All aboard: the effects of port development

César Ducruet, Juhasz Reka, David Krisztian Nagy, Claudia Steinwender

All aboard: The effects of port development*

César Ducruet (CNRS)

Réka Juhász (Columbia, NBER, and CEPR)

Dávid Krisztián Nagy (CREI and CEPR)

Claudia Steinwender (MIT Sloan, NBER, and CEPR)

March, 2021

Abstract

By using local land intensively, ports put pressure on land prices and crowd out other economic activity. Using the introduction of containerized shipping – a relatively land-intensive technology – we find an important role for this effect. At the local level, we find that the causal effect of the shipping boom caused by containerization on population is *zero* – port development increases city population by making a city more attractive, but this well-known market access effect is fully offset by the crowding-out mechanism. At the aggregate level, while we find overall welfare gains from containerization, our quantitative model featuring endogenous port development also implies i) sizeable welfare costs associated with the increased land-usage of ports, and ii) sizeable gains from cities' endogenous specialization across port- and non-port activities. These mechanisms are particularly important for targeted port development policies, which we illustrate using the Maritime Silk Road.

JEL: R40, O33, F6

Keywords: Port development, Containerization, Quantitative Economic Geography

*We thank Treb Allen, David Atkin, Leah Brooks, Arnaud Costinot, Don Davis, Dave Donaldson, Joseph Doyle, Nicolas Gendron-Carrier, Matt Grant, Gordon Hanson, Tarek Hassan, Tom Holmes, David Hummels, Amit Khandelwal, Giampaolo Lecce, Nels Lind, Nina Pavcnik, Giacomo Ponzetto, Jim Rauch, Steve Redding, Roberto Rigobon, Andrés Rodríguez-Clare, Esteban Rossi-Hansberg, Daniel Sturm, Tavneet Suri, Jaume Ventura, Jon Vogel and David Weinstein for helpful comments and discussions. We thank Bruce Blonigen and Mario Martín Antón and their co-authors for kindly sharing data. We thank staff at the Port Authorities of Houston, New Orleans, Long Beach, Los Angeles, Portland, San Francisco and Seattle for helping with our information requests. Olalekan Bello, Sabrina Chen, Naman Garg, Yi Jie Gwee, Hamza Husain, Felix Iglhaut, Rodrigo Martínez Mazza, Emanuela Migliaccio, Shuhua Si, Yue Yu, Howard Zihao Zhang and a team of Columbia University undergraduate students provided outstanding research assistance. César gratefully acknowledges financial support from the ERC (Starting Grant 'World Seastems' No. 313847). Réka is grateful for financial support from the Provost's Office at Columbia.

Introduction

From Sri Lanka to the Netherlands, countries across the income distribution invest heavily in port development.[1] Seaports play a vital role in the global trading system, handling over 80% of world merchandise trade in 2018 in terms of volume (UNCTAD, 2019). Rich and poor countries alike view investments into ports as an integral part of their growth strategy, as modern facilities allowing for the fast flow of cargo through the port are a precondition for a country to participate in global production networks (Rodrigue, 2016, p. 131). Despite this, ports have been understudied relative to other forms of transport infrastructure such as roads or railways.[2] In particular, little is known about the economic effects of port development. What determines the economic geography of ports (i.e., where port activity is located)? What are the gains from port development and how are they distributed across space?

In this paper, we study these questions by examining a major technological shock to port development: the introduction of containerization (the handling of cargo in standardized boxes). We find that cities exogenously more suited to this new technology witnessed a boom in shipping flows after the onset of containerization, but not before. Surprisingly, however, this boom in local shipping *did not* translate into population inflows: we find an effect of shipping on population in our IV estimates that is both economically and statistically insignificant. To conduct this analysis, we use a unique dataset of city populations and shipping flows worldwide for the period 1950-1990 to estimate the local, city-level effects of containerization. To isolate exogenous variation, we build on a previous literature that has shown that access to deep sea ports was an important determinant of a city's suitability for containerization (Brooks et al., 2019; Altomonte, Colantone, and Bonacorsi, 2018). We develop a novel measure of 'naturally endowed' depth (as distinct from depth attained by dredging) using granular data on oceanic depths around each city in our data.

We view the zero local population effects of containerization as an unexpected finding. It is in contrast to standard models that predict an inflow of population as improved market access makes a location more desirable for firms and consumers (Coşar and Fa-

[1] For example, the Port of Rotterdam (Netherlands) undertook the expansion of its container facilities by 110 ha in 2004 at a cost of EUR 657m, 200m of which was financed by the European Investment Bank (Source: https://www.eib.org/en/projects/pipelines/all/20030288). The Port of Colombo (Sri Lanka) has made massive investments in recent years. A single project upgrading harbor infrastructure was undertaken between 2008-2012 at a cost of Rs 42 billion (Source: https://www.slpa.lk/port-colombo/projects).

[2] Redding and Turner (2015) provide an overview of this literature. An exception is Brooks, Gendron-Carrier, and Rua (2019), who study the reduced-form effects of containerization on county-level economic outcomes in the U.S.

jgelbaum, 2016; Nagy, 2020b; Fajgelbaum and Redding, 2018). Indeed, other papers studying similar shocks to a location's accessibility have found a positive effect on population (Bleakley and Lin, 2012; Campante and Yanagizawa-Drott, 2018; Brooks et al., 2019).

We argue that the zero local population effects are driven by a *crowding-out* mechanism. Different to other transport infrastructures such as railways or roads, ports are investments that occupy large amounts of land in the cities in which they are located. Today, the median port in our worldwide sample occupies space equivalent to 250 soccer pitches, while the port at the top decile occupies 1,100.[3] By using locally scarce land resources heavily, ports may drive up land rents and crowd out other economic activity.

Using rich historical evidence, as well as systematic contemporary data on port area and cargo composition, we show that containerization is indeed a much more land-intensive technology than the one it replaced. Our analysis suggests that moving from a fully non-containerized port to a fully containerized one requires 75% larger land area (holding the volume of traffic constant). As such, the higher land intensity of containerized port technology can provide an explanation for the zero population effect. Intuitively, the increased use of scarce local land can counteract the market access effect by driving up land prices and crowding out other economic activity from the city. We provide empirical evidence for this mechanism by showing that after containerization, shipping increased disproportionately more in low land rent cities.

Of course, the land-price mechanism is not the only force that can lead to the crowding out of population. We consider three other prominent mechanisms that could account for our findings; i) the lower labor intensity of containerized port technology, ii) pollution and other disamenities associated with port development, and iii) additional decreases in overland transportation costs caused by containerization. We show that these either do not stand up to more rigorous examination, or are quantitatively too small to explain our results.

Informed by the local, reduced form effects of port development, in the second part of the paper we develop a model to study the effects of port development in general equilibrium. The model is an otherwise standard economic geography model of trading cities to which we add an endogenous port development decision. As such, the model incorporates not only the standard market access effect, but also allows for port development to crowd out other forms of economic activity. This is because in the model,

[3] We use high resolution remote sensing data from Google Earth to delineate the area occupied by a random subset of 236 ports in our dataset. We discuss the methodology in detail in the Supplementary Material.

developing the port (and hence reducing trade costs) requires scarce local land that can be used for other purposes. Whether a city ultimately gains in population is the outcome of the trade-off between the market access and crowding-out mechanisms. Thus, the model has the ability to rationalize the zero population effects of shipping found in the data.

Guided by the model, we re-estimate the causal effect of increased shipping flows on population controlling for market access. In line with the model predictions, our causal estimates point to a *negative* effect of shipping on city population once market access is controlled for. This finding provides further empirical evidence consistent with the crowding-out effect of port development.

Next, we examine the aggregate effects of port development by taking the model to the data. We use data on shipping flows, city GDP and population in 1990 to back out cities' unobserved model fundamentals. We conduct two counterfactual simulations. In our first counterfactual, we simulate the pre-containerization equilibrium in the model by *undoing* the containerization shock. We show that the model-simulated data closely match the zero population effects of shipping using the same IV strategy (based on depth) as in the reduced form. Furthermore, we show that containerization increased shipping more in low land-rent cities, as in the data.

Our estimates suggest that containerization increased world welfare by 3.84%. To better understand how the crowding-out channel affects these welfare gains, we compare the aggregate welfare effects in our model to what a standard model in which transport cost reductions are *exogenous* and *free* (i.e., they do not use scarce resources) would predict. We find a quantitatively meaningful role for two mechanisms. First, we estimate the aggregate resource cost of containerization to be substantial: it reduces the welfare gains arising from a standard model by about 16%. Second, we also find a role for additional welfare gains stemming from endogenous specialization in port- and non-port activities based on comparative advantage. In particular, these gains offset about 58% of the resource cost of containerization. In addition, we find that, unlike in our baseline model, the local population effects of shipping are positive, economically meaningful and statistically significant in the standard model. This result again underscores the link between the zero local population effects of shipping and the endogenous crowding-out mechanism that is absent from standard models.

In our second counterfactual, we examine the effects of targeted port-development policies. We focus on a setting similar to the 'Maritime Silk Road' project – a large set of port investments currently being undertaken by China in South-Asian, African and European ports. Our findings suggest that targeted port development has the potential for large distributional effects triggered by the reallocation of shipping activity.

Most strikingly, we predict a large decline in shipping in Singapore (a non-targeted port which we estimate loses about 50% of its shipping flows), which is driven by the fact that shipping activity reallocates to nearby, targeted ports. The initial shock is then amplified by less endogenous port development in Singapore as demand for port services falls, illustrating the increasing returns to scale at work in our model. However, despite losing a sizeable fraction of its shipping flows, Singapore gains 1% in GDP, as resources reallocate to Singapore's highly productive non-port activities. This illustrates that, because of the resource cost of port development, gains and losses in shipping do not translate directly into gains to real GDP. These findings highlight the importance of accounting for our endogenous port development mechanism when quantifying how the gains from targeted port development are distributed across space. More speculatively, they question the wisdom of highly productive, expensive cities such as Hong Kong and Singapore continuing to specialize heavily in port services.

Related literature. A recent, growing literature provides evidence that better trading opportunities lead to local benefits inducing city development (Bleakley and Lin, 2012; Armenter, Koren, and Nagy, 2014; Nagy, 2020a; Campante and Yanagizawa-Drott, 2018). Some of these studies focus on city development at port locations in particular (Fujita and Mori, 1996; Coşar and Fajgelbaum, 2016; Fajgelbaum and Redding, 2018). We contribute to this literature by showing that trade-induced development can also have substantial local costs. The crowding-out mechanism that drives the cost side in our setting also relates the paper to the 'Dutch disease' literature. This literature shows that booming industries can entail significant costs by putting a strain on scarce local resources and therefore crowding out other (tradable) sectors (Corden and Neary, 1982; Krugman, 1987; Allcott and Keniston, 2017).[4] Relative to this literature, our setting contains the potential for not only costs but also gains, as booming port activities benefit local tradables through improving market access. Thus, one contribution of our paper is to generalize the predictions from these two, seemingly disparate literatures that have focused on either the costs or the benefits from booming sectors.

Our paper is also related to the quantitative international trade literature, which has developed tractable models of trade across multiple countries with various dimensions of heterogeneity (Anderson, 1979; Eaton and Kortum, 2002; Melitz, 2003). These seminal models characterize trade and the distribution of economic activity across countries as a function of exogenous trade costs. A standard prediction of these models is that the relationship between trade flows and costs follows a gravity equation, which has been documented as one of the strongest empirical regularities in the data (Head and Mayer,

[4] Another related paper is Falvey (1976), who discusses how the transportation sector can draw away resources from tradables in particular.

2014). We complement this literature by developing a framework in which trade costs are *endogenous*, in a way that is both tractable and preserves the gravity structure of trade flows. This relates our paper to Fajgelbaum and Schaal (2020) and Santamaría (2020), who consider endogenous road construction in multi-location models of economic geography, as well as Brancaccio, Kalouptsidi, and Papageorgiou (2020), who endogenize trade costs in the non-containerized shipping sector. Unlike these papers, we focus on port development as a source of endogenous shipping costs, and solve for the decentralized equilibrium as opposed to the optimal allocation to quantify the effect of port development on trade, the distribution of population, and welfare.

Finally, our paper is related to a large literature studying the effects of transport infrastructure improvements.[5] In particular, there is a growing empirical literature studying the effects of containerization (Hummels, 2007; Bernhofen, El-Sahli, and Kneller, 2016; Gomtsyan, 2016; Coşar and Demir, 2018; Holmes and Singer, 2018; Altomonte et al., 2018; Brooks et al., 2019) or the role of container shipping networks in world trade (Wong, 2020; Heiland, Moxnes, Ulltveit-Moe, and Zi, 2021; Ganapati, Wong, and Ziv, 2020). Most closely related is Brooks et al. (2019), who study the reduced-form effects of containerization on local economic outcomes across U.S. counties. Our main contribution to this literature is twofold. First, motivated by the evidence that containerization dramatically increased land use in ports, this paper highlights the crowding-out effect of containerization and finds sizeable local and global costs stemming from this effect. Second, to the best of our knowledge, this is the first paper seeking to quantify the aggregate effects of port development on global trade and welfare through the lens of a general equilibrium economic geography model.

The paper is structured as follows. In the next section, we describe the main features of containerized technology. Section 2 discusses the main data sources used in the analysis. Section 3 presents the reduced form empirical strategy and results, while Section 4 introduces the model. Section 5 revisits the empirics guided by the predictions of the model. Section 6 measures the aggregate effects of containerization and Section 7 considers alternative explanations for the crowding-out mechanism. In Section 8, we illustrate the effects of targeted port development policies similar to the 'Maritime Silk Road'. Finally, Section 9 concludes.

1 The increased land-intensity of containerization

The introduction of steamships and railroads in the 19th century substantially reduced both water and overland transportation costs. However, transshipment remained slow and expensive through the middle of the 20th century (Krugman, 2011). As a report by

[5] Redding and Turner (2015) provides an overview of recent developments in this literature.

McKinsey highlighted; “The bottleneck in freight transport has always been the interface between transport modes, especially the crucial land/sea interface” (1972, pp. 1-3). Containerization, that is, the handling of cargo in standardized boxes, was the breakthrough innovation that dramatically reduced transshipment times and costs (Hummels, 2007; Rodrigue, 2016). In this section, we show that while substantial transshipment cost reductions were achieved in shipping as a result of containerization, this came at the cost of needing to dedicate much more land to the port.

1.1 The cost – space trade-off in containerization

As late as the mid-1950s, transshipment at seaports was a costly and slow procedure as it entailed handling cargo item-by-item – a process called breakbulk shipping. The reason for this was that cargo came in many different sizes and needed to be handled individually, despite the widespread use of machinery introduced pre-containerziation (see Panel A of Appendix Figure C.1). The San Francisco Port Commission (1971) estimated that it took 7 to 10 days to merely discharge cargo from a ship using this technology. According to Bernhofen et al. (2016), two-thirds of a ship’s time was spent in port. This led to high costs as the capital utilization of ships was low, and the cost of capital tied up in inventory was high.[6]

U.S. shippers first started placing cargo into containers in the late 1950s.[7] Containerized port technology can be seen in its mature form at the Port of Seattle in 1969 in Panel B of Appendix Figure C.1 (a mere 10 to 15 years after the photos shown in Panel A were taken). Cargo, packed in standardized containers, is loaded onto and off ships using large, purpose-built cranes situated on the wharf. Large, open areas beside the wharf are used to line up containers.

Containerization substantially reduced transshipment costs for a number of reasons. First, as containers could be handled in a uniform way, loading and unloading times were vastly reduced. The San Francisco Port Commission (1971) estimated that a container ship could be unloaded and loaded in 48 hours or less, a tenth of the previous time spent in port. Similarly, using detailed data on vessel turnaround times for one anonymized port, Kahveci (1999) estimates that the average time ships spent in port fell from 8 days to 11 hours as a result of containerization, a reduction of 94%. Second, the reduction in turnaround time justified investment in much larger vessels (Gilman, 1983). The average size of newly-built container ships increased by 402% between

[6]Industry experts estimated that the handling of cargo at the port accounted for a major share of freight costs (Levinson, 2010). As an example, transshipment costs were estimated to account for 49% of the total transport cost on one route from the U.S. to Europe (Eyre, 1964).

[7]Containerized shipping was initially introduced on domestic routes between U.S. ports, but the technology was rapidly adopted and standardized worldwide in 1967 (Rua, 2014).

1960 and 1990.[8] Larger ship sizes made it possible to realize even larger cost reductions through increasing returns to scale in shipping and port handling. Rodrigue (2016, p. 118) estimates that moving from a 2,500 TEU capacity vessel to one with 5,000 TEU reduced costs per container by 50%.

Adapting ports to containerized technology was not without costs, however. Most importantly, faster turnaround times could only be achieved at the cost of building much larger terminals. In discussing the 'challenges' associated with containerization, Rodrigue (2016, p. 118) puts site constraints in the first place, and in particular, the large consumption of terminal space. Containerized terminals need more space as the easy accessibility of containers allows for efficient on- and off-loading. Containers are lined up next to where the ships dock, and space is also needed to rapidly off-load cargo. The increased space requirements of containerized facilities were evident from the start. For example, in a 1971 report, alarm bells were being raised about the inadequacy of San Francisco's finger piers to accommodate new types of cargo handling; "No pier facilities in the Bay Area today are capable of handling the new space requirements on this scale of new and larger container ships. (...) thus more berthing and backup area is needed" (1971, p. 13). Ports in densely built up areas such as Manhattan and San Francisco were almost certainly doomed to decline as one observer noted for San Francisco; "Rows of finger piers adjacent to a densely built up city could not adequately serve container shipping, which involved larger ships that required larger wharves and much larger areas of open space for loading and unloading" (Corbett, 2010, p. 164).

1.2 Evidence for the increased land-intensity of containerization

We present two pieces of quantitative evidence that point to a substantial increase in the land intensity of port technology with the introduction of containerization. First, for the Port of Seattle, we are able to measure the area of the port per volume of traffic handled around the time of containerization.[9] Between 1961 and 1973, the port built a number of new containerized facilities, increasing the area of the port almost fourfold. Consistent with these investments, by 1973, containerized cargo accounted for 43% of total traffic. While the total volume of traffic handled more than doubled, we calculate that area relative to throughput increased by 90%. This suggests that containerized

[8] These calculations are based on data from the *Miramar Ship Index* (Haworth, 2020). More details on these data are provided in the Supplementary Material.

[9] More precisely, we use engineering maps for all properties under the ownership of the Seattle Port Authority to calculate the area of the port (excluding the airport and other land not used for seaport activities) and divide by the five-year moving-average of throughput to smooth out year-to-year fluctuations in capacity utilization. A more detailed discussion, including information on data sources, is included in the Supplementary Material.

technology requires substantially more land per unit of cargo shipped.

Second, we examine the relationship between port area and containerized cargo volume across a large set of cities in our sample. We exploit the availability of high resolution remote sensing data that makes it possible to measure the area of ports in our sample today.[10] We match this with data on the cargo composition handled by each port using data from *Le Journal de la Marine Marchande (JMM)* for 2008-09.[11] Column (2) in Table B.1 shows that, controlling for the total volume of traffic, ports that handle more containerized cargo are typically larger. In columns (3)-(6), we control for additional potential confounders. The coefficient remains significant and very similar in magnitude when we add controls for the volume of non-liquid and solid bulk handled by the port (e.g., oil, coal or grain, the handling of which may be technologically very different), the GDP per capita of the country (to get at differences in the extent of automation), or even country fixed effects.

To get a better sense of magnitudes, in column (7) we use the *share* of containerized cargo (while continuing to control for the total volume of cargo). The coefficient of interest is large and statistically significant. Based on this specification, moving from a fully non-containerized port to a fully containerized port is associated with a 75% increase in port area (i.e., $exp(0.5624) - 1$), holding the volume of traffic fixed.[12]

In summary, the historical and quantitative evidence paint a consistent picture. The cost reductions containerization makes possible through faster transshipment times can only be achieved at the expense of dedicating more land per unit shipped to the port.

