustness of the results.

We also explore whether the treatment effect is heterogeneous across country pairs. One hypothesis is that country pairs with fewer hops along the route will have a greater treatment effect than country pairs with many hops. For example, if the expansion reduces travel time due to less congestion, then the

[10] Our estimation sample covers about 82% of global imports reported to Comtrade. The missing 18% are due to countries not reporting trade data to Comtrade on a monthly basis (which are aggregated to the quarterly level).

Table 4: The impact of the Panama Canal expansion on trade.

Sample:	Unbalanced			Balanced		
	(1)	(2)	(3)	(4)	(5)	(6)
$Post_t \times PanExposure_{ik}$	0.089**	0.090**		0.084**	0.090**	
	(0.042)	(0.043)		(0.037)	(0.038)	
$\times 1[\#hops \leq med]$			0.216**			0.237**
			(0.087)			(0.092)
$\times 1[\#hops > med]$			0.042*			0.074*
			(0.087)			(0.038)
Controls	No	Yes	Yes	No	Yes	Yes
Fixed effects: ik, it, kt	Yes	Yes	Yes	Yes	Yes	Yes
Observations	68,112	68,112	68,112	49,978	49,978	49,978
Exporters/Importers	209/90	209/90	209/90	200/61	200/61	200/61
R^2	0.964	0.964	0.964	0.968	0.968	0.968

Note: The time period is 2015Q3 to 2017Q2. $Post_t = 1$ if $t > 2016Q2$. The control variables are: an FTA indicator and geographical variables (distance, contiguity and common language) interacted with $Post_t$.S.e. in parentheses clustered by exporter and importer. The three first columns include all country-pairs, while the three last columns only include country-pairs with positive trade in all quarters. The triple interaction term in columns (3) and (6) is an indicator variable for whether the number of hops between i and j is below of above the median number of hops. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

time saved in percent will be higher on routes with fewer hops. Columns (3) and (6) in Table 4 interact the main regressor with an indicator variable for whether the number of hops between i and j is below or above the median number of hops. The results strongly support the hypothesis; the treatment effect is roughly twice as high for country pairs with below median number of hops compared to the average treatment effect.

Placebo. To check the robustness of our results, we also re-estimate the specification from column two of Table 4 with placebo treatments in June 2015 or 2017, using four quarters of 2015 and 2017, respectively. Results are reported in Table 5. The estimated coefficients are not significant, suggesting that there are no pre-trends driving our results.

Table 5: Placebo treatments in 2015 and 2017.

Time period:	2015Q1-2015Q4	2017Q1-2017Q4
$Post_t \times PanExposure_{ik}$	-0.035	0.010
	(0.045)	(0.044)
Controls	Yes	Yes
Fixed effects: ik, it, kt	Yes	Yes
Observations	39,449	39,748
Exporters/Importers	208/90	209/88
R^2	0.971	0.971

Note: Column 1 (2): Placebo treatment is $Post_t = 1$ if $t > 2015Q2$ ($Post_t = 1$ if $t > 2017Q2$).The control variables are: an FTA indicator and geographical variables (distance, contiguity and common language) interacted with $Post_t$. S.e. in parentheses clustered by exporter and importer. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

3.2 The Effect of the Panama Canal Expansion on Container Traffic: An Event Study

We supplement the analysis above with an event study design to assess the impact of the Panama Canal on container traffic. Compared to the previous analysis, this section uses the AIS data itself to compute global container cargo flows and compares cargo flows passing through the Canal with cargo flows not passing through.

Global Cargo Flows. To create a comprehensive data set for global cargo flows based on the AIS data we proceed by using technical information about the ships' draught. The draught of a ship is the distance between the surface of the water and the lowest point of the vessel and therefore informs us about the ship's load. Draught is one of the variables the ship reports when the port call is registered by the AIS system. Due to the availability of AIS data, the use of draught-based estimates of ships' cargo has recently emerged in the maritime transport literature, see e.g. Adland et al. (2017). By using information about the movement and draught of every ship, we construct a comprehensive data set for global container cargo flows. Appendix B.1 provides details about the precise measurement of cargo flows.

