Technology with unequal gains: Steamship and globalization

Shotaro Beppu[*]

November 2023

Abstract

Transportation technology allows more market integration. The benefit could be unequal due to the adoption of this technology. This paper studies the rapid adoption of steamships in the late 19th century and its effect on the First Globalization and the Great Divergence. For this, this paper uses a novel deep-learning method to digitize historical shipping data to analyze the transition from sailing to steamships and the change in worldwide shipping patterns. Using the change in duration, this paper shows that the advent of steamships increased trade and growth on average. However, colonized countries experienced insignificant gains. To understand the mechanism, this paper incorporates differences in shipping technology to trade with heterogeneous firms. The estimate points out that the advantage of steamships might not have been captured due to the high cost of adoption for trade links involving colonies. This provides further insights into how shipping technology affects economic activity through trade and how the fixed cost of adoption matters in gains from such technology.

1 Steamship as a promoter and hindrance of growth

How can technology lead to growth? This paper uses the rapid adoption of steamships to illustrate how technology leads to unequal gains and how the fixed cost of adoption matters. The faster and bigger steamships replaced sailing ships within a couple of decades. During this period (1880-1900), we observed an explosion of trade (the First era of globalization) as well as divergent trends in countries' welfare (the Great Divergence). Due to limitations in data, we lacked an understanding of how steamships affected this global trend. At a higher level, steamships-induced changes give insights into how technology that increases integration leads to growth. The availability of technology is not often enough for gains to materialize. Growth disproportionately comes to those wanting to take up the technology, potentially exacerbating inequality. Steamship provides an ideal case study of how different countries decided to adopt one such technology and whether this helps explain global inequality.

To investigate these questions, this paper digitized historical shipping data, Lloyd's Shipping Index. Although the data provides us with comprehensive information on trips of ships, either sailing or steamships, from 1880 onward, the use was limited due to data availability. This paper digitized this index en masse using deep learning. The methodology uses recent advances in optical character recognition, identifying texts in images, and table structure recognition, identifying the table structure in images. The lack of either makes using large historical data in table formats difficult. While the machine learning community made substantial advances in character recognition for historical texts, table structure recognition remains a problem even in contemporary documents. As such, this paper extended recent work in structure recognition using graph neural networks to digitize the shipping index. The methodology's success provides more than 30,000 trip data in 1880, 1890, and 1900 with information on the type and size of each ship, the origin and destination port of the trip and the duration.

The changes brought forth by steamships were substantial in maritime shipping and economic growth. First, steamship usage increased dramatically. In 1880, only 20% of the total tonnage was steamships. In 1900, the share of steamships increased to about 80%. The rapid adoption could be explained by how fast and big those steamships were compared to sailing ships. Steamships drove down the duration between ports by more than half on average. In addition, the size of steamships in tonnage was about three times that of sailing ships. Using this change in duration and gravity equation model in international trade, this paper finds steamships led to a significant increase in countries' GDP per capita. The effect was driven by being closer to larger markets (market access) rather than being closer to larger suppliers (supplier access). This average gain, however, masked an important heterogeneous effect. In particular, while high-income countries benefited from market and supplier access, colonized countries saw no gains from both channels. This led to the worsening of global inequality. An important fact that might explain this is the lower adoption of steamships by ports in colonized countries. Steamships induced a large shipping pattern change, and yet ports in colonized ports were, on average, 15 percentage points less in adoption, and a significant share of ports were still exclusively using sailing ships.

This paper combines the heterogeneous impact and steamship adoption by incorporating the adoption of shipping technology in trade with heterogeneous firms (M. Melitz, 2003). Specifically, firms with different productivity do not export, export using sailing ships, or export using steamships. The sorting is explained by the two differences between sailing and steamships: the duration and the entry cost. If it is relatively harder to export using steamships, only the more productive firms are willing to export using steamships as they can compensate for the higher fixed cost. If the fixed cost is low enough, on the other hand, all the firms previously using sailing ships use steamships. In addition, firms that could not export using sailing ships can now export due to better technology. Thus, the fixed cost of using the steamship can explain the large

differences in trade gains. Using the share of tonnage carried by steamships and the differences in trade volumes in 1880 and 1900, the model estimates the fixed cost between countries. This paper finds that the estimated fixed cost is higher for colonized countries, implying that these countries could not fully benefit from steamships. Further, the fixed cost of trade is significantly lower between the colonized and the country that colonized. The cause of the differences in the fixed cost, such as port investment, competition, and extractive colonial relationships, which can be incorporated into the general equilibrium, is a fruitful area for further studies.

The remainder of the paper is as follows. Section 2 provides historical background of the late 19th century and literature in shipping, trade, and development that helps understand the transition from sail to steamship and the resulting gains. Section 3 details the digitization process of the shipping data as well as the data used both at the shipping and country level. Section 5 provides a theoretical framework to understand how a reduction in duration and fixed cost of adoption affects trade and welfare. Section 6 estimates the fixed cost implied by the model using data. Section 7 concludes.

2 Background and literature

2.1 Period of Change

The period from 1880 to 1914 is called the First Era of Globalization. With the lowest tariff in history, trade between countries increased exponentially (Pascali, 2017). Nations in South America, for example, reversed its protectionism policy and reduced tariffs greatly (Britannica, 2023). Great Britain dominated the sea as it expanded all over the world. At the same time, countries such as the United States and Germany were rapidly expanding challenging the dominance of Great Britain. In parallel, the period also coincides with what we call the Great Divergence where the Western civilization took off while other regions fell behind. This period was coined "Belle Époque" in French representing optimism, peace and economic growth. On the other hand, the European powers moved inland in Africa claiming most of the land during this period (New Imperialism). India was under the rule of Great Britain and lost its place to the U.S. in terms of GDP (Bolt and van Zanden, 2014). The decline of China and other Southeast Asian nations also occurred in this period. Related to this paper, the First Sino-Japanese War which took place in 1894-1895 was the first warfare where modern steam-propelled ironclad warships were used in battle. The demise and the division of China accelerated thereafter.

In this global setting, numerous innovations came out. Telephones, transatlantic cables, light bulbs and automobiles began to be used from 1870 to 1890. The significant change in the 19th century that led to those inventions was the steam engine and electricity. This propelled the Second Industrial Revolution and drastically changed transportation. Railroads and automobiles expanded across the world and changed transportation on land. Steamships combined with triple-expansion engines began to be used in the 1880s, allowing long-distance travel to be feasible for the first time. Until the advent of the diesel engine after World War II, steamships became the dominant mode of transportation at sea. During this era, along with inventions in refrigerators and metal, steamships allowed safer and faster transportation. The main advantage of steamships compared to sailing ships was also due to less concern for wind at sea. As (Pascali, 2017) uses in its analysis, steamships travelled a shorter route where sailing ships had to take a detour due to the wind. In addition, the use of steamships further accelerated due to the opening of the Suez Canal in 1869 as the wind conditions in the Mediterranean and the Red Sea meant steamships were much more feasible than sailing (Fletcher, 1958). This paper, then, investigates how this innovation in the sea affected economic growth and in particular the Great

Divergence.

2.2   Transportation, trade, and development

This paper uses three strands of literature in economics. First is empirical analyses of gains from trade. Second is how transportation matters in trade. Third is how technologies can affect the global pattern of trade. In addition, due to the methodologies used, this paper expands on historical knowledge of the 19th century and provides a novel method to digitize historical documents, which will be explained in the data section.

First, this paper builds on empirical work investigating trade gains. The ideal regression researchers would run is welfare on openness, often proxied by trade volumes. This, however, has omitted variable bias as higher GDP per capita likely leads to more trade for variety and more trade is associated with better institutions to promote industries that promote growth. Numerous research employed instrumental variable approaches to tackle this issue (see (Donaldson, 2015) for further review). The first of such works was (Frankel and Romer, 1999), which uses distance to other countries as an instrument. As distance is likely correlated with other variables that affect welfare (proximity to other large developed countries for instance), subsequent works used transportation investments as a natural experiment. (Feyrer, 2021) uses the switch from ocean shipping to aeroplanes as an exogenous change in distance, (Feyrer, 2019) used the closure and reopening of the Suez Canal, and (Donaldson and Hornbeck, 2016) used the expansion of railroad networks to see gains within the U.S. These provide positive estimates from more openness. However, some research documented the negative consequences of better proximity to other regions. (Faber, 2014) finds that regions in China that had better access to the national highway experienced a decrease in economic activity. (Campante and Yanagizawa-Drott, 2018) uses the discontinuous adoption of large aeroplanes in passenger travel to show that although the overall effect of aeroplanes is positive, the poorest places did not benefit. On the other hand, (Okoye, Pongou, and Yokossi, 2019) shows that the colonial railway benefited the least accessible places. The most related result is from (Pascali, 2017). This used the change in the distance travelled by sailing to steamships and found that the gains from trade only accrued to countries with strong institutions. This paper investigates why we observe those negative effects for select regions. The idea explored in this paper that technology such as steamships does not lead to immediate take-up is applicable to many situations where the benefit of those technologies is related to past economic conditions and does not lead to convergence.

Second, this paper considers port and shipping an important aspect of economic growth. In addition to the papers above using sea distance, numerous papers used changes in shipping and ports to answer this question. (Heiland et al., 2022), for example, used the 2016 Panama Canal expansion as a quasi-natural experiment and showed that expansion not only affected trade flows directly exposed to the canal but also indirect effects on world trade due to the shipping network. An important aspect of shipping is ports. Ports are a source of efficiency gains by investment. Inefficient ports, on the other hand, are as if they impose additional tariffs on goods arriving. (Blum et al., 2019) considers inventory management and shows that improvement, especially in poor countries, can increase trade. (Feenstra and Ma, 2014) showed that port efficiency is important in the extensive margin, i.e. trade facilitation. Other papers used the advent of containerization to see how port improvements affect various outcomes. A number of papers looked into the increase in trade (Bernhofen, El-Sahli, and Kneller, 2016) (Coşar and Demir, 2018). Others looked into local effects around ports that integrated into the world market (Brooks, Gendron-Carrier, and Rua, 2021). Related to the paper's time period, (P. Fajgelbaum and Stephen J. Redding, 2022) uses Argentina during the late 19th century to show how closeness to ports during this time induced more trade and structural transformation. (Galiani, Jaramillo, and Uribe-Castro, 2023) used the opening of the Panama Canal on economic activity

in the U.S. Combining insights from local and global economic activity. (Coşar and P. D. Fajgelbaum, 2016) looked at the interplay of regions that are nearby ports and ports themselves and how more trade induces specialization across cities. With digitized data on shipping data, this paper investigates how steamships changed the port structure.