2 Data

Our analysis builds on a decadal city-level dataset of shipping flows, population, and other economic outcomes for the period 1950-1990. We complement this with GIS data that allows us to calculate geographic characteristics of the city. We review the main variables used in the analysis below and report summary statistics in Appendix Table B.2. Additional details for all the data used in the paper can be found in the Supplementary Material.

Shipping flows. Crucial to our analysis is a dataset of worldwide bilateral ship movements at the port level for the period 1950-1990 from Ducruet, Cuyala, and Hosni

[10] In particular, for a random set of cities in our dataset, we hand-coded polygons from *Google Earth* that we identified as containing port activities. A more detailed description of data and methodology can be found in the Supplementary Material. Unfortunately, the resolution of historical satellite images is not sufficiently high to replicate this exercise for our sample period.

[11] These data were not available for a more recent year.

[12] The binscatter (plotted in Appendix Figure C.2) visualizes the positive relationship between the area occupied by the port and the share of containerized cargo.

(2018). An observation is a ship moving from one port to another at a particular point in time.[13] One week samples of these data were extracted from the *Lloyd's Shipping Index*, a unique source that provides a daily list of merchant vessels and their latest inter-port movements.[14] We are aware of no previous application in the economics literature.

These data provide us with rich variation to study the geography of sea-borne trade through the second half of the 20th century. They cover both domestic and international shipping. Moreover, the data cover a long time period spanning the containerization revolution. We are thus able to compare the effects of port activity on cities both before and after the arrival of the new technology. We know of no other data source that has a similar coverage across time and space, especially at such a detailed level of disaggregation. An important limitation, however, is that we do not observe either the value or the volume of shipment but only bilateral ship movements. From these ship movements, we sum the total number of ships passing through each port, which we call *shipping flows*.

City population. As we are interested in the economic effects of containerization, we use data on city population worldwide for locations with more than 100,000 inhabitants from *Villes Géopolis* (Moriconi-Ebrard, 1994) for each decade between 1950-1990 (Geopolis cities, henceforth). The advantage of these data relative to sources such as the more frequently used *UN World Cities* dataset is that a consistent and systematic effort was made to obtain populations for the urban agglomeration of cities (that is, the number of inhabitants living in a city's contiguous built-up area) as opposed to the administrative boundaries that are often reported in country-specific sources. For example, New York (New York) and Newark (New Jersey) form one 'city' according to this definition. We observe population for cities that reached 100,000 inhabitants in any year throughout this period. For most of these cities, we observe population even when the city had fewer than 100,000 inhabitants, leading to potentially important sampling bias. To address this, we will show that our results are robust to using the subset of cities that had already attained 100,000 inhabitants in the first sample year, 1950.

Ports were hand-matched from the shipping data to cities based on whether the port was located within the urban agglomeration of a city in the Geopolis dataset, allowing for multiple ports to be assigned to one city (Ducruet et al., 2018). We define port cities in a time invariant manner; a port city with positive shipping flows in at least one year

[13]As such, it is similar to contemporary satellite AIS (Automatic Identification System) data that tracks the precise movements of vessels around the globe. These type of AIS data are used in Heiland et al. (2021) and Brancaccio et al. (2020).

[14]The data were entered from issues for the first week of May. The data are discussed in more detail in Ducruet et al. (2018).

will be classified as a port city for all years. Of the 2,636 cities in the Geopolis dataset, 553 have at least one port. We label these as *port cities*. The quantitative estimation covers the full set of 2,636 Geopolis cities (port and non-port cities).

Underwater elevation levels. We use gridded bathymetric data on underwater elevation levels at a detailed spatial resolution (30 arc seconds, or about 1 kilometer at the equator) from the *General Bathymetric Chart of the Oceans (GEBCO)* to measure sea-depth around the city.

Saiz land rent proxy. While we are not aware of any dataset that covers land rents globally going back to the 1950s, Saiz (2010) has proposed a geography-based measure that correlates well with land-rents. This allows us to construct land rent proxies for all cities in our dataset. The 'Saiz-measure' is defined as follows: Take a 50 kilometer radius around the centroid of the city. Exclude all sea cells, all internal water bodies and wetland areas, as well as all cells with a gradient above 15%. The remaining cells, as a share of the total cells, can be used as a proxy for land rents. We replicate the methodology in Saiz exactly, using GIS data that have global coverage.

City-level GDP per capita. Data on city-level income levels are needed for the quantitative estimation only. We are not aware of readily available sources of GDP per capita data for cities worldwide. For this reason, we estimate GDP per capita for the last year in our sample (1990) for the full sample of 2,636 worldwide cities in the following way. First, we use estimates of city GDP from the *Canback Global Income Distribution Database* for a subset of our sample (898 cities) for which data are reported for 1990. We extrapolate GDP per capita for the full sample of cities using the linear fit of the GDP per capita data on nightlight luminosity and country fixed effects, building on a growing body of evidence suggesting that income can be reasonably approximated using nightlight luminosity data (Donaldson and Storeygard, 2016).

3 The reduced form effects of containerization

In this section, we study the local effects of containerization on port cities. To isolate the causal effect, we first develop an exogenous measure of port suitability, and then proceed to discussing our main reduced form findings.

3.1 An exogenous measure of port suitability

Section 1 discussed the fact that containerization led to larger ship sizes, and that this in turn required greater depth at the port. Following the previous literature, we think of *naturally endowed* depth as an exogenous cost-shifter that makes it cheaper for a port to reach a desired depth through costly dredging (Brooks et al., 2019; Altomonte et al., 2018). The empirical challenge is that *observed* port depth is a combination of naturally endowed depth and depth attained by dredging. Our solution to this relies on using contemporary granular data on underwater elevation levels around the port to isolate the naturally endowed component of depth. In particular, we take all sea cells within buffer rings around the geocode of the port and sum the number of cells that are 'very deep,' which we define as depth greater than 30 feet following Brooks et al. (2019). These authors argue that given vessel sizes in the 1950s (pre-containerization), depth beyond 30 feet conferred no advantage to the port. Below, we will test how reasonable this assumption is by examining pre-trends in shipping.

To operationalize our measure, we need to take a stand on which set of cells around the port to consider. Our aim is to measure depth in areas around the port that are used by ships to navigate and wait for their docking time. We examine the location (using exact geocodes) of stationary ships around the port in a one hour window for 100 random ports in our sample using contemporary data.[15] The majority of stationary ships are located within 5 km, which justifies our baseline measure of port suitability: the log of the sum of 'very deep' cells in a buffer ring 3-5 km around the port.[16] We examined the effect of depth measured at various buffers and confirmed that the effects are similar in nearby rings, suggesting that the variation we use from the 3-5 km buffer is a representative measure of depth at the port.

Testing for endogenous dredging. The key assumption behind our ability to isolate naturally endowed depth (from depth attained by dredging) is that when ports need to invest in costly dredging, they typically do not dredge entire areas in our buffers, but narrow channels that ships use to navigate to the port. By calculating depth over many sea cells, the vast majority of depth measurements for each port should reflect naturally endowed depth. We test this assumption in the following way. For 100 random ports in our sample, we obtained access to nautical maps from *marinetraffic.com* which clearly demarcate the dredged channels that ships use to navigate to the port.[17] We then constructed a binary variable, *'Dredging'*, that takes the value 1 if a port has a dredged channel in the 3-5 km buffer ring. Appendix Table B.3 shows the association

[15]These data are from *marinetraffic.com* and refer to *stationary* ships near the port captured between November 4 and 10, 2019, at 12:00-13:00 local time. More details regarding these data and how we choose the buffer around the port are provided in the Supplementary Material. There is a concern that measures of where ships are found around the port *today* is a poor proxy for where ships were located during our sample period. Partly for this reason, we show that depth measured in the same way at different nearby buffers yields similar results.

[16]There are zeros in the data, that is, there are ports with no cells deeper than 30 feet in the 3-5 km buffer around the port. For this reason, in practice, we use $\ln(1 + \sum_i \mathbb{1}(depth_i \geq 30ft))$, where i denotes a cell.

[17]We provide more details on this exercise in the Supplementary Material.

between this measure and the depth measure. The unconditional association (column (1)) is *negative* and statistically significant. That is, ports that we measure to be shallow are more likely to have a dredged channel. This is what we would expect to find if our measure captured naturally endowed depth.[18]

Balancing checks. We examine the extent to which our measure of exogenous port suitability is correlated with other observables pre-containerization in order to assess the types of confounders that may bias the results. Appendix Table B.4 shows the results. If greater depth would have led to more shipping even before containerization, we would expect to see a positive coefficient between depth and shipping flows. However, we see that the unconditional measure of depth is *negatively* correlated with both the level of shipping flows in 1950 (measured in logs), and population in 1950 (measured in logs), indicating that initially small cities had larger depth. In terms of growth rates pre-containerization, depth is weakly positively correlated with population growth between 1950 and 1960 (the coefficient is significant at 10%). This suggests that our depth measure is correlated with small cities that are growing relatively fast, i.e., population convergence. In order to purge our depth measure of this variation, we residualize it on city population in 1950 (measured in logs).[19] We re-examine how the part of the variation in depth that is uncorrelated with 1950 population, 'residualized depth,' correlates with the same observables. Reassuringly, residualized depth is correlated neither with the level of shipping and population in 1950 (the latter by construction), nor with the change in shipping and population between 1950 and 1960. In the empirical analysis, we therefore use the residualized measure of depth as the baseline measure of exogenous port suitability.

Appendix Table B.4 also shows the correlation with other observables. Residualized depth is uncorrelated with all observables we consider, except the Saiz land rent proxy. This is perhaps unsurprising, as arguably similar geographic characteristics determine the overland (Saiz measure) and underwater (depth measure) geographic features around a city. For this reason, we show robustness of all our results to the inclusion of the Saiz land rent proxy interacted with year indicator variables. The final issue concerns potential spatial correlation in the depth measure. We will tackle the issue of spatial correlation in the empirics by testing the robustness of our results to using only

[18]Adding continent or coastline fixed effects (columns (2) and (3), respectively) reduces the size of the negative coefficient and we lose statistical significance in column (3), but the estimated coefficients remain negative.

[19]More precisely, we regress the log of depth on the log of population in 1950 and take the residuals from this regression. Population in 1950 is not observed for 21 out of 553 port cities. For these, we replace 1950 population with the first year in which population is observed, which is generally 1960.

within-region variation, and to adjusting for spatial autocorrelation in the error term by reporting Conley standard errors (Conley, 1999).

3.2 Dealing with land reclamation

Finally, both the depth and the Saiz measure are constructed using contemporary GIS data, which captures natural geography in combination with investments in reclaiming land from the sea. To investigate the extent to which reclamation may introduce systematic measurement error, we use data from Martín-Antón, Negro, López-Gutiérrez, and Esteban (2016) on coastal land reclamation for any purpose.[20] Appendix Table B.5 shows that there is somewhat more land reclamation in cities we measure to be more geographically constrained (and hence have higher land rents). This is what we would expect if the Saiz measure was mostly capturing natural geography. The reason for this seems to be that while land reclamation is fairly common (76 out of 553 ports report *some* land reclamation), it is typically small relative to the area over which the Saiz measure is constructed.[21] Turning to the association between the depth measure and the binary indicator of land reclamation, the estimated coefficient is small and never statistically significant.

3.3 Results

Result 1: Depth predicts shipping, but only after 1960. First, we examine whether depth predicts shipping flows during our sample period. We implement this using the following flexible specification that allows us to examine the timing of when depth started to matter for shipping.

$$\begin{aligned}\ln(Ship_{it}) \quad = \quad & \sum_{j=1960}^{1990} \beta_j * Depth_i * \mathbb{1}(Year = j) + \sum_{j=1960}^{1990} \phi_j * \ln(Pop_{i,1950}) * \mathbb{1}(Year = j) \\ & +\alpha_i + \delta_t + \epsilon_{it}\end{aligned}$$

The outcome variable of interest, $\ln(Ship_{it})$, is the log of shipping flows observed in city i at time t. We need to take a stand on the treatment of zeros in the shipping data.[22]

[20] See the Supplementary Material for a discussion of the data.

[21] The median size of reclaimed area in the sample for the non-zero observations is 13 square kilometers, which pales in comparison to the 7850 square kilometers covered in the Saiz measure.

[22] The data contain zeros for two reasons. First, we may observe zeros because of measurement error: small ports with low shipping flows may not register an inter-port movement during the week in which we capture the data. Second, zeros may appear due to the time-invariant definition of port status that we use. We observe zero shipping flows in a particular year if a port was established in the city only after 1950, or if a port shut down in the city during our sample

In the baseline measure, we annualize the weekly counts of ships from the raw data by multiplying the one-week sample of shipping flows we observe by 52. This is primarily so that our results are comparable to regressions we run using model-simulated data in the quantification exercise in Section 6. Finally, we replace the zeros in the data with ones and take the natural logarithm of this adjusted annualized count.[23] $Depth_i$ is the cross-sectional measure of port suitability defined in the previous subsection. We interact this measure with binary indicators for the decades 1960 – 1990 to estimate the time path of how depth affected shipping flows. In addition, we include the full set of city and year fixed-effects (denoted α_i and δ_t, respectively) as well as the log of population in 1950 interacted with year indicator variables across all specifications. This is equivalent to using the residualized depth measure in a panel setting. We cluster standard errors at the city level in the baseline to account for the serial correlation of shocks. We also report Conley standard errors (in curly brackets).[24] Each β_j in this specification estimates the increase in shipping caused by having a deeper port in a given year relative to 1950.

Table 1 contains the estimated coefficients. Column (1) presents coefficients for the baseline specification. A number of points should be noted. First, deeper ports did not witness differential growth in shipping flows between 1950 and 1960, consistent with this being a decade in which containerization was just being developed in a few ports around the world. Second, we see an effect of depth in each of the following decades, as containerization was adopted worldwide. The coefficient of interest is much larger and significantly different from zero for the interaction of depth and each year indicator including and after 1970. This is consistent with containerization technology being rolled out in the early 1960s across US ports and worldwide later in the decade, as we discussed in Section 1.

A causal interpretation of the estimated effect of depth relies on the identifying assumption that the time-varying effect of depth is uncorrelated with the error term. The timing of when depth started to matter and the lack of pre-trends provide some evidence that this assumption is plausible. Next, we turn to further testing this result with more demanding specifications. One concern is that many determinants of depth may be

period. Overall, we observe zero shipping flows for 16% of the port-year observations. From examining the data, the zeros seem to be more likely driven by mismeasuring small shipping flows rather than the entry and exit of ports.

[23]In robustness checks discussed below, we show that all of the results presented in this section are robust to other standard ways of dealing with the zeros. In these, we do not annualize the data in order to verify that this transformation does not drive the results.

[24]As these are typically very close to the clustered standard errors, we only report them for the main results for easier readability of the tables. We allow for spatial correlation at distances up to 1,000 km and set the spatial decay function to be linear.

Table 1: Depth predicts shipping flows, but only after 1960

	Dependent variable: ln(Shipment)				
Independent variables	(1)	(2)	(3)	(4)	(5)
Depth × post 1970					0.247***
					(0.059)
					{0.052}
Depth × 1960	-0.051	0.029	0.050	-0.055	
	(0.063)	(0.069)	(0.066)	(0.068)	
Depth × 1970	0.222***	0.233***	0.278***	0.213***	
	(0.069)	(0.077)	(0.082)	(0.071)	
Depth × 1980	0.188**	0.212**	0.291***	0.192**	
	(0.079)	(0.085)	(0.090)	(0.081)	
Depth × 1990	0.255***	0.222**	0.312***	0.283***	
	(0.086)	(0.087)	(0.099)	(0.087)	
Observations	2765	2765	2765	2360	2765
R-squared	0.126	0.248	0.131	0.142	0.126
Number of cities	553	553	553	472	553
Year FE	✓	✓	✓	✓	✓
City FE	✓	✓	✓	✓	✓
Population 1950 × Year	✓	✓	✓	✓	✓
Coastline × Year FE	×	✓	×	×	×
Saiz × Year	×	×	✓	×	×
GDP pc (country) × Year	×	×	×	✓	×

Notes: 'Depth' indicates the port suitability measure. It is interacted with decade dummies or an indicator variable for decades including and after 1970, as indicated. Standard errors clustered at the city level in parentheses, Conley standard errors to adjust for spatial correlation in curly brackets. *** p<0.01, ** p<0.05, * p<0.1 (significance refers to clustered standard errors).

spatially correlated and if true, the estimates could be hard to disentangle from broader regional trends. To this end, column (2) adds the full set of 'coastline' by year-fixed effects to examine the extent to which our identifying variation relies on cross-regional variation.[25] Note that this set of fixed effects subsumes continent by year fixed effects.

[25] We define coastlines in the following way. We assign each port to its nearest ocean (e.g.,

Column (3) adds the Saiz land rent proxy interacted with year indicators to capture trends driven by the time-varying effect of land rents. Column (4) adds country GDP per capita (measured in 1960) interacted with year indicators to control for potentially differential growth trends across initially rich and poor countries.[26] The coefficients are remarkably stable.

Based on these results, we introduce a 'containerization' treatment indicator that turns on in years including and after 1970. This yields a single coefficient that estimates the differential effect of depth on shipping after the onset of containerization. Column (5) shows the results. Cities endowed with more depth, and hence more suitable to containerized technologies witnessed disproportionate increases in their shipping flows after containerization. Panel A in Appendix Figure C.3 shows the coefficient of interest remains fairly stable as we drop continents one at a time, underscoring that no single region appears to be driving the results. The coefficient becomes somewhat smaller when we drop North America, which is in line with the United States being the birthplace and an early adopter of containerization.

Panel A in Appendix Table B.6 contains further robustness checks. First, we test robustness to different data construction choices. In particular, we examine different ways of treating zero shipping values, different ways of defining the depth measure for the handful of ports that are located far inland from the coastline and restricting the sample to the subset of cities that had already attained 100,000 inhabitants by 1950 to examine sample selection bias. The coefficient of interest remains similar in magnitude and highly significant across all these checks. We now turn to examining how this boom in shipping affected city population.

Result 2: The local causal effect of shipping on population is not distinguishable from zero. We estimate the effect of shipping on population using the following specification;

$$\ln(Pop_{it}) = \beta * \ln(Ship_{it}) + \sum_{j=1960}^{1990} \phi_j * \ln(Pop_{i,1950}) * \mathbb{1}(Year = j) + \alpha_i + \delta_t + \epsilon_{it} \quad (1)$$

where $\ln(Pop_{it})$ is the natural logarithm of population in city i at time t, and all other variables are as previously defined. The identification challenge is that the shipping flows of a city are endogenous. Our main worry is reverse causality: fast growing cities

'Pacific Ocean') or body of water (e.g., 'Great Lakes') and further disaggregate oceans by continent. This yields 22 coastlines worldwide. Examples are 'Mediterranean – Europe' and 'North America – Atlantic'.

[26]We use the 1960 (pre-containerization) measure of country GDP per capita as this is observed for a larger set of countries than for 1950.

will also witness increases in their shipping flows. Our solution is to isolate variation in shipping using the binary version of our containerization treatment defined in the previous section: we interact the cross-sectional measure of depth with an indicator variable that takes the value of one in years including and after 1970. We cluster standard errors at the city level and we also report Conley standard errors for our main results.