Empirical Model and Identification. We split voyages into two groups: voyages leaving country i that pass through the Panama Canal ($p = 1$) at time t and voyages leaving country i that do not pass through the Panama Canal ($p = 0$). We then sum across voyages by mode p and take logs of the relevant variable. The event study compares all global cargo flows that run through the Panama Canal before and after the Panama Canal expansion, with all global cargo flows that do not pass through the canal. The key identifying assumption is that the two groups of cargo flows would have followed the same trend in the absence of the Panama Canal expansion.

Margins of cargo flows. We measure five different margins of cargo flows: total shipments, number of ships, average shipments (per ship), total ship capacity and average capacity (per ship), all in logs. Shipments is the total metric tons of goods shipped using the draught based cargo measure, while capacity is the total metric tons of goods the ships could carry if fully loaded.[11]

Empirical Results. The results can be summarized by a set of differences-in-differences regressions using the five different margins of container traffic as the dependent variable:

$$x_{ipt} = Post_t * I_p + I_p + I_i + I_t + \epsilon_{ipt} \tag{2}$$

where x_{ipt} refers to the five different margins. $Post_t$ is a dummy equal to one if the voyage took place after the date of Panama Canal expansion, and zero otherwise. I_p is a dummy that equals one if the shipment passes the Panama Canal, and I_t is a day fixed effect.

The results are reported in Table 6. The Panama Canal expansion had a positive effect on total shipments, shipments per ship and capacity per ship, while the number of ships appears to have declined. Therefore, the increase in total shipments is to a large extent driven by larger ships.

[11]See the Appendix for details on the calculation.

Table 6: Results on Container Traffic

Dependent Variables	Shipments (1)	Average Shipments (2)	No. of Ships (3)	Capacity (4)	Ship Size (5)
$Post_t * I_p$	0.07***	0.16***	-0.08***	0.02***	0.11***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Fixed effects	i, p, t	i, p, t	i, p, t	i, p, t	i, p, t
Observations	22,563	22,563	22,563	22,563	22,563
R-squared	0.81	0.57	0.88	0.87	0.71

Note: The dependent variables are log metric tons of shipments, average shipments per ship, number of ships, total shipping capacity and average ship size (per ship) leaving from country i at the date t. Each observation is country-time-p specific. Robust standard errors in parentheses are clustered at the country level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

4 The Model

This section introduces a canonical model of world trade to quantify the general equilibrium effect of the Panama Canal expansion. After presenting the economic framework, we will quantify the effect of the expansion by feeding in (i) changes in bilateral trade costs according to the reduced form estimates from Section 3 along with (ii) the fastest route information from Section 2.3, while holding all other parameters constant. We then calculate the resulting changes in all equilibrium outcomes according to the model.

4.1 World Equilibrium

Consider a global economy of N countries, a continuum of differentiated goods, and a constant elasticity of substitution (CES) aggregator. Several theories of international trade then lead to a gravity equation of the form

$$\chi_{ij} = \frac{T_i w_i^{-\theta} d_{ij}^{-\theta}}{\Phi_j}, \tag{3}$$

where χ_{ij} is country i's share in country j's manufacturing spending and $\Phi_j = \sum_{i' \in N} T_{i'} w_{i'}^{-\theta} d_{i'j}^{-\theta}$. In the Eaton-Kortum (2002) model, T_i denotes country i's average efficiency (absolute advantage), θ is the dispersion in efficiency across goods (comparative advantage), $d_{ij} \geq 1$ is the iceberg trade cost between i and j and w_i is the nominal wage.

There are two sectors, manufacturing (M) and non-manufacturing (N). Only manufacturing goods are traded. Gross production of manufactures in a country, Y_i^M, equals total worldwide spending on goods from country i. This gives the goods market clearing condition

$$Y_i^M = \sum_{j=1}^{N} \chi_{ij} X_j^M,$$

where X_i^M is manufacturing spending in country i.

We allow for the possibility of unbalanced trade, and denote the trade deficit as $D_i^M = X_i^M - Y_i^M$. The trade deficit relative to GDP, D_i^M / Y_i, is assumed to be constant. A constant share α of income is spent on manufacturing goods, so $X_i^M = \alpha X_i$, where total spending is $X_i = Y_i + D_i^M$ and Y_i is total income. Under perfect competition, aggregate income equals labor income, $Y_i = w_i L_i$. We can then manipulate the goods market clearing condition to (see Appendix D):

$$w_i L_i \left(1 - \frac{1-\alpha}{\alpha} \frac{D_i^M}{Y_i}\right) = \sum_j \chi_{ij} w_j L_j \left(1 + \frac{D_j^M}{Y_j}\right). \tag{4}$$

An equilibrium is then a vector of wages w_i that satisfies equations (3) and (4).