Third, this paper uses theoretical frameworks in trade. First, the empirical estimate largely follows the market access approach (S. Redding and Venables, 2004) (Donaldson and Hornbeck, 2016) (Stephen J. Redding and Turner, 2015). (S. Redding and Venables, 2004), in particular, expands the (Krugman, 1980) model of economies of scale and monopolistic competition to show that the closeness to other countries increases welfare. Papers stemming from (M. Melitz, 2003) used (Krugman, 1980) to incorporate heterogeneous firms. As (Helpman, M. Melitz, and Rubinstein, 2008) estimated the trade relationship, this model features a fixed cost of entry and makes only the more productive firms export. This additional extensive margin for trade is shown to dominate compared to the intensive margins of trade when using a Pareto distribution for firm productivity (Chaney, 2008). These canonical models, however, only predict gains from trade using domestic trade flows and trade elasticity. Papers such as (Allen and Arkolakis, 2022) incorporated transport networks and more structure to the local economic activity to derive changes in wages due to transportation investment. This paper, however, uses extended versions of the canonical trade model to see the gains from steamships. Steamships can be considered a technology that reduces variable trade costs but may introduce higher fixed costs upfront. From this perspective, this paper is related to models incorporating market access (Arkolakis, Ganapati, and Muendler, 2021), reduction in duration as a quality upgrade (D. L. Hummels and Schaur, 2013), binary technology adoption (Bustos, 2011). This paper is further related to understanding the unequal gains from trade (M. J. Melitz and Stephen J Redding, 2015) (P. D. Fajgelbaum and Khandelwal, 2016).

3 Digitized shipping and trade data

This paper uses two types of data. One is shipping data at the port level and another is trade and economic development at the country level. For port-level data, this paper digitizes shipping log data. For country-level data, this paper uses data created by (Pascali, 2017) and TRADHIST (Fouquin and Hugot, 2016). This paper combines these data for the years 1880, 1890, and 1900.

3.1 Port level data

3.1.1 Overview

For shipping data, this paper uses a traditional shipping index, "Lloyd's Shipping Index". Since the late 16 century, the merchant company Lloyd's List has recorded maritime flows of merchant vessels. In 1880, the company started publishing Lloyd's Shipping Index, reporting weekly the flows of ships globally for both sailing and steamships. Importantly, these publications detail the movements of most of the world's fleet, including vessels insured by other companies, making it highly representative of global maritime trade. Although this provides us with the most comprehensive movements of ships at this time, it was traditionally used only for exploratory purposes (Ducruet and Itoh, 2022). Figure 1 shows an example of the first few rows of the tables. In this case, the index records, from the left, the association the ship belongs to, the ship name, the master of the ship, the tonnage in brackets, the nationality, the type of ship, the origin and when it left, the destination, and the latest whereabouts often accompanied by indicators of status (e.g., "Ar" means Arrived). With this information across multiple weeks, it is possible to describe how a particular ship moved to different places in detail based on characteristics such

Figure 1: Example of 1900's Index covering sailing ships

Reg.	Ship Master	Ton.	Flag Rig	From	For	Latest Reports
R v	E A O'Brien *Pratt*	(1038)	Br bq	Manilla Apr 4	Boston	ArSept10—ForBuenosAyres
R v	E B Sutton *Carter*	(1639)	Am s	Honolulu Oct 13	New York	
*R	ECMowatt *Hersey*	(1026)	Am bq	Philadelphia Sept6	Table Bay	Pd Marcus Hook Sept 6
*	E JSpence *Stronach*	(519)	Br bq	Singapore July 26	Mauritius	Ar Sept 12
v	E J Spicer *Cochran*	(1268)	Br s	Table Bay July 21	Nestle(NSW)	Ar Spt2—ForWSCAmerica
R	E K Wood *Hansen*	(452)	Am sc	Tacoma Aug 25	Haiphong	
v	ESHocken *Willcock*	(249)	Br bq	Clyde Aug 4	Table Bay	Ar Oct 25
*	EagleCrag *Shimmin*	(1347)	Bı bq	Cardiff Sept 15	CaletaBuena	Pd Barry Island Sept 15
*	Earl Cadogan *Williams*	(1334)	Br bq	Nestle(NSW)Aug24	Antofagasta	Ar Oct 15
*	EarlDerby *Mackintosh*	(961)	Brbq	Brisbane Sept 8	Nestle(NSW)	ArSpt15--ForWSCAmerica

as nationality, ship type, and tonnages. Furthermore, it is possible to observe how the shipping network has evolved over the years. Above all, this data provides the total tonnage from one port to another, allowing port-level analysis combined with country-level data. This research uses three years of data (1880, 1890, and 1900) from October to December due to limitations in data. Each year, this research obtains shipping movements for both sailing and steamships. An important limitation of this data is that it does not contain any information on cargo and where those were unloaded. Thus, even if one ship travelled from one port to another, it may well be possible that that cargo went to another destination. This paper shows later, however, that the total tonnage is highly correlated with the total trade at the country level suggesting that the tonnage is a good proxy for trade between countries.

This paper further extends the database of ports created by (Ducruet, Cuyala, and El Hosni, 2018). Here, for each port, the data includes the longitude and latitude of the port and the nearest city. This paper further categorizes each city by its location to the port: coastal, downstream, inland, and upstream. This category is useful to obtain as the difficulty of ports to be used for steamships depended on the geographical characteristics as this paper later shows in the data. For example, a number of upstream ports were unsuited for steamships as it was hard to expand (Ducruet and Itoh, 2022).

The issue we have when using duration is that there is only data when there is a ship travelling. As duration itself is a product of geographical characteristics (wind and ocean conditions), for the unobserved duration, this paper computed the distance similar to (Pascali, 2017) and used predicted sailing/steam ship time using observed observations. This will be used in understanding the intensive and extensive margin.

3.1.2 Digitization

The data were only available in images. This paper digitized these images extending explored methods in deep learning. Recently, digitizing historical documents using deep learning has started to emerge in the field of economics. LayoutParser (Shen et al., 2021), for example, created an easy interface combining recent advances in optical character recognition to be used to obtain texts from documents. However, these tools are not sufficient to create structured datasets such as tables. Current research still requires intensive post-processing of these texts to associate which text belongs to which row. This is indispensable as, in the case of the shipping data, we need which ship went to which port which is hard to obtain due to distortions in the data. For example, spanning rows as in figure 2 is fairly prevalent in the documents. Matching these variables to one ship is difficult using simple heuristics. Thus, this paper extends the methodologies in table structure recognition (Gatos et al., 2005) (Kasar et al., 2013) (Hao et al., 2016) (Rashid et al., 2017) (Schreiber et al., 2017) (Dai et al., 2017). This field, in

Figure 2: Example of 1900's Index with spanning rows

```
 v  Asia Bjorkman (838)   Ru bq  BuenosAyres Dec19 Liban       Ar Apr 22
*v  Asie LeCorfec (2452)   Fr bq  Port Talbot June13 Valparaiso
*   Askoy Morner (1545)    No s   Antwerp Mar 3    Broadmount(Qnsland)Ar June 9
*   Asnieres Touze (2715)  Fr bq  Clyde Feb 23     Noumea      Ar about June 16
1   Aspasia Elesio (556)   It bq  Cadiz Mar 26     Rio Grande  In FlorianopolisRoadsJne!3
*   Astoria Christensen                                                    [*June23
                (1027) No bq  Galveston Feb 25  BuenosAyresAr June 17
 v  Astraea Svane (228)    Da sc  Laguna June 18     ——
R v Astral Dunham (2987)   Am s   New York Apr 15  San FrnciscoSpMay20,17S38W—All well
*   Atacama Gundersen
                (1113) No bq  Boston June 13    BuenosAyros
G v Atalanta Stendahl
                (998) No bq   Launceston May11 MaldenIs& Bluff
*   Athene Dreier (2360)   Ge bq  Port Talbot Jan25 Iquique     Ar Apr 16—In pt June 27
 v  Atlantic Rasmussen
                (271) Da sc   Hamburg          Sundswall   Ar June 15
 v  Atlantic Kramer(145)   Du sc  Faro June 10     Gloncester
*   Atlantic Moller (1032) No bq  Mobile Feb 11 (cld) BuenosAyresAr May 21
N   Atlantic Lovik (1852)  No s   AtNorthport(NS)Jne24 Mersey  Loading
*v  Atlantique Becquet
```

addition to optical character recognition to recognize texts in documents, uses further machine learning to create tables. This paper specifically explores graph neural networks due to recent success (Qasim, Mahmood, and Shafait, 2019). As past work only explored these methodologies in tables created by Latex or HTML files, this paper is the first to explore this usage in historical documents. In particular, this paper provides a case where a graph neural network is fairly successful in obtaining data from historical tables where there are irregularities in texts as well as structures such as spanning rows.

The structure of the digitization architecture is illustrated in figure 3. This architecture has three main parts: construction of graphs (from image to graph using OCR), learning text representations (BERT and Graph Neural Network), and prediction of relationships (The final proponent using a neural network). This study uses the OCR tool, Google Cloud Vision, to construct graphs to extract texts and their location. By extracting locations, this study obtains the closeness and the direction between pairs of texts, which the later stage uses as edge attributes. The extracted texts are encoded via the BERT model, a prominent natural language processing model representing text well in many instances (Devlin et al., 2019). During training, each text is matched to hand-labelled data, creating a graph for each table. For learning text representations, this study uses a graph neural network to learn the graph's structure and outputs a numerical representation for each text extracted. For the final stage, predicting relationships between texts, this study uses a standard neural network given the text's representations and edge attributes. The output is the probability of whether the pair of texts belong to 1) the same row, 2) the same column, and 3) the same cell. The latter two stages use the PyTorch geometric library (Fey and Lenssen, 2019) built on PyTorch (Paszke et al., 2019) in order to test out multiple models.