Table 2: The local causal effect of shipping on population is not distinguishable from zero

	Panel regression						Long difference			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Indep. Variables	ln(Pop)	ln(Pop)	ln(Ship)	ln(Pop)	ln(Ship)	ln(Pop)	∆ln(Pop)	∆ln(Pop)	∆ ln(Ship)	∆ln(Pop)
ln(Shipment)	0.013***	0.015								
	_0.030_***	_0.035_								
	(0.005)	(0.049)								
	{0.005}	{0.039}								
∆ ln(Shipment)							0.013	0.006		
							_0.052_	_0.022_		
							(0.009)	(0.073)		
							{0.014}	{0.115}		
Depth									0.272***	0.002
									_0.134_***	_0.003_
									(0.086)	(0.020)
Depth × post 1970			0.268***	0.004						
			_0.143_***	_0.005_						
			(0.058)	(0.013)						
Depth × 1960					-0.042	-0.003				
					(0.064)	(0.008)				
Depth × 1970					0.246***	0.007				
					(0.069)	(0.013)				
Depth × 1980					0.213***	-0.002				
					(0.079)	(0.017)				
Depth × 1990					0.280***	0.002				
					(0.086)	(0.020)				
Observations	2734	2734	2734	2734	2734	2734	531	531	531	531
Number of cities	552	552	552	552	552	552				
Year FE	✓	✓	✓	✓	✓	✓	×	×	×	×
City FE	✓	✓	✓	✓	✓	✓	×	×	×	×
Population 1950 × Year	✓	✓	✓	✓	✓	✓	×	×	×	×
Population 1950	×	×	×	×	×	×	✓	✓	✓	✓
Specification	OLS	2SLS	FS	RF	dyn FS	dyn RF	OLS	2SLS	FS	RF
KP F-stat		21.13						9.98		

Notes: 'Depth' indicates the port suitability measure. It is interacted with decade dummies or indicator variables for decades including and after 1970, as indicated. Standardized coefficients in italics underneath the baseline coefficients. Notation for specification as follows: 'FS' refers to the first stage, 'RF' to the reduced form, 'dyn FS' to the fully flexible first stage and 'dyn RF' to the fully flexible reduced form. Standard errors clustered at the city level in parentheses, Conley standard errors to adjust for spatial correlation in curly brackets. *** p<0.01, ** p<0.05, * p<0.1 (significance refers to clustered standard errors).

Table 2 contains the main regression results. In columns (1) to (6) we estimate the effects of interest using all years in the sample, while columns (7) to (10) show the estimates from the long-differenced specification. Turning first to the full panel specification, the OLS specification of equation (1) shows that the association between shipping and population is small, positive and statistically different from zero (coefficient 0.013, se. 0.005). The 2SLS estimate in column (2) shows a similarly sized coefficient but we cannot reject zero (coefficient 0.015, se. 0.049). To assess magnitudes, we report the standardized 'beta' coefficients for our effects of interest in italics underneath the estimated regression coefficients. These make clear that while the OLS may be statistically significant, the magnitudes of both the OLS and the 2SLS estimates are economically negligible. A one standard deviation increase in shipping leads to a 0.03 (OLS) or 0.035 (2SLS) standard deviation increase in population. Columns (3) and (4) show the first stage and reduced form respectively. These make clear why the results are indistinguishable from zero. While the first stage is strong (the Kleibergen-Paap F-statistic is 21.13), there is no reduced form relationship between depth and population (the reduced form coefficient is 0.004, se 0.013).[27] Column (6) shows the full time path of effects for the reduced form. These make clear that the statistically insignificant coefficient in the 2SLS estimate does not stem from the fact that population is sluggish to adjust. The time path of the coefficients shows no discernible trend, and there is no clear difference in population growth post-containerization for deeper ports. All of the coefficients are estimated to be very close to zero (the one 'furthest' away from zero is 0.007), the coefficients are never close to statistical significance and in two of the five decades, the estimated reduced form coefficient is negative, suggesting that if anything, deeper ports were growing at a slower rate than shallower ones in some decades.

While the large standard errors typical of 2SLS estimation make a definitive answer difficult, there are several reasons why we believe that the most reasonable interpretation of our results is that shipping booms caused by containerization led to no discernible effects on population. First, the standardized 'beta' coefficients make clear that the magnitudes of both the OLS and the 2SLS estimates are economically negligible. Second, if we examine the long-differenced specification in columns (7) to (10), neither the OLS nor the 2SLS estimate is significantly different from zero, and both standardized beta coefficients again show an economically negligible effect. In fact, the 2SLS coefficient estimate is smaller – it is less than half the size estimated in the full panel, consistent with the fact that it was population observations from *earlier* years that drove the point estimate in the full panel specification. While the long-differenced specification has the disadvantage of using fewer observations, it has the advantage that

[27] The specification here is identical to that in Table 1, but the sample size shrinks slightly as we lose those observations where population is unobserved in some years (1% of the sample).

it examines the long-run effects of the shipping boom on population, once the latter has had time to adjust.

Third, we subject the 2SLS specification to the same set of robustness checks conducted above: the inclusion of coastline by year fixed effects, controlling for the time varying-effect of the the Saiz land-rent proxy and GDP per capita (Appendix Table B.7). Despite the demanding nature of these specifications, the first stage remains sufficiently strong (the Kleibergen-Paap F-statistic is always above 10) and the estimated 2SLS coefficient is never statistically different from zero. In fact, in two out of three cases, the estimated coefficient is *negative*. Fourth, no single continent drives this result (Appendix Figure C.3, panel B, plots the estimated coefficient dropping continents one at a time). Appendix Table B.6, panel B, shows that the results are robust to the same set of additional robustness checks to data choices performed for Result 1.

We view the null effect on population as a surprising finding. Intuition and standard models (Coşar and Fajgelbaum, 2016; Nagy, 2020b; Fajgelbaum and Redding, 2018) would both suggest that a boom in shipping should make a location more attractive for households and firms, as they can access consumers and producers more cheaply (the 'market access effect'), leading to an inflow of population. Indeed, the past literature has found that these types of positive shocks to a city's accessibility tend to lead to a boom in local population (Bleakley and Lin (2012); Brooks et al. (2019); Campante and Yanagizawa-Drott (2018).[28] While comparing the *economic* size of the effect in these papers relative to ours is difficult given the different contexts and different 'treatments,' these papers all show that their effect is economically meaningful, while making the same claim with our results would be difficult.

What can explain the difference between our findings and previous work? One notable difference in our setting is the increased land intensity of port activities induced by containerization discussed in Section 1. In the last part of this section, we examine the extent to which we can detect the effects of this in our data. We also note that in Section 6, we will use our quantified model to return to the question of what magnitude one would expect in our setting *in the absence* of the crowding out mechanism.

Result 3: Containerization increased shipping more in low rent cities. To the extent that

[28]The paper closest to our setting is Brooks et al. (2019), who study the effect of containerization on the population of U.S. counties located nearby. They find a positive and statistically significant effect of containerization on local population. Though the two settings are difficult to compare as we study cities around the world, we think one crucial difference is that while we examine the effects on cities, their unit of analysis is a county. As we argue below, the most likely mechanism driving the null result is that the land intensity of port technology acts as an important opposing force crowding out population. This mechanism is more likely to be detectable at the generally finer level of spatial resolution that we examine.

land prices affect where port development takes place, we would expect low land-rent cities to be more attractive places for containerized ports all else equal, as the opportunity cost of port development in these locations is low. We test for this by examining the heterogeneity of the depth-shipping relationship from Result 1 using the following specification;

$$\begin{aligned}\ln(Ship_{it}) \quad = \quad & \beta * Depth_i * \mathbb{1}(Year \geq 1970) + \gamma * Depth_i * Rent_i * \mathbb{1}(Year \geq 1970) \\ & + \eta * Rent_i * \mathbb{1}(Year \geq 1970) + \sum_{j=1960}^{1990} \phi_j * \ln(Pop_{i,1950}) * \mathbb{1}(Year = j) \\ & + \alpha_i + \delta_t + \epsilon_{it} \qquad (2)\end{aligned}$$

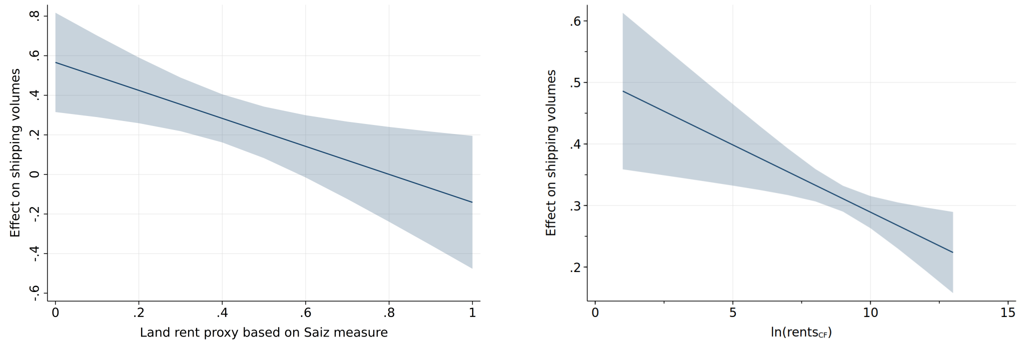
where $Rent_i$ is the Saiz land rent proxy for city i, and all other variables are as defined above.[29] The coefficient of interest is γ – that is, we are interested in the interaction between our depth measure and the Saiz land rent proxy (interacted with the 'containerization' treatment variable that turns on in 1970). We have defined the Saiz measure such that higher values correspond to less area that can be developed, implying high land rents. Note that this is a fully saturated specification in that we allow both depth and the Saiz measure to have their own time trend break in 1970. We plot the marginal effect of depth at different values of the Saiz measure in Figure 1a (the corresponding estimates are presented in Appendix Table B.8). Consistent with the land intensive nature of containerized technology, the coefficient of interest, γ, is negative, large and statistically different from zero (coefficient -0.707, se. 0.323). Cities with exogenously deeper ports witnessed increased shipping flows after 1970, but disproportionately more so in low land rent cities.

Panel C in Appendix Figure C.3 explores the heterogeneity of the result by dropping continents one at a time. The effect is consistently negative. We perform the same set of robustness checks for this result as for previous ones (see Appendix Tables B.6, panel C, and B.8). The results are largely robust to these specifications, as our coefficient of interest, γ, remains negative and economically large throughout all these checks, though in two especially demanding specifications the level of significance drops slightly below 10%.

In Appendix Table B.9 we provide additional evidence that land prices matter for where port activity takes place by examining the exact location of ports *within* cities.[30]

[29]We report standard errors clustered at the city level, as well as Conley standard errors in curly brackets.

[30]We use information from the *World Port Index* on the geocodes of ports in our sample in



(a) Data

(b) Model-simulated data

Figure 1: Containerization increased shipping more in low land-rent cities

Notes: Panel A shows the estimated γ coefficient from equation (2) evaluated at different values of the Saiz land rent proxy. Panel B shows the same estimated coefficient using model-simulated data evaluated at different values of the counterfactual land rents.

We show that during this time period, ports moved further from the centroid of the city towards the outskirts, where land prices are typically lower (Duranton and Puga, 2019). This is particularly striking for the subset of cities in which a new port was built (e.g., in Sydney, Australia). In these cases, the new port was located on average 9 km further from the centroid of the city than the old port.

The land rent heterogeneity result suggests that the land-intensive nature of containerized technology is an empirically important determinant of *where* containerized port infrastructure was developed. Armed with this evidence, we now turn to writing down a quantitative spatial model that captures many realistic features of port infrastructure development, including, but not limited to, the land price mechanism. We note that of course there are mechanisms other than the land-intensity of containerization that could account for the crowd-out of population. We explore three prominent ones in Section 7, after presenting both the empirics and the quantified model.

4 A model of cities and endogenous port development

To measure the aggregate effects of port development, we develop a rich and flexible quantitative general equilibrium model of trade across cities that captures both the benefits and costs of port development.

4.1 Setup

The world consists of $S > 0$ cities, indexed by r or s. An exogenously given subset of cities are port cities, while the rest are non-port cities. We make the Armington assumption that each city produces one variety of a differentiated final good that we

1953 and 2017. We provide details on the data used for this exercise, and in particular, on how we calculate city centroids in the Supplementary Material.

also index by r or s (Anderson, 1979). Each city belongs to one country, and each country is inhabited by an exogenous mass of workers who choose the city in which they want to live. Mobility across cities is, however, subject to frictions.

4.1.1 Workers

Each worker owns one unit of labor that she supplies in her city of residence. The utility of a worker j who chooses to live in city r is given by

$$u_j(r) = \left[\sum_{s=1}^{S} q_j(r,s)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} a(r)\, b_j(r) \tag{3}$$

where $q_j(r,s)$ is the worker's consumption of the good made in city s, $a(r)$ is the level of amenities in city r, and $b_j(r)$ is an idiosyncratic city taste shifter. $\sigma > 1$ is the elasticity of substitution across goods.

The dispersion of $b_j(r)$ represents the severity of cross-city mobility frictions that workers face, similar to Kennan and Walker (2011) and Monte, Redding, and Rossi-Hansberg (2018). For tractability, we assume that $b_j(r)$ is drawn from a Fréchet distribution with shape parameter $1/\eta$ and a scale parameter normalized to one. Hence, a larger value of η corresponds to more severe frictions to mobility.

4.1.2 Landlords

Each city r is also inhabited by a positive mass of immobile landlords who own the exogenously given stock of land available in the city. We normalize the stock of land available in each city to one.[31] Landlords have the same preferences over goods as workers. They do not work but finance their consumption from the revenues they collect after their stock of land.

Each landlord is small relative to the total mass of landlords in the city and hence thinks that she cannot influence prices. Yet the mass of landlords is small enough that the population of each city can be approximated well with the mass of workers who choose to reside in the city.

In non-port cities, landlords rent out their land to firms that produce the city-specific good. In port cities, landlords can also use part of their land to provide transshipment services. The more land they use for transshipment services, the more the cost of trans-

[31] We could allow the stock of available land to vary across cities. This more general setup is isomorphic to our current model, except that, instead of productivity in the city-specific good sector, a combination of the stock of land and productivity enters the model's equilibrium conditions. In other words, the city productivity levels we identify from our current model reflect not only productivity per se, but also the stock of available land. This fact, however, does not affect our quantitative results as we keep productivity levels fixed in our model simulations.

shipping a unit of a good decreases. The landlord can charge a price for the transshipment service she provides. Competition among port city landlords drives down this price to marginal cost. Hence, profits from transshipment services are zero in equilibrium.[32]

4.1.3 Production

Firms can freely enter the production of the city-specific good. Hence, they take all prices as given and make zero profits. Production requires labor and land. The representative firm operating in city r faces the production function

$$q(r) = \tilde{A}(r)\, n(r)^{\gamma} (1 - F(r))^{1-\gamma}$$

where $q(r)$ denotes the firm's output, $\tilde{A}(r)$ is total factor productivity in the city, $n(r)$ is the amount of labor employed by the firm, and $F(r)$ is the share of land that landlords in the city use for transshipment services (thus, $F(r) = 0$ in non-port cities). Hence, $1 - F(r)$ is the remainder of land that landlords rent out to firms for production, and γ and $1 - \gamma$ correspond to the expenditure shares on labor and land, respectively.

We incorporate agglomeration economies by allowing total factor productivity to depend on the population of the city, $N(r)$:

$$\tilde{A}(r) = A(r)\, N(r)^{\alpha}$$

where $A(r)$ is the exogenous fundamental productivity of the city, and $\alpha \in [0, 1 - \gamma]$ is a parameter that captures the strength of agglomeration economies.[33] The representative firm does not internalize the effect that its employment decision has on local population. Hence, it takes $N(r)$ as given.

4.1.4 Shipping and port development

Firms in city r can ship their product to any destination $s \in S$. Shipping is, however, subject to iceberg costs: if a firm i from city r wants to ship its product over a route ρ that connects r with s, then it needs to ship $T(\rho, i)$ units of the product such that one unit arrives at s. Shipping costs consist of a component common across firms $\bar{T}(\rho)$, as

[32] In Section 6, we show that the aggregate gains from containerization remain similar in an alternative framework in which landlords have market power and thus can make profits. We provide a detailed description of this alternative framework in the Supplementary Material.

[33] We make the assumption $\alpha \leq 1 - \gamma$ to guarantee that agglomeration forces are not overwhelmingly strong in the model. Estimates of the land share, $1 - \gamma$, tend to be substantially above estimates of agglomeration externalities α. In particular, our calibration involves setting α to 0.06 (a standard value used in the literature) and $1 - \gamma$ to 0.16 based on Desmet and Rappaport (2017).

well as a firm-specific idiosyncratic component $\epsilon(\rho, i)$ that is distributed i.i.d. across firms and shipping routes:[34]

$$T(\rho, i) = \bar{T}(\rho)\,\epsilon(\rho, i)$$

For tractability, we assume that $\epsilon(\rho, i)$ is drawn from a Weibull distribution with shape parameter θ and a scale parameter normalized to one. Firms only learn the realizations of their idiosyncratic cost shifters after making their production decisions. Therefore, they make these decisions based on the expected value of shipping costs,

$$\mathbf{E}\left[T(\rho, i)\right] = \bar{T}(\rho)\,\mathbf{E}\left[\epsilon(\rho, i)\right] = \bar{T}(\rho)\,\Gamma\left(\frac{\theta+1}{\theta}\right).$$

After learning $\epsilon(\rho, i)$, they choose the route that minimizes their total shipping costs.

Certain shipping routes involve land shipping only (*land-only*), while others involve a combination of land and sea shipping through a set of ports (*land-and-sea*). Land-only shipping is only available between cities that are directly connected by land. The common cost of land-only shipping between cities r and s is an increasing function of the minimum overland distance between the two cities, $d(r, s)$:

$$\bar{T}(\rho) = 1 + \phi_\varsigma\left(d(r, s)\right)$$

The cost of land-and-sea shipping depends on the set of ports en route. In particular, the common cost of shipping from r to s through port cities $p_0, ..., p_M$ takes the form

$$\bar{T}(\rho) = \left[1 + \phi_\varsigma\left(d(r, p_0)\right)\right]\left[1 + \phi_\varsigma\left(d(p_M, s)\right)\right]\prod_{m=0}^{M-1}\left[1 + \phi_\tau\left(d(p_m, p_{m+1})\right)\right]\prod_{m=0}^{M}\left[1 + O(p_m)\right]$$

where $\phi_\varsigma(d(r, p_0))$ corresponds to the overland shipping cost between the origin and the first port en route p_0, and $\phi_\varsigma(d(p_M, s))$ corresponds to the overland shipping cost between the last port en route p_M and the destination. $\phi_\tau(d(p_m, p_{m+1}))$ denotes the sea shipping cost between ports p_m and p_{m+1}, a function of the minimum sea distance between the two ports, $d(p_m, p_{m+1})$. Finally, $O(p_m)$ denotes the price that the firm needs to pay for transshipment services in port city p_m.[35]

[34] The assumption of idiosyncratic shipping cost shifters follows Allen and Atkin (2016) and Allen and Arkolakis (2019), and allows us to tractably characterize shipping flows with a large number of cities. In the alternative case with no idiosyncratic shifters, applied in Allen and Arkolakis (2014) and Nagy (2020a), finding optimal shipping flows is computationally more demanding.

[35] Note that this formulation does not allow for land shipping between two subsequent ports

Transshipment costs are central to our analysis as these are the costs that port city landlords can lower by developing the port, that is, by allocating more land to the port. In particular, we assume that the landlord's cost of handling one unit of a good at port p_m equals

$$[\nu(p_m) + \psi(F(p_m))] \, Shipping(p_m)^{\lambda}$$

where $\nu(p_m)$ is an exogenous cost shifter capturing the fundamental efficiency of port p_m, $\psi(F(p_m))$ is a non-negative, strictly decreasing and strictly convex function of $F(p_m)$, the share of land allocated to the port, and $Shipping(p_m)^{\lambda}$ captures congestion externalities arising from the fact that handling one unit of cargo becomes more costly as the total amount of shipping, $Shipping(p_m)$, increases for a given port size.[36] As each port city landlord is atomistic, she takes the price of transshipment services $O(p_m)$ and the total port-level shipping $Shipping(p_m)$ as given when choosing $F(p_m)$. Moreover, perfect competition among port city landlords ensures that the price of transshipment services is driven down to marginal cost and therefore

$$O(p_m) = [\nu(p_m) + \psi(F(p_m))] \, Shipping(p_m)^{\lambda} \tag{4}$$

in equilibrium.