4.2 The General Equilibrium Effect of the Panama Canal Expansion

Consider relative changes from an initial to a counterfactual equilibrium and denote the relative change by $\hat{x} = x'/x$, where x' and x are the new and initial equilibria. The change in trade shares are then

$$\hat{\chi}_{ij} = \frac{\hat{w}_i^{-\theta} \hat{d}_{ij}^{-\theta}}{\hat{\Phi}_j}, \tag{5}$$

where $\hat{\Phi}_j = \sum_{i \in N} \chi_{ij} \hat{w}_i^{-\theta} \hat{d}_{ij}^{-\theta}$ (see Appendix D). The goods market clearing conditions can be written as

$$\hat{w}_i = \sum_j \frac{\chi_{ij} X_j^M}{Y_i^M} \hat{w}_j \hat{\chi}_{ij}. \tag{6}$$

As is well known (Caliendo and Parro, 2015), in this class of models the change in real income is simply

$$\frac{\hat{w}_i}{\hat{P}_i} = \hat{\chi}_{ii}^{-\alpha/\theta},$$

where $\hat{P}_i$ is the relative change in consumption prices.

We now ask what would happen to the world equilibrium if we change trade costs d_{ij} for the country pairs affected by the Panama Canal expansion. The data requirements for this exercise are relatively modest: First, it requires data on X_i^M, Y_i^M, χ_{ij} and α in the initial equilibrium. All these variables are available from Eora MRIO global supply chain database; we gather data for 189 countries for the 2015 cross-section.[12] Second, it requires data on the trade elasticity θ. Since our analysis does not identify the trade elasticity, we rely on estimates from the previous literature, and choose the value $\theta = 5$, close to the aggregate elasticity estimated in Caliendo and Parro (2015).

Finally, it requires data on the change in trade costs $\hat{d}_{ij}$. Recall that we know from Section 2.3 which country pairs are exposed to the canal, i.e. that the fastest route from i to j is using the canal. Country-pairs not trading through the canal therefore have $\hat{d}_{ij} = 1$. For country pairs using the canal, the reduced from results from Section 3 gave us an estimate β of the impact on trade caused by the expansion. We assume the following functional form for

[12] α (the share of spending on manufacturing goods) is only required for the calculation of the change in the consumer price index. We choose a value of 0.35, which is the average share across countries in Eora MRIO.

trade costs: $\hat{d}_{ij} = \alpha^{-Panama_{ij}}$, where $Panama_{ij} = 1$ if i and j are connected by the Panama Canal, and $= 1$ otherwise. For canal-connected countries, the relative change in trade costs d_{ij} is then $1/\alpha$. The structural interpretation of the coefficient estimate β from equation (1) is then simply $\theta \ln \alpha$. Given the estimate $\beta = 0.09$ and $\theta = 5$, we get $\alpha = exp\,(0.09/5) \approx 1.02$, or $\hat{d}_{ij} = 0.98$ for canal-connected countries.

4.3 Results

Figure 5 shows the change in the domestic trade share, $\hat{\chi}_{ii}$, plotted against the change in manufacturing exports for all the countries in our sample (both in percent). As expected, countries close to the Panama Canal increase their share of imports in total spending (i.e., the percent change in χ_{ii} is negative). This includes the U.S, Canada, Mexico and Panama itself. These countries also increase their total exports, as their market access to other countries improves. For the world as a whole, global trade increases by roughly 1 percent.