This paper tests the feasibility of this digitization process by comparing hand-coded tables for each year in a different month. The detailed methodology, result and validation using historical facts are devoted in the appendix. In the end, this methodology was able to capture 60% of the total rows of shipping movements. Whether this is a good result is conditional on future progress on table structure recognition for historical documents. Yet, the general conclusion is that the data obtained is accurate enough to observe yearly port-to-port level tonnage as well as average duration. Since Lloyd's Index is separated by sailing and steamships, the extent to which we see the adoption of steamships is accurate.

Figure 3: Overview of Digitization Architecture
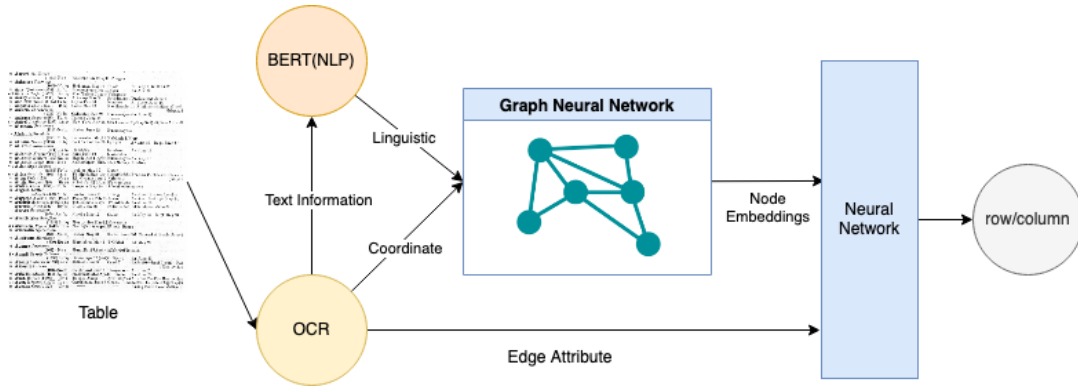
3.2 Country level data

The second data is trade. For this, this research uses data compiled from (Pascali, 2017) and (Fouquin and Hugot, 2016). The former is said to be the most comprehensive trade data for this period (1880 to 1900); it is a database for bilateral trade covering the second half of the nineteenth century constructed from primary sources and, overall, the data consist of almost 24,000 bilateral trade observations for nearly 1,000 distinct country pairs (Pascali, 2017). On the other hand, "TRADHIST" (Fouquin and Hugot, 2016) has an advantage in length due to the purpose of exploring the two modern waves of globalization: the First Globalization of the nineteenth century and the post-World War II Second Globalization (Fouquin and Hugot, 2016). Overall, they provided about 1.9 million bilateral trade observations from 1827 to 2014. After matching countries in the shipping data described in section 3.1, the former had more coverage; hence, this research primarily uses the data from (Pascali, 2017). This research uses the latter for analyzing the long-term and robustness checks.

In addition to trade volumes, this research also uses 344 entries on total exports and 154 entries on the share of exports in non-agricultural products (37 countries every 5 years from 1845 to 1905, with gaps) compiled by (Pascali, 2017). (Pascali, 2017) writes, "approximately half a million entries on exports by product were collected from primary sources; a SIC (rev1) code was then assigned to each of these products; and, finally, the share of exports that did not belong to the SIC categories 0 (food and live animals), 1 (beverages and tobacco), 4 (animal and vegetable oils and fat) was computed."

Lastly, this research uses various economic indicators to gain insights into how steamships affect development. This paper uses data compiled by (Pascali, 2017) and (Fouquin and Hugot, 2016). Summary statistics of the variables collected are described in the table 1. In addition to these variables, this research also used geographic information such as whether a pair of countries is contiguous and the latitude/longitude of countries (Pascali, 2017).

4 Changes in shipping, trade, and development

Using both port-level and country-level datasets, this paper first documents 1) how steamships were adopted, 2) the effect on countries' welfare, and 3) the change in port-level trade that leads to the theoretical framework.

Table 1: Variables for economic development

Name	Description	Source
GDP	Country wealth	*
Population	Country population	*, ***
Per Capita Income	Individual wealth	*, ***
Urbanization +50	% population in cities with more than 50,000	*
Urbanization +100	% population in cities with more than 100,000	*
Colony	Colonized by a country or not	*, **
Tariffs	Tariff rate for each trading partner or overall	**
Constraints on the executive	Institutionalized constraints on power by POLITY IV	*

* indicates (Pascali, 2017), **(Fouquin and Hugot, 2016), and *** (Bolt and van Zanden, 2014).
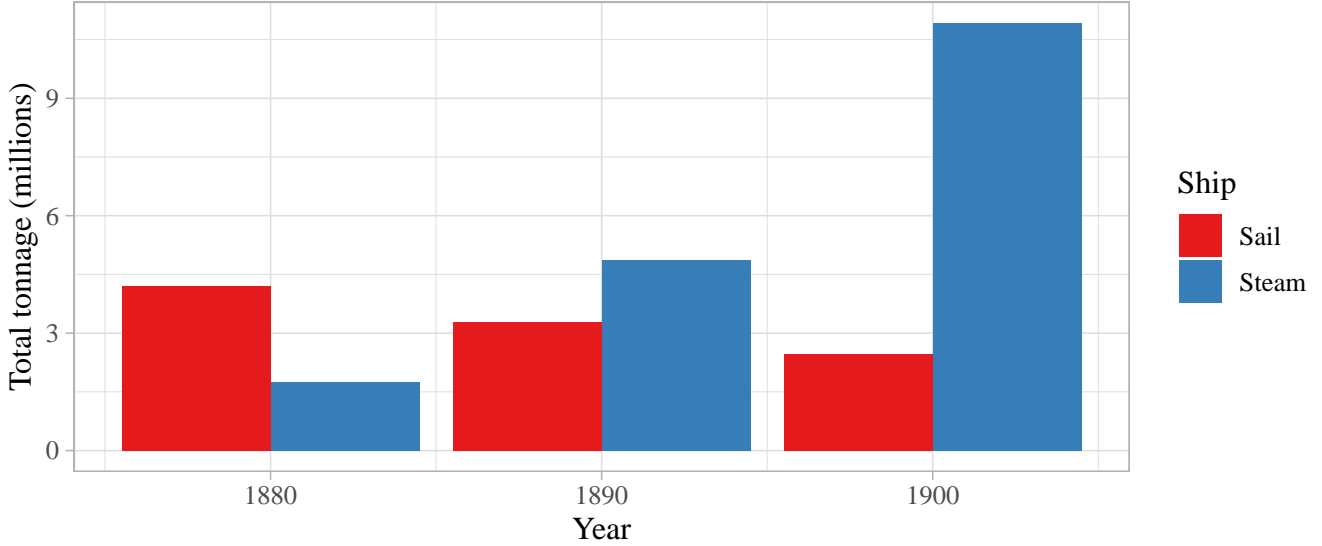

Figure 4: Total tonnage by year

4.1 Changes in geography

This paper is the first to provide a global analysis of the transition from sailing to steamships in a gravity model manner[1]. As such, this paper first presents a number of facts on the characteristics of steamships and how it was used.

4.1.1 Adoption of steamships

The Lloyd's Shipping Index provides us with the travel of ships categorized into either sailing or steamships. We can thus therefore see which travels (port to port) were most used by steamships. Note that the advent of steamships in international trade was in the 1860s. As the time period we have is from 1880, we still observe a number of steamships at this time. Figure 4 shows the total tonnage of ships travelled in the data set. The sailing ship was still the major mode of shipping in 1880 but steamship took over in 1890 and 1900. Steamship use exploded in 1900 surpassing 10 million tonnage travelled. Sailing ships, on the other hand, declined.

Figure 5 shows how this adoption differs by continent. First, the left figure shows that ships are predominantly involved in Europe or America (in particular North America). The share of tonnage going through European ports is more than half for all years and combined with

---

[1](Kosowska-Stamirowska, Ducruet, and Rai, 2016) and (Ducruet and Itoh, 2022) uses a month for each decade from 1880 to 1900 with no analysis on duration due to data limitation. Past research such as (Fletcher, 1958) only focused on a subset of developed countries.
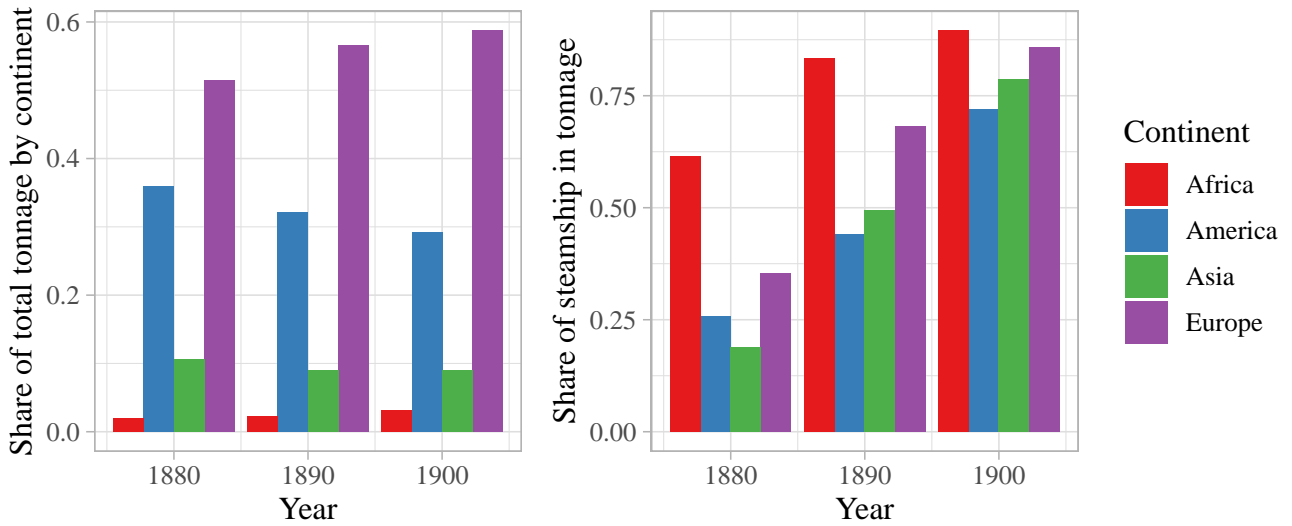
Figure 5: Tonnage and steamship share by continent

America's ports, the share goes up to about 80%. The adoption of steamships, on the other hand, is fairly similar across continents. This shows the share of the tonnage of steamships compared to the total tonnage that passed through ports in each continent. Across all continents, the share of steamships is trending upward. One noteworthy point is that steamship adoption was faster in Africa and Europe. The other continents, Asia and America, caught up in 1900 to a similar level. We see that the adoption of steamships in 1900 was above 75% from about 25% in 1880.