One concern is that, according to our formulation, land is required for transshipment services while labor is not. In reality, ports employ labor. To address this concern, the Supplementary Material presents an extension of our model in which a combination of land and labor must be employed in transshipment. This appendix also shows that the model with transshipment labor, although more complex in its structure, delivers qualitative predictions that are extremely similar to the predictions of our baseline model.

4.1.5 Equilibrium

In equilibrium, workers choose their consumption of goods and residence to maximize their utility, taking prices and wages as given. Landlords choose their consumption and land use to maximize their utility, taking prices, land rents and shipping flows as given. Firms choose their production of goods, employment and land use to maximize their profits, taking prices, land rents and wages as given. Competition drives profits from

along the route. In practice, this is extremely unlikely to arise as land shipping is substantially more expensive than sea shipping.

[36] To be precise, $Shipping(p_m)$ is defined as the dollar amount of shipping flowing through port p_m, excluding the price of transshipment services at p_m. We exclude the price of transshipment services from the definition of $Shipping(p_m)$ as it simplifies the procedure of taking the model to the data.

production and profits from transshipment services down to zero. Markets for goods, land and labor clear in each city, and markets for transshipment services clear in each port city. The Supplementary Material provides a formal definition and characterization of the equilibrium.

4.2 Predictions of the model

In equilibrium, the share of land allocated to the port in port city r is the solution to the equation

$$-\psi'(F(r)) = \frac{R(r)}{Shipping(r)^{1+\lambda}} \tag{5}$$

where $R(r)$ denotes land rents in city r, given by

$$R(r) = \frac{1-\gamma}{\gamma}\frac{w(r)\,N(r)}{1-F(r)} \tag{6}$$

such that $w(r)$ is the wage in city r.[37] As the left-hand side of equation (5) is decreasing in $F(r)$ by the convexity of ψ, we have the following two propositions.

Proposition 1. *Land allocated to the port is increasing in the amount of shipping flows.*

Proposition 1 is the consequence of two forces in the model. The first is increasing returns to scale in transshipment technology: as shipping flows increase, it becomes profitable to lower unit costs by allocating more land to the port. The second force is congestion: an increase in shipping flows makes landlords allocate more land to the port to palliate congestion.

Proposition 2. *Land allocated to the port is decreasing in land rents.*

Proposition 2 highlights that the cost of port development differs across cities. Cities that have high land rents do not allocate much land to the port as the opportunity cost of land is very high. As a result, everything else fixed, port development primarily takes place in low-rent cities, consistent with what we document in the data.

Finally, the model delivers the spatial distribution of population $N(r)$ as the solution to the following equation:

$$N(r)^{[1+\eta\sigma+(1-\gamma-\alpha)(\sigma-1)]\frac{\sigma-1}{2\sigma-1}} = \gamma^{\sigma-1}\tilde{a}(r)^{\frac{\sigma(\sigma-1)}{2\sigma-1}}A(r)^{\frac{(\sigma-1)^2}{2\sigma-1}}(1-F(r))^{(1-\gamma)\frac{(\sigma-1)^2}{2\sigma-1}}MA(r) \tag{7}$$

[37]The derivation of all equations presented here are included in the Supplementary Material.

where $MA(r)$ is the *market access* of city r, given by

$$MA(r) = \sum_{s=1}^{S} \frac{\tilde{a}(s)^{\frac{(\sigma-1)^2}{2\sigma-1}} A(s)^{\frac{\sigma(\sigma-1)}{2\sigma-1}} (1-F(s))^{(1-\gamma)\frac{\sigma(\sigma-1)}{2\sigma-1}} N(s)^{[1-\eta(\sigma-1)-(1-\gamma-\alpha)\sigma]\frac{\sigma-1}{2\sigma-1}}}{\mathbf{E}[T(r,s)]^{\sigma-1}} \tag{8}$$

and $\tilde{a}(r)$ can be obtained by scaling amenities $a(r)$ according to

$$\tilde{a}(r) = \aleph_c a(r)$$

where the endogenous country-specific scaling factor $\aleph_c$ adjusts such that the exogenously given population of country c equals the sum of the populations of its cities.

How is the population of a port city affected by the development of its port? Our last proposition shows that the net effect on population is the outcome of two opposing forces: the *market access effect* that increases the population of the city, and the *crowding-out effect* that leads to a decrease in the city's population.

Proposition 3. *An increase in the share of land allocated to the port in city r, $F(r)$, decreases shipping costs $\mathbf{E}[T(r,s)]$, thus increasing $MA(r)$. Everything else fixed, an increase in $MA(r)$ increases the population of the city (market access effect). At the same time, holding $MA(r)$ fixed, an increase in $F(r)$ decreases the share of land that can be used for production, $1-F(r)$, thus decreasing the population of the city (crowding-out effect).*

Proof. These results follow directly from equation (7). □

Proposition 3 sheds light on the fact that, to measure the net effect of port development, it is essential to consider both its benefits and its costs. On the one hand, port development lowers shipping costs. On the other hand, it requires scarce local land that needs to be reallocated from other productive uses. The model, and equation (7) in particular, provide a structure that allows us to capture these opposing forces. The next section is aimed at looking for evidence on these opposing forces in the data.

5 Empirical evidence for the model's mechanisms

Due to the lack of time-varying data on port sizes, we cannot directly take equation (7) to the data. However, we can estimate a simplified version of the equation to understand whether we can disentangle the positive and negative effects of port development predicted by the model. We estimate the following relationship, referred to as *model-inspired empirical specification*:

$$\ln(Pop_{it}) = \phi_1 * \ln(Ship)_{it} + \phi_2 * \ln(MA)_{it} + \alpha_i + \delta_t + \epsilon_{it} \tag{9}$$

where $\ln(MA)_{it} = \ln\left(\sum_{s=1}^{S} \frac{Pop_{st}^{[1-\eta(\sigma-1)-(1-\gamma-\alpha)\sigma]\frac{\sigma-1}{2\sigma-1}}}{T_t(i,s)^{\sigma-1}}\right)$ is the empirical equivalent of the model-based market access term, and all other variables are as previously defined in Section 3. According to the mechanism described in the model, we expect ϕ_1 to be negative and ϕ_2 to be positive.

We estimate time-varying bilateral trade costs $T_t(i, s)$ between origin and destination in the following way. As in the model, we assume that these bilateral costs consist of a combination of three possible components: first, the cost of shipping overland; second, the cost of sea shipping; and third, the cost of transshipment at seaports. We use the fast marching algorithm to calculate the lowest overall shipping cost between any given pair of cities. Following Allen and Arkolakis (2014), we assume that overland shipping costs ϕ_ς and sea shipping costs ϕ_τ take the form

$$\phi_\varsigma(d) = e^{t_\varsigma d} \qquad\qquad \phi_\tau(d) = e^{t_\tau d}$$

where d is (point-to-point) distance traveled. We take the values of t_ς and t_τ from the road and sea shipping cost elasticities estimated by Allen and Arkolakis (2014).[38]

There are no readily available measures of transshipment costs that we are aware of. To construct these, we use the following approach. Both the model and the transportation literature on ports argue that there are important increasing returns to scale in port technologies, rendering larger ports more cost-efficient (Rodrigue, 2016). We use estimates of port costs, available for a subset of our ports from Blonigen and Wilson (2008), to estimate the empirical relationship between port costs and shipping flows at the port level in our data using a simple linear OLS specification.[39] Consistent with increasing returns to scale in shipping, we find a negative and statistically significant association between port costs and the size of shipping flows. We use the estimated coefficient from this regression to predict port efficiency for all the ports in our data for each decade. Note that changing transshipment costs are the only source of time series variation in our estimated trade costs.

The model-based measure of market access requires taking a stand on the values of the parameters η, σ, γ and α. Table B.10 contains the parameter values we use and their source. We use the same values of these parameters when taking the full model to the

[38] Allen and Arkolakis (2014) also allow for costs of inland and sea shipping that are fixed with respect to distance. However, they set the fixed costs of road shipping to zero. In the case of sea shipping, our aim is to define transshipment costs incurred at the seaport in a broad sense, such that they include any cost that is not a function of shipping distance, such as the fixed costs of sea transportation.

[39] The Supplementary Material provides the full details and results of the estimation.

data. Section A.1.2 discusses the calibration of all structural parameters in detail.

Both regressors in the model-inspired specification (9) are potentially endogenous, requiring two sources of exogenous variation. We use depth as an instrument for shipping, as explained in Section 3. In addition, we use an exogenous population-growth shifter based on regional climate to construct an instrument for market access. This IV is based on insights from the urban economics literature, which has found that people have moved to places with warm winters over the course of the 20th century – a phenomenon attributed to the invention of air conditioning (e.g., Oi (1996); Rappaport (2007)).

We use the average number of frost free days, $frostfree_i$, during the years between 1961-1990 in each city to predict population growth during our time period.[40] In order to predict population, we estimate the following specification:

$$\ln(Pop)_{it} = \sum_{k=1960}^{1990} \beta_k * frostfree_i * \mathbb{1}(Year = k) + \alpha_i + \delta_{ct} + \epsilon_{it}$$

where β_k estimates the effect of warmer winters on population in each decade, α_i denotes city-specific fixed effects, and δ_{ct} allows for the full set of country by year fixed effects. Inclusion of these implies that we only use *within-country* variation in climatic conditions when estimating the effect of frost-free days on population growth. We do this to address the concern that climatic conditions vary across regions in ways that may correlate with unobserved drivers of population growth, confounding our estimates of interest. Appendix Table B.11 shows the result of this estimation and presents some robustness checks. To construct our second instrument, we predict population for each city-year pair based on the estimated effects of frost free days and the estimated city fixed effect (we do not use the estimated country-year fixed effects to predict population). Using these predictions for city-level population, we define our second instrument as follows:

$$\ln(MAIV_{it}) = \ln\left(\sum_s \frac{\exp(\ln\widehat{(Pop)}_{it})}{(T_{1950}(i,s))^{\sigma-1}}\right)$$

where $T_{1950}(i,s)$ is the transport cost between cities i and s in 1950. We hold bilateral transport costs fixed throughout all years in order to make sure that potentially endoge-

[40]The Supplementary Material contains a description of the data on the number of frost free days. This second instrument uses a source of exogenous variation that is orthogonal to port depth, as the number of frost free days and port depth (both residualized and un-residualized) are uncorrelated with each other. The correlation between the number of frost free days and unresidualized depth is 0.04 (p-val: 0.40), and the correlation between the number of frost free days and residualized depth is -0.02 (p-val: 0.68).

nous changes in trade costs over time are not used in the instrument. The specifications are estimated on the set of *port* cities in our dataset. Importantly, however, the market access of port cities is calculated using the full set of (port and non-port) cities.

Table 3 presents the estimation results. Columns (1) and (2) report the baseline reduced form OLS and 2SLS estimates for comparison. Columns (3) and (4) add the measure of market access as a control. The OLS estimate in column (3) shows a very small negative effect of shipping on population relative to column (1) that is not distinguishable from zero. Column (4) shows the 2SLS specification; consistent with the predictions of the model, once we control for market access, shipping has a negative, statistically significant effect on population.[41] The instruments yield a combined Kleibergen-Paap F-statistic of 9.63 which is just below the often recommended value of 10; however, it is larger than the critical value of 7.03 that the Stock-Yogo weak ID test suggests for 10% maximum bias (Stock and Yogo, 2002).[42] Columns (5) and (6) report the first stages of the regression. Reassuringly, depth is a strong predictor of shipping, while the market access IV predicts market access strongly. Appendix Table B.12 shows that the pre-trends check with respect to depth holds (for both first stages) in this more complex specification that adds market access.[43]

We test the robustness of this result in a number of ways in Appendix Table B.13. First, we show that the results are remarkably robust to dropping cities in the close vicinity of the city in the market access IV, suggesting that much of the identifying variation is coming from population movements further away from the city itself. Moreover, the sign of the effects are robust to the same set of controls used in Section 3, though in the case of these demanding specifications, we don't always retain statistical significance at 10%. We conclude that this lends well-identified evidence for the model mechanism. In the next section, we therefore turn to taking the full model to the data.

[41] As expected, market access has a significant positive effect on population. It is difficult to compare the size of the market access effect to existing estimates (Donaldson and Hornbeck, 2016; Jedwab and Storeygard, 2021; Maurer and Rauch, 2021) because different papers construct market access in different ways. Jedwab and Storeygard (2021) are the only paper we are aware of that report standardized coefficients that allow for a comparison. They estimate that a one standard deviation increase in market access leads to a 0.43 – 0.85 standard deviation increase in population. Relative to that paper, our estimate is slightly larger (1.13), but within the same ballpark.

[42] With the usual caveat that Stock and Yogo (2002) values have been derived only for i.i.d. errors, whereas we allow for autocorrelated or spatially correlated standard errors.

[43] As there is no similar 'pre-treatment period' for the market access IV, it is not possible to conduct a similar exercise for this IV.

Table 3: Model-inspired specification: Disentangling market access effect and crowding out effect

Independent variables	(1) ln(Population)	(2) ln(Population)	(3) ln(Population)	(4) ln(Population)	(5) ln(Shipment)	(6) ln(Market Access)
ln(Shipment)	0.015***	0.014	-0.001	-0.159**		
	(0.005)	(0.048)	(0.006)	(0.065)		
	{0.005}	{0.038}	{0.005}	{0.051}		
ln(Market Access)			1.512***	7.103***		
			(0.536)	(0.795)		
			{0.317}	{0.854}		
Depth × post 1970					0.275***	0.007***
					(0.058)	(0.001)
					{0.051}	{0.001}
Market Access IV					7.188	1.927***
					(5.428)	(0.140)
					{5.748}	{0.188}
Observations	2696	2696	2696	2696	2696	2696
Number of cities	544	544	544	544	544	544
Year FE	✓	✓	✓	✓	✓	✓
City FE	✓	✓	✓	✓	✓	✓
Population 1950 × Year	✓	✓	✓	✓	✓	✓
Specification	OLS	2SLS	OLS	2SLS	FS	FS
KP F-stat		22.07		9.63		

Notes: 'Depth' indicates the port suitability measure. It is interacted with an indicator variable for decades including and after 1970. 'ln(Market Access)' is the empirical counterpart of the market access term defined in Section 5. 'Market access IV' is the instrument for the market access term defined in Section 5. 'FS' refers to the first stage. Standard errors clustered at the city level in parentheses, Conley standard errors to adjust for spatial correlation in curly brackets. *** p<0.01, ** p<0.05, * p<0.1 (significance refers to clustered standard errors).

6 The aggregate effects of containerization

We use our model to measure the aggregate effects of containerization in this section. Here, we provide an overview of the main steps involved and refer the reader to Appendix A for details. We also test whether the model can replicate our reduced form results and discuss the aggregate effects of containerization.

6.1 Taking the model to the data

Taking the model to the data consists of three steps. In the first step, we calculate inland and sea shipping costs across cities following Allen and Arkolakis (2014) as described above, and choose a functional form for endogenous transshipment costs as a function of land use, $\psi(F)$. We choose the function such that

$$\psi'(F) = 1 - F^{-\beta} \tag{10}$$

where $\beta > 0$ is a parameter driving the shape of endogenous transshipment costs as a function of land use.

In the second step, we choose the values of the model's seven structural parameters based on estimates from the literature (Appendix Table B.10 reports the values and their sources): agglomeration externalities α, the labor share γ, the migration elasticity η, the elasticity of substitution across tradable goods σ, the dispersion of idiosyncratic shipping costs θ, congestion externalities λ, and the shape parameter of endogenous transshipment costs, β.

Finally, in the last step, we back out the values of unobserved city fundamentals that rationalize the data under these parameters: amenities $a(r)$, productivities $A(r)$ and exogenous transshipment costs $\nu(r)$. To this end, we combine the structure of the model with a cross section of city characteristics that we observe in 1990 (the only year for which all three variables are available): population, shipping flows and city GDP.

6.2 Counterfactual: rolling back containerization

Our counterfactual involves *rolling back containerization*: i.e., changing port technologies to pre-containerization ones. We keep all other fundamentals of the model fixed (i.e., city amenities and productivities, inland and sea shipping costs and country populations). Hence, comparing the 1990 equilibrium to the counterfactual equilibrium allows us to measure the aggregate effects of containerization on the world economy.

We incorporate the two technological aspects of containerization: lower costs, particularly in deep ports, and the increased land-intensity of transshipment. To capture the higher transshipment costs of pre-containerization technologies, we increase exogenous transshipment costs $\nu(r)$ uniformly across ports relative to the 1990 values of these costs. More precisely, we increase $\log \nu(r)$ by 0.280 at each port to match the estimated 25% average change in the sum of exogenous *and* endogenous transshipment costs as a result of containerization.[44] To capture the fact that containerization made port depth relevant for transshipment, we also offset the negative relationship between $\nu(r)$ and depth that we observe in 1990. Finally, to capture the lower land intensity of

[44] Rodrigue (2016, p. 117) estimates that containerization led to an overall 70% to 85% reduction in maritime transport costs by 2010; "While before containerization maritime transport costs could account for between 5 and 10 percent of the retail price, this share has been reduced to about 1.5 percent, depending on the goods being transported." A reduction from 5% to 1.5% of retail price equals a 70% cost reduction ($= 1 - 1.5/5$); similarly, a reduction from 10% to 1.5% equals an 85% cost reduction. We estimate that 36% of the total cost reduction took place up to 1990, by assuming that cost reductions are proportionate to ship size increases. These calculations are based on data from the *Miramar Ship Index* (Haworth, 2020). More details on these data are provided in the Supplementary Material. Using the more conservative estimate of 70%, this gives us a 25% decrease in average transshipment costs.

pre-containerization technologies, we decrease the shape parameter of our endogenous transshipment cost function, β, from its calibrated 1990 value of 0.031 to 0.018. This replicates the 75% difference in mean port size between uncontainerized and containerized ports that we found in column (7) of Appendix Table B.1.

Overall, according to our simulation, these changes in transshipment technology lead to an increase in the international trade to world GDP ratio by 4.7 percentage points from the counterfactual to the 1990 equilibrium. As a reference point, the trade to world GDP ratio increased by 15 percentage points between 1960 and 1990. This suggests that containerization was responsible for about *one-third* of the overall increase in trade to world GDP during these three decades.

The fraction of land occupied by ports (i.e., the port share) increases in most port cities from the counterfactual to the 1990 equilibrium. Port shares become larger for two reasons. First, the increase in β increases the incentive to invest more land in port development, mimicking the changing land-intensity of port technologies caused by containerization. Second, the reduction in trade costs leads to increased demand for shipping, encouraging yet more investment in port development. Appendix Figure C.5 presents the full distribution of port share changes across cities. The median change is 3 percentage points, while the 5th percentile is zero pp and the 95th percentile is 35 pp.

6.3 Test of the model: the reduced-form effects of containerization

In this section, we examine whether our quantified model can replicate the two key reduced-form facts related to containerization estimated in Section 3. First, to examine the local population effects of shipping in the model, we consider the long-differenced version of equation (1). Following the same identification strategy, we instrument the change in shipping with residualized port depth. Table 4 presents the results of this exercise. The model produces the same striking null result (column (2)) as the data (column (1)). The coefficient is neither statistically, nor economically significant.[45] The results from column (2) confirm that the crowding-out effect of increased land use caused by port development is sufficient to offset the positive local population effect of increased shipping due to the market access effect.

Second, we examine whether containerization induces shipping activity to reallocate toward low-rent cities in the model. To this end, we consider the long-differenced version of the rent heterogeneity result (specification 2). $Depth_i$ is our depth measure residualized on population in 1950, as in the data. The difference is that, while we had to rely on a proxy of city-level rents in specification (2), we can use model-implied (pre-containerization) rents $R_{i,CF}$ here. We evaluate the coefficient of interest, γ, at

[45] We discuss columns (3) and (4) of Table 4 in Section 6.4.