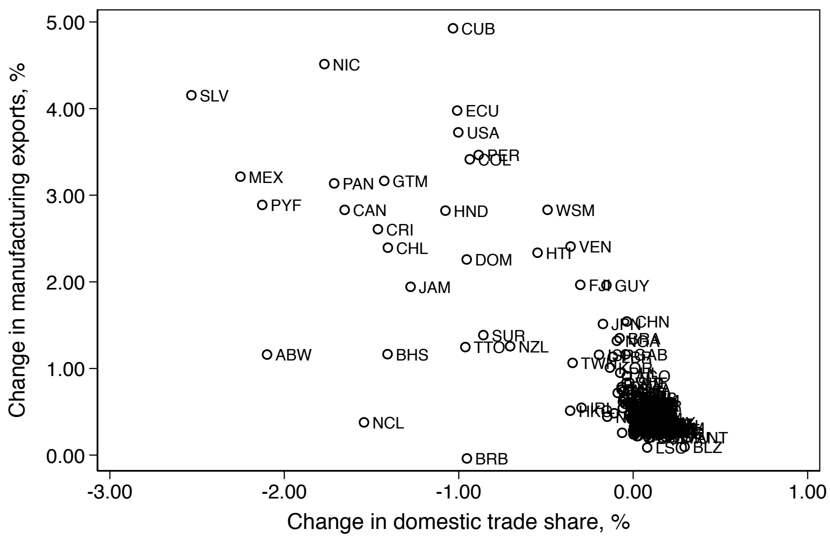
Table 7 reports the changes to manufacturing exports, imports, the domestic trade share and the real wage for the top 5 countries with the largest real wage gains. Countries close to the canal emerge as the top winners from the expansion, with increases in real wages of around 0.20 percent. For the large majority of countries, the real wage gains are close to zero. A few countries, including Austria, Hungary and Zimbabwe, experience a (small) real wage loss due to the expansion. These are landlocked countries that do not themselves get improved market access, but might compete with other exporters that do get better market access. The weighted average of the real wage change across all countries is 0.02%, or roughly 20 billion USD.[13]

5 Concluding remarks

We exploit novel satellite data on all global port calls made by container ships in order to construct a global and comprehensive data base of the shipping net-

[13]Using PPP adjusted GDP as weights.

Figure 5: Counterfactual: The Impact of the Panama Canal Expansion.



Note: The figure shows the counterfactual change in the domestic trade share on the horizontal axis plotted against the change in manufacturing exports on the vertical axis (both in percent).

Table 7: Counterfactual: The Impact of the Panama Canal Expansion.

	Exports	Imports	$\hat{\chi}_{ii}$	Real wages
El Salvador	4.15	4.31	-2.53	0.18
Mexico	3.21	4.55	-2.25	0.16
Nicaragua	4.51	4.32	-1.77	0.13
Panama	3.14	3.55	-1.71	0.12
Canada	2.83	3.83	-1.65	0.12

Note: The table shows the change in outcomes for the countries with the largest increase in real wages from the Panama Canal Expansion. Small island states are excluded. All values in percent.

work. This allows for the construction of a new data set on optimal shipping routes. We apply this dataset to analyze how shocks hitting a node or an edge in the shipping network affect all trading partners to varying degrees based on their exposure to the shock. Using the 2016 Panama Canal expansion as a natural experiment, we show that the expansion not only had a direct effect on shipments traveling through the canal, but importantly also affected trade flows between countries using the Panama Canal intensively. The expansion produced sizable gains from trade according to a model-based counterfactual analysis. The results highlight the importance of trade networks for the quantification of the gains from trade.

References

Adland, R., H. Jia, and S. P. Strandenes (2017). Are ais-based trade volume estimates reliable? the case of crude oil exports. *Maritime Policy and Management 44*(5), 657–665.

Bernhofen, D. M., Z. El-Sahli, and R. Kneller (2016). Estimating the effects of the container revolution on world trade. *Journal of International Economics 98*, 36 – 50.

Brancaccio, G., M. Kalouptsidi, and T. Papageorgiou (2017). Geography, search frictions and endogenous trade costs. Technical report, NBER Working Paper No. 23581, forthcoming in Econometrica.

Brooks, L., N. Gendron-Carrier, and G. Rua (2018). The local impact of containerization. Technical report, manuscript.

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

David, M. (2015). Vessels and ballast water. In M. David and S. Gollasch (Eds.), *Global Maritime Transport and Ballast Water Management*. Springer.

Donaldson, D. and A. Storeygard (2016). The view from above: Applications of satellite data in economics. *Journal of Economic Perspectives 30*(4), 171–198.