4.1.2 Steamships were faster and bigger

What drove the use of steamships? Figure 6 illustrates that steamships were both faster and bigger. Figure 6a shows the relative average duration of a steamship compared to the average duration of a sailing ship for the same origin-destination pair. Thus a value below 1 (the vertical dotted line) means that steamships were faster than sailing ships for a particular route. We can see that, in all but a few of the routes, steamships are faster to reach. In fact, the median duration change is 50%. In addition, figure 6b shows the density of tonnage of ships by ship type. We can see steamships were much bigger compared to sailing ships.

4.2 Changes in trade and development

How did this change in geography affect the trade and growth of countries? For this, this paper uses the exogenous change in duration induced by steamships. Specifically, this paper employs the market access approach (S. Redding and Venables, 2004) (Donaldson and Hornbeck, 2016).

First, this paper tests the assumption of exogeneity in duration change. The duration change came from various geographical characteristics, such as the wind and ocean currents. Yet, it is possible that, by chance, this duration is correlated with observable that affect trade and growth. Running a univariate regression of the relative change in duration for steamships on various port-level and country-level shows that the duration change can be considered good-as-random.

To estimate the impact of the steamship, this paper first runs the following regression

$$y_{c,t} = \alpha_t + \Delta duration_{c,1880} + \varepsilon_{c,t} \tag{1}$$

where $y_{c,t}$ is the outcome (GDP per capita) for country c at year t, α_t is the time-fixed effect and $\Delta duration_{c,1880}$ is the duration change for country c defined by
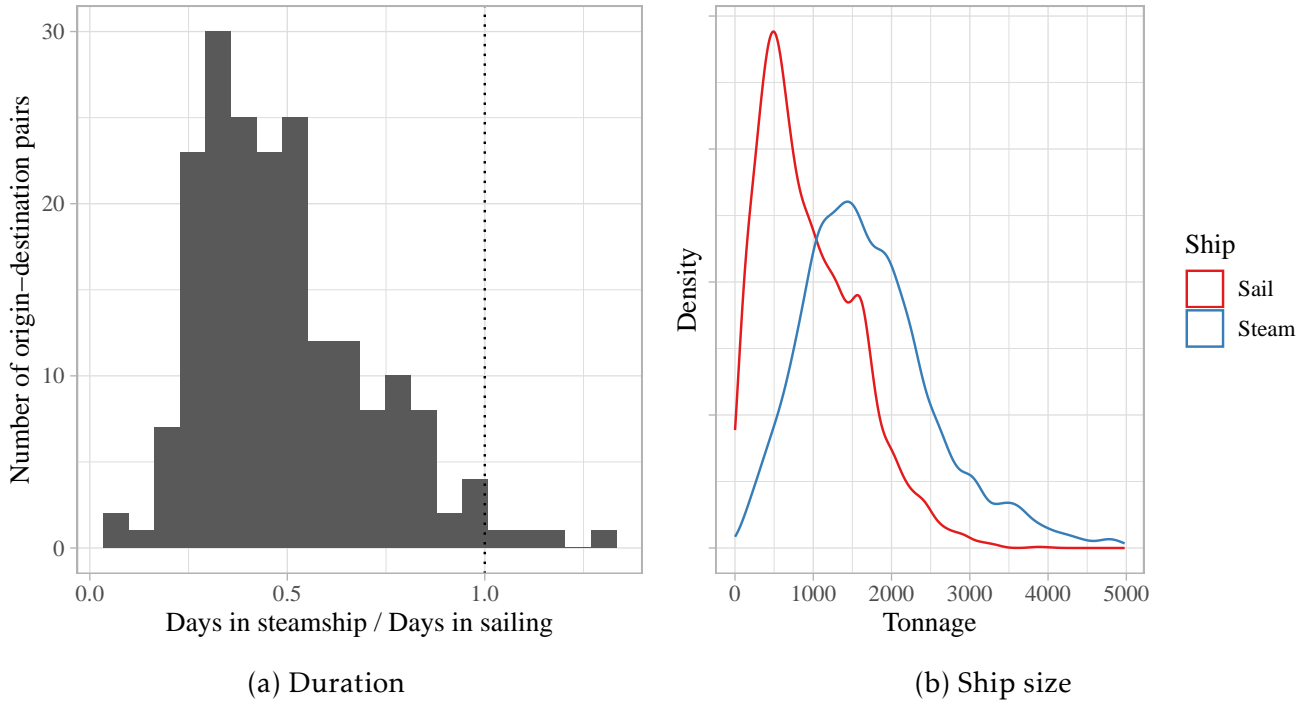
(a) Duration (b) Ship size

Figure 6: Differences between sailing and steamship

$$\Delta duration_{c,1880} = \sum_j (duration_{sail,1880} - duration_{steam,1880}) gdp_j \tag{2}$$

This is akin to (Pascali, 2017) but using the actual duration as a measure of closeness between country c and j. Column (1) of table 2 shows that being closer to the rest of the world led to positive growth in GDP per capita. As shipping duration was asymmetric (duration from country i to j was not the same for the duration from j to i, particularly for sailing ships), we can define a similar measure, but instead of the change in duration from country c but to the country c. Column 2 reports that while being closer to exporting had a positive effect, being closer to importing decreased welfare. The positive effect can be associated with better access to markets. If firms produce differentiated products and consumers love variety, firms now closer to other countries become more attractive. As those firms now offer a greater variety at lower prices to other countries, wages in that country go up. The negative effect is associated with lower-priced competition. With cheaper goods coming to one's own country, this increases competition and lowers wages. This negative effect, however, is counteracted by having cheaper and more varieties to consume. The results indicate that access to a larger market dominated the negative effect of opening up. Past research such as (S. Redding and Venables, 2004), (Donaldson and Hornbeck, 2016), and (Faber, 2014) used the market access approach to capture those effects precisely following the gravity model. Specifically, this paper defines market access and supplier access as

$$MarketAcess_{c,t} = \sum_j duration_{cj,t}^{1-\sigma} m_j$$

$$SupplierAcess_{c,t} = \sum_j duration_{jc,t}^{1-\sigma} s_i \tag{3}$$

This follows from the gravity equation (S. Redding and Venables, 2004) and here m_j and s_i are proxied by the country's GDP. This paper makes an important assumption to estimate

Table 2: Effect of change in duration

Dependent Variable:	ln(GDP per capita)		
Model:	(1)	(2)	(3)
Variables			
ln(Duration Change Exp)	0.728**	1.021**	
	(0.020)	(0.039)	
ln(Duration Change Imp)		-0.492*	
		(0.060)	
ln(Market Acess)			14.041***
			(0.000)
ln(Supply Access)			-4.651***
			(0.000)
Fixed-effects			
Year	Yes	Yes	Yes
Country			Yes
Fit statistics			
R^2	0.158	0.198	0.975
Observations	58	58	58

Clustered (Year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

σ, the trade elasticity with respect to trade cost. That is, trade cost is equal to duration. Past papers such as (Eaton and Kortum, 2002), (Donaldson, 2018), and (P. Fajgelbaum and Stephen J. Redding, 2022) use price information to estimate the elasticity. However, this was not possible in this paper due to data limitations. As such, this paper runs the conventional gravity equation with origin, exporter, and time-fixed effect with duration to estimate σ. Using PPML due to many trade flows with 0, the estimated σ is 1.02. This relatively small value is consistent with (Jacks and Pendakur, 2010) that documents low sensitivity of trade volumes with respect to freight rates. The value, however, does not change the magnitude of the subsequent results. Column (3) of table 2 presents the result. Consistent with using change of duration in 1880, larger market access benefits while larger supplier access decreases the country's welfare.

The positive result of the market access, however, only accrues to selected countries. In particular, table 3 shows that colonized countries lost out on this increase in world integration. Interacting both market access and supplier access with whether a country is a colony shows that colonized countries fared worse the more supplier access they had (column 2). This cannot be explained by the institution's quality, as in column 3, where the effect is not precise. On the other hand, countries in the top quantile in terms of GDP in 1850 (as in (Pascali, 2017)) benefitted from increased supplier access but not from market access. The conclusion from (Pascali, 2017) holds with this data that only a subset of countries seems to have benefitted from steamships. Yet, this paper points to the importance of colonial relationships in explaining the difference in welfare.

4.3 Changes in shipping pattern

What could drive the heterogeneous effect of market and supplier access? Using port-level data, this paper investigates the possible mechanism of steamships- specifically, the fact that ports

Table 3: Heterogeneous effect of change in duration

Dependent Variable:	ln(GDP per capita)			
Model:	(1)	(2)	(3)	(4)
Variables				
ln(Market Acess)	14.041***	-133.324	-56.009	133.000
	(0.000)	(117.167)	(33.204)	(174.775)
ln(Supply Access)	-4.651***	162.752	89.897	-102.799
	(0.000)	(107.995)	(24.760)	(143.361)
Colony		22.758		
		(13.294)		
ln(Market Acess) × Colony		299.054*		
		(31.956)		
ln(Supply Access) × Colony		-300.041*		
		(32.541)		
Institution			0.139	
			(0.805)	
ln(Market Acess) × Institution			22.490	
			(9.810)	
ln(Supply Access) × Institution			-22.490	
			(9.846)	
above top 25pc GDP				4.928
				(7.083)
ln(Market Acess) × above top 25pc GDP				-83.532***
				(0.813)
ln(Supply Access) × above top 25pc GDP				83.326***
				(1.119)
Fixed-effects				
Year	Yes	Yes	Yes	Yes
Country	Yes			
Fit statistics				
R^2	0.975	0.340	0.495	0.175
Observations	58	58	58	45

Clustered (Year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 4: Steamships impact on link destruction/creation

Dependent Variables:	Ship link destroyed	Ship link created
Model:	(1)	(2)
Variables		
Duration change between 1900 and 1880	-1.22***	1.65***
	(0.251)	(0.325)
Steamship advantage in 1900	-1.03***	0.782***
	(0.135)	(0.112)
Steamship advantage in 1880	-1.13***	0.342
	(0.319)	(0.362)
Fixed-effects		
Port origin	Yes	Yes
Port destination	Yes	Yes
Fit statistics		
Observations	6,314	6,314
R^2	0.5092	0.5475
Within R^2	0.0634	0.0395

Clustered (Port origin) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

must be invested in for steamships to be adopted. Colonized countries could benefit from increased steamships due to market access because ports in the destination country were easy for steamships to use and allowed them to export. On the other hand, high-GDP countries could utilize their ports to import goods much cheaper using steamships. The null effect for supply access for colonized countries and market access for high GDP countries could be rationalized because steamships could not be utilized due to low investment in ports in and outside their country. This is based on historical anecdotes that the price of goods exported to colonized countries did not change (Harley, 2015).