Table 4: The causal effect of shipping on local population in the data, 'baseline model,' and 'benchmark models'

	Δln(Population)			
	Data	Model	Benchmark 1	Benchmark 2
Independent variables	(1)	(2)	(3)	(4)
Δln(Shipment)	0.006	0.004	0.016**	0.018***
	0.022	*0.033*	*0.125*	*0.146*
	(0.073)	(0.007)	(0.006)	(0.006)
Observations	531	553	553	553
Specification	2SLS	2SLS	2SLS	2SLS
KP F-stat	9.98	585.24	669.19	647.41

Notes: Column (1) uses depth as IV for shipping, controlling for population in 1950, which is equivalent to using residualized depth as an IV. Columns (2) to (4) use residualized depth as IV, which is the variation that we feed into the model to simulate the counterfactual. Standardized coefficients in italics underneath the baseline coefficients. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

different values of log rents $\ln(R_{i,CF})$ in Figure 1b.[46] As the figure shows, the effect of land rents on shipping is negative, large and statistically significant, as in the data. This provides further evidence that the land price mechanism is present in the model not only in a qualitative sense (as we showed in Section 4.2), but it is a significant driver of where port development takes place.

6.4 The aggregate welfare effects of containerization

We estimate that aggregate world welfare increased by 3.84% as a result of containerization.[47] The welfare gains from containerization stem from a combination of three factors in the model: lower shipping costs, which increase welfare; the increased cost of land use, i.e., the *resource costs* of containerization, which lower the gains; and the gains from increased specialization of cities in port or non-port activities, i.e., the *specialization gains* from containerization.

To assess the quantitative importance of each of these margins, we develop two

[46] Appendix Table B.14 shows the corresponding estimates.

[47] We define the change in aggregate world welfare as the average of changes in country-level welfare between the counterfactual and the 1990 equilibrium, weighted by country population. Within each country, labor mobility equalizes welfare across cities, as in Redding (2016). However, we do not allow for mobility across countries, hence different countries experience different welfare effects.

simple benchmark models that will allow us to isolate the three mechanisms at work. 'Benchmark 1' is closest to a standard model, as it assumes that transshipment costs are *exogenous* and *free* – that is, land is solely used for the production of the city-specific good. Thus, the welfare gains from containerization only stem from shipping cost reductions in this benchmark model. 'Benchmark 2,' on the other hand, requires land to be used to reduce transshipment costs. However, we restrict land use to be identical across port cities (and equal to the mean port share in our baseline).

As Benchmark 2 only differs from Benchmark 1 in land being used for port activities, a comparison between these two models reveals the resource costs of increased land use due to containerization. As our baseline model only differs from Benchmark 2 in the potential specialization of port cities in port or non-port activities (through each city choosing the allocation of land between the two), a comparison between these two models reveals the endogenous specialization gains from containerization.

To implement the decomposition of the aggregate welfare effects, we first take Benchmark 1 and Benchmark 2 to our 1990 data. Next, we conduct the containerization counterfactual in each benchmark model. In particular, we conduct the counterfactual such that the world trade to GDP ratio changes to the same extent ($+4.7$ pp) in each benchmark as in our baseline model. Hence, differences in the welfare effects across the models do not stem from trade changing to a different extent.[48]

We find that containerization leads to welfare gains of 4.12% in Benchmark 1. In Benchmark 2, the gains from containerization reduce to 3.45%. The difference between Benchmark 1 and Benchmark 2, 0.67 percentage points, captures the resource costs of containerization. These costs are sizeable: they eat up as much as 16.3% of the gains from the shipping cost reduction. Finally, the difference between Benchmark 2 and our baseline model, 0.39 percentage points, captures the specialization gains from containerization. Note that these gains are able to offset about 58% of the resource costs of containerization, but they do not fully compensate for all the costs. Based on this exercise, relative to a standard model in which transport cost reductions are exogenous and free, both model mechanisms – the resource cost and the endogenous specialization effect – lead to quantitatively meaningful effects on welfare.

We can also use the benchmark models to provide another test of whether it is indeed our endogenous crowding-out mechanism that leads to the null effect of shipping on population in the model. To this end, we estimate the causal effect of shipping on population in the two benchmarks (columns (3) and (4) of Table 4, respectively). Unlike in our baseline model, shipping leads to a significant increase in city population in both

[48]We provide a detailed description of each benchmark model and their quantitative estimation in the Supplementary Material.

benchmarks. This is intuitive: while better market access draws people into the city in all three models, increased land use in transshipment does not have a differential impact on city population in the benchmarks.[49] The standardized coefficients demonstrate that a one standard deviation increase in shipping translates into substantially larger (0.125 and 0.146 standard deviation) increase in population in the benchmarks than in the baseline model (0.033) or in the data (0.022). This underscores that the crowding out effect is driving the zero local population effect of shipping in the model. It also points to the fact that the crowding-out effect is sizeable – not just in terms of the effect it has at the aggregate level, but also in terms of its local effect.

In Appendix A.3, we show that the aggregate and local effects of containerization implied by the model are robust to different values of the containerization shock and some alternative modeling choices. These alternative specifications include different values of transshipment cost shape parameter β in 1990 or in the counterfactual, different changes in exogenous transshipment costs, and a model in which landlords make profits from the provision of transshipment services.

7 Alternative explanations for the crowding-out effect

Thus far, we have shown both reduced form and structural evidence that the land-price mechanism can account for the crowding out of population that we observe in the data. Of course, there are other explanations that could also explain these findings. In this section, we consider a number of them.

Declining labor-intensity of port technology. Containerization is a labor-saving technology as the standardization of cargo-handling allows for much more extensive automation (Bridgman, 2014). If the job destruction that occurred in port-related activities was sufficiently large, it could account for a substantial part of the population crowding-out identified in our empirics. We assess the magnitude of this channel using readily available historical data on U.S. employment in water transportation (the industry to which longshoremen belong) from the *Bureau of Economic Analysis (BEA)*.[50]

There were approximately 222,000 full-time equivalent employees in water transportation in the U.S., accounting for 0.12% of the population in 1960.[51] Between 1960

[49]In Benchmark 2, land used for transshipment increases equally across port cities. Hence, land used for transshipment does not react endogenously to shipping, leading to no differential impact on the population of cities with different changes in shipping. In Benchmark 1, no land is used for transshipment by assumption.

[50]We use 'Water Transportation' employees from the BEA's 'Full-Time Equivalent Employees by Industry' series (Table 6.5). Accessed February, 2021 at https://apps.bea.gov/iTable/iTable.cfm?reqid=19&step=2&isuri=1&1921=survey.

[51]Source: https://en.wikipedia.org/wiki/1960_United_States_census. Employees in water-

and 1987 (the last year in the series), employment fell by 23%. We conduct a back-of-the envelope calculation to assess the magnitude of population crowding out that this channel could account for. We make the following conservative assumptions: i) we assign *all* U.S. water-transportation workers pre-containerization to our sample of U.S. cities; ii) we assume *all* of these jobs were made redundant by containerization (when in fact, only 23% of jobs disappeared); and iii) we assume *all* workers moved out of the port city. This would result in a reduction of the population in port cities by 0.34% relative to 1960 population levels.[52]

The implied population losses are about an order of magnitude too small relative to i) the causally estimated average crowding-out effect, which is 3.8%,[53] or ii) the average crowding out implied by the quantitative model, which is 2.25%.[54] Both of these numbers are about an order of magnitude larger than the entire size of the water transport sector in 1960. Based on this, we conclude that the declining labor intensity of port technology can at most account for a very small fraction of the crowding-out effect.

Inland transport cost reductions. While we focus on a specific aspect of containerization (the transshipment cost reduction at *seaports*), in reality, containerization had broader effects on transport costs. Most importantly, it arguably also reduced overland transport costs as intermodal transshipping between trucks and railways became cheaper. This raises the concern that the population outflows estimated in port cities are driven by the fact that these overland transport cost reductions made inland cities more attractive, leading to population outflows from port cities that counteract the standard market access effect.

We investigate this channel by using our quantitative model to examine how adding an inland cost increase to our counterfactual simulation (i.e., the rolling back of containerization) changes the results. Rows (10) and (11) of Appendix Table B.15 show that adding this additional shock does not meaningfully alter i) the estimated aggregate resource costs and specialization gains from containerization, ii) the zero local popu-

transportation accounted for 0.34% of all full-time equivalent employees.

[52]The population of the sample of U.S. Geopolis cities was 64,951,000 in 1960.

[53]Shipping increased on average by 24% between 1960 and 1990. Multiplying this by the elasticity of population w.r.t. shipping estimated in Table 3, column (4) (-0.159) yields a 3.8% population loss.

[54]In the quantitative estimation, the size of the non-shipping sector decreases on average by 8.98% across port cities with the introduction of containerization. The log-linearized equation determining equilibrium city population (equation 7) implies a population loss of 8.98%*0.25 = 2.25%, where the elasticity of 0.25 is based on the parameter values used in the quantitative estimation (and reported in Appendix Table B.10).

lation effects, iii) the reallocation of shipping toward low-rent cities.[55] The reason for this is that the overland cost reduction has two opposing effects of roughly similar magnitude. On the one hand, overland transport cost reductions make these routes more attractive relative to sea routes, leading to less endogenous port development. However, they also increase the overall volume of shipping, which increases port development.[56] Put differently, we find no evidence of missing interaction effects between the aspect of containerization we are interested in, and the overland transport cost reductions, which we do not account for in the main analysis of this paper.

Pollution and other negative amenities associated with port development. A different explanation relies on the argument that port development is a 'dirty' activity that leads to population outflows as people flee the disamenities of port activities. While pollution may have increased due to the larger volume of trade, there are two main reasons why this is unlikely to account for the crowding-out effects that we estimate.

First, the public has only recently become aware of the polluting effects of port activities; "until recently, the environmental consequences of port operations were largely unrecognized by the public and ignored by government policymakers. As other sources of pollution have been reduced in the decades-long battle to improve urban air quality, pollution from ports is becoming more obvious, as visible as it is worrisome" (Cannon, 2008, p. 7). In fact, the first studies estimating pollution from marine ships only appeared in the mid-1990s, after our sample period (e.g., Carlton, Danton, Gawen, Lavender, Mathieson, Newell, Reynolds, Webster, Wills, and Wright (1995), Corbett and Fischbeck (1997), Corbett, Fischbeck, and Pandis (1999)).

We also systematically examined historical annual reports from multiple U.S. ports to study whether their accounts are consistent with this.[57] Indeed, while we found evidence of some environmental concerns related to dredging and the (noise) pollution associated with airport operations (but not seaport operations), we found no paper trail for concerns regarding air quality or pollution more generally from port activities. Consistent with the timing of public awareness around the issue, the Port of Houston started reporting and undertaking mitigation related to air quality only in the 2000s.[58]

Second, note that for this explanation to confound our mechanism, it must be the case that containerization is disproportionately more polluting or associated with more

[55] Adding the inland cost reduction obviously changes the overall welfare gains. We provide further details on this robustness exercise in Appendix A.3.

[56] The canceling out of these two forces is also reflected in the fact that mean port size changes are very similar with and without overland cost changes; in both cases they increase by about 8.6 percentage points.

[57] The Supplementary Material contains a detailed list of all ports examined and the sources.

[58] This is the only port for which we had material past our sample period.

disamenities. There is no evidence of this. In fact, we can only find examples of the opposite in the historical annual reports of ports. This may be because the cargo is more exposed and can therefore more easily cause pollution. For example, in the 1970s, the Port of Seattle undertook efforts to mitigate dust emitted by its grain terminal operations, serving as a model for EPA guidelines on dust control.

In summary, while we present multiple pieces of evidence consistent with the increased land-intensity of ports driving the population crowd-out, some of the most promising alternative explanations are either far too small to account for our results, or do not stand up to more rigorous examination.

8 The effects of targeted port development

We use our estimated model to illustrate the effects of targeted port development policy in this section. Our model is particularly well-suited to such an exercise, as it allows the development of both targeted and non-targeted ports to endogenously respond to port development policy, highlighting rich distributional impacts. We study a large-scale port development policy similar to the Chinese government's 'Maritime Silk Road' project, which is part of the 'Belt and Road Initiative.'[59] In particular, we study the effects of a 10% reduction in exogenous transshipment costs (recovered in Section 6.1) in 24 port cities in Asia, Africa and Europe targeted by Chinese investment (see Appendix Figure C.6 for the set of targeted ports).[60]

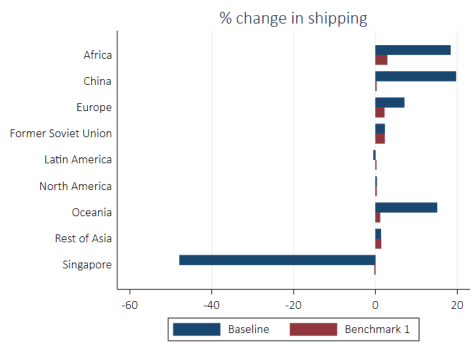
Appendix Table B.16 examines the effects of this policy on treated and untreated port cities, and inland cities. We compare the effects generated by our model ('Baseline') to those of a more standard model ('Benchmark 1' – introduced in Section 6.4). As column (1) demonstrates, targeted port cities see a significant and large increase in shipping activities, primarily at the expense of non-targeted port cities in the same country. This local reallocation of shipping is more pronounced in the baseline model than in Benchmark 1 (column (5)). To see why this is the case, in columns (2) and (6) we examine the effect on port costs (the sum of exogenous and endogenous transshipment costs, $\nu(r) + \psi(F(r))$). In Benchmark 1, endogenous transshipment costs are absent, implying that targeted port cities see an exact 10% (0.105 log point) decline in their transshipment costs, while non-targeted cities see no effect. By contrast, in the baseline model, the direct effect of the policy is amplified by an endogenous reallocation of land within the city. This results in a decline in endogenous transshipment costs in

[59] The simulation we conduct is *similar* to the Maritime Silk Road project, as we analyze effects relative to the 1990 equilibrium, not today. Moreover, the absence of specific details on the size of the actual investments precludes us from matching exactly what the project entails.

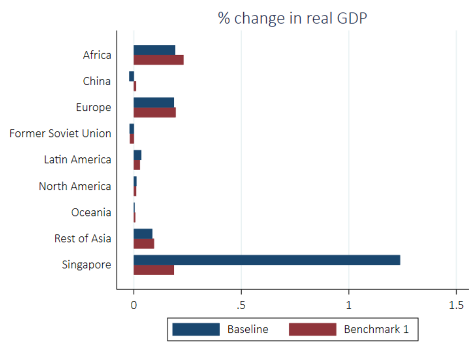
[60] We take the targeted ports from OECD (2018) and choose the decrease in $\nu(r)$ to be 10% to illustrate the effects of a sizeable, but not dramatic decrease in transshipment costs.

targeted ports (where more land is allocated to the port) and an increase in endogenous transshipment costs in non-targeted ports (where less land is allocated to the port). This endogenous port development response to the policy is precisely what draws additional shipping into targeted cities and away from non-targeted ones.[61]

We study the effects on cities' market access (as defined by equation (8)) and population across both models. The effect on market access is similar in both simulations, as we would expect: all cities gain on average, particularly targeted port cities and inland cities located in the same country. In terms of population responses, however, the similar improvement in market access results in strikingly different population responses – highlighting the crowding out mechanism at work in our model. In the baseline (column (4)), endogenous port development in targeted port cities moves people out of the city through increased land use, primarily to non-targeted port cities. In contrast, in Benchmark 1 (column (8)), targeted ports gain population at the expense of non-targeted ones. These findings highlight the importance of accounting for the crowding-out effect when evaluating how targeted port development affects the spatial distribution of population.



(a) Change in shipping, percent



(b) Change in real GDP, percent

Figure 2: Simulated changes across regions, Maritime Silk Road

Notes: Panel A (B) shows the change in total shipping (total real GDP) of each region between the model inversion and the maritime silk road counterfactual. When delineating these regions, we roughly follow the world's continents. An exception is 'Rest of Asia', which is Asia except China, Singapore, and the former Soviet Union. We treat China separately as we are naturally interested in the effects that this Chinese government policy has on China itself. We treat Singapore separately as we find strikingly large effects on this port city, which we discuss in the text.

We examine how targeted port development redistributes shipping and real GDP across regions of the world in Figure 2. We find the most dramatic distributional effects in Asia. Strikingly, we see a dramatic reallocation of shipping to China and away from Singapore (which we estimate loses 50% of its shipping flows).[62] Neither countries

[61] The results remain similar if we include country fixed effects. We present these results in the Supplementary Material.

[62] It must be noted, however, that China's percentage change in shipping does not correspond to an equally large absolute change, as China had relatively little shipping back in 1990.

have targeted ports in this simulation. While these effects are also present in the benchmark model, they are far more muted. In our model, the initial reallocation of shipping is amplified by increasing returns to scale. As shipping moves away from Singapore towards targeted ports, incentives to develop the port of Singapore decrease, which ultimately leads the city to cut back substantially on its port activities by reallocating land away form the port. However, Singapore sees a more than 1% gain in real GDP in our baseline model, as the city's declining port frees up land that can be used profitably outside the shipping sector. This is particularly true in the case of Singapore, where the non-port sector is very productive.[63]

As the example of Singapore illustrates, increasing returns have the potential to substantially amplify changes in shipping and real GDP in our baseline model relative to a standard trade model such as Benchmark 1. This is also evident from the fact that changes in shipping and real GDP vary to a much larger degree across regions in the baseline model than in the benchmark. In Benchmark 1, changes in shipping range from -0.25% to 2.88% across our nine regions; in the baseline model, the corresponding numbers are -47.93% and 19.68%. In Benchmark 1, changes in real GDP range between -0.02% and 0.23%, while they vary between -0.02% and 1.24% in the baseline. These results underscore that modeling endogenous port development is essential to correctly assess the magnitude of the distributional effects of port development policies.

9 Conclusion

The containerization shock studied in this paper allows us to shed light on the economic effects of port development. We have shown that the land-intensive nature of port development is an empirically strong force that has the potential to matter for the local, aggregate and distributional economic effects of port development. Though the analysis in this paper is positive, it offers some tentative implications for where port development is likely to have the biggest beneficial impact. On the one hand, the recent aggressive port development strategy followed by some developing country cities such as Colombo, Sri Lanka seems promising. These are cities where the opportunity cost of land remains relatively low given the low productivity of non-port activities. On the other hand, our findings cast some doubt on the wisdom of continuing to specialize in port services for some of the world's most productive and expensive cities such as Hong-Kong and Singapore. While these cities arguably benefited enormously from

[63] According to our model, Singapore is at the 98th percentile in the world productivity distribution. Of course, the economic benefits from dismantling a port may be not the only factor considered by decision-makers in reality. Governments' objective functions may include geopolitical advantages from maintaining a central position in the global shipping network. In our analysis, we focus on the economic effects and do not consider these additional factors.

their position as important ports historically (at a time when they were also far poorer relative to the rest of the world), subsequent productivity growth *outside* the port sector has made the opportunity cost of the land occupied by the port extremely high. Our findings suggest that the 'Hong-Kongs' and 'Singapores' of the world may benefit from following the path of cities such as London (United Kingdom) – a city at the center of world trade for many decades, but one that now houses Canary Wharf, an important second financial district, on redeveloped land once occupied by the port.

References

Allcott, H. and D. Keniston (2017). Dutch Disease or Agglomeration? The Local Economic Effects of Natural Resource Booms in Modern America. *Review of Economic Studies 85*(2), 695–731.

Allen, T. and C. Arkolakis (2014). Trade and the Topography of the Spatial Economy. *Quarterly Journal of Economics 129*(3), 1085–1140.