Ducruet, C., R. Juhasz, D. Nagy, and C. Steinwender (2019). All aboard: The aggregate effects of port development. Technical report, manuscript.

Feyrer, J. (2009). Distance, trade, and income - the 1967 to 1975 closing of the suez canal as a natural experiment. Technical report, NBER Working Paper No. 15557.

Levinson, M. (2010). *The Box: How the Shipping Container Made the World Smaller and the World Economy Bigger*. Princeton University Press.

Maurer, S. and F. Rauch (2019). Economic geography aspects of the panama canal. Technical report.

Rajkovic, R., N. Zrnic, O. Cokorilo, S. Rajkovic, and D. Stakic (2014). Multi-objective container transport optimization on intermodal networks based on mathematical model.

Rua, G. (2014). Diffusion of containerization. Technical report, Finance and Economics Discussion Series 88.

UNCTAD (2016). Review of maritime transport. Technical report.

Appendix

A Constructing the container data set

Our point of departure are the AIS data containing all port calls made in 2016 that has been provided to us by Marine Traffic. We also use the World Fleet register data base constructed by Clarkson, which has vessel specific information on a range of ship characteristics. Based on the ship categories used by data supplier, Marine Traffic, we use the ships categorized as "container ship" and "Cargo/containership". Marine Traffic gives each ship a unique identifier (Ship ID). We start out with close to 5,300 ships based on this identifier. We use this to identify each ship's travel history. A ship also has an IMO number and an MMSI number as well as a Ship Name. Ideally there should be a perfect match between ship identifiers (IMO, MMSI and Ship ID). However, for around 5% of the ships this is not the case. The mismatch could either because of misreporting, or changing of owners (containerships typically change their MMSI number when changing the owner). We correct for both misreporting and the change of identifiers by cross checking a ship's IMO and MMSI number, as well as ship's characteristics, like its deadweight tons (dwt) provided by Clarkson and the office site of Marine Traffic. We are able to correct for most of the misreporting and end up with 5,165 distinct containerships.

Then we proceed by cleaning the routes of each container ship. We first sort trips for each ship by their time stamp, so that their travel records are listed as Arrival-Departure-Arrival-Departure, etc. In the case that a ship reports multiple arrivals at the same port before the following departure (e.g., Arrival-Arrival-Departure), we choose the first arrival as the actual arrival time at the port; similarly, we treat the last departure timestamp as the "true" departure time from a port, when a ship reports multiple departures from the same port before the next arrival (e.g., Departure-Departure-Arrival). A "trip" is defined as a direct port-to-port trip. We ignore the case when ships are in transit or stopped at an anchorage, hence if a ship "properly" stopped at a non-anchorage port A, followed several "in transit" stops or stops at anchorage ports, and finally arrived at a non-anchorage port B, we define the travel from A to B as one trip of the ship. We use the draught reported when the ship reaches the arrival port as the draught of the trip.

In the very last step, we merge the AIS data with the World Fleet register data from Clarkson to add information on ships' scantling draught. Scantling draught is the draught the ship will have when it is fully loaded, and it is also referred to as design draught, as it is this draught it is build for. We add scantling draught in order to be able to construct a draught based measure of cargo. The AIS data from Marine Traffic also report a ship's maximum draught "draught max", which is the highest draught value reported by a ship. If the value of maximum draught from Marine Traffic exceeds the one on scantling draught from Clarkson, we use the greater value of the two as the ship's scantling draught in our data set. Finally, we introduce a threshold of 15,800 deadweight tons. We are able to match 4,908 out of 5165 ships using either their IMO or MMSI number.

B Calculating Cargo

B.1 AIS based Cargo Flows

A ship sailing without cargo is referred to as a ship sailing in ballast. In practice, a ship sails in ballast if its draught is smaller than a given threshold value, which we refer to as ballast draught (H_B). Specifically, we define $H_B = 0.55H_S$, where H_S is the ship's scantling draught. We have access to technical information on ships' scantling (design) draught from the Clarkson World Fleet Database (see Appendix A). We use 0.55 as the weight to define ballast draught based on the maritime engineering literature.[14] We calculate the metric tons of cargo, from now on referred to as effective dwt (B), as*:*

$$B = dwt * (H_A - H_B) / (H_S - H_B), \tag{7}$$

[14]The threshold for ballast water is chosen based on information from Marine Traffic supported e.g. David (2015).

where H_A is the draught reported by the ship en route and *dwt* is the deadweight tons (i.e., the vessel's weight carrying capacity) from the Clarkson data.