First, this paper shows the substantial impact of steamships on shipping patterns by simply analyzing whether shipping patterns were destroyed or created due to the introduction of steamships. For this, table 4 shows the following regression results.

$$I_{ij} = \alpha_0 + \alpha_1 \Delta_{ij} + \alpha_2 \Delta_{ij,1900}^{steam-sail} + \alpha_3 \Delta_{ij,1880}^{steam-sail} + \alpha_i + \alpha_j + \varepsilon_{ij,t} \quad (4)$$

Here, Δ_{ij} is the difference in duration from port i to j between 1900 and 1880. $\Delta_{ij,t}^{steam-sail}$ is the difference in duration by steamships minus that of sailing ships in year t (either 1880 or 1900). The outcome variable I_{ij} is either a dummy variable whether 1) there were no ships in 1900 while there were in 1880 (destroyed) and 2) ships in 1900 while there were no ships in 1880 (created). First, we can observe that links that did not benefit from a shorter duration time are likely to be destroyed (column (1)), while those that benefited from a shorter duration were created. Furthermore, links in which steamships were quicker than sailing ships in 1880 or 1900 were less likely to be destroyed and more likely to be created (column (2)).

Given this change, however, the adoption was not equal for all countries. Specifically, steamship adoption was slower for colonized countries. Table 5 describes the rate of adoption. Column 1 reports the share of ports that only have sailing, column 2 is for ports with steamships only, and column 3 is for the share of steamships in total tonnage. We can see that, compared to 1880, there are about 40 percentage points fewer ports only with sailing ships, about 30 percentage

Table 5: Changes in port and ship composition

Dependent Variables:	Only sailing	Only steam	Share steam
Model:	(1)	(2)	(3)
Variables			
Constant	0.649***	0.067***	0.177***
	(0.025)	(0.021)	(0.021)
Year 1900	-0.379***	0.279***	0.424***
	(0.032)	(0.028)	(0.027)
Year 1880 × Colony	0.050	0.010	-0.030
	(0.052)	(0.045)	(0.043)
Year 1900 × Colony	0.173***	-0.082**	-0.133***
	(0.044)	(0.038)	(0.037)
Fit statistics			
R^2	0.135	0.097	0.217
Observations	1,096	1,096	1,096

IID standard-errors in parentheses
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

points more ports with steamships only, and an increase of 40 percentage points in terms of share of steamships. However, focusing on ports in colonized countries, there were 17 percentage points more likely to have sailing ships, eight percentage points fewer ports with steamships only, and 13 percentage points less steamship share in 1900.

5 Steamship adoption: Theory

To further understand the differences in gains from steamships and the adoption rate, this paper uses the canonical (M. Melitz, 2003) with two different shipping technologies. Specifically, the marginal cost of using steamships and sailing ships is determined by the duration, but the fixed cost of using it is also different.

5.1 Trade with heterogeneous firms

This paper first illustrates how geography and fixed cost matter in the trade pattern observed. This follows (M. Melitz, 2003), (Chaney, 2008), and lecture notes by Treb Allen. This paper then builds upon this canonical model to analyze how steamships affects trade patterns and welfare.

5.1.1 Set up

World Denote the set of countries as S and denote $i \in S$ as an origin country and j as a destination country. In each country $i \in S$, there is an exogenous measure L_i of workers who supply a unit of labour inelastically. Let the wage of a worker in i w_i. Further, assume that labour is the only factor of production.

Demand A representative consumer in j gets utility U_j from the consumption of goods shipped by all other firms in all other countries with CES preferences over varieties.

$$U_j = \left(\sum_{i \in S} \int_{\Omega_{ij}} (q_{ij}(\omega))^{\frac{\sigma}{\sigma-1}} d\omega \right)^{\frac{\sigma-1}{\sigma}} \tag{5}$$

Here, $q_{ij}(\omega)$ is the quantity of ω shipped from country j and σ is the elasticity of substitution.

Supply For the supply side, following (Krugman, 1980), there is a continuum Ω of possible varieties the world can produce, and suppose that every firm in the world produces a distinct variety $\omega \in \Omega$. A firm uses $\frac{1}{\varphi}$ unit of labour to produce a unit of its variety. φ is drawn from a cumulative distribution function $G_i(\varphi)$ for firms in i. Let the set of varieties produced by firms in country i be denoted $\Omega_i \subset \Omega$. There is a mass M_i of firms from country i and that firm must incur a fixed cost $f_{ij} > 0$ to export to destination $j \in S$. In addition to the fixed cost, firms are subject to iceberg trade costs $\{\tau_{ij}\}_{i,j \in S}$. This means that destination i must ship $\tau_{ij} \geq 1$ for a good to arrive in j. Equality holds when the good is shipped to its own country $i = j$.

5.1.2 Obtaining the gravity equation

Demand The properties of the CES function imply that a consumer in country $j \in S$ demands good $\omega \in \Omega$ following

$$q_{ij}(\omega) = p_{ij}(\omega)^{-\sigma} Y_j P_j^{\sigma-1} \tag{6}$$

where Y_j is the income of country j and

$$P_j = \left(\sum_{i \in S} \int_{\Omega_i} p_{ij}(\omega)^{1-\sigma} d\omega \right)^{\frac{1}{1-\sigma}} \tag{7}$$

is the Dixit-Stiglitz price. To obtain the total trade flows, we aggregate across all firms in country i and since the amount spent on a variety ω from j is $p_{ij}q_{ij}(\omega)$, we get

$$X_{ij} = Y_j P_j^{\sigma-1} \int_{\Omega_i} p_{ij}(\omega)^{1-\sigma} d\omega \tag{8}$$

Supply A firm with productivity φ maximizes profits by solving

$$\max_{q_{ij}(\varphi)_{j \in S}} \sum_{j \in S} \left(p_{ij}(\varphi)q_{ij}(\varphi) - \frac{w_i}{\varphi}\tau_{ij}q_{ij}(\varphi) - f_{ij} \right) \quad \text{s.t. } q_{ij}(\varphi) = p_{ij}(\varphi)^{-\sigma} Y_j P_j^{\sigma-1}. \tag{9}$$

From the first order condition, the firm will set the price as follows conditional on selling to destination j

$$p_{ij}(\varphi) = \frac{\sigma}{\sigma-1}\frac{w_i}{\varphi}\tau_{ij}. \tag{10}$$

The total revenue becomes

$$x_{ij}(\varphi) = \left(\frac{\sigma}{\sigma-1}\frac{w_i}{\varphi}\tau_{ij} \right)^{1-\sigma} Y_j P_j^{\sigma-1} \tag{11}$$

and the profit is

$$\pi_{ij}(\varphi) = \left(p_{ij}(\varphi) - \frac{w_i}{\varphi}\tau_{ij}\right)q_{ij}(\varphi)$$

$$= \frac{1}{\sigma}x_{ij}(\varphi) \tag{12}$$

Trade flow   Denote $\mu_{ij}(\varphi)$ and M_{ij} to be the equilibrium probability density function of the productivities of and the number of firms from country i selling to j. Then the average price can be written as

$$
\begin{aligned}
\int_{\Omega_i} p_{ij}(\omega)^{1-\sigma}d\omega &= \int_0^\infty M_{ij}\left(\frac{\sigma}{\sigma-1}\frac{w_i}{\varphi}\tau_{ij}\right)^{1-\sigma}\mu_{ij}(\varphi)d\varphi \\
&= \left(\frac{\sigma}{\sigma-1}w_i\tau_{ij}\right)^{1-\sigma}M_{ij}\int_0^\infty \varphi^{\sigma-1}\mu_{ij}(\varphi)d\varphi \\
&= \left(\frac{\sigma}{\sigma-1}w_i\tau_{ij}\right)^{1-\sigma}M_{ij}\tilde{\varphi}_{ij}^{\sigma-1}
\end{aligned}
\tag{13}
$$

where $\tilde{\varphi}_{ij}$ captures the average productivity of products from i selling to j. From equation 8,

$$X_{ij} = \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma}(\tau_{ij}w_i)^{1-\sigma}M_{ij}\tilde{\varphi}_{ij}^{\sigma-1}Y_jP_j^{\sigma-1} \tag{14}$$

Selection in exporting   A firm in i will only export to j if and only if its profit is at least the fixed cost of entering, i.e. $\pi_{ij}(\varphi) \geq f_{ij}$. From the equilibrium revenue and profit, the threshold productivity of firms can be written as

$$
\begin{aligned}
\frac{1}{\sigma}\left(\frac{\sigma}{\sigma-1}\frac{w_i}{\varphi}\tau_{ij}\right)^{1-\sigma}Y_jP_j^{\sigma-1} &\geq f_{ij} \Leftrightarrow \\
\varphi \geq \varphi_{ij}^* &\equiv \left(\frac{\sigma f_{ij}\left(\frac{\sigma}{\sigma-1}w_i\tau_{ij}\right)^{\sigma-1}}{Y_jP_j^{\sigma-1}}\right)^{\frac{1}{\sigma-1}}.
\end{aligned}
\tag{15}
$$

Using the distribution of productivity, the trade flow becomes

$$X_{ij} = \left(\frac{\sigma}{\sigma-1}w_i\tau_{ij}\right)^{1-\sigma}M_i\left(\int_{\varphi_{ij}^*}^\infty \varphi^{\sigma-1}dG_i(\varphi)\right)Y_jP_j^{\sigma-1}. \tag{16}$$

This shows that as the threshold productivity of exporting firms decreases, we will observe equal or larger trade flows than before.