Allen, T. and C. Arkolakis (2019). The Welfare Effects of Transportation Infrastructure Improvements.

Allen, T. and D. Atkin (2016). Volatility and the Gains from Trade.

Altomonte, C., I. Colantone, and L. Bonacorsi (2018). Trade and Growth in the Age of Global Value Chains. Technical report, BAFFI CAREFIN Centre Research Paper.

Anderson, J. (1979). A Theoretical Foundation for the Gravity Equation. *American Economic Review 69*(1), 106–116.

Armenter, R., M. Koren, and D. Nagy (2014). Bridges.

Bernhofen, D., Z. El-Sahli, and R. Kneller (2016). Estimating the Effects of the Container Revolution on World Trade. *Journal of International Economics 98*, 36–50.

Bleakley, H. and J. Lin (2012). Portage and Path Dependence. *Quarterly Journal of Economics 127*, 587–644.

Blonigen, B. and W. Wilson (2008). Port Efficiency and Trade Flows. *Review of International Economics 16*(1), 21–36.

Brancaccio, G., M. Kalouptsidi, and T. Papageorgiou (2020). Geography, transportation, and endogenous trade costs. *Econometrica 88*(2), 657–691.

Bridgman, B. (2014). Why Containerization Did Not Reduce Ocean Trade Shipping Costs.

Brooks, L., N. Gendron-Carrier, and G. Rua (2019). The Local Impact of Containerization.

Campante, F. and D. Yanagizawa-Drott (2018). Long-range Growth: Economic Development in the Global Network of Air Links. *Quarterly Journal of Economics 133*(3), 1395–1458.

Cannon, J. S. (2008). Us container ports and air pollution: A perfect storm. Technical report, Energy Futures Inc.

Carlton, J., S. Danton, R. Gawen, K. Lavender, N. Mathieson, A. Newell, G. Reynolds, A. Webster, C. Wills, and A. Wright (1995). Marine Exhaust Emissions Research Programme. *Lloyd's Register Engineering Services, London 63*.

Coşar, A. and P. Fajgelbaum (2016). Internal Geography, International Trade, and Regional Specialization. *American Economic Journal: Microeconomics 1*(8), 24–56.

Conley, T. (1999). Gmm Estimation with Cross Sectional Dependence. *Journal of Econometrics 92*(1), 1–45.

Corbett, J. J. and P. Fischbeck (1997). Emissions from ships. *Science 278*(5339), 823–824.

Corbett, J. J., P. S. Fischbeck, and S. N. Pandis (1999). Global Nitrogen and Sulfur Inventories for Oceangoing Ships. *Journal of Geophysical Research: Atmospheres 104*(D3), 3457–3470.

Corbett, M. (2010). *The History and Transformation of the Port of San Francisco, 1848-2010*. San Francisco Architectural Heritage.

Corden, W. and J. Neary (1982). Booming sector and de-industrialisation in a small open economy. *Economic Journal 92*(368), 825–848.

Coşar, A. and B. Demir (2018). Shipping Inside the Box: Containerization and Trade. *Journal of International Economics 114*, 331–345.

Desmet, K. and J. Rappaport (2017). The Settlement of the United States, 1800–2000: The Long Transition Towards Gibrat's Law. *Journal of Urban Economics 98*, 50–68.

Donaldson, D. and R. Hornbeck (2016). Railroads and American Economic Growth: A "market access" Approach. *Quarterly Journal of Economics 131*(2), 799–858.

Donaldson, D. and A. Storeygard (2016). The View from Above: Applications of Satellite Data in Economics. *Journal of Economic Perspectives 30*(4), 171–98.

Ducruet, C., S. Cuyala, and A. E. Hosni (2018). Maritime Networks as Systems of Cities: The Long-term Interdependencies Between Global Shipping Flows and Urban Development. *Journal of Transport Geography 66*, 340–355.

Duranton, G. and D. Puga (2019). Urban Growth and its Aggregate Implications.

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

Eyre, J. (1964). Shipping Containers in the Americas. In *Pan American Union: Recent Developments in the Use and Handling of Unitized Cargoes*, pp. 38–42.

Fajgelbaum, P. and S. Redding (2018). Trade, Structural Transformation and Development: Evidence from Argentina, 1869-1914.

Fajgelbaum, P. D. and E. Schaal (2020). Optimal transport networks in spatial equilibrium. *Econometrica 88*(4), 1411–1452.

Falvey, R. E. (1976). Transport Costs in the Pure Theory of International Trade. *The Economic Journal 86*(343), 536–550.

Fujita, M. and T. Mori (1996). The Role of Ports in the Making of Major Cities: Self-agglomeration and Hub-Effect. *Journal of Development Economics 49*, 93–120.

Ganapati, S., W. F. Wong, and O. Ziv (2020). Entrepôt: Hubs, Scale, and Trade Costs.

Gilman, S. (1983). *The Competitive Dynamics of Container Shipping*. Gower Publishing Company.

Gomtsyan, D. (2016). Rise of the Machines: Evidence from the Container Revolution.

Haworth, R. B. (2020). Miramar ship index.

Head, K. and T. Mayer (2014). *Gravity Equations: Workhorse, Toolkit, and Cookbook*, Volume 4 of *Handbook of international economics*. Elsevier.

Heiland, I., A. Moxnes, K. H. Ulltveit-Moe, and Y. Zi (2021). Trade From Space: Shipping Networks and The Global Implications of Local Shocks.

Holmes, T. and E. Singer (2018). Indivisibilities in Distribution. *National Bureau of Economic Research* (Working Paper 24525).

Hummels, D. (2007). Transportation Costs and International Trade in the Second Era of Globalization. *Journal of Economic Perspectives 21*(3), 131–154.

Jedwab, R. and A. Storeygard (2021). The Average and Heterogeneous Effects of Transportation Investments: Evidence from Sub-Saharan Africa 1960-2010. *National Bureau of Economic Reserach* (Working Paper 27670).

Kahveci, E. (1999). Fast Turnaround Ships and Their Impact on Crews.

Kennan, J. and J. Walker (2011). The Effect of Expected Income on Individual Migration Decisions. *Econometrica 79*(1), 211–251.

Krugman, P. (1987). The Narrow Moving Band, the Dutch Disease, and the Competitive Consequences of Mrs. Thatcher: Notes on Trade in the Presence of Dynamic Scale Economies. *Journal of Development Economics 27*(1-2), 41–55.

Krugman, P. (2011). Comparative Advantage, Growth, And The Gains From Trade And Globalization: A Festschrift in Honor of Alan V Deardorff. Technical report, Citigroup Foundation Special Lecture.

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

Martín-Antón, M., d. C. J. M. Negro, V., J. S. López-Gutiérrez, and M. D. Esteban (2016). Review of coastal land reclamation situation in the world. *Journal of Coastal Research 75*, 667 – 671.

Maurer, S. and F. Rauch (2021). Economic Geography Aspects of the Panama Canal.

McKinsey & Company (1972). Containerization: A Five-Year Balance Sheet. Technical report, McKinsey & Company.

Melitz, M. J. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica 71*(6), 1695–1725.

Monte, F., S. Redding, and E. Rossi-Hansberg (2018). Commuting, Migration, and Local Employment Elasticities. *American Economic Review 108*(12), 3855–90.

Moriconi-Ebrard, F. (1994). *Geopolis*.

Nagy, D. (2020a). Hinterlands, City Formation and Growth: Evidence from the U.S. Westward Expansion.

Nagy, D. (2020b). Trade and Urbanization: Evidence from Hungary.

OECD (2018). The Belt and Road Initiative in the Global Trade, Investment, and Finance Landscape. Technical report, OECD Publishing Paris.

Oi, W. (1996). The Economics of New Goods. In *The Welfare Implications of Invention*. University of Chicago Press.

Port of San Francisco (1971). San Francisco Port Needs, Shipping and Area Requirements. Technical report, Port of San Francisco.

Rappaport, J. (2007). Moving to Nice Weather. *Regional Science and Urban Economics 37*(3), 375–398.

Redding, S. (2016). Goods Trade, Factor Mobility and Welfare. *Journal of International Economics 101*, 148–167.

Redding, S. and M. Turner (2015). *Transportation Costs and the Spatial Organization of Economic Activity*, Volume 5 of *Handbook of regional and urban economics*, pp. 1339–1398. Elsevier.

Rodrigue, J. (2016). *The Geography of Transport Systems*. Taylor Francis.

Rua, G. (2014). Diffusion of Containerization.

Saiz, A. (2010). The Geographic Determinants of Housing Supply. *Quarterly Journal of Economics 125*(3), 1253–1296.

Santamaría, M. (2020). The Gains from Reshaping Infrastructure: Evidence from the division of Germany.

Stock, J. and M. Yogo (2002). Testing for Weak Instruments in Linear IV Regression. In *Essays in Honor of Thomas Rothenberg*. Cambridge University Press.

UNCTAD (2019). Review of Maritime Transport. Technical report, United Nations Publication.

Wong, W. F. (2020). The Round Trip Effect: Endogenous Transport Costs and International Trade.

Online Appendix

All aboard: The effects of port development

César Ducruet   Réka Juhász   Dávid Krisztián Nagy   Claudia Steinwender

A Further details on taking the model to the data and the counterfactuals

A.1 Taking the model to the data

A.1.1 Calculating shipping costs

In the first step of taking the model to the data, we calculate inland and sea shipping costs and choose a functional form for the endogenous transshipment cost function. To calculate inland and sea shipping costs across cities[1] as a function of distance d, we follow our strategy outlined in Section 5 and assume

$$\phi_\varsigma(d) = e^{t_\varsigma d} \qquad\qquad \phi_\tau(d) = e^{t_\tau d}$$

and set the elasticities t_ς and t_τ to the corresponding estimates in Allen and Arkolakis (2014).[2]

Next, we choose endogenous transshipment costs as a function of the share of land allocated to the port (*port share*, F), $\psi(F)$. The existing literature provides us with little guidance on this, as ours is the first paper that argues for the relevance of this relationship in a quantitative trade and geography framework. Hence, our goal is to keep the functional form of ψ as simple as possible. That said, the functional form needs to satisfy our theoretical restrictions ($\psi \geq 0, \psi' < 0, \psi'' > 0$) and needs to be numerically tractable in the model inversion and counterfactual simulations. In particular, the range of ψ' should ideally span the entire $(-\infty, 0)$ interval over its domain $(0, 1)$, as otherwise it would be potentially impossible to obtain port shares that rationalize the GDP and shipping data in every port city from equations (5) and (6). One simple function that satisfies all these restrictions is the one in equation (10):

$$\psi'(F) = 1 - F^{-\beta}$$

where we restrict $\beta > 0$ to guarantee $\psi' < 0$. We can then obtain ψ by integrating

[1] We have 553 port and 2,083 non-port cities in our data. For details, see Section 2.

[2] See Section 5 for details on these estimates.

equation (10) as

$$\psi(F) = \frac{F^{\beta} + (\beta - 1)^{-1}}{F^{\beta-1}} + \kappa$$

where we restrict $\kappa \geq \bar{\kappa} = -\left[1 + (\beta - 1)^{-1}\right]$ to guarantee $\psi \geq 0$.[3]

A.1.2 Choosing the values of structural parameters

We also need to choose the values of the model's seven structural parameters. On the production side, we take the estimate of the strength of agglomeration externalities, $\alpha = 0.06$, from Ciccone and Hall (1993). This estimate has performed well in the literature for various countries and time periods. $\alpha = 0.06$ implies that doubling city size increases city productivity by 6%. Still on the production side, the expenditure shares on labor and land equal γ and $1 - \gamma$, respectively. Unfortunately, we are not aware of any study that measures the land share for the entire world. Thus, we base our benchmark value of γ on Desmet and Rappaport (2017), who estimate a value of 0.10 for the difference between the land share and the agglomeration elasticity in the United States between 1960 and 2000, a period that corresponds to our sample period. Given we set $\alpha = 0.06$, this suggests choosing $\gamma = 0.84$.[4]

On the consumption side, we have two structural parameters: the migration elasticity, which we set to $\eta = 0.15$ based on Kennan and Walker (2011), and the elasticity of substitution across tradable final goods, which we set to $\sigma = 4$ based on Bernard, Eaton, Jensen, and Kortum (2003).

Finally, there are three structural parameters that influence shipping costs. One is the dispersion of idiosyncratic shipping costs, which – together with the functional form of these costs – we take from Allen and Arkolakis (2019), setting $\theta = 203$. Another is the elasticity of transshipment costs to total shipping at the port (congestion externalities), which we take from the empirical estimates of Abe and Wilson (2009), setting $\lambda = 0.074$. Table B.10 summarizes the calibration of our structural parameters.

The last structural parameter to choose is β from the endogenous transshipment function. Given the role that this parameter plays in driving the relationship between the value of shipping flows and the port share through equation (5), we calibrate it to match the correlation between these two variables in the data.[5] We have been able to

---

[3]As total transshipment costs in city r equal $[\nu(r) + \psi(F(r))]\, Shipping(r)^{\lambda}$, κ is isomorphic to a uniform shifter in exogenous port costs $\nu(r)$ and therefore cannot be identified separately from them. Thus, we set κ to its theoretical lower bound $\bar{\kappa}$ without loss of generality.

[4]Another advantage of using the land share estimate by Desmet and Rappaport (2017) is that it also accounts for the share of land embedded in housing, which is absent from our model but could matter for the quantitative results.

[5]To calculate this correlation, we first transform the number of ships, which is what we directly observe in the data, into the value of shipments, which is what enters equation (5). This

find high-quality, consistent data for the land area occupied by ports for only 7 port cities in 1990, as these data are typically not recorded. We define the port share as the ratio of the land area occupied by the port and the total land area of the city. We find that the correlation between shipping and port share for these seven cities is 0.474 in the data.

In the model, we compute the correlation between shipping and port share in the following way. First, for each port city, we numerically solve equations (5) and (6) for the port share that rationalizes shipping flows, $Shipping\,(r)$, and city GDP, $\gamma^{-1} w\,(r)\,N\,(r)$. As we explain in the Supplementary Material, our theoretical restrictions on ψ' guarantee that this procedure identifies a unique port share $F\,(r) \in (0, 1)$ for each port city. Next, we calculate the correlation between $Shipping\,(r)$ and $F\,(r)$ for our set of port cities.

Under higher values of β, the endogenous port development mechanism plays a stronger role in the model. This is because, under higher β, the endogenous transshipment cost function is more responsive to changes in the port share:

$$\frac{d\,|\psi'\,(F)|}{d\beta} = -F^{-\beta} \log\,(F) > 0$$

Hence, everything else fixed, landlords have an incentive to increase the port share further if β is high. As a consequence, we expect a stronger correlation between shipping and port share under higher values of β. This is precisely what we find. Appendix Figure C.4a plots the values of the correlation for a range of β between 0.020 and 0.046. Within this range, $\beta = 0.031$ is the one that implies the correlation found in the data, 0.474.[6] Hence, we use this value of β in our baseline calibration.[7]

A.1.3 Recovering post-containerization fundamentals

In the final step of taking the model to the data, we use observed data on city populations, shipping flows and city level GDP per capita together with the structure of the model to find the set of city amenities $a\,(r)$, productivities $A\,(r)$ and exogenous transshipment costs $\nu\,(r)$ that rationalize the data.

As city-level GDP data are only available for 1990, we choose to back out the model fundamentals based on the 1990 distribution of population, shipping and GDP.

procedure is described in detail in Section A.1.3.

[6] Instead of calculating the model-implied correlation over the entire set of port cities, we can compute it for the same set of seven port cities where we observe the port share. Reassuringly, for $\beta = 0.031$, this gives us a correlation of 0.463, essentially identical to the one found for the whole set of port cities.

[7] In Section A.3, we investigate robustness of the aggregate gains from containerization to alternative values of β.

Since this year is after the advent of containerization, our counterfactual will involve *rolling back*, or undoing, the containerization shock. Hence, the aggregate effect of containerization can be assessed by comparing the counterfactual equilibrium (pre-containerization) to our 1990 equilibrium (post-containerization).

We transform the number of ships observed in the data in port city r in 1990, $Ship\left(r\right)$, into the value of shipments, $Shipping\left(r\right)$, according to

$$Shipping\left(r\right) = V \cdot Ship\left(r\right)$$

where we choose V to match the ratio of shipping to world GDP. The rationale behind choosing this particular moment is that it can be calculated as a simple linear function of V:

$$\frac{\sum_r Shipping\left(r\right)}{\sum_r GDP\left(r\right)} = V \cdot \frac{\sum_r Ship\left(r\right)}{\sum_r GDP\left(r\right)}$$

where $Ship\left(r\right)$ and $GDP\left(r\right)$ are both observable in the data. This procedure gives us a value of $V = 364$.[8]

Using city-level GDP data, we can obtain wages as

$$w\left(r\right) = \gamma \frac{GDP\left(r\right)}{N\left(r\right)}$$

according to the model, where the structural parameter γ is calibrated to 0.84, as explained in Section A.1.2.

Once population $N\left(r\right)$ and wages $w\left(r\right)$ are available for each city and the value of shipments, $Shipping\left(r\right)$, is available for each port city, the equilibrium conditions of the model can be inverted to back out city amenities up to a country-level scale, $\tilde{a}\left(r\right)$, fundamental city productivities $A\left(r\right)$, and each port city's exogenous transshipment costs $\nu\left(r\right)$. We provide the details of this inversion procedure in the Supplementary Material.[9]

A.2 Counterfactuals

In this section, we describe the two counterfactuals conducted in the paper: the one simulating the roll-back of containerization (Section 6.2) and the one simulating targeted

[8]As not all our port cities have a positive shipping flows in 1990 but the model cannot rationalize zero shipping flows under finite positive values of city-specific fundamentals, we change $Ship\left(r\right)$ from zero to one in these cities.

[9]The complex structure of the model does not allow us to prove that the inversion procedure identifies a unique set of $\tilde{a}\left(r\right), A\left(r\right)$ and $\nu\left(r\right)$. Nonetheless, we have experimented with various different initial guesses, and the inversion algorithm converges to the same fixed point, suggesting that the vector of city-specific fundamentals that rationalize the data is likely unique.

port development (Section 8). The Supplementary Material provides further details on the procedure of simulating the model in these counterfactual scenarios.

A.2.1 Rolling back containerization

One of the major characteristics of pre-containerization transshipment technology was its lower land-intensity (Section 1). In the model, we can capture this lower land intensity by decreasing the shape parameter of transshipment technology, β. As we argued in Section A.1.2, a decrease in β makes the endogenous transshipment cost function less responsive to changes in the port share, $F(r)$. Hence, under lower values of β, port city landlords have less incentive to increase $F(r)$, and port sizes will be generally smaller.

To choose the value of the parameter in the counterfactual, β_{CF}, we use the empirical evidence of Section 1. In particular, we argued in Section 1 that containerized ports occupy on average 75% larger area if we hold the volume of traffic fixed. In our model, this means that the average port share would have increased by 75% if we *held the non-technological determinants of the port share*, i.e., shipping and land rents, fixed:

$$\frac{\sum_{r\in P} F(r)}{\sum_{r\in P} \hat{F}(r)} - 1 = 0.75 \tag{A.1}$$

where P is the set of port cities, $F(r)$ is the port share of port city r in 1990, given by

$$-\left[1 - F(r)^{-\beta}\right] = \frac{R(r)}{Shipping(r)^{1+\lambda}} \tag{A.2}$$

which we obtain by combining equations (5) and (10), and $\hat{F}(r)$ is the port share implied by the *same* rents and shipping but shape parameter β_{CF}:

$$-\left[1 - \hat{F}(r)^{-\beta_{CF}}\right] = \frac{R(r)}{Shipping(r)^{1+\lambda}} \tag{A.3}$$

To back out β_{CF}, we first solve equation (A.2) for $F(r)$. Next, we solve equation (A.3) for $\hat{F}(r)$ for a range of β_{CF}. Finally, we pick the β_{CF} under which equation (A.1) holds. Appendix Figure C.4b plots the increase in mean port share against β_{CF}, highlighting that the increase in mean port share is monotonic in β_{CF} and hence the parameter is identified. The value at which mean port share increases by 75% is $\beta_{CF} = 0.018$.