A ship's draught as well as estimated cargo relates to one specific trip, i.e. to a voyage between two ports.

Table 8 shows that, based on our draught based estimates, on average container ships do merely 0.2% of their trips without cargo (in ballast). This stands in sharp contrast to other types of vessels that are typically involved in very different trades, and do not operate on "bus routes" like container ships. Brancaccio et al. (2017) focus on dry bulk ships and report that 42% of the ships travel without cargo. We also observe that there is substantial variation across trips with respect to draught, effective dwt, and across ports with respect to total incoming and outgoing cargo.

Table 8: Ships, Trips and Port

Variable:	Obs	Mean	Sd	Min	Max
Ships:					
Share of trips in ballast (<55%)	4,908	0.002	0.04	0	1
Trips:					
Actual draught (% of scantling draught)	331,265	0.80	0.10	0.55	1
Effective dwt on loaded trips	331,265	26,113	24,560	1.23	199,744
Ports:					
Total incoming effective dwt in millions	515	16.80	44.30	0	498.70
Total outgoing effective dwt in millions	515	16.80	44.30	0	499.98

Note: Summary statistics are based on the port calls made by container ships in 2016. Effective dwt is calculated based on dwt and draught and is used as a measure for cargo. Only ships with deadweight tons>15,800 and trips with non-zero duration are used. Summary statistics include only routes taken by at least 5 ships and only routes between ports that appear both as arrival and departure ports.

B.2 Calculating cargo flows pre and post the Panama Canal expansion

To compute cargo flows before and after the Panama Canal expansion, we first identify cargo flows that passed through the Panama Canal. Container ships typically travel through the Canal in two ways. If a container ship has a light draught (when the ship is small or has few loads), it can directly pass through the Canal without stop (in transit) or stop only at anchorage ports if it was waiting in a queue. In other cases, a container ship typically makes a stop at a non-anchorage port in Panama, unload some containers to lift its draught before passing through the Canal. Those unloaded containers will then be transported over land and then loaded back on the container ship.

We identify both types of trips using the timestamps of each ship. The former is identified in our data as trips that departed from a third country, then passed through the Panama Canal in transit or at anchorage ports. We identify the latter as trips that left from one side of Panama to a third country, followed after a previous stop at the other side of the Canal. Altogether, 982 out of 2,660 trips that passed through the Panama Canal directly or at anchorage ports (37% of the pass-through). In the remaining cases, container ships made stops at non-anchorage ports before crossing the Canal. The ships passing through directly are indeed smaller (average deadweight tonnage 50,842 versus 60,506 of those that stopped).

We calculate departure time as the last time a pass-through ship appeared at Panama before heading to a foreign destination port. Over 99% of the cases ships left from the same anchorage (Panama anchorage Atlantic or Panama anchorage Pacific, depending on their journey direction). The tonnage of goods that passed the Panama Canal is measured at the moment when a ship first arrived at the other side of the Canal for non-stopped trips. For trips that made stops, to take into account reloading, we measure tonnage of goods when the ship is headed to a foreign destination port.[15]

[15] Implicitly, we assume that when a container ship is traveling across Panama, the number of newly added containers at Panama is sufficiently small compared to the number of containers it carried through the Canal.

Summing the tonnage identified above by date/week/month yields the total tonnage passing though the Canal. To construct our control group, we use the rest of the trips in our data and sum over the departed tonnage by country and by time. This also gives us cargo tonnage that left from Panama but did not pass through the Panama Canal. Global shipping activities tend to spike in early December due to Christmas and fall sharply in January due to the Chinese New Year; hence, we drop January and December from our analysis. The daily data has lots of fluctuations, which remains high when we aggregate the data to week and bi-week level. Hence when we plot our daily or weekly figures, we smooth our data first using a 31-day simple moving average throughout the paper.