The Pareto distribution   (Chaney, 2008) uses the Pareto distribution to get crisp results on the extensive and the intensive margin. Suppose $\varphi \in [1,\infty]$ and

$$G_i(\varphi) = 1 - \varphi^{-\theta_i} \quad \theta_i > \sigma - 1. \tag{17}$$

Then the gravity equation becomes

$$X_{ij} = C_1(\tau_{ij}w_i)^{-\theta_i}f_{ij}^{\frac{\sigma-\theta_i-1}{\sigma-1}}M_i(Y_iP_j^{\sigma-1})^{\frac{\theta_i}{\sigma-1}} \tag{18}$$

where $C_1 = \sigma^{\frac{\sigma-\theta_i-1}{\sigma-1}}\left(\frac{\sigma}{\sigma-1}\right)^{-\theta_i}\frac{\theta_i}{\theta_i+1-\sigma}$

17

5.2 Gains from steamships

5.2.1 Introducing steamships

This paper now considers sailing and steamships in this canonical model. Let $s = \{sail, steam\}$ denote the type of ship used. The difference in ship type manifests in two ways. One is the change in the iceberg trade cost τ_{ij}. As steamships can move faster, $\tau_{ij}^{steam} \leq \tau_{ij}^{sail}$ in most cases. The second is the change in the fixed cost of entry. As steamships were bigger and also required additional investments (such as storage for coal and suitable geography) the fixed cost to arrive at the port was different by ship type. Thus, $f_{ij}^{steam} \neq f_{ij}^{sail}$. Which one is lower depends on port investment and thus, the year. For example, if i invested heavily in port infrastructure at year t, $f_{ij,t}^{steam} > f_{ij,t+1}^{steam}$. In this case, if the investment is large enough $f_{ij,t+1}^{steam} < f_{ij,t+1}^{sail}$.

Given this framework, firms will decide which mode to use. The profit using s and exporting is

$$\pi_{ij}^s(\varphi) = \frac{1}{\sigma}\left(\frac{\sigma}{\sigma-1}\frac{w_i}{\varphi}\tau_{ij}^s\right)^{1-\sigma} Y_j P_j^{\sigma-1} - f_{ij}^s \tag{19}$$

Let $f_{ij}^{steam} = \rho_{ij} f_{ij}^{sail}$ and $\tau_{ij}^{steam} = \xi_{ij}\tau_{ij}^{sail}$. Then the thresholds for exporting using sailing ships and steamships and the productivity in which the profit using steamships and sailing ships is equal are

$$\varphi_{ij,sail}^* = \left(\frac{\sigma f_{ij}^{sail}\left(\frac{\sigma}{\sigma-1}w_i\tau_{ij}^{sail}\right)^{\sigma-1}}{Y_j P_j^{\sigma-1}}\right)^{\frac{1}{\sigma-1}} \tag{20}$$

$$\varphi_{ij,steam}^* = \varphi_{ij,sail}^*\left(\rho_{ij}\xi_{ij}^{\sigma-1}\right)^{\frac{1}{\sigma-1}} \tag{21}$$

$$\varphi_{ij,steam=sail}^* = \varphi_{ij,sail}^*\left(\frac{\rho_{ij}-1}{\xi_{ij}^{1-\sigma}-1}\right)^{\frac{1}{\sigma-1}} \tag{22}$$

Given these productivity thresholds, we need to consider two cases. For simplicity, the origin destination is omitted. Importantly, this paper assumes $\rho \leq 1$, i.e., steamships trade cost is weakly lower in all origin destination links. This comes from the empirical fact that the duration is almost always less for steamships and the obtain tractability in estimating trade flow by ships.

(I) When using steamship is harder to export: $\varphi_{sail}^* \leq \varphi_{steam}^*$
In this case, it must be that $\xi \geq 1$ as $\rho \leq 1$. That is, the fixed cost is higher for steamships than sailing ships. Given this, we obtain $\left(\frac{\rho-1}{\xi^{1-\sigma}-1}\right) \geq \rho\xi^{\sigma-1}$. Therefore, the firms that exports and which ships they use will be as in figure 8. As shown, this implies that even if steamships can be used, due to lower profitability steamships are only used by firms that can satisfy the profit condition.

Figure 7: Export and ship used when cost of using steamship is high

18
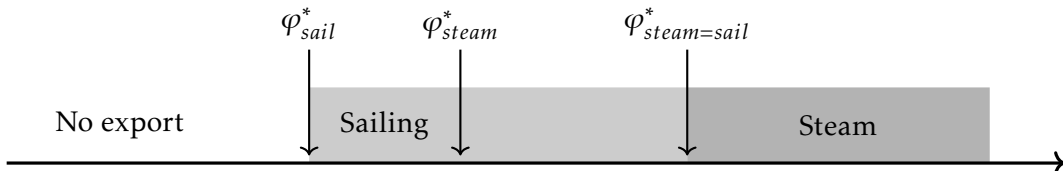
(II) When using steamship is easier to export: $\varphi^*_{sail} > \varphi^*_{steam}$

Here, we consider two cases: $\xi < 1$ and $\xi \geq 1$. $\xi < 1$ is when the profit condition is lower than the entry condition of steamships. Thus, firms with productivity above φ^*_{steam} all export using steamships. Similarly, when $\xi \geq 1$, since the entry condition implies that $\xi < \left(\frac{1}{\rho}\right)^{\frac{1}{\sigma-1}}$ this also means $\varphi^*_{steam=sail} < \varphi^*_{steam}$. Therefore, when steamship is easier to export, all firms with productivity above φ^*_{steam} exports using steamships.

Figure 8: Export and ship used when cost of using steamship is low

These cases illustrate how lowering the fixed cost for steamships have on export patterns. First, lowering the fixed cost allows firms that did not export before to export. Specifically firms with productivity $\varphi \in [\varphi^*_{steam}, \varphi^*_{sail}]$ now exports. Second, more firms can benefit from steamships. Those who could earn positive profit using steamships now fully uses steamships.

5.2.2 Gains

To compute the gains from trade, this paper considers two cases in period time $t+1$ from a state where only sailing ships are used at time t. First is changes in fixed cost f^{steam}_{ij} within case (I), i.e. when sailing ships are still used for all destinations. Another is a complete transition from sailing to steamships as in case (II). This is based on the empirical fact that while countries such as Great Britain mostly use steamships, other countries use both sailing and steamships. The key to computing welfare gains is the Price index P_i. With the reduction in trade cost, more firms will enter as profit is weakly better for every productivity. This, therefore, raises wages, and thus less productive firms will have to exit. This entry of firms can also be captured in the Price index as in (Bustos, 2011). To compute welfare, this paper makes a number of assumptions. First, welfare is computed in the symmetric equilibrium with n identical countries. Second, $f^{sail}_{ij,t} = f^{sail}_{ij} \, \forall t$. This is required to estimate the fixed cost. Third, productivity is Pareto distributed with shape parameter θ_i. To compute welfare, this paper introduces a free entry condition. Specifically, firms must pay an entry cost f_e to know their productivity φ and decide whether to exit or export. Free entry condition implies that subject to exogenous exit rate δ,

$$f_e = [1 - G(\varphi^*_{entry})]\frac{1}{\delta}\bar{\pi} \tag{23}$$

where $[1 - G(\varphi^*_{entry})]$ is the probability of survival and $\bar{\pi}$ is expected profits of surviving firm. After solving for f_e, we can compute the price index from the zero-profit condition for the marginal firm selling at a domestic market.

(I - Gain) When sailing ships are still usable

In this case, the increase in profit takes place for firms with productivity above $\varphi^*_{ij,steam=sail}$ and the number of firms exporting does not change. Therefore, the price will be given by

$$p_{ij,t+1}(\varphi) = \begin{cases} \frac{\sigma}{\sigma-1}\frac{w_i}{\varphi}\tau_{ij,steam}, & \text{if } \varphi > \varphi^*_{ij,steam=sail} \\ \frac{\sigma}{\sigma-1}\frac{w_i}{\varphi}\tau_{ij,sail}, & \text{otherwise} \end{cases} \tag{24}$$

19

Given this, the expected profit $\bar{\pi}$ is given by the sum of domestic profit, the probability of using a sailing ship and the profit when doing so, and the probability of using steamships and the profit associated with the use. Similar to the computation illustrated in the appendix of (Bustos, 2011), the price index will, therefore, be

$$P = \tag{25}$$

(II - Gain) When steamships are the only ship used

In this case, the price for all previously exported goods and a fraction of goods previously not exported gets affected.

Comparing the two, we can see that the welfare when steamship is easier to use is larger. This paper now turns to whether this fact explains the difference in welfare due to steamships.

6 Role of fixed cost in adoption: Estimation

6.1 Estimating the fixed cost of adoption

Here, this paper estimates the fixed cost of entry f_{ij}. This paper considers case (I) from the theory. To use the panel structure of the data, recall that the share of steamships in terms of quantity in period $t+1$ is given by

$$
\begin{aligned}
\mu_{ij} &= \frac{\int_{\varphi^*_{steam=sail}}^{\infty} q_{ij}(\varphi) dG(\varphi)}{\int_{\varphi^*_{sail}}^{\infty} q_{ij}(\varphi) dG(\varphi)} \\
&= \frac{\xi_{ij}}{1 - \left(\frac{\rho_{ij}-1}{\xi_{ij}^{1-\sigma}-1} \right)^{\frac{1}{1-\sigma}}}
\end{aligned}
\tag{26}
$$

In the data, we observe both μ_{ij}, the share of steamships in terms of tonnage, and ξ_{ij}, the relative change in trade cost under a number of assumptions. First, the tonnage of ships is sufficient to know the quantity sold from i to j. This is a restrictive assumption. It is possible that the ship contains a fraction of goods that will be transported to another place, and ships may not be used in full capacity proportional to their ship size. Another assumption is trade cost is proportional to duration. This is also a restrictive assumption as the marginal costs of sailing ships and steamships differ. For example, the iceberg cost itself is subject to change depending on the cost of coal. With additional considerations, such as endogenous freight rates, more work is needed to understand the full effect of steamships. This paper considers subsequent estimates as a baseline for future work. Given those assumptions, we can compute the relative fixed cost between sailing and steamships by solving equation 26. The resulting distribution of ρ_{ij} and the relative productivity cutoff for steamship and sailing ship is shown in figure 9a. The estimate shows that the difference in fixed cost is about 2% more for steamships.