To capture the fact that depth was not relevant for transshipment prior to containerization, we also offset the relationship between exogenous transshipment costs and depth in the counterfactual. To this end, we run the regression

$$\log \nu(r) = \omega_0 - \omega_1 * Depth(r) + \varepsilon(r)$$

on our sample of port cities, where $\nu(r)$ is the exogenous transshipment cost of city r recovered in Section A.1.3, and $Depth(r)$ is our residualized depth measure, defined in Section 3. In line with the fact that depth lowers transshipment costs after containerization, we find $\widehat{\omega_1} = 0.048$ (se. 0.025, p-value 0.053). We then undo the dependence of exogenous transshipment costs on depth by adding $\widehat{\omega_1} * Depth(r)$ to $\log \nu(r)$.[10]

Finally, we incorporate the overall reduction in transshipment costs due to containerization by increasing $\log \nu(r)$ uniformly across ports. In particular, we add a constant $\nu_{CF} > 0$ to all $\log \nu(r)$ such that the sum of exogenous and endogenous transshipment costs, $\nu(r) + \psi(F(r))$, is on average 25% higher in our counterfactual than in 1990. Naturally, higher values of ν_{CF} yield a larger change in transshipment costs, suggesting that there should be a unique ν_{CF} at which we meet our 25% target. This procedure identifies $\nu_{CF} = 0.280$.

A.2.2 Targeted port development

In the counterfactual exercise aimed at estimating the effects of targeted port development, we decrease exogenous transshipment costs $\nu(r)$ in the 24 port cities targeted by the Maritime Silk Road project by 10% relative to their exogenous transshipment costs recovered in Section 6.1. We keep all other fundamentals of the model fixed at their levels recovered in Section 6.1.

A.3 Aggregate and local effects of containerization: robustness

In Table B.15, we examine the sensitivity of the model-implied aggregate and local effects of containerization to different values of the containerization shock and some alternative modeling choices. We focus on the sensitivity of our five headline findings: the aggregate welfare gains from containerization; the aggregate resource costs; the aggregate specialization gains; the local population effects of shipping; and the reallocation of shipping toward low-rent cities. Row (1) of Table B.15 repeats these five results in our baseline model calibration, while rows (2) to (11) report them for each of our ten robustness exercises.

In the exercises of rows (2) and (3), we use higher and lower values of our transshipment cost parameter β, respectively. We showed in Section A.1.2 that our endogenous port development mechanism is stronger under higher values of β. Thus, it is not surprising that we obtain higher resource costs and specialization gains from container-

[10] To avoid outliers influencing the results of this step, we trim the values of $\nu(r)$ at 0.01 before estimating ω_1. The inversion algorithm assigns very small ν's to some cities. Due to lack of machine precision in the inversion algorithm for very small values of ν's, very small differences in ν's may be exaggerated greatly when taking logs of ν's in the estimation. This affects 19 port cities, for which we identify $\nu(r)$ below 0.01. As these 19 port cities are deeper than average, we obtain a slightly higher regression coefficient, $\widehat{\omega_1} = 0.056$, without the trimming.

ization, net population effects closer to zero as a result of more crowding-out, and more reallocation of shipping toward low-rent cities in row (2). The opposite is true under the lower β of row (3).

In rows (4) and (5), we use alternative values of our counterfactual β: one that implies a smaller (65%) increase in the mean port share, and one that implies a larger (85%) increase. As expected, a smaller increase in land use leads to slightly higher welfare gains, lower resource costs and lower specialization gains from containerization.

In row (6), we do not offset the relationship between exogenous transshipment costs and port depth in the counterfactual. As depth no longer plays a role in the model in this case, we cannot estimate the local population effects and the reallocation toward low-rent cities, as estimating these reduced-form coefficients relies on depth as an instrument for shipping changes. Nevertheless, we find aggregate welfare gains, resource costs and specialization gains that are fairly close to our baseline estimates.

In rows (7) and (8), we choose ν_{CF} to target different (30% and 20%, respectively) changes in the sum of exogenous and endogenous transshipment costs. Unsurprisingly, a larger change in total transshipment costs is associated with higher aggregate gains, resource costs and specialization gains. The opposite is true if we assume that total transshipment costs changed less. The estimated local effects of containerization hardly change, however.

To study how the assumption of perfect competition in transshipment influences our results, we develop a model in which the provision of transshipment services is subject to monopolistic competition. The key difference relative to our baseline setup is that transshipment activity involves positive profits in this monopolistic competition model.[11] Row (9) reports the aggregate and local effects of containerization in the monopolistic competition model. While the results obviously change to some extent, they remain close to our baseline model, both qualitatively and quantitatively.

Finally, rows (10) and (11) add a uniform 10% and 20% change in the elasticity of inland shipping costs to distance, respectively. This amounts to making inland shipping costs higher in the counterfactual, mimicking a decline in inland shipping costs brought about by containerization besides the change in transshipment costs. Obviously, adding an inland shipping cost reduction increases the estimated aggregate gains. However, the resource costs and specialization gains from containerization, as well as the local effects, remain very similar to our baseline model. Overall, we conclude that the estimated aggregate and local effects of containerization are fairly stable across these different model specifications.

[11] We describe the monopolistic competition model in the Supplementary Material.

B Tables

Table B.1: Relationship between containerization and port area

	Ln(Port area, km2)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ln(Container traffic, TEUs)	0.288***	0.127***	0.133***	0.151***	0.153***	0.144***	
	(0.049)	(0.045)	(0.044)	(0.047)	(0.046)	(0.053)	
Ln(Total merchandise traffic, tons)		0.375***	0.283*	0.311***	0.247	0.356***	0.506***
		(0.080)	(0.166)	(0.080)	(0.161)	(0.118)	(0.069)
Ln(Non-bulk traffic, tons)			0.014		0.008		
			(0.099)		(0.096)		
Ln(Country GDP/capita)				0.311***	0.292**		
				(0.108)	(0.134)		
Container traffic share							0.562***
							(0.209)
Observations	123	123	73	122	73	123	123
R-squared	0.287	0.395	0.327	0.431	0.352	0.672	0.398
% change							0.75
Country FEs	×	×	×	×	×	✓	×

Notes: Container traffic share is defined as (container traffic in TEUs * 12 tons per TEU)/total merchandise traffic in tons. Non-bulk traffic is all traffic net of liquid and solid bulk. Container traffic and total merchandise traffic are averaged across 2008 and 2009 in order to maximize the sample size. Country level GDP per capita and non-bulk traffic are for 2009. Data sources: *Google Earth* and *Le Journal de la Marine Marchande*. See the Supplementary Material for further details. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B.2: Summary statistics

	Observations	Mean	Standard Deviation
Shipment (annualized)	2,765	2,913	7,051
Population (in ’000s): *All Cities*	12,698	386	1,086
Population (in ’000s): *Port Cities*	2,735	724	1,886
Depth	553	2.19	1.49
Saiz Land Rent Proxy	553	0.44	0.19

Notes: Shipment reports the annualized flow of shipments across all port city-year pairs (in levels). Population refers to the level of the population of each city-year pair in thousands. Depth and the Saiz Land Rent Proxies are time invariant measures and are defined in Sections 2 and 3, respectively.

Table B.3: Relationship between dredging and measured depth

	Dredging		
Independent variables	(1)	(2)	(3)
Depth	-0.058**	-0.042*	-0.028
	(0.025)	(0.024)	(0.028)
Observations	100	100	100
R-squared	0.059	0.138	0.250
FE	none	continent	coastline

Notes: This table tests the extent to which the baseline measure of depth captures naturally endowed depth (as opposed to depth attained by dredging). Dredging is a binary indicator that takes the value of one if nautical maps from *marinetraffic.com* show the presence of a dredged channel. Depth is the baseline measure of port suitability used in the paper. The sample consists of 100 randomly selected ports from the baseline sample. See the Supplementary Material for more details on data construction and sources. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B.4: Balancing checks for candidate depth measures

Independent variables	(1) ln(Shipping flows 1950)	(2) ln(Population 1950)	(3) Δln(Shipping flows)	(4) Δln (Population)	(5) ln(GDP pc country)	(6) Latitude	(7) Longitude	(8) Saiz land rent proxy
Depth	-0.2308**	-0.1953***	-0.0351	0.0135*	-0.0215	-0.4541	1.7585	0.0625***
	(0.0955)	(0.0389)	(0.0606)	(0.0076)	(0.0301)	(0.7176)	(2.1507)	(0.0051)
Residualized depth	-0.0416		-0.0507	-0.0003	0.0065	0.3900	2.4352	0.0587***
	(0.0977)		(0.0636)	(0.0082)	(0.0308)	(0.7160)	(2.1867)	(0.0053)
Observations	553	553	553	532	472	553	553	553

Notes: This table regresses observables pre-containerization on two candidate measures of port suitability: ‘Depth’ and ‘Residualized depth’. The former is defined as the log of 1 + the sum of cells deeper than 30 feet within 3-5 kilometers of the port. The latter takes the depth measure and residualizes it on the log of population in 1950. Δln(Shipping flows) and Δln(Population) are the growth rates between 1950 and 1960 for the respective variables. GDP per capita, measured at the country level in 1960, is from the *Penn World Tables*. Latitude and Longitude are observed at the city level. The ‘Saiz measure’ is the land-rent proxy defined in Saiz (2010). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B.5: Relationship between coastal land reclamation and Saiz measure

	Coastal land reclamation (indicator)					
Independent variables	(1)	(2)	(3)	(4)	(5)	(6)
Saiz	0.1296*	0.1356**	0.1146			
	(0.0686)	(0.0678)	(0.0754)			
Depth				0.0008	0.0038	-0.0003
				(0.0093)	(0.0096)	(0.0106)
Observations	553	553	553	553	553	553
R-squared	0.00534	0.08521	0.13287	0.00001	0.07991	0.12925
FE	none	continent	coastline	none	continent	coastline

Notes: Dependent variable is equal to one in case coastal land reclamation was reported and zero otherwise. See the Supplementary Material for more details on data construction and sources. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B.6: Robustness to data choices

	Panel A: Depth predicts shipping flows					
	ln(Shipment)					
	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline	Shipment +1	IHST	Port Cities	Depth=0	100K
Depth × post 1970	0.247***	0.144***	0.164***	0.218***	0.247***	0.285***
	(0.059)	(0.029)	(0.034)	(0.060)	(0.059)	(0.071)
Observations	2765	2765	2765	2640	2765	1565
R-squared	0.126	0.156	0.155	0.133	0.126	0.139
Number of cities	553	553	553	528	553	313
	Panel B: The local causal effect of shipping on population					
	ln(Population)					
	(1)	(2)	(3)	(4)	(5)	(6)
Independent variables	Baseline	Shipment +1	IHST	Port Cities	Depth=0	100K
ln(Shipment)	0.015	0.027	0.024	0.025	0.015	0.045
	(0.049)	(0.086)	(0.076)	(0.053)	(0.049)	(0.052)
Observations	2734	2734	2734	2609	2734	1563
R-squared	0.717	0.719	0.719	0.720	0.717	0.606
Number of cities	552	552	552	527	552	313
	Panel C: Containerization increased shipping more in low rent cities					
	ln(Shipment)					
	(1)	(2)	(3)	(4)	(5)	(6)
Independent variables	Baseline	Shipment +1	IHST	Port Cities	Depth=0	100k
Depth × post 1970	0.566***	0.318***	0.368***	0.583***	0.566***	0.348**
	(0.152)	(0.079)	(0.090)	(0.147)	(0.152)	(0.177)
Depth × Saiz × post 1970	-0.707**	-0.408**	-0.472***	-0.779**	-0.707**	-0.203
	(0.323)	(0.159)	(0.183)	(0.315)	(0.323)	(0.376)
Saiz × post 1970	0.975	0.740**	0.814*	0.950	0.975	0.694
	(0.804)	(0.376)	(0.436)	(0.802)	(0.804)	(0.963)
Observations	2765	2765	2765	2640	2765	1565
R-squared	0.129	0.161	0.159	0.137	0.129	0.139
Number of cities	553	553	553	528	553	313

Notes: 'Baseline' reports the baseline specification for comparability. Columns (2)-(3) examine robustness to different ways of dealing with zero shipping flows. Column (2) uses $\ln(Shipment + 1)$ as dependent variable – that is, we take the raw shipping variable and replace the zeros with ones and then take the natural logarithm. Column (3) uses the inverse hyperbolic sine transformation (IHST) for shipment. Different to the baseline, neither of these transformations annualizes the data. Columns (4) - (5) examine robustness to different ways of dealing with 'inland cities'. Column (4) drops them, reducing the sample size. Column (5) assigns depth equal to zero for these cities. Column (6) uses the subset of cities that already attained 100,000 inhabitants in 1950 to examine the effect of sample selection bias. 'Depth' indicates the port suitability measure interacted with indicators for decades including and after 1970. Standard errors clustered at the city level. *** p<0.01, ** p<0.05, * p<0.1.

Table B.7: The local causal effect of shipping on population – robustness

	ln(Population)			
Independent variables	(1)	(2)	(3)	(4)
ln(Shipment)	0.015	-0.071	0.018	-0.015
	(0.049)	(0.060)	(0.051)	(0.051)
Observations	2734	2734	2734	2338
R-squared	0.717	0.759	0.717	0.756
Number of cities	552	552	552	471
Year FE	✓	✓	✓	✓
City FE	✓	✓	✓	✓
Population 1950 × Year	✓	✓	✓	✓
Coastline × Year FE	×	✓	×	×
Saiz × Year	×	×	✓	×
GDP pc (country) × Year	×	×	×	✓
Specification	2SLS	2SLS	2SLS	2SLS
KP F-stat	21.13	13.71	16.26	19.48

Notes: All specifications are 2SLS, using the depth measure as an instrument for shipping (interacted with a dummy for decades including and after 1970). Standard errors clustered at the city level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table B.8: Containerization increased shipping more in low rent cities

	ln(Shipment)			
Independent variables	(1)	(2)	(3)	(4)
Depth × post 1970	0.464***	0.566***	0.437***	0.497***
	(0.138)	(0.152)	(0.142)	(0.166)
Depth × Saiz × post 1970	-0.408*	-0.707**	-0.431	-0.586*
	(0.220)	(0.323)	(0.308)	(0.331)
	{0.153}	{0.237}		
Saiz × post 1970		0.975	-0.052	1.176
		(0.804)	(0.811)	(0.749)
Observations	2765	2765	2765	2360
R-squared	0.128	0.129	0.250	0.143
Number of cities	553	553	553	472
Year FE	✓	✓	✓	✓
City FE	✓	✓	✓	✓
Population 1950 × Year	✓	✓	✓	✓
Coastline × Year FE	×	×	✓	×
GDP pc (country) × Year	×	×	×	✓

Notes: 'Depth' indicates the port suitability measure. 'Saiz' is the Saiz land rent proxy defined in Saiz (2010). Each measure is interacted with an indicator for decades including and after 1970. Standard errors clustered at the city level in parentheses. Conley standard errors to adjust for spatial correlation in curly brackets. *** p<0.01, ** p<0.05, * p<0.1 (significance refers to clustered standard errors).

Table B.9: Within cities, ports moved further towards the outskirts of the city

	Distance to city centroid (km)					
	1953	N	2017	N	Difference	t-statistics
All ports	5.98	494	7.02	511	1.05	2.08
Movers	7.43	18	16.35	18	8.92	4.38

Notes: This table reports the average distance of the port from the centroid of the city (in km) for each city in our sample in 1953 and 2017. 'All ports' reports the average for the entire sample of Geopolis port cities for which data are available. 'Movers' refers to the subsample of cities which moved the location of the port completely by setting up a new port. Additional details on data construction are discussed in the Supplementary Material.

Table B.10: Calibration of structural parameters

Parameter	Target
$\alpha = 0.06$	Agglomeration externalities (Ciccone and Hall, 1993)
$\gamma = 0.84$	Non-land share in production (Desmet and Rappaport, 2017)
$\eta = 0.15$	Migration elasticity (Kennan and Walker, 2011)
$\sigma = 4$	Elasticity of substitution across tradables (Bernard et al., 2003)
$\theta = 203$	Idiosyncratic shipping cost dispersion (Allen and Arkolakis, 2019)
$\lambda = 0.074$	Congestion externalities in ports (Abe and Wilson, 2009)

Table B.11: Prediction of population based on the number of frost free days

	ln(Population)								
			Africa	North America	Latin America	Asia	Europe	Oceania	USSR
Independent variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Frost Free Days × 1960	0.0007***	0.0003*	-0.0005	0.0015***	0.0008***	-0.0001	-0.0001	0.0014	0.0001
	(0.0001)	(0.0002)	(0.0013)	(0.0003)	(0.0002)	(0.0002)	(0.0001)	(0.0012)	(0.0007)
Frost Free Days × 1970	0.0017***	0.0006***	-0.0006	0.0026***	0.0013***	0.0008***	0.0004**	0.0023	0.0012
	(0.0001)	(0.0003)	(0.0021)	(0.0005)	(0.0003)	(0.0002)	(0.0002)	(0.0020)	(0.0010)
Frost Free Days × 1980	0.0028***	0.0012***	-0.0006	0.0039***	0.0017***	0.0005*	0.0011***	0.0036	0.0010
	(0.0001)	(0.0003)	(0.0020)	(0.0006)	(0.0004)	(0.0003)	(0.0003)	(0.0027)	(0.0011)
Frost Free Days × 1990	0.0039***	0.0013***	-0.0009	0.0047***	0.0015***	0.0014***	0.0013***	0.0049	0.0011
	(0.0002)	(0.0004)	(0.0026)	(0.0007)	(0.0005)	(0.0003)	(0.0003)	(0.0031)	(0.0012)
Observations	12368	12368	1184	987	1532	4784	2410	104	1367
R-squared	0.729	0.839	0.798	0.686	0.852	0.756	0.556	0.647	0.762
Number of cities	2568	2568	245	198	308	1038	482	21	276
City FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Country year FE	×	✓	×	×	×	×	×	×	×
Year FE	✓	×	✓	✓	✓	✓	✓	✓	✓

Notes: Column (2), which includes country-year fixed effects, is our preferred specification used for predicting population. Column (1) only controls for year fixed effects. The time pattern in the effect of frost free days on population is similar; however, country-year fixed effects take out some of the explanatory power of temperature and are therefore a more conservative measure. Columns (3) to (9) estimate column (1) for countries in different regions of the world. Reassuringly, the number of frost free days has the strongest effect in North America, where air conditioning is arguably most prevalent, has medium effects in Latin America, Asia, and Europe, and has no detectable effects in Africa, Oceania and USSR, where air conditioning was probably less wide-spread. Frost Free Days indicates the average number of frost free days per year in the city between 1961-1990. All regressions include city fixed effects. Standard errors clustered at the city level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table B.12: Model-inspired specification: fully flexible first stages

Independent variables	(1) ln(Shipment)	(2) ln(Market Access)
Depth × 1960	-0.038	0.003
	(0.064)	(0.002)
Depth × 1970	0.251***	0.007***
	(0.069)	(0.002)
Depth × 1980	0.226***	0.009***
	(0.079)	(0.002)
Depth × 1990	0.290***	0.009***
	(0.085)	(0.002)
Market Access IV	7.169	1.929***
	(5.413)	(0.140)
Observations	2696	2696
R-squared	0.125	0.728
Number of cities	544	544
Year FE	✓	✓
City FE	✓	✓
Population 1950 × Year	✓	✓

Notes: 'Depth' indicates the port suitability measure. It is interacted with dummy variables for decades in order to examine the time path of when depth started to matter for $\ln(Shipment)$ and $\ln(MarketAccess)$. '$\ln(MarketAccess)$' is the empirical counterpart of the market access term defined in Section 5. 'Market access IV' is the instrument for the market access term defined in Section 5. Standard errors clustered at the city level. *** p<0.01, ** p<0.05, * p<0.1.