B.3 Summary Statistics Panama Canal Exposure

Table 9: Summary statistics on Panama Canal exposure

Rank	Importer	Share of total imports passing PC	Share in world imports	Exporter	Share of total exports passing PC	Shar e
1	USA	52.7	14.0	USA	30.1	
2	MEX	10.1	2.5	CHN	17.9	
3	CAN	9.6	2.7	MEX	11.6	
4	CHN	5.0	7.7	CAN	9.7	
5	JPN	2.8	3.7	JPN	5.7	
6	KOR	1.8	2.6	DEU	3.3	
7	CHL	1.2	0.4	KOR	3.2	
8	COL	1.2	0.3	GBR	1.3	
9	HKG	1.1	3.4	FRA	1.3	
10	BRA	1.0	0.9	ITA	1.2	
11	NLD	1.0	3.0	CHL	1.0	
12	PAN	0.9	0.2	BRA	0.9	
13	FRA	0.9	3.8	IRL	0.9	
14	AUS	0.8	1.2	PER	0.7	
15	PER	0.8	0.2	COL	0.7	

C Estimation Data: Summary Statistics

Table 10: Summary Statistics of the Estimation Sample

Variable	N	Mean	Std. Dev	Min	Max	Source
ln Value (in $)	68,112	15.73	3.31	2.83	25.64	monthly COMTRADE
ln Qty (in kg)	65,701	14.39	4.13	0	35.98	monthly COMTRADE
FTA	68,112	.29	.45	0	1	WTO RTA database
ln Distance	68,112	8.57	.88	4.11	9.89	CEPII
Contiguity	68,112	.03	.16	0	1	CEPII
Common Language	68,112	.13	.34	0	1	CEPII
Pan Exposure	68,112	*.08*	.24	0	1	AIS data

Note: Export data in rows 1 and 2 is aggregated from monthly to quarterly frequency and covers the period 2015Q3 - 2017Q2.

D Theory Appendix

Market potential. The change in the market potential term Φ_j is

$$\begin{aligned}\hat{\Phi}_j &= \frac{\sum_{i\in N} T_i w'^{-\theta} d'^{\,-\theta}_{ij}}{\sum_{i\in N} T_i w_i^{-\theta} d_{ij}^{-\theta}} \\ &= \sum_{i\in N} \frac{T_i w_i^{-\theta} d_{ij}^{-\theta}}{\sum_{i\in N} T_i w_i^{-\theta} d_{ij}^{-\theta}} \hat{w}_i^{-\theta} \hat{d}_{ij}^{-\theta} \\ &= \sum_{i\in N} \chi_{ij} \hat{w}_i^{-\theta} \hat{d}_{ij}^{-\theta}.\end{aligned}$$

Goods market clearing. Manufacturing gross production can be written as

$$\begin{aligned}Y_i^M &= X_i^M - D_i^M \\ &= \alpha X_i - D_i^M \\ &= \alpha\left(Y_i + D_i^M\right) - D_i^M \\ &= \alpha w_i L_i + \alpha D_i^M - D_i^M \\ &= \alpha w_i L_i \left(1 - \frac{1-\alpha}{\alpha}\frac{D_i^M}{Y_i}\right).\end{aligned}$$

The market clearing condition then becomes

$$\begin{aligned}
\alpha w_i L_i \left(1 - \frac{1-\alpha}{\alpha} \frac{D_i^M}{Y_i}\right) &= \sum_j \chi_{ij} X_j^M \\
&= \sum_j \chi_{ij} \alpha X_j \\
&= \sum_j \chi_{ij} \alpha \left(Y_j + D_j^M\right) \\
&= \sum_j \chi_{ij} \alpha \left(w_j L_j + D_j^M\right) \\
&= \sum_j \chi_{ij} \alpha w_j L_j \left(1 + \frac{D_j^M}{Y_j}\right).
\end{aligned}$$

This can be re-written to

$$w_i L_i \left(1 - \frac{1-\alpha}{\alpha} \frac{D_i^M}{Y_i}\right) = \sum_j \chi_{ij} w_j L_j \left(1 + \frac{D_j^M}{Y_j}\right).$$

Holding the trade deficit constant relative to GDP, we can write the market clearing condition in changes as:

$$\hat{w}_i = \sum_j \frac{\chi_{ij} X_j^M}{Y_i^M} \hat{w}_j \hat{\chi}_{ij}.$$