Second, consider case (II). For this, this paper looks not between modes but between years and looks at trade links that switched from all sailing ships in 1880 to all steamships in 1900. The share of such change is about 50% of all trade links. The change in trade volume implies

$$\frac{x_{ij,t+1}}{x_{ij,t}} = \Delta w_i^{-\theta_i} \Delta Y_j^{\frac{\theta_i-1}{\sigma-1}} \Delta P_j^{\theta_i-1} \rho_{ij}^{\frac{\sigma-\theta_i}{\sigma-1}} \xi_{ij}^{-\theta_i} \tag{27}$$
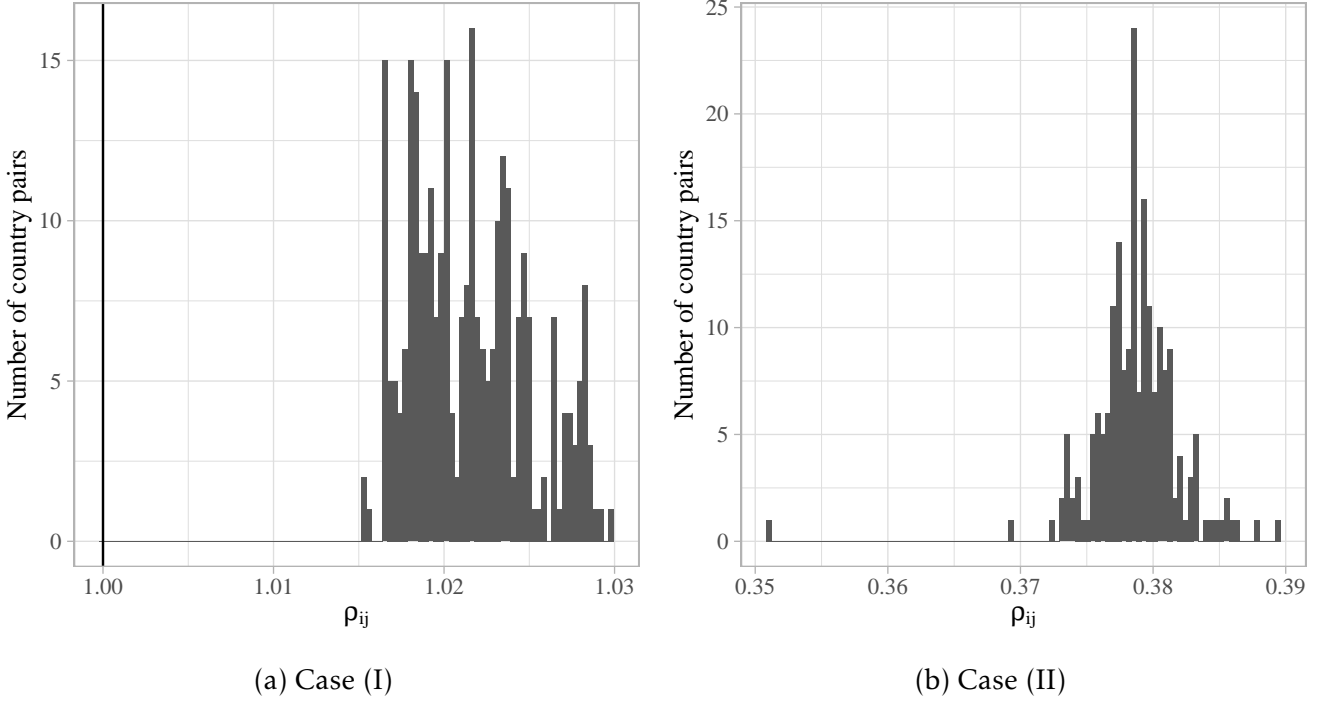
(a) Case (I)

(b) Case (II)

Figure 9: Distribution of ρ_{ij}

Taking the log gives us the following specification, which we can run using origin and destination fixed effects and setting the shape parameter, θ_i, of the Pareto distribution as 4.25 as in (M. J. Melitz and Stephen J Redding, 2015).

$$\log x_{ij,t+1} - \log x_{ij,t} = \alpha_i + \alpha_j - \theta_i \log \xi_{ij} + \frac{\sigma - \theta_i}{\sigma - 1} \log \rho_{ij} \tag{28}$$

Finally, this paper considers when links switch from using both sailing and steamship to only steamships. Similarly using the difference in quantity traded, we obtain

$$\log x_{ij,t+1} - \log x_{ij,t} = \alpha_i + \alpha_j + \log \frac{\rho_{ij}^{\frac{\sigma-\theta_i}{\sigma-1}} \xi_{ij}^{-\theta_i}}{(1 + \xi_{ij}^{-\sigma})\left(\frac{\rho_{ij}-1}{\xi_{ij}^{1-\sigma}-1}\right)^{\frac{\sigma-\theta_i}{\sigma-1}} - 1} \tag{29}$$

The resulting distribution of fixed entry costs relative to sailing ships for links fully transitioned to steamships are shown in figure 9b. We can see that even though the full transition is still possible with a higher ρ due to a decrease in variable cost (duration), the fixed cost is much lower than sailing ships. Specifically, the fixed cost is centred around 0.38, implying those ports are substantially easier for steamships to enter.

With these estimates, we can now test whether it is indeed the case that colonized countries have higher costs for steamships to enter. First, table 6 shows the correlation between colonized and countries in the top quartile of GDP in 1880 and the fixed cost of steamship relative to sailing. We can see that the fixed cost is generally higher for colonized countries while lower for high-income countries. This suggests the importance of the fixed cost in explaining the heterogeneous impact of opening up to trade. As the theoretical framework illustrated, even if variable cost falls, the country will not gain from increased openness if adoption is prevented by high fixed entry.

Table 6: Correlation between fixed cost and country characteristics

Dependent Variable:				log(rho)		
Year	Full sample	1880	1900	Full sample	1880	1900
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Constant	-1.063***	-0.507*	-1.323***	-1.081***	-0.212	-1.299***
	(0.169)	(0.245)	(0.199)	(0.216)	(0.196)	(0.240)
Colony	0.442	0.297	0.589*			
	(0.271)	(0.448)	(0.306)			
above top 25pc GDP				-0.135	-1.264**	0.600
				(0.529)	(0.311)	(0.865)
Fit statistics						
R^2	0.072	0.052	0.134	0.004	0.847	0.042
Observations	36	10	26	18	5	13

IID standard-errors in parentheses
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

6.2 Discussion

This section explored whether the differences in fixed cost affected steamship adoption and whether this could explain the differences in gains from changes in duration across countries. The key idea is that even if variable costs decrease, allowing countries to export and import more easily, some countries could not access this technology fully due to the high barrier to use. This paper discusses some hypotheses as to why this might be the case for steamships and shipping at large.

First is port investments. Similar to container shipping (Ducruet, Juhász, et al., 2020), ports for steamships required modernization. The decision to modernize and improve ports is endogenous. One would consider the effect on the local economy and allocation of land to other sectors. It also requires coordination among countries to adopt steamships. Including in the model whether to invest in ports and thus lower f_{ij}^{steam} will be able to close the model and provide a valid counterfactual analysis.

The second is markups in the transportation sector. A number of papers in recent years have shown (Djankov, Freund, and Pham, 2010) (D. Hummels, Lugovskyy, and Skiba, 2009) (D. Hummels, 2007) (Asturias, 2020) (Brancaccio, Kalouptsidi, and Papageorgiou, 2020) the importance of price setting by the shipping sector on trade volumes. Although this has been largely ignored in the past with the "iceberg" cost, the price setting is especially important in the case of new shipping technology. As historical data shows (Jacks, Meissner, and Novy, 2008), the adoption of steamships did not immediately lead to a decline in freight rates. Market power by the steamships likely influenced this result. If, for example, the entry of steamships were high for certain links, steamships could command a higher markup in this case. Thus, the endogenous price setting is affected by the fixed cost of entry and affects the variable trade cost.

Last is search friction. Even if variable trade costs decreased, colonized countries did not export more to other countries (shown in the appendix). In fact, for countries colonized by Great Britain, more ships travelled to Great Britain and ships owned by British companies were used more often. This suggests that past relationships before the adoption partially decided the pattern of trade afterwards. If the colonizers could enter the market disproportionately easier, this may allow them to take advantage of it.

These hypotheses illustrate the possible role of shipping and ports in promoting more trade and extortion through high entry barriers. This effect could be particularly pronounced in the age of disruptive technologies such as steamships. A further investigation into how this invention changed the shipping industry and, henceforth, trade patterns and inequality is a fruitful area of research.

7 Conclusion

This paper used the adoption of steamships to investigate what could drive unequal gains from a reduction in trade cost. Using the exogenous changes in duration due to the adoption of steamships, this paper first shows that increased integration to large markets benefits countries, but there are welfare losses when a country becomes closer to a large supplier. This effect, however, is driven by colonized countries. The hypothesis this paper tested is whether steamship adoption was unequal, leading to such results. For this, this paper implemented a novel deep-learning method to digitize historical shipping data en masse for the first time. Combining the canonical trade model with technology adoption, which varies in variable trade costs and fixed costs for exporting, this paper shows that colonized countries had a higher barrier to using steamships. This may have led to producers in those countries losing out on the benefits of market integration. The results and the data introduced have larger implications for inequality, technology adoption, and trade. The period of steamships coincides with the Great Divergence, where the Western civilization took off while other regions stagnated. How technologies affected this result and what would be a better way for trade benefits to be equal has important implications today.

References

Allen, Treb and Costas Arkolakis (Feb. 2022). "The Welfare Effects of Transportation Infrastructure Improvements". In: *The Review of Economic Studies*, rdac001.

Arkolakis, Costas, Sharat Ganapati, and Marc-Andreas Muendler (Oct. 2021). "The Extensive Margin of Exporting Products: A Firm-Level Analysis". In: *American Economic Journal: Macroeconomics* 13.4, pp. 182–245.