Table B.13: Model-inspired specification – robustness

	ln(Population)							
Independent variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln(Shipment)	-0.159**	-0.158**	-0.156**	-0.153**	-0.147**	-0.084**	-0.080	-0.164**
	(0.065)	(0.066)	(0.067)	(0.068)	(0.072)	(0.041)	(0.058)	(0.074)
ln(Market Access)	7.103***	7.078***	6.982***	6.874***	6.613***	0.588	7.111***	5.692***
	(0.795)	(0.806)	(0.844)	(0.899)	(1.043)	(2.918)	(0.713)	(1.354)
Observations	2696	2696	2696	2696	2696	2696	2696	2303
R-squared	0.417	0.419	0.429	0.440	0.467	0.755	0.507	0.544
Number of cities	544	544	544	544	544	544	544	464
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
City FE	✓	✓	✓	✓	✓	✓	✓	✓
Population 1950 $\times$ Year	✓	✓	✓	✓	✓	✓	✓	✓
Coastline $\times$ Year FE	×	×	×	×	×	✓	×	×
Saiz $\times$ Year	×	×	×	×	×	×	✓	×
GDP pc (country) 1960 $\times$ Year	×	×	×	×	×	×	×	✓
Specification	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Drop Cities in Market Access IV	none	$\leq$ 100	$\leq$ 200	$\leq$ 300	$\leq$ 500	none	none	none
KP F-stat	9.63	9.51	9.16	8.78	7.75	4.02	8.64	8.43

Notes: This table reports the same specification as Table 3, column (4). All specifications are 2SLS. The instruments used are the depth measure and the market access IV as defined in Section 5. Nearby cities are dropped from the market access IV in columns (2)-(5). 'Drop Cities in Market Access IV' defines the point-to-point distance buffer (in km) for the set of cities to be dropped. Columns (6) - (8) examine robustness of the results to the inclusion of the same set of controls we added in Section 3. Standard errors clustered at the city level. *** p<0.01, ** p<0.05, * p<0.1.

Table B.14: Heterogeneity of effect of land rents on shipping – model-simulated data

	(1)
Independent Variables	Δ ln(Shipment)
Depth	0.508***
	(0.087)
ln(Rent$_{CF}$) $\times$ Depth	-0.022**
	(0.010)
ln(Rent$_{CF}$)	0.014
	(0.013)
Observations	553
R-squared	0.508

Notes: 'Depth' (residualized) indicates the port suitability measure. $\ln(Rent_{CF})$ refers to the logarithm of the counterfactual (pre-containerization) land rents. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B.15: The aggregate welfare effects of containerization – sensitivity analysis

Model	Welfare effect (%)	Resource cost (pp)	Specialization gains (pp)	Local population effects	Rent heterogeneity
1. Baseline	3.84	-0.67	0.39	0.033 (0.052)	-0.073** (0.032)
2. 20% higher β in inversion	3.86	-0.82	0.46	0.017 (0.053)	-0.076** (0.032)
3. 20% lower β in inversion	3.84	-0.53	0.33	0.049 (0.051)	-0.068** (0.032)
4. Counterfactual β implying 65% increase in port share	3.88	-0.65	0.38	0.025 (0.052)	-0.074** (0.032)
5. Counterfactual β implying 85% increase in port share	3.81	-0.70	0.42	0.040 (0.052)	-0.072** (0.032)
6. No depth-dependent change in $\nu(r)$	4.20	-0.75	0.43	n/a	n/a
7. Larger ν_{CF}: implies 30% change in total transshipment costs	4.52	-0.71	0.42	0.034 (0.055)	-0.068** (0.033)
8. Smaller ν_{CF}: implies 20% change in total transshipment costs	3.13	-0.63	0.38	0.029 (0.049)	-0.078** (0.031)
9. Monopolistic competition	4.20	-0.72	0.46	0.049 (0.058)	-0.076** (0.034)
10. Additional 10% inland cost transport reduction	5.38	-0.68	0.40	0.056 (0.051)	-0.081*** (0.030)
11. Additional 20% inland cost transport reduction	6.86	-0.69	0.40	0.076 (0.052)	-0.085*** (0.030)

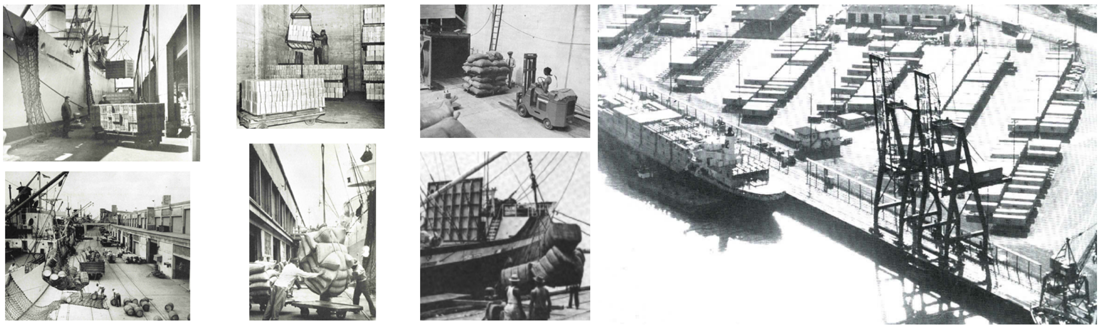
Notes: Welfare effect refers to the gain in welfare due to containerization in our baseline model. Resource cost refers to the difference in welfare gains between Benchmark 1 (with exogenous and free transshipment cost reductions) and Benchmark 2 (with identical land use across port cities). Specialization gains refer to the difference in welfare gains between the baseline and Benchmark 2. Local population effects report the (standardized) causal effect of shipping on local population, using residualized depth as an IV, as in column (2) of Table 4. Rent heterogeneity reports the coefficient on the interaction of rent and depth on changes in shipping, as in Table B.14. The latter two coefficients cannot be estimated in model 6, as depth is not used in that counterfactual. *** p<0.01, ** p<0.05, * p<0.1.

Table B.16: The effects of targeted port development: The Maritime Silkroad

	Baseline				Benchmark 1			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Δ ln(Shipment)	Δ ln(Port cost)	Δ ln(Market access)	Δ ln(Population)	Δ ln(Shipment)	Δ ln(Port cost)	Δ ln(Market access)	Δ ln(Population)
Treated port city	0.78638***	-0.14973***	0.02945***	-0.01393***	0.60825***	-0.10536***	0.02733***	0.01150***
	(0.12000)	(0.03147)	(0.00556)	(0.00524)	(0.09082)	(0.00000)	(0.00581)	(0.00249)
Untreated port city in treated country	-0.41491***	0.01510**	0.01340***	0.00621*	-0.29875***	0.00000	0.01519***	-0.00602***
	(0.08221)	(0.00705)	(0.00286)	(0.00370)	(0.05001)	(0.00000)	(0.00283)	(0.00216)
Port city in untreated country	0.01042	-0.00204	0.00076**	-0.00116	0.00291***	0.00000	0.00084***	0.00023***
	(0.01108)	(0.00796)	(0.00031)	(0.00086)	(0.00034)	(0.00000)	(0.00004)	(0.00003)
Inland city in treated country			0.02066***	0.00189***			0.02015***	-0.00050**
			(0.00123)	(0.00061)			(0.00122)	(0.00021)
Inland city in untreated country			0.00041***	0.00003			0.00041***	-0.00004***
			(0.00006)	(0.00003)			(0.00005)	(0.00001)
Observations	553	544	2,636	2,636	553	553	2,636	2,636
R-squared	0.19261	0.02911	0.44286	0.01915	0.43060	1.00000	0.46368	0.12880

Notes: The regressors are dummy variables that divide the cities into 5 mutually exclusive groups as indicated, the regression is estimated without the constant. Treated port indicate the 24 treated ports of the Maritime Silkroad counterfactual. Treated country are countries that have at least one treated port. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

C Figures



Break-bulk shipping, 1950s Container shipping, 1967

Figure C.1: Illustration of changes in port technology

Notes: Sources: Annual reports for the Port of Seattle and the Port of New Orleans (1950, 1951, 1952, 1955).

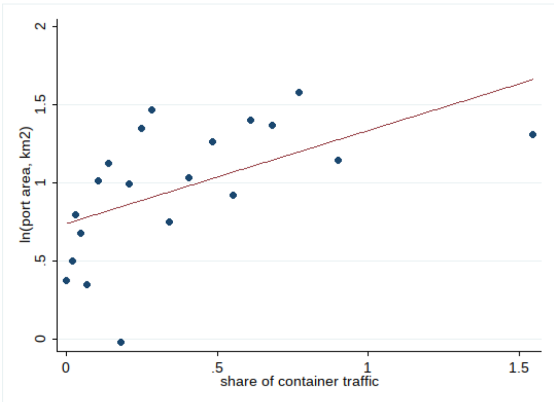


Figure C.2: Relationship between port area and share of container traffic

Notes: The figure shows the correlation between the area of ports and the share of container traffic at the port; the latter is defined as (container traffic in TEUs * 12 tons per TEU)/total merchandise traffic in tons.

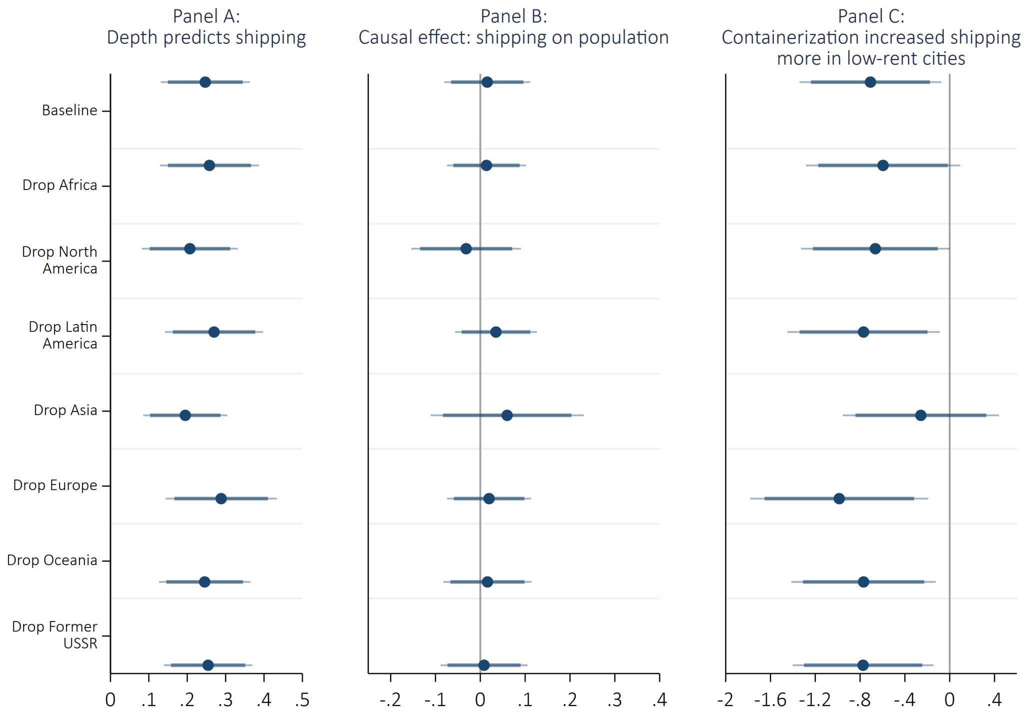
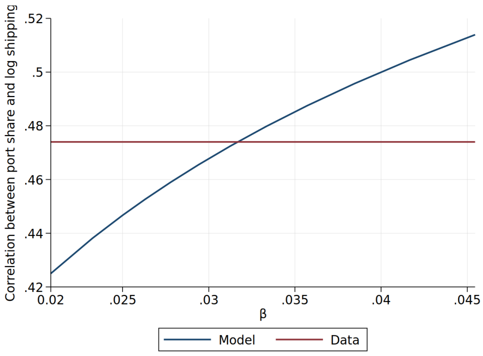
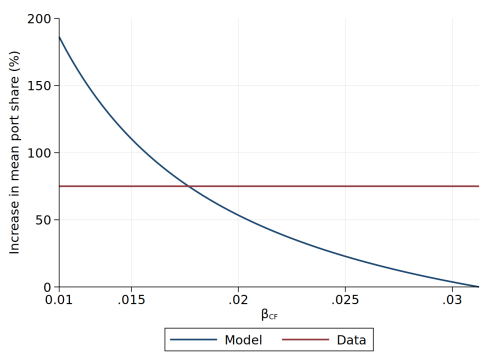


Figure C.3: Dropping continents one at a time

Notes: The plotted coefficients for Panel A are based on the specification in Table 1, column (5). The plotted coefficients for Panel B are based on the specification in Table 2, column (2). The plotted coefficients for Panel C are based on the specification in Appendix Table B.8, column (2). 'Baseline' uses the full sample, while the remaining rows drop continents one at a time as labelled.



(a) Correlation between port share and shipping as a function of β



(b) Increase in mean port share as a function of β_{CF}

Figure C.4: Calibration of transshipment cost parameter β in 1990 (left) and in the no-containerization counterfactual (right)

Notes: The left panel shows the correlation between the port share and log shipping flows in the model as a function of the transshipment cost parameter β (blue line). It also shows the value of this correlation based on 7 ports in the data (red line). The right panel shows the increase in mean port share in the model, keeping the non-technological determinants of the port share fixed, as a function of β_{CF} (blue line). It also shows the increase in port share induced by containerization in the data (red line).

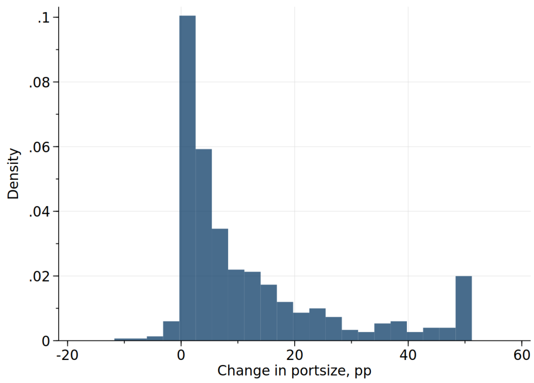


Figure C.5: Histogram of changes in port share between the counterfactual and the 1990 equilibrium, in percentage points

Notes: The figure shows the histogram of the percentage point change in port shares between the model-simulated counterfactual (pre-containerization) and the 1990 equilibrium (after containerization, also model-simulated data).

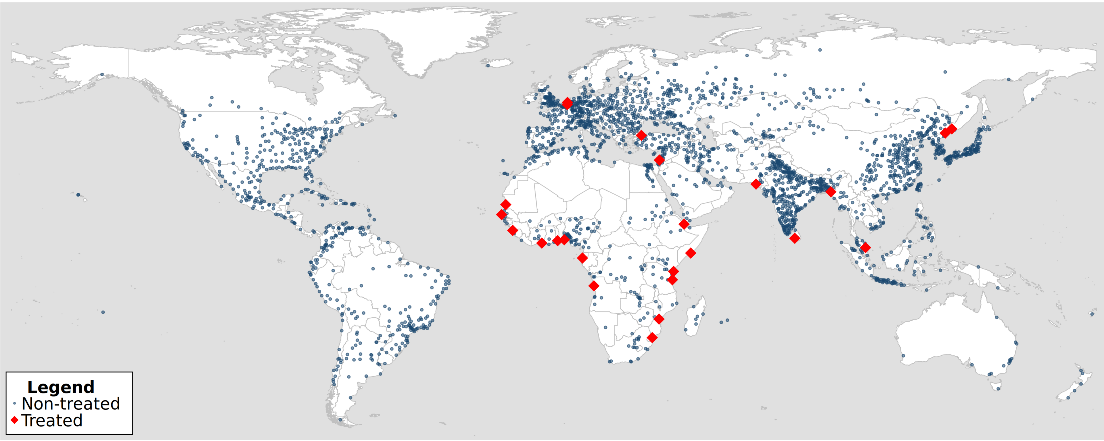


Figure C.6: Maritime Silk Road: targeted ports

Notes: The targeted cities are: Abidjan (Côte d'Ivoire), Antwerpen (Belgium), Beira (Mozambique), Chittagong (Bangladesh), Chongjin (North Korea), Colombo (Sri Lanka), Conakry (Guinea), Dakar (Senegal), Dar Es Salaam (Tanzania), Djibouti (Djibouti), Haifa (Israel), Istanbul (Turkey), Karachi (Pakistan), Kuantan (Malaysia), Lagos (Nigeria), Libreville (Gabon), Lome (Togo), Maputo (Mozambique), Mogadisho (Somalia), Mombasa (Kenya), Nouakchott (Mauritania), Rotterdam (Netherlands) and Vladivostok (Russia). Source: OECD (2018).

References

Abe, K. and J. Wilson (2009). *Weathering the Storm: Investing in Port Infrastructure to Lower Trade Costs in East Asia*. World Bank.

Allen, T. and C. Arkolakis (2014). Trade and the Topography of the Spatial Economy. *Quarterly Journal of Economics 129*(3), 1085–1140.

Allen, T. and C. Arkolakis (2019). The Welfare Effects of Transportation Infrastructure Improvements.

Australian Competition and Consumer Commission (2017). Container Stevedoring and Monitoring Report 2016-17. Technical report, ACCC.

Baragwanath, K., R. Goldblatt, G. Hanson, and A. Khandelwal (2019). Detecting Urban Markets with Satellite Imagery: An Application to India. *Journal of Urban Economics* (103173).

Bernard, A., J. Eaton, J. Jensen, and S. Kortum (2003). Plants and Productivity in International Trade. *American Economic Review 93*(4), 1268–1290.

Blonigen, B. and W. Wilson (2008). Port Efficiency and Trade Flows. *Review of International Economics 16*(1), 21–36.

Ciccone, A. and R. Hall (1993). Productivity and the Density of Economic Activity. *National Bureau of Economic Research* (Working Paper 4313).

Desmet, K. and J. Rappaport (2017). The Settlement of the United States, 1800–2000: The Long Transition Towards Gibrat's Law. *Journal of Urban Economics 98*, 50–68.

Donaldson, D. and A. Storeygard (2016). The View from Above: Applications of Satellite Data in Economics. *Journal of Economic Perspectives 30*(4), 171–98.

Haworth, R. B. (2020). Miramar ship index.

Kennan, J. and J. Walker (2011). The Effect of Expected Income on Individual Migration Decisions. *Econometrica 79*(1), 211–251.

Kirchner, M. (2006). Container Vessels and Risk Aggregation: The Cargo Underwriter's View. Technical report, AXA Corporate Solutions.

Lloyd's of London Press (1990). Lloyd's Ports of the World.

Martín-Antón, M., d. C. J. M. Negro, V., J. S. López-Gutiérrez, and M. D. Esteban (2016). Review of coastal land reclamation situation in the world. *Journal of Coastal Research 75*, 667 – 671.

OECD (2018). The Belt and Road Initiative in the Global Trade, Investment, and Finance Landscape. Technical report, OECD Publishing Paris.

Port Authority of Seattle (1989). Seaport Properties Book. Technical report, Port Authority of Seattle.

Port of Long Beach (1971). Harbor Highlights 1971 Report. Technical report, Port Authority of Long Beach.

Port of New Orleans (1984). 88th Annual Report. Technical report, Port of New Orleans Board of Commissioners.

Riffenburgh, R. (2012). A Project History of the Port of Long Beach 1970 to 2010. Technical report, Port of Long Beach.

Saiz, A. (2010). The Geographic Determinants of Housing Supply. *Quarterly Journal of Economics 125*(3), 1253–1296.