Asturias, Jose (Apr. 2020). "Endogenous Transportation Costs". In: *European Economic Review* 123, p. 103366.

Bernhofen, Daniel M., Zouheir El-Sahli, and Richard Kneller (Jan. 2016). "Estimating the Effects of the Container Revolution on World Trade". In: *Journal of International Economics* 98, pp. 36–50.

Blum, Bernardo S. et al. (Jan. 2019). "Inventory Management, Product Quality, and Cross-Country Income Differences". In: *American Economic Journal: Macroeconomics* 11.1, pp. 338–388.

Bolt, Jutta and Jan Luiten van Zanden (2014). "The Maddison Project: Collaborative Research on Historical National Accounts". In: *The Economic History Review* 67.3, pp. 627–651.

Brancaccio, Giulia, Myrto Kalouptsidi, and Theodore Papageorgiou (2020). "Geography, Transportation, and Endogenous Trade Costs". In: *Econometrica* 88.2, pp. 657–691.

Britannica (2023). *History of Latin America - Colonialism, Revolution, Independence — Britannica.* https://www.britannica.com/place/Latin-America/The-new-order-1850-1910.

Brooks, Leah, Nicolas Gendron-Carrier, and Gisela Rua (Nov. 2021). "The Local Impact of Containerization". In: *Journal of Urban Economics* 126, p. 103388.

Bustos, Paula (Feb. 2011). "Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinian Firms". In: *American Economic Review* 101.1, pp. 304–340.

Campante, Filipe and David Yanagizawa-Drott (Aug. 2018). "Long-Range Growth: Economic Development in the Global Network of Air Links*". In: *The Quarterly Journal of Economics* 133.3, pp. 1395–1458.

Chaney, Thomas (Sept. 2008). "Distorted Gravity: The Intensive and Extensive Margins of International Trade". In: *American Economic Review* 98.4, pp. 1707–1721.

Coşar, A. Kerem and Banu Demir (Sept. 2018). "Shipping inside the Box: Containerization and Trade". In: *Journal of International Economics* 114, pp. 331–345.

Coşar, A. Kerem and Pablo D. Fajgelbaum (Feb. 2016). "Internal Geography, International Trade, and Regional Specialization". In: *American Economic Journal: Microeconomics* 8.1, pp. 24–56.

Dai, Jifeng et al. (Oct. 2017). "Deformable Convolutional Networks". In: *2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 764–773.

Devlin, Jacob et al. (May 2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*.

Djankov, Simeon, Caroline Freund, and Cong S Pham (Feb. 2010). "Trading on Time". In: *The Review of Economics and Statistics* 92.1, pp. 166–173.

Donaldson, Dave (Aug. 2015). "The Gains from Market Integration". In: *Annual Review of Economics* 7.1, pp. 619–647.

— (Apr. 2018). "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure". In: *American Economic Review* 108.4-5, pp. 899–934.

Donaldson, Dave and Richard Hornbeck (May 2016). "Railroads and American Economic Growth: A "Market Access" Approach*". In: *The Quarterly Journal of Economics* 131.2, pp. 799–858.

Ducruet, César, Sylvain Cuyala, and Ali El Hosni (Jan. 2018). "Maritime Networks as Systems of Cities: The Long-Term Interdependencies between Global Shipping Flows and Urban Development (1890–2010)". In: *Journal of Transport Geography* 66, pp. 340–355.

Ducruet, César and Hidekazu Itoh (May 2022). "The Spatial Determinants of Innovation Diffusion: Evidence from Global Shipping Networks". In: *Journal of Transport Geography* 101, p. 103358.

Ducruet, César, Réka Juhász, et al. (Nov. 2020). "All Aboard: The Effects of Port Development". In: *National Bureau of Economic Research Working Paper Series* No. 28148.Journal Article.

Eaton, Jonathan and Samuel Kortum (Sept. 2002). "Technology, Geography, and Trade". In: *Econometrica* 70.5, pp. 1741–1779.

Faber, Benjamin (2014). "Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System". In: *The Review of Economic Studies* 81.3 (288), pp. 1046–1070.

Fajgelbaum, Pablo and Stephen J. Redding (May 2022). "Trade, Structural Transformation, and Development: Evidence from Argentina 1869–1914". In: *Journal of Political Economy* 130.5, pp. 1249–1318.

Fajgelbaum, Pablo D. and Amit K. Khandelwal (Aug. 2016). "Measuring the Unequal Gains from Trade *". In: *The Quarterly Journal of Economics* 131.3, pp. 1113–1180.

Feenstra, Robert C. and Hong Ma (2014). "Trade Facilitation and the Extensive Margin of Exports". In: *The Japanese Economic Review* 65.2, pp. 158–177.

Fey, Matthias and Jan Eric Lenssen (Apr. 2019). *Fast Graph Representation Learning with PyTorch Geometric*.

Feyrer, James (Oct. 2019). "Trade and Income—Exploiting Time Series in Geography". In: *American Economic Journal: Applied Economics* 11.4, pp. 1–35.

— (Nov. 2021). "Distance, Trade, and Income — The 1967 to 1975 Closing of the Suez Canal as a Natural Experiment". In: *Journal of Development Economics* 153, p. 102708.

Fletcher, Max E. (1958). "The Suez Canal and World Shipping, 1869-1914". In: *The Journal of Economic History* 18.4, pp. 556–573.

Fouquin, Michel and Jules Hugot (2016). "Two Centuries of Bilateral Trade and Gravity Data: 1827-2014". In: p. 39.

Frankel, Jeffrey A. and David H. Romer (June 1999). "Does Trade Cause Growth?" In: *American Economic Review* 89.3, pp. 379–399.

Galiani, Sebastian, Luis F. Jaramillo, and Mateo Uribe-Castro (Aug. 2023). *Market Access and Migration: Evidence from the Panama Canal Opening during the First Great Migration*. Working Paper.

Gatos, Basilios et al. (2005). "Automatic Table Detection in Document Images". In: *Pattern Recognition and Data Mining*. Ed. by Sameer Singh et al. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, pp. 609–618.

Hao, Leipeng et al. (Apr. 2016). "A Table Detection Method for PDF Documents Based on Convolutional Neural Networks". In: pp. 287–292.

Harley, Charles K. (Mar. 2015). "6. The Shift from Sailing Ships to Steamships, 1850-1890: A Study in Technological Change and Its Diffusion". In: *6. The Shift from Sailing Ships to Steamships, 1850-1890: A Study in Technological Change and Its Diffusion*. Princeton University Press, pp. 215–238.

Heiland, Inga et al. (2022). "Trade From Space: Shipping Networks and The Global Implications of Local Shocks *". In: Generic.

Helpman, Elhanan, Marc Melitz, and Yona Rubinstein (May 2008). "Estimating Trade Flows: Trading Partners and Trading Volumes *". In: *Quarterly Journal of Economics* 123.2, pp. 441–487.

Hummels, David (July 2007). "Transportation Costs and International Trade in the Second Era of Globalization". In: *Journal of Economic Perspectives* 21.3, pp. 131–154.

Hummels, David, Volodymyr Lugovskyy, and Alexandre Skiba (May 2009). "The Trade Reducing Effects of Market Power in International Shipping". In: *Journal of Development Economics* 89.1, pp. 84–97.

Hummels, David L and Georg Schaur (Dec. 2013). "Time as a Trade Barrier". In: *American Economic Review* 103.7, pp. 2935–2959.

Jacks, David S, Christopher M Meissner, and Dennis Novy (Apr. 2008). "Trade Costs, 1870–2000". In: *American Economic Review* 98.2, pp. 529–534.

Jacks, David S and Krishna Pendakur (2010). "Global Trade and the Maritime Transport Revolution". In: *The Review of Economics and Statistics* 92.4, p. 11.

Kasar, T. et al. (Aug. 2013). "Learning to Detect Tables in Scanned Document Images Using Line Information". In: *Proceedings of the 2013 12th International Conference on Document Analysis and Recognition*. ICDAR '13. USA: IEEE Computer Society, pp. 1185–1189.

Kosowska-Stamirowska, Zuzanna, César Ducruet, and Nishant Rai (Dec. 2016). "Evolving Structure of the Maritime Trade Network: Evidence from the Lloyd's Shipping Index (1890–2000)". In: *Journal of Shipping and Trade* 1.1, p. 10.

Krugman, Paul (1980). "Scale Economies, Product Differentiation, and the Pattern of Trade". In: *The American Economic Review* 70.5, pp. 950–959.

Melitz, Marc (2003). "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity". In: *Econometrica* 71.6, pp. 1695–1725.

Melitz, Marc J and Stephen J Redding (2015). "New Trade Models, New Welfare Implications". In: *THE AMERICAN ECONOMIC REVIEW* 105.3.

Okoye, Dozie, Roland Pongou, and Tite Yokossi (Sept. 2019). "New Technology, Better Economy? The Heterogeneous Impact of Colonial Railroads in Nigeria". In: *Journal of Development Economics* 140, pp. 320–354.

Pascali, Luigi (Sept. 2017). "The Wind of Change: Maritime Technology, Trade, and Economic Development". In: *American Economic Review* 107.9, pp. 2821–2854.

Paszke, Adam et al. (2019). "Automatic Differentiation in PyTorch". In.

Qasim, Shah Rukh, Hassan Mahmood, and Faisal Shafait (July 2019). *Rethinking Table Recognition Using Graph Neural Networks*.

Rashid, Sheikh Faisal et al. (Nov. 2017). "Table Recognition in Heterogeneous Documents Using Machine Learning". In: pp. 777–782.

Redding, Stephen and Anthony J. Venables (Jan. 2004). "Economic Geography and International Inequality". In: *Journal of International Economics* 62.1, pp. 53–82.

Redding, Stephen J. and Matthew A. Turner (2015). "Transportation Costs and the Spatial Organization of Economic Activity". In: *Handbook of Regional and Urban Economics*. Vol. 5. Elsevier, pp. 1339–1398.

Schreiber, Sebastian et al. (Nov. 2017). "DeepDeSRT: Deep Learning for Detection and Structure Recognition of Tables in Document Images". In: *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*. Vol. 01, pp. 1162–1167.

Shen, Zejiang et al. (2021). *LayoutParser: A Unified Toolkit for Deep Learning Based Document Image Analysis*.